Investigations into the Role of the Medial Temporal Lobes in Autism

Claire Helen Salmond
Institute of Child Health
University College London

Thesis submitted for examination for the PhD degree of the University of London
October 2001
Abstract

Autism is a neurodevelopmental psychiatric syndrome characterised by impairments in three domains: social interaction, communication, and restricted and repetitive behaviours and interests. The neuropathology associated with Autism is unclear.

A recently developed animal model implicates the medial temporal lobes in Autism. Bachevalier and colleagues have demonstrated that bilateral ablation of the medial temporal lobes of neonatal monkeys leads to the development of a constellation of symptoms similar to those of Autism.

This thesis investigates the hypothesis that medial temporal lobe abnormality is responsible for some of the cognitive and behavioural impairments seen in individuals with Autism. A number of different techniques are used to compare the brain structure and cognitive and behavioural function of children with Autism with those of normal controls.

The neuropsychological investigations described in Part I revealed an impairment in episodic memory and no evidence of impaired semantic memory or recognition memory. Additionally, a protocol of tests of executive functions revealed a deficit in a task sensitive to the functioning of the orbitofrontal cortex and a motor checklist revealed co-ordination difficulties consistent with cerebellar abnormality.

Part II describes investigations of brain structure, using a variety of magnetic resonance techniques. A new analysis technique was developed specifically to examine bilateral abnormalities in developmental disorders. This technique revealed bilateral abnormalities in the amygdala, hippocampal formation, orbitofrontal cortex, superior temporal gyrus and cerebellum in children with Autism.
In Part III, the functional integrity of three event related potentials was investigated. These components were selected as they have been shown to be disrupted by medial temporal lobe abnormality. The results suggest that any functional abnormality of these waveforms is of a more subtle nature than could be detected through the paradigms used in this thesis.

In summary, convergent evidence of abnormality in the medial temporal lobes, orbitofrontal cortex and the cerebellum in Autism was obtained.
Part IV Discussion

Chapter 10 General Discussion and Future Studies

10.1 Summary of Findings

10.2 Generalisation of the Results to the Autistic Population

10.3 The Medial Temporal Lobe Hypothesis of Autism

10.4 Future Studies and Directions

Appendix A Diagnostic Criteria: DSM-IV

A.1 Autistic Disorder [DSM-IV 1994]

A.2 Asperger’s Syndrome [DSM-IV 1994]

Appendix B Missing Data

Appendix C Statistical Results

Appendix D Voxel-based Morphometry

D.1 Precision of Spatial Normalisation in Voxel-Based Morphometry

D.2 Distributional assumptions in Voxel-Based Morphometry

D.3 Independence of Residuals

Appendix E Abbreviations and Definitions

Appendix F References
Figure 6:3  Comparison of Probability Values 151
Figure 6:4  Conjunction Analysis Results 152
Figure 7:1  Centre of Small Volume Correction 167
Figure 7:2  ±27–18–18 Hippocampal Formation (4mm, SVC) 168
Figure 7:3  ±20–10–24 Amygdala (+ rostral hippocampal formation) (12mm, SVC) 169
Figure 7:4  ±24 21–18 Orbitofrontal cortex (12mm) 169
Figure 7:5  ±52 18–21: Rostral Superior Temporal Gyrus (12mm) 169
Figure 7:6  ±24 –74–45 Cerebellum (Inferior region (including VIIB, VI A and Crus II)) (4mm) 170
Figure 7:7  ±28 –66–22 Cerebellum (Superior, including areas VI, V, IV) (4mm) 170
Figure 7:8  ±16 –4–20 Amygdala (12mm) 176
Figure 7:9  ±30 –26–27 Parahippocampal region (4mm) 176
Figure 7:10 ±44 –2 48 Superior precentral gyrus (12mm) 177
Figure 7:11 ±42 –10 14 Operculum bank of the subcentral gyrus (12mm) 177
Figure 7:12 ±39 –34–28 Fusiform gyrus (12mm) 177
Figure 7:13 ±26 32–14 Posterior orbitofrontal cortex (12mm) 178
Figure 7:14 –24 –32 27 Corona radiata at the level of the caudal putamen (4mm) 179
Figure 7:15 ±44 –4 46 Superior precentral gyrus (12mm) 180
Figure 7:16 Posterior orbitofrontal cortex (±27 22–16) (12mm) 181
Figure 7:17 Rostral superior temporal gyrus (±45 10–26) (12mm) 182
Figure 7:18 Cerebellum (Inferior including VIII B) (± 21 –48–51) (4mm) 182
Figure 7:19 Cerebellum (Inferior, including VII B, VIII A, VIII B) (±30 –54–51) (12mm) 182
Figure 7:20 Claustral Insula ±36 0–9 (4mm) 183
Figure 7:21 Hippocampal Formation (±24 –15–18) (4mm) 184
Figure 7:22 Amygdala (±18 –6–21) (12mm) 184
Figure 7:23 Rostral Superior Temporal Gyrus (±45 24–24) (12mm) 184
Figure 7:24 Cerebellum (Superior including VI) (±28 –63–24) (4mm) 185
Figure 7:25 Postcentral gyrus (somatosensory motor cortex) (±40 –30 66) (12mm) 185
Figure 7:26 Volumes of the Hippocampal Formation (Mean ± SEM) 199
Figure 7:27 Volumes of the Amygdala (Mean ± SEM) 199
Figure 7:28 Intracranial Volumes (Mean ± SEM) (cm³). 200
Figure 7:29 Scattergraph of AT2 values 204
Figure 7:30 Scattergraph of HCT2 values 204
Figure III:1 A typical electrode montage 209
Figure III:2 A typical ERP component 210
Figure 8:1 ERP waveforms evoked by the novel and standard stimuli for adult control group. 219
Figure 8:2 ERP waveforms evoked by the target and standard stimuli for adult control group. 220
Figure 8:3 ERP waveforms evoked by the novel and standard stimuli for child control group. 223
Figure 8:4 ERP waveforms evoked by the target and standard stimuli for child control group. 225
Figure 8:5 a) Percentage of targets correctly identified; b) False alarms 228
Figure 8:6 Reaction times for target detection 229
Figure 8:7 ERP waveforms evoked by the novel stimuli for all 3 groups 230
Anterior–Posterior Topographical differences in response to novel stimuli

Effect of repeated presentation of novel sounds (Mean ± SEM) Negative is down

Mean Voltage of target P3b on midline electrodes (Mean ± SEM) Negative is up

ERP waveforms evoked by the target for all three groups.

Anterior–Posterior Variations in response to related and unrelated target words in adult control group

ERP waveforms evoked by the related and unrelated targets for adult control group.

Difference in magnitude of N400 effect in adults, averaged by the social nature of the pair.

Anterior–Posterior Responses along midline in child control group (Negative is up)

ERP waveforms evoked by the related and unrelated targets for child control group.

Difference in magnitude of N400 effect in control children, averaged by the social nature of the pair.

Behavioural results from N400 paradigm

Percentage of responses ‘Yes’ to semantic congruency question (Mean ± SEM)

Reaction Times for related and unrelated word pairs (Mean ± SEM)

Percentage of correct behavioural responses divided according to the social nature of the word pairs.

Anterior–Posterior variation of N4 effect in the three groups (Mean ± SEM)

Difference ERP waveforms (related – unrelated) for all three groups

Difference in magnitude of N400 effect, averaged by the social nature of the pair. (Mean ± SEM)

Anatomical Landmarks

Standard deviation from mean position of landmarks using medium regularisation

Standard deviation from mean position of landmarks using varying levels of regularisation

Significant areas of decreased grey matter in the amnesic children versus controls

Significant areas of decreased grey matter in the amnesic children versus controls

QQ Plots: The proportion of data points significantly violating the assumptions of normality

Number of analyses with one or more false positives at 12mm, 8mm, and 4mm

Number of analyses with one or more false positives at 12mm, 8mm and 4mm:

Number of analyses with one or more false positives at 12mm, 8mm, and 4mm:

Number of analyses with one or more false positives at 12mm, 8mm and 4mm:

Table 2:1 Group characteristics

Table 2:2 Handedness results RH refers to Right Handed, LH refers to Left Handed

Table 2:3 Intelligence Quotients of groups, as measured on the appropriate Wechsler Intelligence Scale

Table 3:1 Statistical Results from Verbal Recognition Tests of the Children's Memory Scale

Table 3:2 Statistical Results from Semantic Subtests of the Wechsler Intelligence Scales

Table 3:3 Statistical Results from Cued and Repeated Presentation Verbal Subtests

Table 3:4 Statistical Results from Single Presentation Recall Subtests

Table 4:1 Mean number of trials with no blink response and mean number of rejected trials

Table 6:1 Location of all points significant in the conjunction analysis at the corrected level of p = 0.05

Table 7:1 MR Protocol Details

Table 7:2 Summary of results from clinical neuroradiological assessment
Acknowledgements

My supervisors, Professor Faraneh Vargha-Khadem and Professor David Gadian, have been key in the creation of this thesis. I have been exceedingly fortunate to be given the freedom to guide the direction of my research in a supportive environment. Throughout my learning curve, they have created enviable opportunities, provided gentle guidance, offered valuable insights and nurtured my independence.

I have also been extremely privileged to be supervised (formally and informally) by a number of other people, to whom I am enormously grateful. Professor Mort Mishkin has generously shared his time and knowledge throughout my studies. All the voxel-based morphometry work has been completed under the direct supervision of Professor Karl Friston with Dr John Ashburner and Dr Alan Connelly. The ERP work was carried out with Dr Torsten Baldeweg, whilst Dr Michelle de Haan was indispensable in the work with the Visual Paired Comparison and the emotional modulation of the startle response paradigms. Professor Tim Cole advised on the statistical analysis of the ERP and behavioural data.

Collection of the MR data would not have been possible without the skill and patience of the radiographers in MR2. Many thanks to all of you, especially Lindy McWhirter and Jane Ho. Additional control MR data were generously provided by Dr Elizabeth Isaacs. Dr Kling Chong, consultant paediatric neuroradiologist, very kindly reviewed the scans of all the children.

Colleagues and fellow students at the Wolfson and MR2 have been highly tolerant of my endless questions, and generous with their time, advice and friendship. Leasha Lillywhite, Anna Adlam, Alison Rowan and Frederique Liegeois deserve a special mention for helping to screen the control children.
My family and friends have not only humoured my obsessional pursuit of my PhD, but also provided endless encouragement and sympathetic ears. In the last six months, three people in particular have supported me through the emotional rollercoaster of writing up: Mark, Fiona and Jo – I couldn’t have done it without you!

Finally, this thesis would not exist at all without the hard work and enthusiasm of the children and their families. May this thesis reflect your courage, determination and dedication. Thank you.
Chapter 1  Introduction

This chapter provides a broad overview of Autism and explains the motivation for investigating the medial temporal lobes in Autism. Unanswered questions regarding the role of the medial temporal lobes in Autism are highlighted.

There are three sections to this chapter. In the first section, Autism is defined and the characteristics of the disorder outlined. A number of different theories of Autism are discussed, including those implicating the medial temporal lobes. Evidence for the role of the medial temporal lobes in social behaviour is then reviewed in the second section. In the final section, the structure and hypotheses of this thesis are detailed.

1.1 Autism

Autism is a neurodevelopmental psychiatric syndrome that affects approximately 1 in 1000 children with varying degrees of severity [Fombonne 1999]. The syndrome is characterised by impairments in three domains: social interaction, communication, and restricted and repetitive behaviours and interests.

Impairments in social interaction in individuals with Autism are gross and sustained. These may be characterised by a relative absence of non-verbal behaviours (eye gaze, facial and body expression); failure to develop appropriate peer relationships; lack of spontaneous sharing of interest or activity with others; and little appreciation of social convention or the needs of others.

The impairment in communication affects both verbal and non-verbal expression and comprehension. The degree of deficit can vary enormously, with some individuals having no useful language and others having good basic language, yet marked difficulties with the more subtle aspects (such as the use of metaphor, sarcasm, etc.). When language does develop, it may be repetitive, idiosyncratic or have unusual pitch, intonation or rhythm.
Restricted and repetitive behaviours and interests may be characterised by one or more stereotyped or narrow interests which are abnormal in focus or intensity. For example, a child with Autism might be particularly interested in the London Underground and have a wealth of knowledge about the stations, lines and trains. Others may show stereotypies (stereotyped motor mannerisms) such as hand flapping. Often individuals with Autism show inflexible adherence to routines, insistence on sameness and resistance to change. For example, some children with Autism become very distressed if the route taken to school is altered, or sandwiches are not cut on the diagonal.

The first report of Autism was by Kanner in 1943. He described a group of 11 children who all showed persistent echolalia (speech repetition), pronoun reversal, a failure to use speech to communicate, an anxious desire to preserve sameness, repetitious behaviours, a general lack of awareness of other people’s existence or feelings and a lack of ability to play imaginatively with other children [Kanner 1943]. These deficits can all be subsumed by deficits in the three categories of impairment described above. Around the same time, Asperger also described a group of children who shared many of the same characteristics as Kanner's children (Asperger 1944; English translation Asperger 1991).

There are no known environmental or neurobiological aetiologies common to all individuals with Autism, and thus diagnosis is based exclusively on the individual’s historical and current behavioural symptomatology. There is considerable heterogeneity in the profile of individuals with Autism and this has led to a number of nosologies being developed. The most commonly used diagnostic frameworks are DSM-IV and ICD-10 [ICD-10 1992; DSM-IV 1994] (see Appendix A). Both frameworks have attempted to subdivide individuals with Autism into more homogeneous groups. These divisions are discussed below.

1.1.1 Subtypes Of Autism

The traditional concept of Autism is sometimes termed 'Kanner' Autism. Typically, children with Kanner Autism are aloof and indifferent to others, are mute or have abnormal
language, display little use of gesture, and have stereotyped routines involving people or objects. Other individuals with Autism are quite different: they are more passive, have good language, use gestures inappropriately and have interests that involve the amassing of facts or mathematical constructs. Wing introduced the term Asperger’s syndrome to describe these more able children with Autism [Wing 1981a]. She hoped that the term Asperger’s syndrome would aid recognition of the difficulties these individuals experience.

Whilst the term Asperger’s syndrome (AS) has proved popular in the clinical setting, the distinction between AS and Autism is unclear. Both ICD-10 and DSM-IV require that individuals with AS have no clinically significant general delay in language or cognitive development (other than in social interaction). However clinicians may interpret ‘clinically significant’ differently, and many clinicians believe that language delays should not lead to automatic exclusion of a diagnosis of AS [Eisenmajer et al. 1996]. Such inconsistencies within the clinical field may lead to the same child receiving different diagnoses from two clinicians [Howlin and Asgharian 1999].

A number of experimental studies have attempted to characterise differences between individuals diagnosed with Asperger’s syndrome and high functioning individuals with Autism (HFA) (e.g. Szatmari et al. 1990; Eisenmajer et al. 1996). Methodological limitations, including unclear attempts to exclude individuals with AS from the HFA group and failure to match groups on IQ (e.g. Ozonoff et al. 1991a), compromise interpretation. Additionally, cross-study comparisons are hampered by the use of different diagnostic criteria (see discussion in Ghaziuddin et al. 1992).

Wing argued that the difference between Asperger’s syndrome and HFA is severity [Wing 1981a]. The view that HFA and AS are variants of the same condition (or along the same continuum) is shared by many [Bosch 1970; Ozonoff et al. 1991a; Aicardi 1992]. It is accepted that at least some cases of AS are mild cases of Autism [Howlin and Asgharian 1999], although it is disputed whether this is true for all cases (see Szatmari et al. 1986; Happé 1994).
Introduction

Further, the behavioural symptoms of individuals with Autism are likely to show dynamic changes with time, and thus the appropriate diagnostic label may also change with age. For example Wing suggests that a child may look like a 'Kanner' child in infancy and yet develop into a more AS-like adolescent and adult [Wing 1981a]. A number of family studies have shown that there is a higher co-occurrence of Autism and AS than would be expected by chance (e.g. Burgoine and Wing 1983; Bowman 1988; DeLong and Dwyer 1988; Gillberg 1991). Most importantly for this study, which is concerned with identifying the neurobiological substrate of Autism, there is no compelling evidence to suggest that the bases of Autism and Asperger’s syndrome are distinct [Gillberg et al. 1987; Gillberg 1989].

In this thesis, Asperger’s syndrome and high functioning Autism will therefore be considered to represent the same entity.

1.1.2 Psychological Profile Of Autism

1.1.2.1 Socioemotional Profile

The nature of the socioemotional profile of individuals with Autism has been the focus of much research. It has been shown that the deficits are pervasive throughout many aspects of social behaviour. These difficulties include: poor interpretation of social cues (especially if multiple cues are present) [Pierce et al. 1997]; failure to respond to gestural behaviours [Loveland and Tunali-Kotoski 1997]; odd or inappropriate intonation and prosody [Fine et al. 1991; Boucher et al. 1998]; difficulties anticipating others’ knowledge or beliefs (known as poor Theory of Mind – [Baron-Cohen et al. 1985a; Loveland 1991; Critchley et al. 2000]); lack of interest in peers [Loveland et al. 1997]; and difficulties initiating communication [Loveland et al. 1997].

In addition, individuals with Autism have difficulty with the following: recognising; understanding; labelling; and matching emotional facial expressions (e.g. Hobson 1986a; Hobson 1986b; Weeks and Hobson 1987; Hobson et al. 1988; Tantam et al. 1989; Davies
Matching emotional facial expressions with appropriate contexts, gestures or vocalisations [Hobson 1986a; Hobson 1986b], and perception of eye gaze direction [Ellis et al. 1994; Howard et al. 2000] have also been shown to be impaired, as well as matching emotional vocalisations to emotional gestures [Hobson 1986a].

The degree of impairment in the recognition of emotional expression varies widely in children with Autism. Some studies have reported normal recognition of emotion when verbal ability is accounted for (e.g. Ozonoff et al. 1990; Prior et al. 1990; Loveland et al. 1997; Buitelaar et al. 1999a; Teunisse and Gelder 2001). It has been suggested that some of this variation may be due to high functioning individuals implementing cognitive strategies to compensate for their inability to perceive emotion more directly [Loveland et al. 1997]. This suggestion is in part supported by the finding that individuals with high functioning Autism are more impaired than controls at sorting faces by emotional expression when parts of the face are covered [Hobson et al. 1988; Teunisse and Gelder 2001]. There is also a suggestion that specific emotional expressions have been reported as affected (especially fear [Howard et al. 2000] and surprise [Baron-Cohen et al. 1993]).

Individuals with Autism have been reported to have abnormal production of emotional expressions. Both clinical and experimental evidence has highlighted deficits and idiosyncrasies in facial [Loveland et al. 1994; Loveland and Tunali-Kotoski 1997], vocal [Ricks 1975; Ricks 1979] and gestural [Attwood 1984] expression of emotion.

It should be noted that individuals with Autism with significant impairment on emotional tasks often perform above chance, indicating that they do have some understanding of emotion [Hobson 1986a]. Similarly, children with Autism respond differently to different people and in different situations [Sigman and Ungerer 1984; Landry and Loveland 1988], suggesting that they are aware to some extent that different individuals hold different significance for them [Loveland and Tunali-Kotoski 1997].
1.1.2.2 Memory Profile
The memory profile of individuals with Autism has not been as extensively studied as the social difficulties described above. There are many components of memory and these interact to produce normal memory and learning. Few studies have attempted to assess memory skills comprehensively, with most focussing on just one component. However, the results in the literature (e.g. Boucher 1981; Boucher and Lewis 1989; Farrant et al. 1998) are consistent with the hypothesis that individuals with high functioning Autism have a selective impairment in episodic (contextual) memory with intact semantic (factual) and recognition memory. Individuals with Autism and learning disabilities show a more generalised pattern of deficit in memory and learning (see Chapter 3).

1.1.2.3 Executive Function Profile
Executive function refers to mental operations which enable an individual to disengage from the immediate context in order to guide behaviour by reference to mental models or future goals. Individuals with Autism have been noted to have difficulty with tasks assessing different aspects of executive function, particularly those requiring planning and set shifting [Ozonoff et al. 1991b; Hughes et al. 1994]. Executive function deficits have been proposed to explain the characteristic cognitive and behavioural profile of Autism (see Chapter 5).

1.1.3 Genetics Of Autism
Familial and twin studies have demonstrated the important role that genetics play in the aetiology of Autism [Risch et al. 1999; Maestrini et al. 2000]. The risk of siblings of individuals with Autism also being affected is 75 times higher than the incidence in the general population [Bolton et al. 1994]. Concordance rates in monozygotic twins are considerably higher than those found in dizygotic twins [Bailey et al. 1995].

Our understanding of the role of genetics in Autism is limited. Autism does not follow a simple mode of inheritance [Risch et al. 1999] and it is thought that several genes (at least 3 or 4) are involved in the pathogenesis of the syndrome [Folstein et al. 1999; Lamb et al. 2000]. It is believed that each locus confers an increased risk of Autism but alone is
insufficient to cause the disorder [International Molecular Genetic Study of Autism Consortium 1998].

The locations of these genes are still being investigated. One chromosome that has been repeatedly reported as abnormal is chromosome 15 [Phillippe et al. 1999; Turner et al. 2000]. A second locus on chromosome 7q has been identified by a number of genome studies [Lamb et al. 2000; Maestrini et al. 2000].

Whilst genetics may be an important factor in the behavioural profile of an individual, it should be noted that two individuals with identical genotypes do not necessarily have identical phenotypes [Kates et al. 1998; Pickles et al. 2000]. This is likely to be a reflection of environmental factors modulating expression of genetic predispositions [McInnes et al. 1998].

1.1.4 Neuropathological Basis Of Autism

Critical to our understanding of the aetiology of Autism is the underlying neuropathology associated with this disorder. A large number of studies have highlighted different areas of the brain as abnormal in Autism, though replication has proved difficult in many cases.

A number of different techniques have highlighted three main regions of the brain as abnormal in individuals with Autism: the medial temporal lobes, the prefrontal cortex and the cerebellum. The details of these findings are discussed in Chapter 7, and only briefly summarised here.

The pathology affecting the medial temporal lobes has been revealed through post-mortem studies showing increased cell density in the amygdala and hippocampal formation (e.g. Raymond et al. 1989; Bauman and Kemper 1994). Similarly, imaging studies have detected abnormal volumes of the amygdala and hippocampal formation (e.g. Aylward et al. 1999; Howard et al. 2000), while cognitive and behavioural profiles characteristic of
individuals with medial temporal lobe abnormality have been described in Autism (e.g. Howard et al. 2000).

Functional and structural imaging studies have highlighted abnormalities in the frontal lobes [Minshew et al. 1993; Carper and Courchesne 2000] whilst behavioural studies have found similarities between individuals with frontal lobe injury and individuals with Autism (e.g. Hughes et al. 1994). Cellular abnormalities have also been reported in the frontal lobes [Bailey et al. 1998].

Finally, findings consistent with cerebellar abnormality include a reduction in Purkinje cell numbers observed at post-mortem [Bauman and Kemper 1985], abnormalities in vermal lobule volumes observed in imaging studies [Courchesne et al. 1988; Courchesne et al. 1994a], and poor motor co-ordination [Manjiviona and Prior 1995].

1.1.5 Theories Of Autism

The complex and elusive nature of Autism has led to a number of diverse theories being proposed to account for the spectrum of this disorder. These are briefly outlined here (see Chapter 10 for further discussion).

Perhaps the most widely known psychological account of Autism is the Theory of Mind hypothesis. Theory of Mind is a generic term that refers to the individual’s ability to think about a range of mental states (intentions, desires, thoughts, beliefs, dreams, pretence, etc.) in others [Baron-Cohen et al. 1994]. Frith and colleagues have argued that deficits in Theory of Mind underlie the social and communication difficulties of individuals with Autism (e.g. Baron-Cohen et al. 1985b).

A second and more recent psychological account of Autism is the weak ‘central coherence’ theory. A characteristic of normal information processing appears to be the tendency to draw together diverse information to construct higher-level meaning in context. Frith refers to this as ‘central coherence’. It has been argued that disturbances of central coherence
would parsimoniously explain the troughs and peaks of cognition in Autism [Frith 1989; Happé 1999].

Although these psychological theories have advanced our understanding of the difficulties experienced by individuals with Autism, they offer little insight into the neural basis of the disorder. A number of neuropathological theories of Autism have implicated different regions of the brain. Whilst some have proposed the frontal lobes [Russell 1997] or cerebellum (e.g. Courchesne et al. 1994a) as key sites, others have hypothesised about the important role of the medial temporal lobes in Autism (e.g. Bachevalier 1994).

For example, Damasio and Maurer proposed that Autism resulted from dysfunction in the mesolimbic system [Damasio and Maurer 1978]. Explicitly included in this system are the hippocampal formation, the amygdala and the parahippocampal region. Brothers proposed that the central and primary pathology in Autism is a lack of empathy (defined as interpersonal phenomena) and argued that the amygdala plays a central role in the neural basis of empathy [Brothers 1989]. Fotheringham also argued that the amygdala is central to the neuropathology of Autism. He proposed that abnormalities in the amygdala cause the central deficit in Autism: a failure to attach appropriate emotional value to stimuli [Fotheringham 1991].

DeLong argued that Autism is the result of the failure of a central cognitive processor which is necessary for flexible multidimensional association of sensorial stimuli, memory and motivational states. He argued that this function is critically dependent on the hippocampal formation, thus proposing that Autism is a developmental syndrome of hippocampal dysfunction [DeLong 1992]. Waterhouse et al. proposed that Autism resulted from abnormal neuronal organisation during brain development in the amygdala, hippocampal formation, temporal and parietal association cortices, and from dysfunction of neuropeptide pathways [Waterhouse et al. 1996a].

Although these researchers placed emphasis on different components of the medial temporal lobes, all acknowledge the wealth of evidence that the medial temporal lobes play an important role in social behaviour.
1.2 The Medial Temporal Lobes And Social Behaviour

The medial temporal lobes have been implicated in social behaviour in a large range of studies, including animal and human lesion studies. Damage to the medial temporal lobes (and in particular the amygdala) results in deficits in social behaviour. These deficits are similar to those experienced by individuals with Autism and in this section the evidence is reviewed.

1.2.1 The Anatomy Of Social Behaviour In Adult Monkeys

The role of the medial temporal lobes in social behaviour is highlighted by work in adult monkeys. In 1888, Brown and Schäfer reported that monkeys with bilateral temporal cortex lesions became tame and indifferent [Brown and Schäfer 1888]. Klüver and Bucy later replicated this work and described in detail the resulting emotional and behavioural changes. The animals displayed psychic blindness (approaching animate and inanimate objects with no fear), oral tendencies, hyperresponsiveness to visual stimuli, visual agnosia and blunted aggression and fear [Klüver and Bucy 1937; Klüver and Bucy 1939]. This syndrome has been termed the Klüver-Bucy syndrome.

More selective bilateral lesion studies have shown that in particular the amygdala is involved in many aspects of social behaviour. Animals with amygdalectomies lose their social standing, are unresponsive to group members, fail to show appropriate social signals, are more submissive, demonstrate abnormal levels of aversive and aggressive behaviour (both increases and decreases have been reported depending on social context) and fail to respond maternally to their infants (for review see Kling and Brothers 1992; Aggleton and Young 2000; Bachevalier 2000; Emery and Amaral 2000; Barton and Aggleton 2001). Abnormal sexual behaviour has also been described (see Kling and Brothers 1992).

Interestingly, whilst monkeys with medial temporal lobe lesions do show abnormal social skills, they still appear to possess some residual social awareness. Lesioned animals differentiate between other medial temporal lobe lesioned monkeys and controls. In the presence of other lesioned monkeys, they are more confident, more affiliative and less
avoidant than when paired with control monkeys. This pattern is also seen in control monkeys paired with lesioned monkeys versus other control monkeys [Emery et al. 1998].

Although the lesions in these studies were centred on the amygdala, many were either aspiration lesions or radiofrequency lesions. These lesion types are likely to cause damage to areas beyond the amygdala. Aspiration lesions necessarily remove part of the entorhinal cortex, and both radiofrequency and aspiration lesions damage all the fibres that pass through the amygdala and some adjacent to the amygdala (see Baxter and Murray 2000). Indeed, retrograde tracers have revealed damage to fibres of passage following amygdala aspiration lesions. These pathways include the projections of the entorhinal and perirhinal cortices and area TE to the medial thalamus [Goulet et al. 1998] and to the orbital frontal cortex [Baxter et al. 1998]. It is also difficult to exclude the possibility of damage to surrounding structures of the amygdala such as the hippocampal formation. Changes in behaviour following such lesions can not therefore be attributed solely to the amygdala.

In the light of this caveat, it should be noted that there is evidence that the hippocampal formation is also important in social behaviour. Animals with hippocampal lesions display abnormal social behaviour [Beauregard et al. 1995]. For example they have been shown to be inflexible and rigid in their behavioural responses (e.g. Devenport et al. 1988). They also show abnormal sensitivities to context (increased dependence on irrelevant context and decreased dependence on relevant context) (e.g. Winocur and Olds 1978; Winocur et al. 1987; Good and Honey 1991).

More recent studies have shown that many of the social and emotional changes initially reported with aspiration and radiofrequency lesions are also seen in animals with neurotoxic lesions, although some of the effects are less dramatic [Bachevalier 2000]. Reduced aggression, fear, increased submission and increased manual and oral exploration have all been reported in animals with neurotoxic lesions of the amygdala [Meunier et al. 1996; Emery et al. 1998; Meunier et al. 1999].
1.2.2 Neonatal Lesions Of The Medial Temporal Lobes In Monkeys

The effects of neonatal lesions of the medial temporal lobe were first examined by Kling and Green [1967]. They reported grossly normal somatic and affective development over the first two years of life in monkeys raised in naturalistic conditions. Thompson undertook a series of systematic studies investigating the effects of bilateral aspiration lesions of the amygdala in 3 month old monkeys who were subsequently housed individually. In contrast to Kling et al., he found changes in social affiliation very similar to those found in monkeys with amygdala lesions in adulthood [Thompson et al. 1969; Thompson and Towfighi 1976; Thompson et al. 1977; Thompson 1981].

Bachevalier and colleagues have extensively investigated the behavioural and cognitive sequelae of neonatal bilateral medial temporal lobe lesions (see e.g. Bachevalier 1994; Bachevalier 2000; Bachevalier et al. 2001). The monkeys were individually housed, and played in pairs or triads each day [Bachevalier 2000]. Monkeys with lesions encompassing the amygdala, hippocampal formation, periamygdaloid cortex and portions of the parahippocampal region showed an increasing number of socioemotional abnormalities as they matured. By the age of 6 months, they spent significantly less time in social interaction, actively avoided social contacts, developed locomotor stereotypies (significantly more than control monkeys) and self-directed behaviours, had poor facial and body expression and displayed little eye contact. When tested at 11-14 months, animals with neonatal medial temporal lesions had a reduced ability to modulate their vocalisations as a means of conveying levels of affect [Newman and Bachevalier 1997]. Wide individual variation in the severity of symptom expression has been noted following medial temporal lobe damage [Bachevalier 1994]. Such individual variation of symptom severity has also been noted in Autism.

When compared in adulthood to animals with adult onset medial temporal lobe lesions, animals with neonatal lesions were significantly more impaired. The neonatal lesioned animals showed substantially less social interaction than adult lesioned monkeys, and self-
directed behaviours. Furthermore, only the neonatal lesions animals showed increased object manipulation. In contrast to animals with comparable lesions made in adulthood, the monkeys with neonatal ablations did not have a lack of fear responses and did not display signs of hyperorality [Nawla and Bachevalier 1991].

Bachevalier et al. have further demonstrated that more selective medial temporal lobe lesions also result in aberrant social behaviour [Bachevalier 1994; Bachevalier et al. 1995]. For example, monkeys with aspiration lesions to the amygdala (but sparing the hippocampal formation) display a similar pattern to that described above, with the magnitude of the disturbances being smaller. However the animals did not show less acceptance of approach, stereotypic behaviours, or loss of facial and body expressions, and the monkeys were more passive compared to animals with bilateral lesions to both the amygdala and hippocampal formation. Neonatal lesions to the hippocampal formation have been shown to result in less significant socioemotional changes during infancy, but in adulthood the monkeys display a significant loss of social interactions and stereotypic behaviours [Beauregard et al. 1995].

It is important to note that neonatal neurotoxic lesions of the amygdala have not been reported in monkeys, and so it is difficult to ascertain the role of the amygdala in the socioemotional disturbances described above compared to the role of surrounding cortex and fibres of passage. Additionally, the effects of the medial temporal lobe lesions may be more pronounced in the work of Bachevalier and colleagues, as the animals are peer-reared, not reared within a naturalistic family setting [Bachevalier 2000]. However, whilst this might lead to greater symptom severity, the control animals are also peer-reared, and so peer-rearing alone can not explain the behaviour changes observed.

### 1.2.3 Effects Of Medial Temporal Lobes Lesions In Humans

Social behaviour of humans with medial temporal lobe damage has not been systematically investigated. In general, however, the severity of the socioemotional disturbance in humans is considerably less than that reported in primates. Klüver-Bucy syndrome is rarely
Introduction

reported even in individuals with extensive bilateral damage (Rossitch et al. 1989; Aggleton and Young 2000; but see Terzian and Ore 1955). For example, HM has extensive bilateral medial temporal lobe damage (encompassing the amygdala bilaterally [Corkin et al. 1997]), and yet only shows emotional indifference [Corkin 1984].

Difficulties with social interaction have been reported in individuals with medial temporal lobe damage (including the amygdala) (e.g. Tranel and Hyman 1990; Broks et al. 1998; Adolphs et al. 1998; Fine et al. 2001). The research on patients with amygdala damage has focussed primarily on interpretation of emotional expression. Facial expressions are an outward display of emotional state and are important cues in social communication.

For example, many individuals with amygdala damage are unable to recognise certain emotionally loaded facial expressions (in particular of fear) (Adolphs et al. 1994; Young et al. 1995a; Calder et al. 1996; but see Adolphs et al. 1995; Hamann et al. 1996; Young et al. 1996; Broks et al. 1998). A study of 9 subjects with bilateral amygdala damage (albeit not selective) found impairment of recognition of some negative facial expressions (including but not restricted to fear). Considerable intersubject variability was noted [Adolphs et al. 1999]. In some individuals, difficulties in recognising emotional auditory expression have also been noted (Scott et al. 1997 but see Anderson and Phelps 1998; Adolphs and Tranel 1999), as well as reports of difficulties detecting emotional similarity in facial expressions [Adolphs et al. 1994].

Other studies have noted abnormal eye gaze direction detection in individuals with amygdala damage [Young et al. 1995b; Broks et al. 1998], although once again this has not been found in all individuals [Adolphs et al. 1998; Broks et al. 1998]. Consistent with the animal literature, humans with amygdala damage have been found to have abnormal fear conditioning [Bechara et al. 1995; Labar et al. 1995].

Individuals with amygdala damage also show impaired memory. There is abundant evidence that an emotionally arousing experience is recalled more accurately, more easily and for a longer time than a mundane experience (see Holmes 1970; Cahill 2000). Humans with amygdala damage show no enhanced memory when the material is emotionally
A loss of automatic social interpretation has been documented in a patient (SM) with selective bilateral amygdala damage. Normal subjects were shown a video depicting three geometric shapes moving on a plain white background and they were asked to describe what they saw. Normal subjects ascribed social meaning to what they saw, whilst SM described the stimuli in purely geometric terms [Heberlein et al. 1998]. Replications with other patients with bilateral amygdala damage have not been reported.

The contradictory nature of the socioemotional profile of humans with medial temporal lobe damage encompassing the amygdala has been attributed to different extent of the lesion, different IQ and varying aetiologies. An alternative hypothesis is that early acquired amygdala damage leads to a different cognitive profile compared to damage to the amygdala acquired in adulthood [Hamann et al. 1996]. For example, perhaps late acquired amygdala damage leads to a sparing of fear recognition. However a number of subsequent reports have found fear recognition deficits in adults with late onset damage [Calder et al. 1996; Broks et al. 1998].

1.2.4 Electrophysiology: Stimulation

Stimulation of the amygdala produces behavioural and autonomic responses that are characteristic of changes in emotion. For example, electrical or chemical stimulation of the central nucleus induces fear behaviours and defensive reactions that are typical of the species [Hitchcock and Davis 1986; Rosen and Davis 1988; Balaban and Taussig 1994; Davis 2000], such as increases in freezing and cardiovascular changes [Rosen et al. 1996]. In humans, stimulation of the amygdala elicits feelings of fear and anxiety as well as autonomic reactions [Gloor 1997] and hallucinations [Halgren et al. 1978]. Hallucinations can also be elicited by stimulation of the hippocampal formation [Halgren et al. 1978]. Stimulation of the periamygdaloid area has been occasionally reported to evoke rage reactions in monkeys and man [Scoville et al. 1953; Kaada et al. 1954].
The role of the amygdala in social interaction is further highlighted by the finding that electrical stimulation of the amygdalae of monkeys living in groups yielded significant changes in vocalisations [Robinson 1967; Jurgens and Ploog 1970; Jurgens 1982]. Monkeys will work to obtain electrical stimulation of the amygdala and some human studies have reported that stimulation of the amygdala is rewarding (e.g. Sem-Jacobsen 1968; Rolls et al. 1980; Wilson and Rolls 2000; Rolls 2001).

1.2.5 Electrophysiology: Recordings

Neuronal recording studies in both monkeys and humans have established that there are neurons in the amygdala that respond selectively to the sight of faces and bodies of conspecifics [Perrett et al. 1982; Hasselmo et al. 1989; Washsmuth et al. 1994]. Neurons have been reported to encode facial identity [Perrett et al. 1984; Baylis et al. 1985], expression [Perrett et al. 1984; Hasselmo et al. 1989], hairline, mouth and eyes [Perrett et al. 1982; Desimone et al. 1984; Leonard et al. 1985], as well as more dynamic aspects of social stimuli (such as approach) [Barton and Aggleton 2001]. Wilson and Rolls recently demonstrated that there are amygdala neurons that encode reward and punishment related to visual stimuli [Wilson and Rolls 2000].

Neurons in the amygdala are active during social communication and interaction. For example, radiotelemetry activity recordings found that neurons in the amygdala respond most to ambiguous or threatening situations and least to tension lowering behaviours (such as grooming or huddling) [Kling et al. 1979]. A hierarchy of responses in the amygdala was recorded in response to social communications (both auditory and visual) which correlated with the emotional significance of the stimulus (see Kling and Brothers 1992 for review). Similar studies have shown that the amygdala responds differentially depending on the context in which intra-species calls are heard [Kling and Brothers 1992]. The activity of the neurons during social behaviour in the hippocampal formation has been less well studied, although there are reports of inhibition of activity during social behaviour in rats [Garritano et al. 1996].
1.2.6 Anatomy And Pharmacology Of The Medial Temporal Lobes

The anatomy and pharmacology of the amygdala and the hippocampal formation are consistent with their role in social and emotional behaviour. For example, olfactory signals are important in a wide range of social behaviours in mammals (e.g. Liebenauer and Slotnick 1996). The importance of the amygdala in social behaviour is therefore reflected in the direct projections from the olfactory bulbs to the amygdala: the only modality with a projection from a primary sensory area [Aggleton and Mishkin 1986]. The olfactory bulbs also project to the hippocampal formation via the entorhinal cortex [Suzuki 1996].

Chemical infusions to the amygdala and hippocampal formation cause behavioural and autonomic changes associated with changes in emotional states. For example, injections of the GABA_A antagonist bicuculline into the basolateral nucleus result in an increase in heart rate and blood pressure [Sanders and Shekhar 1991]. Local infusions of benzodiazepines into the amygdala have anxiolytic effects on measures of freezing, shock probe avoidance or the elevated plus-maze (see Davis 2000 for review). Anxiety as measured by the social interaction tests is also reduced following infusion of benzodiazepines into the dorsal hippocampal formation [Gonzalez et al. 1998].

1.2.7 Imaging Studies

The role that the medial temporal lobes play in social cognition has been further highlighted by functional imaging studies. The amygdala is activated following presentation of visual (e.g. Breiter et al. 1996; Morris et al. 1996; Morris et al. 1998), auditory (e.g. Morris et al. 1999) and gustatory (e.g. Zald and Pardo 1997) stimuli that signal unpleasant and arousing emotions [Adolphs and Tranel 2000]. Amygdala activation also occurs during the encoding of emotional material [Cahill et al. 1996; Taylor et al. 1998a; Hamann et al. 1999]. In addition, both amygdala and hippocampal formation activity has been found to be correlated with recall of arousing stimuli [Hamann et al. 1999].
A number of functional imaging studies have demonstrated amygdala involvement in recognition of emotions from facial expressions and vocalisations (specifically negative emotions, such as fear) (e.g. Breiter et al. 1996; Morris et al. 1996; Morris et al. 1999). The amygdala is activated in response to fearful faces, even when there is no conscious perception of the stimuli [Whalen et al. 1998; Adolphs and Tranel 2000]. It should be noted that the amygdala is not activated in isolation. Imaging studies suggest that the neural network also includes the inferior frontal cortex and parts of the parietal and temporal cortex.

1.2.8 Summary

As reviewed above, extensive research has highlighted the similarities in the social profile of animals and humans with medial temporal lobe damage (especially encompassing the amygdala and hippocampal formation) and individuals with Autism. Whilst this points strongly to the involvement of the medial temporal lobes in the neuropathology of Autism, a number of questions remain.

Is the cognitive profile of Autism reminiscent of medial temporal lobe abnormality in areas of functioning other than social behaviour? For example, is the memory profile of individuals with Autism also consistent with medial temporal lobe abnormality? Further, can all aspects of the Autistic profile be explained by medial temporal lobe abnormality? In particular, can the executive function deficits be attributed solely to medial temporal lobe abnormality? Development is an interdependent and dynamic process – how selective is the medial temporal lobe abnormality in Autism? Although medial temporal lobe abnormalities have been detected in individuals with Autism (both at post-mortem and in neuroimaging studies), a comprehensive assessment of the medial temporal lobes using structural magnetic resonance imaging has yet to be completed.

This thesis seeks to answer some of these questions and attempts to find convergent evidence of medial temporal lobe abnormality in individuals with Autism.
1.3 Hypotheses And Structure Of Thesis

There is abundant evidence that individuals with Autism and individuals with medial temporal lobe damage share many difficulties in social behaviour (see above). This provides strong support for the role of the medial temporal lobes in Autism. However, the medial temporal lobes are known to have additional functions in human behaviour and these behaviours have not been systematically investigated in Autism. The integrity (or otherwise) of these functions has important consequences for the hypothesis of the role of the medial temporal lobes in Autism. Other aspects of medial temporal lobe functioning are therefore investigated in individuals with Autism in experiments described in this thesis (such as memory) (see Chapters 3 and 4).

The structure of the medial temporal lobes is investigated comprehensively using a variety of complementary magnetic resonance techniques including T2 maps, volumetric measurements and voxel-based morphometry (see Chapters 6 and 7). Functional measures of neural activity are also obtained using event-related potentials, to provide further convergent evidence of medial temporal lobe abnormality. Two event-related potential components are studied: the P300 and N400. Both of these are thought to be dependent on the integrity of the medial temporal lobes (see Chapters 8 and 9).

Development is a dynamic and interdependent process, and abnormalities in one area are therefore likely to impact on the maturation of other regions. Given the complex genetic and developmental nature of the disorder of Autism, it is therefore highly likely that the medial temporal lobe is not the sole area of abnormality responsible for the full spectrum of the disorder. This thesis therefore additionally investigates the structural and functional integrity of other candidate regions likely to be affected by abnormal medial temporal lobe development; namely the frontal lobes and the cerebellum.

The thesis is divided into three parts. The first is concerned with the behaviour and cognitive profiles of individuals with Autism. Part II describes magnetic resonance
investigations looking at the structure of the brains of individuals with Autism. This second part begins with a technical chapter describing and validating a technique using voxel-based morphometry. Part III describes analyses of the neural activity of individuals with Autism using event related potentials.

Below are detailed the central hypotheses:

Part 1:
Children with Autism will have impaired episodic with relatively preserved semantic memory and recognition memory.

Children with Autism will show an abnormal emotional modulation of the startle response.

Children with Autism will have executive function difficulties consistent with orbitofrontal abnormality.

Part 2:
Children with Autism will have bilateral medial temporal lobe abnormalities as detected by voxel-based morphometry.

Part 3:
Children with Autism will show abnormalities in the oddball P300 and semantic integration N400, consistent with medial temporal lobe abnormality.
Part I  Neuropsychological Investigations

Part I of this thesis describes a series of investigations concerning the neuropsychological profile of the children with Autism. Chapter 2 describes the children who took part in the research described in this thesis and details some baseline neuropsychological assessments. Chapters 3 and 4 describe investigations into functions thought to be dependent on the integrity of the medial temporal lobes. In Chapter 3, a comprehensive memory assessment is described with particular focus on episodic memory (a cognitive correlate of the hippocampal formation). In the experiments described in Chapter 4, an experimental paradigm (emotional modulation of the startle response) was used to assess the function of the amygdala.

In the final chapter of Part I, the executive functions of the children with Autism were investigated. The aim of this chapter was to find evidence of selective deficits in tasks thought to depend on the orbitofrontal cortex. The extensive connections between the orbitofrontal cortex and medial temporal lobes mean that abnormality in the medial temporal lobes may also affect the functions of the orbitofrontal cortex.
Chapter 2 Study Participants and Baseline Neuropsychological Assessment

In this chapter, the selection criteria and characteristics (such as sex, age and verbal IQ) of the groups of children who participated in this study are described. This is followed by the results from a number of baseline neuropsychological tests. Some of the baseline assessments were selected to index the profiles of the children on measures known to be affected in Autism (such as social behaviour, communication, etc.). In addition, the children’s attention was assessed. This was included in order to allow the role of any attention deficits to be accounted for in interpretation of task performance reported in future chapters.

2.1 Selection Criteria

Children with Autism (aged between 8-18) were recruited through parental support groups (including the National Autistic Society) and from schools specialising in the education of children with Autism. The children were all diagnosed with high functioning Autism or Asperger’s syndrome by independent clinicians (including paediatricians, clinical psychologists and psychiatrists).

Children with Autism were invited to take part in the study following a telephone screening procedure. Only children who were described as having good verbal skills, age appropriate academic performance and/or average verbal IQ were included. Children were excluded from the study if they had additional neurological or psychiatric diagnoses (including fragile X, epilepsy and Attention Deficit Hyperactivity Disorder), if they were taking medication or had a history consistent with a diagnosis of secondary Autism (such as rubella).
Normally developing control children (aged between 8-18) were recruited from local London schools. These children were required to meet the same inclusion and exclusion criteria as the children with Autism, with the additional requirement that there was no family history of Autism. All volunteer control children meeting these criteria were seen and their verbal IQ was measured. Control children with similar chronological ages and verbal IQs (see below) to the children with Autism were then invited back to complete the research protocol. A total of 28 control children were assessed. The results from 18 of these children were included in the final group analyses.

All participants were assessed at Great Ormond Street Hospital over a period of 2-3 days (not necessarily consecutively). Most volunteers completed the neuropsychology protocol unaccompanied (although 2 control children and 3 children with Autism were accompanied by a parent at the child’s request). A few of the children did not complete the full assessment, for a variety of reasons. A comprehensive list of missing data (and the reasons) is included in Appendix B.

2.1.1 Matching Children to Controls

2.1.1.1 Introduction

75% of individuals with Autism also have some degree of mental retardation [Rutter 1979]. Some researchers have argued that Autism is therefore a concomitant of mental handicap. However Autism rarely occurs in individuals with Down's syndrome (despite their severe mental retardation) [Wing and Gould 1979] and familial loading associated with Autism mainly applies to language and social abnormalities in individuals with normal intelligence. There is therefore little support for the notion of familial loading of mental retardation [Lord and Rutter 1994]. Autism is therefore not just an automatic characteristic of mental retardation.

When studying individuals with Autism and mental retardation, it is difficult to unravel the role of Autism over and above the mental retardation. Matching control children is also a
complex issue – the often uneven verbal and non-verbal intelligence profile makes it virtually impossible to match on both factors. Rutter argued that in order to best understand the basic and essential nature of Autism, it is necessary to investigate individuals with Autism and normal intelligence (in whom Autism occurs in its most pure form) [Rutter 1983].

It should be noted, however, that the IQ of a child does not necessarily index the severity of Autism (Szatmari and Jones 1991; although see Waterhouse et al. 1996b). Additionally, some have argued that verbal cognitive impairment is an intrinsic (although not quite invariant) part of Autism [Lord and Rutter 1994]. If this is the case, it could be argued that comparing groups of controls and children with Autism matched on verbal ability may not be an appropriate comparison. However, in order to establish that deficits seen in cognitive domains assessed verbally are a manifestation of Autism, it is necessary to rule out the possibility that the children with Autism are performing worse than controls due to a generalised verbal cognitive impairment.

In this study, the Autistic children were divided into two groups: those with verbal IQ > 85 (referred to throughout this thesis as the High group) and those with verbal IQ < 85 (referred to as the Low group). This classification is likely to be robust, as intelligence has been shown to be a stable construct in individuals with Autism [Rutter 1979; Rutter 1983].

In addition to intelligence, age, sex, diagnosis, handedness and receptive language skills of the three groups were measured to characterise the groups further. Receptive language skills were measured in order to index any disadvantage the Low group might have experienced due to poor comprehension of language.

2.1.1.2 Methods

Handedness Questionnaire
The hand preference of the children was determined by the Handedness Inventory adapted from Crovitz and Zener [1962] (as described in Isaacs et al. 1996). This test provides
ratings of both the left and right hands in carrying out unimanual as well as bimanual tasks. The child was asked to indicate he normally used to perform each of 18 common actions. For each action, the subject was also asked to rate the frequency of use of the two hands, with a rating of one being the normal response of a strongly right-handed person and a rating of five the response of a strongly left-handed person. Total scores thus ranged from 18 to 90. In keeping with previous studies, a score of < 30 was considered to be indicative of strong right-handedness and a score of > 55 of strong left-handedness. Scores of 30-55 were taken to indicate some degree of ambidexterity.

Intelligence Quotient
The age appropriate Wechsler Intelligence Test [Wechsler 1991; Wechsler 1997] was administered to provide verbal, performance and full scale IQs. Briefly this assessment consists of a number of subtests designed to tap a range of cognitive skills. Subtests that make up the verbal IQ score include Vocabulary (in which the child is asked to define a series of words), Similarities (in which the child is asked to explain why two objects or concepts are similar), Information (consisting of a series of general knowledge questions) and Comprehension (in which the child is asked to explain social conventions or solve everyday problems).

Receptive Language
The British Picture Vocabulary Scale II (BPVS II) [Dunn et al. 1997] was administered to provide a standardised measure of receptive vocabulary and linguistic concepts. The child was shown a series of four pictures and was asked to point to the picture that best depicted the meaning of a word. Scoring was carried out in accordance with the manual (which provides age normative data) to obtain standard scores (with a mean of 100 and standard deviation of 15) and an age equivalent score.

Statistical analysis throughout the thesis was carried out as described in Appendix C. Planned comparisons were Control versus High (CvH) and High versus Low (HvL).
2.1.1.3 Results

The global characteristics of the groups are shown below in Table 2:1, Table 2:2 and Table 2:3.

Table 2:1 Group characteristics

<table>
<thead>
<tr>
<th>Group</th>
<th>Size</th>
<th>Mean Age</th>
<th>Diagnosis</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>18</td>
<td>12.6</td>
<td>N/A</td>
<td>12 Female; 6 Male</td>
</tr>
<tr>
<td>High</td>
<td>14</td>
<td>12.9</td>
<td>3 HFA; 11 AS</td>
<td>1 Female; 13 Male</td>
</tr>
<tr>
<td>Low</td>
<td>12</td>
<td>11.6</td>
<td>3 HFA; 9 AS</td>
<td>3 Female; 9 Male</td>
</tr>
</tbody>
</table>

Table 2:2 Handedness results (RH refers to Right Handed, LH refers to Left Handed)

<table>
<thead>
<tr>
<th></th>
<th>Strongly RH</th>
<th>Ambidextrous</th>
<th>Strongly LH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>12</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>High</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Low</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2:3 Intelligence Quotients of groups, as measured on the appropriate Wechsler Intelligence Scale (Mean ± SEM)

<table>
<thead>
<tr>
<th>Group</th>
<th>Verbal IQ</th>
<th>Performance IQ</th>
<th>Full Scale IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>104 (2)</td>
<td>101 (3)</td>
<td>103 (3)</td>
</tr>
<tr>
<td>High</td>
<td>102 (4)</td>
<td>96 (4)</td>
<td>101 (3)</td>
</tr>
<tr>
<td>Low</td>
<td>73 (2)</td>
<td>79 (5)</td>
<td>73 (3)</td>
</tr>
</tbody>
</table>
The Control and High groups did not differ significantly on age ($t = -0.3; df = 30; p = 0.7$) or verbal IQ ($t = 0.6; df = 30; p = 0.6$). There was a significant difference between the Control and Low groups in verbal IQ ($t = 9; df = 28; p < 0.001$) but not in age ($t = 1; df = 28; p = 0.3$).

Receptive Language

Figure 2:1 shows the results from the test of receptive language. Statistical analysis showed that there was a significant difference between the three groups ($F(2, 41) = 10; p < 0.001$: Planned Comparison: CvH: $t = -0.009$, df = 41, $p = 0.9$; HvL: $t = 4$, df = 41, $p < 0.001$), reflecting a significantly worse performance by the Low group. This was accounted for by verbal IQ (ANCOVA: $F(2, 40) = 0.2; p = 0.9$). All the children achieved an age equivalent score of 8 years or more.

2.1.1.4 Discussion

The ratio of males to females in the Low group approximated to that found in the literature (around 4:1 [Volkmar et al. 1993]). The higher ratio in the High group is consistent with findings in individuals with comparatively high IQ [Wing 1981b; Aicardi 1992]. Unfortunately the sex composition of the Control group of children did not match the male bias of the two Autistic groups. This was a reflection of a larger proportion of girls than boys volunteering to take part in the study. The implications of this difference for the findings of this thesis will be discussed later.
There was an increased incidence of left handedness and evidence of ambidexterity in the two Autistic groups. This is consistent with Geschwind’s proposal of a higher incidence of left handedness in groups with developmental pathology [Geschwind and Behan 1982]. However the small numbers in each group preclude any conclusive statements about the validity of this suggestion.

No significant group discrepancy between verbal and performance IQ was found in the Autistic groups. Whilst a Wechsler intelligence profile characterised by a higher performance IQ than verbal IQ has been associated with Autism [Yirmiya and Sigman 1991; Rumsey 1992], this pattern is not found consistently in all Autistic individuals (see Siegel et al. 1996).

The results from the receptive language assessment reflected a similar pattern to that of the verbal IQ, suggesting that the children had similar expressive and receptive language skills. This is consistent with the literature which finds that receptive language is strongly correlated with verbal IQ (e.g. see Sparks et al. 1996).

2.2 Baseline Assessments

In this section, a series of tests are described which quantify the cognitive and behavioural profiles of the children on measures known to be affected in Autism (such as social behaviour, communication, restricted and repetitive behaviours, and motor co-ordination). Additionally the children’s attention skills are assessed. Attending is an essential part of successful completion of many of the cognitive and behavioural tasks and needs to be considered as a potential explanation of poor performance. For example, if a child has poor attentional skills, this needs to be taken into account when evaluating memory performance (as poor memory may be simply a reflection of poor attention).
2.2.1 Ratings of Autistic Behaviours

2.2.1.1 Introduction

The Autism Behavioural Checklist, a rating scale completed by parents, was used to confirm behavioural evidence of Autism in the two Autistic groups and establish the absence of such difficulties in the control group.

2.2.1.2 Methods

Autism Behaviour Checklist [Krug et al. 1993] (ABC)

This checklist is part of the Autism Screening Instrument for Educational Planning. The ABC provides a rating of the severity of a child's Autistic symptomatology and is divided into 5 subscales: Sensory, Relating, Body and Object Use, Language, and Social and Self-Help skills. Parents indicated whether each of 57 behavioural statements applied to their child at the time of study. A score over 67 suggests a high probability of a classification of Autism.

2.2.1.3 Results

Figure 2:2 and Figure 2:3 show the results from the Autism Behaviour Checklist. Figure 2:2 shows that both Autistic groups scored higher than controls on all the subscales, indicating greater impairment in all domains. Statistical analysis on the total scores (see Figure 2:3) revealed that there was a significant difference between the groups (ANOVA: F(2, 38) = 34, p < 0.001: Planned comparisons: CvH: t = -8, df = 14.7, p < 0.001; HvL: t = -2, df = 18.4, p = 0.1). This difference is a reflection of the control group scoring significantly lower than both Autistic groups. Analyses of covariance revealed that these differences were not accounted for by verbal IQ differences (ANCOVA F(2, 37) = 22, p < 0.001). Figure 2:3 also indicates that on average the Low group did exceed the threshold of 67 (Autism cut-off, see Methods), whilst the High group did not.
Study Participants and Baseline Neuropsychological Assessment

Figure 2:2 Autism Behaviour Checklist Subscale Scores. A: Sensory; B: Relating; C: Body and Object Use; D: Language; E: Social and Self-Help skills (Mean ± SEM)

Figure 2:3 Autism Behaviour Checklist: Total Score (Mean ± SEM) (Line indicates score of 67)

Figure 2:4 Autism Behaviour Checklist: Individual Scores (Line indicates a score of 67)
2.2.1.4 Discussion

The results of the Autism Behaviour Checklist demonstrated that both the High and Low group have Autistic characteristics. The High group did not score on average above the threshold of 67 on the Autism Behaviour Checklist. This is consistent with previous results using the ABC with high functioning children (see Gillberg et al. 1987) and is likely to be a reflection of the fact that it was devised for severely handicapped children and therefore is less applicable to the more intellectually able individuals.

2.2.2 Social Relationships

2.2.2.1 Introduction

As detailed in Chapter 1, individuals with Autism experience pervasive difficulties in many aspects of social behaviour. These difficulties include poor comprehension of social rules, norms and conventions, and difficulty anticipating others' knowledge and beliefs. High functioning individuals with Autism are often able to give socially appropriate responses in theoretical situations, but fail to apply this knowledge in everyday practical situations. In this section, the theoretical social skills of the participants are characterised.

2.2.2.2 Methods

British Abilities Scale: Social Reasoning [Elliott 1983]

The child was presented with seven moral dilemmas and asked in turn what he thought about each and why. For example, one item describes a situation where a mother breaks her promise to take her child to the zoo because she says that she is busy. The child's response to such items was scored according to the breadth of comprehension of the problem. The moral correctness or otherwise of the child's response did not influence the score obtained. Children were scored according to which stage of development their response indicated, as detailed in the manual. If the response was only in terms of immediate reactions or consequences of an act, the child obtained a score of 1; if the child
progressed to a broader grasp of one side of the problem only, a score of 2 was awarded; if the child attempted to find reasons or explanations for a situation from both sides, a score of 3 was awarded; or finally, if the child ceased to be bound by the particulars of the problem and was able to see it with context and articulate it as an example of more general difficulties, a score of 4 was awarded. The median score was taken to be a reflection of the estimated developmental stage for the child.

Social Judgement
Fifteen short stories describing social situations incorporating behaviour that was either normative or a violation of social norms were read to the child. These stories were adapted by Blair from the original version by Dewey (see Chap 6 in Frith 1991). For example, in one story a boy falls asleep on the carpet of his headmaster’s office whilst he is waiting to see the headmaster. At various points in each story, the child was asked to comment on how appropriate the behaviour was, giving a score from A to C. A rating of ‘A’ meant that the child judged the situation as normative. ‘B’ and ‘C’ scores meant that he judged the situation as a norm violation and indexed the extent of the violation (B being mild and C being serious).

Blair has collected normative data on this task which have led to consistent identification of which situations are normative and which contain violations (see Blair and Cipolotti 2000). The percentage of social norms correctly labelled as ‘A’ and the percentage of social violations correctly labelled by the child as either ‘B’ or ‘C’ were calculated.

2.2.2.3 Results

British Abilities Scale: Social Reasoning

Figure 2:5 shows the results from the social reasoning of the British Abilities scale. Statistical analysis revealed that there was a significant difference between the groups (ANOVA: F(2, 39) = 8, p = 0.001: Planned comparisons: CvH: t = 1, df = 41, p = 0.2; HvL: t = 2, df = 41, p = 0.02). This was a reflection of significantly poor performance by the
Low group. Analyses of covariance revealed that these differences were accounted for by age and verbal IQ differences (ANCOVA F(2, 38) = 1, p = 0.4).

Figure 2:5 British Abilities Scale: Social Reasoning (Mean ± SEM)

Social Judgement

As Figure 2:6 shows, there was no significant difference between the percentage of correctly categorised social norms between the three groups (ANOVA: F(2, 41) = 0.7, p = 0.5). There was a significant group effect between the percentage of correctly categorised social violations (ANOVA: F(2, 41) = 6, p = 0.004; Planned comparisons: CvH: t = 1, df = 18.1, p = 0.3; HvL: t = 2, df = 14.8 p = 0.07). This reflected a significant difference between the control group and the Low group, with the Low group correctly categorising significantly fewer situations. This was not accounted for by verbal IQ or age.
Figure 2.6 Dewey Stories a: Percentage of correctly categorised social norms; b: Percentage of correctly categorised social violations

2.2.2.4 Discussion

The British Abilities Scale Social Reasoning subtest did not detect any gross differences between the three groups, when VIQ and age were taken into account. However, using a more subtle assessment, the Dewey stories, the Low group was characterised as having difficulties in the areas of social relationships. These results suggest that both Autistic groups had good basic theoretical social knowledge, and the High group was able to apply this knowledge even in complex hypothetical situations. It should be noted that in both the British Abilities Social Reasoning subscale and in the categorisation of the social norms of the Dewey stories, the control group performed worse than might have been predicted. This may be partly responsible for the lack of differences found between the groups.

The Dewey stories have been used in another study of young people with high functioning Autism. This study found that the Autistic group performed differently from the control group [Ellis et al. 1994]. However, the use of the adult version of the stories and a more arbitrary scoring method renders across-study comparisons difficult.
These results should not be taken to suggest that the children with Autism have no difficulty with social situations. It is a common finding that theoretical knowledge in high functioning individuals with Autism is good, yet there is a failure to apply this knowledge in everyday situations (e.g. Bowler 1992). Detecting deficits in social behaviour using laboratory tasks in such individuals requires more complex assessments. Evidence of the social difficulties in practical situations that the children in the Autistic groups experience is documented by the Autism Behaviour Checklist (see above) and Children’s Communication Checklist (see below).

2.2.3 Communication

2.2.3.1 Introduction

Whilst the children included in this study had good basic language (as evidenced by their receptive and expressive language (see above)), high functioning individuals with Autism often have difficulty with the more complex aspects of language (such as pragmatics, metaphor and humour). In the following section, these aspects of the children’s language skills are documented.

2.2.3.2 Methods

Right Hemisphere Language Battery [Bryan 1989]

Metaphor Picture Test A list of 11 sentences, each containing a common metaphor, was presented. The metaphors were either psychological-physical (i.e. an adjective drawn from the physical world is used to express a psychological state: she left the scene of the accident with a heavy heart) or cross-sensory (i.e. an adjective drawn from one sensory modality modifies an element from a different sensory domain: he must have green fingers as his garden is lovely). A set of four pictures was supplied for each sentence: a picture representing the correct metaphorical meaning, a picture representing the literal meaning and two control pictures depicting one aspect of the sentence. The child was asked to point to the picture which matched the meaning of the sentence.
Written Metaphor Test    In this task, 11 common metaphors are incorporated into 11 contextual sentences (The leader gave the group a tall order). Each sentence is printed on a card and is followed by three randomised sentences which represent possible meanings of the target. These are the genuine metaphorical meaning, a primitive metaphorical meaning that only focuses on an incidental aspect of the metaphor and a metonymical meaning that merely replaces the sentence so that two terms of the metaphor are interpreted literally without defying realism. The card was placed in front of the child and the sentences read out by the examiner. The patient was asked to point to the sentence that explained the target sentence.

Inference Test    This test assesses the ability to comprehend aspects of inferential meaning in four short paragraphs printed on separate cards which describe a situation or event. The answers all require comprehension of information not made explicit in the passage. The card was placed in front of the child, and the examiner read the passage. The child was then asked each question in turn and was able to refer to the card during this time.

Joke Test    In this task, 11 jokes with clear punchlines are printed separately on cards with a choice of four punchlines. The choice of the punchlines is: the actual punchline; a straight ending of neutral content; a straight ending of emotional content and a surprise ending that does not relate to the body of the joke. The card was placed in front of the child and the examiner read the joke and the alternative endings. The child was told that the joke needed finishing and asked to point to the ending that would make it funny.

Children’s Communication Checklist [Bishop 1998]    The Children’s Communication Checklist assesses aspects of communicative impairment not usually evaluated by other contemporary standardised language tests. These predominantly include pragmatic abnormalities seen in social communication, although restricted interests and difficulties with social relationships are also included. The checklist consists of a series of statements describing aspects of children’s behaviour and each statement is rated as “does not apply”, “applies somewhat”, “definitely applies” or “unable to judge”. The checklist is divided into subscales: A: Speech output: Intelligibility and fluency; B: Syntax; C: Inappropriate initiation; D: Coherence; E: Stereotyped conversation;
Study Participants and Baseline Neuropsychological Assessment

F: Use of conversational context; G: Conversational rapport; H: Social relationships; I: Interests.

The checklist was completed by a parent. The majority of the statements referred to the child’s difficulties. Scores of 2 points were awarded for responses of “definitely applies”, 1 point for “applies somewhat”. Statements describing a child’s strengths were scored negatively (-2 points were awarded for “definitely applies” responses, etc.). 30 points were added to each subscale to ensure that all subscale scores were positive (see Bishop 1998). “Unable to judge” answers were prorated within each subsection.

2.2.3.3 Results

Right Hemisphere Language Battery

Figure 2:7 shows the results from the Right Hemisphere Language Battery. Statistical analysis revealed that there was a significant difference between the groups on all four subtests (Picture Metaphor: ANOVA: F(2, 40) = 4, p = 0.02; Planned Comparison: CvH: t = -0.6, df = 40, p = 0.6; HvL: t = 3, df = 40, p = 0.01; Written Metaphor: ANOVA: F(2, 41) = 8, p = 0.002; Planned Comparison: CvH: t = 1, df = 41, p = 0.2; HvL: t = 3, df = 41, p = 0.02; Inference: ANOVA: F(2, 40) = 11, p < 0.001; Planned Comparison: CvH: t = 2, df = 40, p = 0.07; HvL: t = 3, df = 40, p = 0.007; Joke: ANOVA: F(2, 40) = 7, p = 0.002: Planned Comparison: CvH: t = 0.08, df = 40, p = 0.9; HvL: t = 3, df = 40, p = 0.002). In each case this is a reflection of significantly poorer performance by the Low group. ANCOVA analyses revealed that these differences could be accounted for by age, verbal IQ and basic reading skills differences (Picture Metaphor: ANCOVA F(2, 39) = 0.4, p = 0.7; Written Metaphor: ANCOVA F(2, 39) = 1, p = 0.3; Inference: ANCOVA F(2, 38) = 3, p = 0.06; Joke: ANCOVA F(2, 38) = 0.1, p = 0.9).
Study Participants and Baseline Neuropsychological Assessment

Figure 2:7 Right Hemisphere Language Battery Results: a) Picture Metaphor; b) Written Metaphor; c) Inference; d) Joke (Mean ± SEM)

![Graphs showing results for different groups: (a) Picture Metaphor, (b) Written Metaphor, (c) Inference, (d) Joke.](image)

Figure 2:8 and Figure 2:9 show the results from the Children's Communication Checklist. Figure 2:8 shows that whilst the three groups were similar in the subtests of speech output, syntax and use of conversational context, both Autistic groups scored higher (which indicates greater impairment) on the subscales measuring speech initiation, coherence, stereotyped conversation, conversation rapport, social relationships and interests. Statistical analysis on the total scores (see Figure 2:9) revealed that there was a significant difference between the groups (ANOVA: F(2, 39) = 58, p < 0.001: Planned Comparison: CvH: t = -8, df = 39, p < 0.001; HvL: t = -2, df = 39, p = 0.03). This finding is a reflection of significant differences between the High and Control groups, and the High and Low groups. ANCOVA analyses revealed that these differences were not accounted for by age or verbal IQ differences (ANCOVA F(2, 38) = 37, p < 0.001).
Study Participants and Baseline Neuropsychological Assessment

Figure 2:8 Subtests of Children's Communication Checklist
A: Speech output: Intelligibility and fluency; B: Syntax; C: Inappropriate initiation; D: Coherence; E: Stereotyped conversation; F: Use of conversational context; G: Conversational rapport; H: Social relationships; I: Interests (Mean ± SEM)

![Graph showing scores for each subtest across control, high, and low groups.]

Figure 2:9 Children's Communication Checklist: Total Score (Mean ± SEM)

![Bar graph showing total scores across control, high, and low groups.]

2.2.3.4 Discussion

The children with Autism did not have difficulty with interpretation of metaphors, joke completion or inference from a story, once VIQ and age had been taken into account. The same battery of assessments has been used with adults with Asperger's Syndrome [Ellis and Gunter 1999]. They only reported abnormalities in the joke completion task. Thus
deficits in humour appreciation in the Autistic groups may be masked by developmental trends. i.e. the performance of the control children might be expected to improve with age, leaving the children with Autism further behind.

Although it might be considered unexpected that the children in the Autistic groups were able to complete the metaphor tasks, a number of factors should be considered. Firstly, many of the children with Autism indicated spontaneously during the assessment that they had been taught what metaphors meant. This suggests that earlier in their development, they may indeed have had difficulty with metaphor interpretation. Subtle difficulties with literal interpretation were also noted during testing, for example a number of children interpreted instructions starting ‘In a minute I want you to …’ literally. This once again stressed the children’s difficulty with applying their theoretical knowledge in practical situations.

Although quantitative behavioural assessment of the children did not reveal difficulties in communication, the results of the CCC suggested that they had widespread difficulties with many aspects of pragmatics and social communication. Further, the results from the social relationships subtest underline the social difficulties they experience despite their theoretical knowledge of social rules and conventions (see above).

### 2.2.4 Repetitive and Restricted Behaviours

#### 2.2.4.1 Introduction

The third diagnostic feature of Autism (repetitive and restricted behaviours and interests) has received considerably less attention than the first two features. The obsessions and compulsions may include rigid routines, and narrow interests and preoccupations as well as more frankly obsessional symptoms such as touching compulsions (see Baron-Cohen 1989 for discussion). In this section a parental checklist was used to index the obsessions and compulsions of the cohort of children included in this thesis.
2.2.4.2 Methods

The Children’s Obsessions and Compulsions Inventory [Shaffran 2000] (ChOCI) provides ratings of incidence and severity of obsessions and compulsions in children and adolescents. The checklist was completed by a parent.

2.2.4.3 Results

Figure 2:10 shows the results from the ChOCI. Both Autistic groups scored higher (indicating greater impairment) on the obsessions and compulsions subscales. Statistical analysis on the scores revealed that there was a significant difference between the groups (Compulsions: ANOVA: F(2, 38) = 14, p < 0.001: Planned Comparison: CvH: t = -4, df = 16.5, p = 0.001; HvL: t = -1, df = 22.1, p = 0.3; Obsessions: ANOVA: F(2, 38) = 8, p = 0.001: Planned Comparison: CvH: t = -4, df = 13.6, p = 0.001; HvL: t = -0.2, df = 17.8, p = 0.8). This is a reflection of the control group scoring significantly lower (indicating less impairment) than both Autistic groups. ANCOVA analyses revealed that these differences were not accounted for by age or verbal IQ differences.

Figure 2:10 ChOCI Results a) Compulsions; b) Obsessions (Minimum score = 10; Mean ± SEM)
2.2.4.4 Discussion

Both groups of children with Autism demonstrated the anticipated pattern of abnormal levels of obsessions and compulsions. No previous studies using the ChOCI with children with Autism have been reported. However, the findings of this section are consistent with numerous references to the obsessions and compulsions of children with Autism (e.g. ICD-10 1992; Lord and Rutter 1994; DSM-IV 1999).

2.2.5 Motor Coordination

2.2.5.1 Introduction

Poor motor co-ordination in children with Autism has been noted by several researchers (e.g. Mawson et al. 1985; Tantam 1988a; Cox 1991; Manjiviona and Prior 1995). This motor impairment may be indicative of neural dysfunction (such as cerebellum abnormality, see Courchesne et al. 1988) and may be suggestive of underlying aetiological factors [Manjiviona and Prior 1995].

2.2.5.2 Methods

The Movement Assessment Battery for Children Checklist [Henderson and Sugden 1992] is part of the Movement ABC which evolved from the Test of Motor Impairment [Stott et al. 1972], a reliable and valid test of clumsiness [Sugden and Wann 1987; Wall 1982]. The test is divided into four sections assessing skills in a stable environment with either the child stationary or moving, and skills in a moving environment with either the child stationary or moving. Each item consists of a description of a task and is rated on a scale of 0-3 (0 = child can perform task very well and 3 = child is not close to performing task). The checklist was completed by a parent.
2.2.5.3 Results

Figure 2:11 and Figure 2:12 show the results from the M-ABC. Figure 2:11 shows that both Autistic groups scored higher (which indicates greater impairment) on all the subscales, with the Low group scoring consistently higher than the High group. Statistical analysis on the total scores (see Figure 2:12) revealed that there was a significant difference between the groups (ANOVA: F(2, 36) = 20, p < 0.001: Planned Comparison: CvH: t = -4, df = 12.2, p = 0.002; HvL: t = -2, df = 18.8, p = 0.04). This is a reflection of significant differences between the High and Control groups, and between the High and Low groups. ANCOVA analyses revealed that these differences were not accounted for by age and verbal IQ differences (ANCOVA F(2, 35) = 12, p < 0.001).

Figure 2:11  M-ABC: Subscales (Mean±SEM) CS/ES: Child stationary, Environment stationary; CM/ES: Child moving, Environment stationary; CS/EM: Child stationary, Environment moving; CM/EM: Child moving, Environment moving
2.2.5.4 Discussion

Both groups of children with Autism were reported by parents to have significant problems with many aspects of motor co-ordination. This is consistent with recent reports of motor difficulties in the Autistic spectrum [ICD-10 1992; Teitelbaum et al. 1998].

2.2.6 Attention

2.2.6.1 Introduction

A number of studies have shown that Autistic children have poor attention (e.g.: sustaining attention: [Garretson et al. 1990]; shifting attention: [Courchesne et al. 1994b; Courchesne et al. 1994c]; joint attention: [Loveland and Landry 1986]). Indeed some researchers have suggested that deficits in attention are a primary feature of Autism (e.g. Courchesne et al. 1994b; Pierce et al. 1997). If a child is not attending to a task, interpretation of any deficit in performance is difficult to attribute to the cognitive skill assessed by the task. Obtaining measures of attention will allow assessment of the role of this confounding factor in the children’s performance of other tasks.
One particular aspect of attention that has recently been highlighted as abnormal in Autism is local-global attention. Individuals with Autism have good (or superior) local (or detail) processing, with relatively poor global (or gestalt) processing (see Chapter 10 of Happé 1994 for discussion). Attending to local information at the expense of global information may confound performance on some cognitive tasks. A measure of attention to local visual information is therefore also included.

2.2.6.2 Methods

Attention

The Test of Everyday Attention in Children (TEA-Ch) [Manly et al. 1999] was used to obtain a standardised measure of attention. This test assesses the following domains: selective attention, sustained attention and divided attention. In the selective attention subtest, the child had to find as many ‘target’ spaceships as possible on a sheet filled with very similar distractor spaceships. In the second part of the task there are no distractors. Subtracting part 2 from part 1 gives a measure of a child’s ability to make this selection that is relatively free from the influence of motor speed. In the sustained attention subtest, the child had to keep count of the number of 'scoring' sounds they heard on a tape, as if they were keeping score on a computer game. This task seems easy and due to the long gaps between the sounds does little to grab the child’s attention. In the divided attention subtest, the child was asked to combine the two above tasks of finding the spaceships and keeping a count of scoring sounds. Scoring was done in accordance with the manual, which provides age and sex matched norms.

Local Visual Attention

The Embedded Figures Test [Benton and Spreen 1969] was used to obtain a measure of visual perception and to detect differences in global versus local visual processing.
2.2.6.3 Results

Attention

Figure 2:13 shows the results from the TEA-Ch. Statistical analysis showed that there was no significant difference between the three groups on the subtest of selective attention (ANOVA: F(2, 39) = 2, p = 0.2). Statistical differences between the three groups were found on the subtests of sustained and divided attention (Sustained: F(2, 41) = 6; p = 0.008: Planned Comparison: CvH: t = 3, df = 41, p = 0.005; HvL: t = -0.3, df = 41, p = 0.8; Divided: F(2, 39) = 13, p < 0.001: Planned Comparison: CvH: t = 2, df = 39, p = 0.06; HvL: t = 3, df = 39, p = 0.004). The differences were due to the controls scoring significantly higher in the sustained attention subtest and the Low group scoring significantly lower in the divided attention subtest. In contrast to the sustained attention differences, the divided attention subtest differences could be accounted for by Verbal IQ (ANCOVA F(2, 38) = 0.3, p = 0.7).
Study Participants and Baseline Neuropsychological Assessment

Figure 2:13 Results from Attention Assessment a) Selective attention; b) Sustained attention c) Divided attention (Mean ± SEM)

![Graph showing results from Attention Assessment](image)

Local Visual Attention

Figure 2:14 shows the results from the Embedded Figures Test. There was no significant difference between the three groups on the time taken to find the embedded figure in each stimulus. The Low group performed significantly worse than the remaining two groups on total number of figures completed correctly (ANOVA: F(2, 41) = 7, p = 0.003: Planned Comparison: CvH: t = 0.2, df = 26.4, p = 0.9; HvL: t = 2, df = 12.9, p = 0.004). Analysis of covariance determined that this could be accounted for by verbal IQ (ANCOVA F(2, 40) = 0.8; p = 0.4).
2.2.6.4 Discussion

These results suggest that both the High and Low groups had good selective attention, and poorer sustained attention. Deficits in sustained attention have also been reported in other studies investigating children with Autism [Garretson et al. 1990]. The Low group also had significantly poorer divided attention than the High group. In addition, the High group found divided attention tasks difficult compared with controls, although this did not quite reach significance. Previous reports have also found that children with Autism perform worse than controls on divided attention tasks (e.g. Ciesielski et al. 1995).

In view of the weak central coherence theory recently proposed, it is perhaps surprising that the results from the Embedded Figures Tests do not demonstrate superior local attention in individuals with Autism. However, in the literature there are mixed reports of the performance of individuals with Autism on this task. Whilst some reports describe superior skill in both high and low functioning individuals with Autism [Shah and Frith 1983; Jolliffe and Baron-Cohen 1997], others have failed to replicate this finding [Brian and Bryson 1996]. It is possible that individuals with Autism only show superior disembedding when the figures are meaningful (perhaps because they are less distracted by meaning than controls). If this is the case, the use of complex meaningless figures in this study would
explain the results. Alternatively, differences in disembedding skills may only be detectable at certain developmental windows, with the controls initially lagging behind the children with Autism.

2.3 **Summary**

Three groups of children were studied in this thesis: the Control group, the High group and the Low group. The High group consisted of children with a diagnosis of high functioning Autism (or Asperger's Syndrome) and a verbal IQ above 85. The Low group was comprised of children also with a diagnosis of high functioning Autism (or Asperger's Syndrome) but with a verbal IQ below 85. The Control group did not differ from the High group on age or verbal IQ.

On baseline assessments, the High group were significantly different from the control group on measures of Autistic behaviours, obsessions and compulsions, motor coordination and pragmatic and social communication. The children in the High group had good language skills (understanding metaphors, etc.) and a good theoretical knowledge of social norms and rules. They also had good selective attention, were non-significantly impaired at divided attention and showed poor sustained attention compared to the Control group.

The Low group were more impaired than the High group at most of the baseline tasks. However, when verbal IQ was included in the analysis, they showed a very similar overall pattern to the High group. The only additional deficit displayed that was not accounted for by linear correlations in verbal IQ was poor theoretical understanding of social norms and violations.
Chapter 3 Memory

A widely documented consequence of medial temporal lobe damage, in humans and animals alike, is a deficit in explicit memory. Despite the evidence supporting medial temporal lobe abnormality in Autism, few studies have systematically investigated explicit memory function in this disorder. In this chapter a comprehensive assessment of explicit memory is reported, with each section detailing investigations into a different aspect of explicit memory. Prior to these sections, the anatomical basis of explicit memory is reviewed.

3.1 Introduction

A wide variety of different terminologies and categorisations have been applied to the study of memory. Whilst the term ‘memory’ usually refers to memory for events and facts, it can also include memories for motor procedures, such as riding a bike or playing a musical instrument. These memories tend to be unconscious (i.e. we are not aware of them) and may be termed ‘implicit’ [Graf and Schacter 1985].

This thesis will only discuss explicit memory (i.e. memory for events, places, objects, facts, etc.). These memories are explicit inasmuch as they are accompanied by conscious awareness.

3.1.1 Categories Of Explicit Memory

One of the most important components of explicit memory is episodic memory. First introduced by Tulving, this term refers to memory for events from the past in spatial and temporal context [Tulving 1972]. Explicit memories that are non-episodic include much of our semantic memory (i.e. memory for facts and general knowledge about the world).
Explicit memories can be assessed on the basis of familiarity and/or recollection. Thus, recognition that a stimulus has been encountered before may be made on the basis of familiarity (i.e. ‘knowing’ the stimulus) and/or on the basis of recollection of the contextual or episodic information about the stimulus (i.e. ‘remembering’ or recollection the stimulus) (see Mandler 1980). It should be noted that recollection and familiarity are hierarchical constructs: a stimulus that is recollected (with contextual information) must necessarily also be known (on the basis of familiarity).

The relationship between these components of memory (episodic and non-episodic memory, familiarity and recollection) is complex. Episodic memory may rely mainly on the process of recollection, whereas non-episodic memory may recruit mainly familiarity processes. Clearly recollection and episodic memories share many similar features and one aspect of non-episodic memory is likely to be familiarity-based recognition.

### 3.1.2 Assessment Of Explicit Memory

Assessing different aspects of explicit memory independently is problematic. There are two main methods of measuring explicit memory: recognition and recall. In recognition paradigms, subjects are asked if they have seen (or heard, etc.) stimuli currently being presented to them. Successful recognition performance may not necessarily require recollection and may be achieved on the basis of familiarity alone. Poor performance on recognition tasks may therefore suggest deficits in familiarity processes.

Recall paradigms usually require subjects to describe previously presented stimuli with little or no prompting (or cueing). The importance of recollection of an episode in recall paradigms can be manipulated according to the number of presentations of the stimulus prior to recall and its cueing. Repeated presentations and cueing are likely to reduce the role of recollection in episodic memory [Tulving 2001].

Recently, new behavioural paradigms have been developed to assess recollection versus familiarity processes in isolation (i.e. ‘remember-know’ paradigms or levels of processing)
In these paradigms, the subject is asked if he recognises the stimulus on the basis of recall of contextual information pertaining to its presentation (remember) or if he merely recognises the stimulus because he 'knows' that it is familiar. These paradigms were not used in this study. Previous pilot studies at Great Ormond Street Hospital have revealed difficulties in explaining to children the distinction between 'remember' and 'know'.

3.1.3 Neural Basis Of Explicit Memory

Investigations into the neural basis of explicit memory began over a century ago. The temporal lobes were first suggested to play a critical role in memory in 1900 by Bekhterev [Bekhterev 1900 quoted in Kolb and Whishaw 1996]. In 1954 the importance of the medial temporal lobes was clearly demonstrated by the case of HM [Scoville 1954] who underwent bilateral temporal lobe resections for the relief of epilepsy [Corkin et al. 1997]. Following this surgery he suffered from severe anterograde amnesia [Scoville 1954]. Since then a wealth of clinical and experimental data have accumulated to demonstrate the importance of the medial temporal lobe structures in memory (e.g. see Milner 1970).

Different components of the medial temporal lobes are important for different aspects of explicit memory functions [Squire and Knowlton 2000]. This is indicated by results from lesion studies (in both humans and animals), functional imaging studies, and electrophysiological studies, as well as from the anatomy of the medial temporal lobes. Although the exact roles of the individual regions that are necessary and/or sufficient for different aspects of explicit memory are currently unresolved, the hippocampal formation is thought to be critical for episodic memory, whilst the parahippocampal region (particularly the perirhinal cortex) is hypothesised to support non-episodic memory, including familiarity judgements and simple associations [Mishkin et al. 1997].
3.1.3.1 The Parahippocampal Region And Non-Episodic Explicit Memories

Many research studies support the hypothesis that the parahippocampal region (but not the hippocampal formation) is necessary for many aspects of non-episodic explicit memories, including familiarity judgements and semantic knowledge.

Lesions to the parahippocampal region in monkeys lead to severe deficits in recognition memory [Horel et al. 1986; Zola-Morgan et al. 1989a; Meunier et al. 1993; Elliott et al. 1997]. As indicated earlier, this would suggest a deficit in familiarity-based recognition. The severity of impairment in recognition is the same even when the lesion is restricted to only the perirhinal cortex [Meunier et al. 1993], underlining the particular role of this region in recognition. In contrast, damage to the hippocampal formation alone usually results in a much milder (and in some cases no) recognition deficit [Gaffan et al. 1984; Bachevalier et al. 1985; Zola-Morgan et al. 1989b; Murray and Mishkin 1998] suggesting that familiarity-based recognition does not require the hippocampal formation. However, it should be noted that Squire and colleagues have found impaired recognition following hippocampal damage in monkeys [Zola et al. 2000]. It is not understood why such different results have been obtained. Methodological variations (such as removal of monkeys from the apparatus between presentation and test, and differences in lesion extent) may play a role.

The effects of bilateral lesions of the parahippocampal regions in humans have not been reported, as such selective damage rarely occurs. However two patient groups (patients with semantic dementia and patients with lesions encompassing both the parahippocampal region and the hippocampal formation) offer some insights into the neural basis of non-episodic explicit memory.

Semantic dementia is a progressive neurodegenerative disease. The early cognitive profile of individuals with semantic dementia is a selective loss of semantic knowledge [Hodges et al. 1992]. Performance on recognition tasks is normal (unpublished data reported in
Graham et al. 1999; Simons et al. 2000). The underlying neuropathology of semantic dementia is the degeneration of the cortical, inferior and possibly ventral portions of the temporal lobes. Although initially thought to be selectively affecting inferolateral temporal cortex with sparing of the hippocampal formation [Hodges et al. 1992], recent studies suggest that this degeneration also affects the hippocampal formation, even in the early stages of the disease [Galton et al. 2000].

The complex nature of a neurodegenerative process means that the functions of the areas affected by the disease need not necessarily be totally disrupted. Similarly, the integrity of areas apparently unaffected can not be assumed. Interpretation of the implications of the profile of semantic dementia with regards to the anatomy must therefore be guarded. With this caveat in mind, two important points should be made. In drawing comparisons with patients with Alzheimer's disease, it might be argued that the parahippocampal region is involved in semantic knowledge. Patients with Alzheimer’s disease show selective degeneration of the hippocampal formation, and the typical cognitive profile includes grossly intact semantic knowledge [Hodges and Patterson 1995]. Secondly, the intact recognition skills of individuals with semantic dementia underlines the possibility that recognition and semantic knowledge are independently supported, but should not be taken as evidence that recognition is independent of the medial temporal lobes.

The second type of patients who offer some insight into the neural underpinnings of non-episodic explicit memory are those with lesions to the hippocampal formation and surrounding cortex. These individuals have been reported to have significantly worse recognition memory than individuals with selective hippocampal damage [Aggleton and Shaw 1996; Reed et al. 1997; Buffalo et al. 1998]. This finding is consistent with the parahippocampal region supporting recognition. However it might also be a reflection of the compounding effects of lesions to the parahippocampal region and the hippocampal formation. Selective damage to either may (at least in theory) leave recognition intact, whilst damage to both produces a recognition impairment.

Physiological studies monitoring the activity of neurons in the perirhinal cortex support the hypothesis that the perirhinal cortex is important in familiarity-based judgements.
Perirhinal neurons show encoding of the necessary information to allow solution of a wide range of recognition tasks not requiring spatial and contextual discriminations (see Brown and Xiang 1998). In contrast, no other area in the medial temporal lobe has such extensive encoding capacity. For example, the cells in the hippocampal formation show no evidence of specific stimulus encoding during recognition tasks [Brown et al. 1987; Sakurai 1990; Riches et al. 1991].

Many functional imaging studies of recognition per se report both hippocampal formation and parahippocampal region activation (e.g. Stern et al. 1996; Stark and Squire 2000). This might be taken as an indication that the hippocampal formation is involved in familiarity-based judgements. However, successful performance on recognition tasks in normal controls is likely to be supported by both recollection and familiarity. Recently the ‘remember-know’ paradigm has been used in imaging studies to show that hippocampal activation is only associated with recollection, and not familiarity judgements [Eldridge et al. 2000]. These findings therefore support the role of the parahippocampal region in familiarity judgements.

In summary, research findings are consistent with the hypothesis that the parahippocampal region plays an important role in non-episodic explicit memory. At the present time, however, the necessity of the parahippocampal region for non-episodic explicit memories has not been conclusively demonstrated.

3.1.3.2 The Hippocampal Formation Is Necessary For Contextual Memory

There is a wealth of evidence implicating the hippocampal formation in contextual memory. Many different techniques have shown that retrieval of spatial and temporal information (critical elements in contextual memory) is highly dependent on the integrity of this structure.

Damage to the hippocampal formation results in deficits in association of events across time. This is clearly demonstrated by the differential effects of hippocampal lesions on two
different types of conditioning. In delayed conditioning, a tone (CS) precedes and overlaps with an unconditioned stimulus (UCS) that evokes a reflexive eyeblink [Pavlov 1927]. The subject learns to blink during the presentation of the CS. In trace conditioning, the CS onset is abbreviated so there is a short (~500 ms) delay between the end of the CS and onset of the UCS [Pavlov 1927], thus introducing into trace conditioning a temporal context. Lesions to the hippocampal formation in animals interrupt trace but not delay conditioning (animals: Solomon et al. 1986; Moyer et al. 1990; Kim et al. 1995; McEchron et al. 1998; humans: McGlinchey-Berroth et al. 1997; Clark and Squire 1998).

Memory for spatial information is also impaired in animals with lesions to the hippocampal formation. Rats have difficulty finding the platform in the Morris Water Maze when the cues are spatially distributed throughout the maze [Save and Poucet 2000]. Information about spatial arrays (but not individual object recognition) is also disrupted in monkeys with hippocampal damage [Gaffan 1994].

Animals with hippocampal formation lesions process contextual information abnormally, showing unusual dependence on irrelevant contexts and inefficient processing of relevant contexts. Firstly, they tend to process irrelevant or non-predictive information more than controls do. For example, in control animals, the strength of the conditioning to the background is predicted by the reliability with which the background predicts the unconditioned stimulus. However, rats with hippocampal lesions condition strongly to context, even when the existence of an explicit conditioned stimulus precludes context conditioning in the control rats [Winocur et al. 1987].

Animals with hippocampal lesions also have difficulty learning new tasks in the same context as a previously learnt task [Winocur and Gilbert 1984], and performing a previously learnt task in a new context [Winocur and Olds 1978]. Gaffan reported that monkeys with fornix transection were able to perform an object discrimination learning task with varying backgrounds, but were unable to perform the same task with a unique background [Gaffan 1994]. This suggests further that abnormal contextual processing can follow disruption of the hippocampal formation.
Animals with hippocampal lesions also process relevant contextual cues less efficiently. Good and Honey showed that animals with hippocampal lesions failed to learn that a cue in context A signalled a reward, whilst the same cue in context B signalled no reward [Good and Honey 1991].

Cellular recordings have demonstrated that there are cells in the hippocampal formation that encode spatial location (so called ‘place’ cells) (e.g. O'Keefe 1976). A number of cellular protein changes in the hippocampal formation have been associated with performance on contextual memory tasks in animals (e.g. Atkins et al. 1998; Woolf et al. 1999). Such contextual learning paradigms have also been shown to increase neuronal cell generation in the adult animal’s dentate gyrus. This effect is not seen in non-contextual learning paradigms (such as delay conditioning) [Gould et al. 1999].

Individuals with damage to the hippocampal formation are unable to remember episodic information, but may learn at least some semantic information. In particular, individuals with developmental hippocampal atrophy have been shown to have a selective episodic impairment with relatively intact semantic memory [Vargha-Khadem et al. 1997a]. One individual with such damage has been shown to have intact electrophysiological markers of familiarity but not recollection [Duzel et al. 2001]. A number of functional imaging studies have shown that contextual memory retrieval is associated with hippocampal formation activation (e.g. Eldridge et al. 2000; Burgess et al. 2001).

3.1.4 Explicit Memory In Autism

Although there have been a large number of studies investigating explicit memory skills in Autism, few have examined comprehensively different aspects of explicit memory, with most instead focussing on only one aspect of explicit memory, such as list learning [Minshew and Goldstein 1993]. This, combined with the wide variety of ability levels in those with Autism, makes it difficult to establish the characteristic explicit memory profile of such individuals. However, the evidence is consistent with selective impairment in episodic memory in individuals with Autism. For example, recognition, cued recall and
recall tasks based on repeated presentation of stimuli are intact, whilst recall after single 
presentation and recall of contextual memory are impaired (e.g. Boucher and Warrington 
1976; Boucher 1978; Boucher 1981; Ameli et al. 1988; Boucher and Lewis 1989; 
Minshew and Goldstein 1993; Bennetto et al. 1996; Brian and Bryson 1996; Bowler et al. 
1997; Farrant et al. 1998). Individuals with Autism and severe mental retardation show a 
less consistent and more generalised pattern of deficit (e.g. Barth et al. 1995).

The following sections describe the results obtained from a comprehensive memory 
protocol for children participating in this study. Each section focuses on a different aspect 
of explicit memory. As described earlier, few memory paradigms differentiate between 
episodic and non-episodic memories. However, the various measures are ordered, such that 
there is an increasing dependence on episodic memory. Thus, the first sections deal with 
tasks that may be successfully completed on the basis of familiarity judgements (such as 
recognition and semantic memory). These sections are followed by those on cued recall, 
recall of information with repeated presentation of stimuli, recall of information with single 
presentation of stimuli and finally recall of contextual information.

Each section begins with a review of findings using a particular paradigm in individuals 
with Autism. In recognition of the importance of the functioning levels of the study 
participants, wherever possible details of the intellectual abilities of the participants of the 
studies reviewed are included in brackets. Where available, chronological age (CA), verbal 
IQ (VIQ) and nonverbal IQ (NVIQ) (or performance IQ (PIQ)) are quoted. Some studies 
do not provide IQ details, and so verbal mental age (VMA) and nonverbal mental age 
(NVMA) are quoted instead.
3.2 Recognition

3.2.1 Introduction

Individuals with high functioning Autism have been shown to have good recognition skills for a variety of different types of stimuli. For example, Farrant et al. showed that children with Autism (CA 12:7, VMA 7:8) were able to correctly recognise words that were previously presented to them [Farrant et al. 1998]. Similarly, Bennetto et al. found that children with Autism (CA 16, VIQ 82) were able to recognise pictures and words that had been previously presented to them, even though they were unable to temporally order them [Bennetto et al. 1996]. Good picture recognition has also been reported by Ameli et al. [1988] and Brian and Bryson [1996].

However, comparisons between recognition of objects with and without agency (defined as objects capable of self-propelled motion, both animate and inanimate) have led to suggestions that individuals with Autism have a recognition deficit for objects with agency. For example Blair et al. reported deficits in recognition of stimuli with agency (such as faces) but intact recognition for objects without agency (such as leaves and buildings) (CA 30, VIQ 88, PIQ 95) [Blair et al. 2001]. Ellis et al. showed normal word recognition but impaired face recognition in a group of individuals with high functioning Autism (CA 15.6, VIQ 96, NVIQ 86) [Ellis et al. 1994]. Boucher and Lewis showed that children with Autism (CA 13:6, VMA 6/7) were able to perform a delayed matching to sample task with buildings as well as controls, but were impaired on the recognition of unfamiliar faces [Boucher and Lewis 1992].

Findings are less consistent concerning the recognition of familiar faces in Autism, with some studies finding an impairment (e.g. Boucher et al. 1998), and others reporting no difference compared to controls [Langdell 1978]. Such inconsistency may be due to different levels of familiarity with the face stimuli and/or intellectual abilities of the participants.
The recognition impairment in individuals with low functioning Autism has been found to extend to non-agency objects. For example, Boucher and Warrington found significant impairment in a group of children with Autism (CA 13, VMA 5, NVMA 9) for recognition of pictures of common objects [Boucher and Warrington 1976]. Similarly, Barth et al. found impaired performance on a delayed match-to-sample task with brightly coloured shape stimuli, in the low functioning Autism group (CA 5, VIQ 66, NVIQ 640) but not in the high functioning Autism group (CA 5, VIQ 82, NVIQ 1013) [Barth et al. 1995].

Recognition tests using a variety of stimuli (including words, faces, doors) were therefore included in the present memory protocol.

### 3.2.2 Methods

The tests used to assess recognition memory were taken from two test batteries: the Children’s Memory Scale [Cohen 1997] and the Doors and People Test [Baddeley et al. 1994]. Both of these were completed in full and the results are reported in the relevant sections.

**Doors and People Test** [Baddeley et al. 1994]

This test is comprised four subtests (two visual and two verbal), and is designed to compare also recall and recognition ability. Norms are not available for children between the ages of 10 and 16 years, so raw scores are reported (i.e. total number of correct trials). Although the lack of norms prevents identification of recall-recognition discrepancies on an individual basis, between-group comparisons allow this question to be addressed on a group basis (see Section 3.8). The visual and verbal recognition tests are described below. The visual and verbal recall tests are described later.

**Doors**

The child was shown 12 pictures of doors one at a time, for 3 seconds each. As the doors were presented, the examiner gave a verbal label for the door (e.g. stable door). Then the child was shown 12 pictures, each depicting four doors (3 distractors and 1 target) and asked to point to the door previously seen. Upon completion of this set, the
procedure was then repeated with a second set of 12 doors. This second set was harder to complete than the first set, with the distractors being very similar to the target door.

Names
The child was shown a list of 12 names (each with forename and surname) printed on cards, one at a time and asked to read the name out loud. Then the child was shown 12 cards each with four names (all with the same first name) (1 target, 3 distractors) and asked to point to the name presented earlier. As with the Doors, this was repeated with a second harder list of names with more similar distractors used in the test section.

Recognition subtests of the Children's Memory Scale
The Children's Memory Scale [Cohen 1997] consists of a series of subtests providing an overall indication of children's memory. All scoring was carried out according to the test manual (which provides age norms for children aged 6-16). Standard scores were obtained for each subtest (with a mean of 10 and standard deviation of 3). For the children aged 17 and over, the Wechsler Memory Scale [Wechsler 1998], which is the adult equivalent of the Children's Memory Scale, was used. Many of the subtests are identical, allowing comparison of scores across the different scales. When the subtests of the adult scale differ from those in the children's scale, the adult scores are omitted (see Appendix B). After each immediate memory condition, the child was asked to remember following a delay interval the material presented earlier.

The recognition subtests of the Children's Memory Scale used were: Faces, recognition of stories, word pairs and word list.

Faces
In this subtest the child was shown a series of faces, one at a time and asked to remember each one. Faces were presented at a rate of one face every two seconds. Then the child was shown another series of faces, one at a time, and asked to identify each one as either a face shown earlier or a new one. After a filled delay of approximately 30 minutes, the child was presented with a second series of faces one at a time and asked to say if each face was one he was asked to remember or not. Distractor faces in the immediate recognition condition were not the same as those in the delay recognition condition.
The details of the stories, word pairs and word list subtests are described in detail in later sections. Brief descriptions are provided below.

**Stories** Following the presentation of two stories and attempts at immediate and delayed free recall, the child was asked a series of ‘yes-no’ questions about each story.

**Word Pairs** The child was asked to learn a series of word pairs. Three learning trials preceded immediate and delayed free recall. The child was then read a list of word pairs and asked to identify those word pairs he was asked to remember earlier. All word pairs in the recognition subtest consisted of a target pair or 2 novel words.

**Word List** Following presentation of a word list (with selective reminding) four times, the child completed immediate and delayed free recall trials. The child was then presented with a list of words and asked to identify those which had been on the list he was asked to remember.

### 3.2.3 Results

There was a significant group difference on both recognition measures of the Doors and People Test (see Figure 3:1) (Doors: ANOVA: F(2,41) = 4; p = 0.02: Planned comparisons CvH: t = 2, df = 41, p = 0.1, HvL; t = 1, df = 41, p = 0.3; Names : ANOVA: F(2, 38) = 7; p = 0.003: Planned comparisons CvH: t = 0.2, df = 38, p = 0.9; HvL; t = 3, df = 38, p = 0.003). This is a reflection of the Low group performing significantly worse than the control group (Doors and Names) and the High group (Names). The group difference in the Doors score was accounted for by VIQ and age (ANCOVA: F(2, 39) = 2, p = 0.1)$^1$. The difference in Names score was not accounted for by verbal IQ or age. When analysed by subsets of Names, a significant difference was only found for set 1 (Set 1: ANOVA: F(2, 38) = 11, p < 0.001; Set 2: ANOVA: F(2, 38) = 1, p = 0.3).

---

$^1$ There is considerable evidence that children’s recognition skills improve with age (e.g. Dirks and Neisser 1977; Mandler and Robinson 1978; Newcombe et al. 1977; Sophian and Stigler 1981), consistent with the significant correlation between age and recognition of the doors found here.
There was no significant difference between the scores of the three groups on the immediate or delayed faces recognition subtest of the CMS (Immediate: $F(2, 41) = 2, p = 0.1$; Delay: $F(2, 41) = 2, p = 0.1$) (see Figure 3:2). Wide variation in performance was seen in all three groups (Range of scores in controls: 2-15; High: 1-12; Low: 1-12).

There were significant group effects on all the results from the verbal delayed recognition subtests of the CMS (see Table 3:1 and Figure 3:3). The group differences on the Stories recognition were accounted for by verbal IQ differences (ANCOVA $F(2, 40) = 0.2, p = 0.8$). The group differences on the Word List recognition remained significant when verbal
IQ was included as a covariate in the analysis. The group differences in the Word Pair recognition subtest were accounted for by VIQ.

Table 3:1 Statistical Results from Verbal Recognition Tests of the Children's Memory Scale (ANOVA and Planned Comparisons)

<table>
<thead>
<tr>
<th>Recognition Subtest</th>
<th>F(2,41)</th>
<th>p</th>
<th>CvH t</th>
<th>CvH df</th>
<th>CvH p</th>
<th>HvL t</th>
<th>HvL df</th>
<th>HvL p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stories</td>
<td>6</td>
<td>0.004</td>
<td>0.9</td>
<td>41</td>
<td>0.4</td>
<td>3</td>
<td>41</td>
<td>0.01</td>
</tr>
<tr>
<td>Word Pairs</td>
<td>8</td>
<td>0.002</td>
<td>0.9</td>
<td>26.3</td>
<td>0.4</td>
<td>3</td>
<td>19.1</td>
<td>0.02</td>
</tr>
<tr>
<td>Word List</td>
<td>8</td>
<td>0.001</td>
<td>0.2</td>
<td>41</td>
<td>0.9</td>
<td>3</td>
<td>41</td>
<td>0.002</td>
</tr>
</tbody>
</table>
3.2.4 Discussion

Results from the High group
The results show that the High group did not perform significantly differently from the control group on any of the recognition assessments, consistent with much of the literature. There was no evidence of selective difficulty in visual recognition of objects with agency as proposed in the literature, even at a delay (see faces results) (e.g. see Blair et al. 2001). Wide individual variation in the groups may in part account for the absence of a deficit. Considerable inter-subject variability in face processing in individuals with Autism has
been noted previously [Davies et al. 1994]. Perhaps difficulty with recognition of agents is a non-universal feature of Autism. Additionally, studies which have found agent selective recognition deficits have included subjects with below average intellectual ability (e.g. the group in the Blair et al. study included subjects with verbal IQ of 67). Perhaps there is a gradually increasing impairment in recognition in Autism, with high functioning individuals showing no recognition deficit, below average individuals showing a selective deficit in agent recognition and low functioning individuals showing a more general recognition deficit. Further research is required to address this hypothesis.

Results from the Low group

The Low group performed significantly worse than the High group on recognition of word pairs, word list, names and stories as well as on the doors subtest. Interpretation of these results must be guarded, however, due to the lack of controls matched to the Low group for verbal ability. With this caveat, analysis of covariance suggests that performance on the stories, word pairs and doors subtests was a reflection of the group’s low verbal IQ. Performance on the word list recognition subtest, however, was not correlated with verbal IQ. Examination of the individual scores demonstrates that there was a bimodal distribution of scores within the Low group on this subtest, with some children scoring in the very impaired range (e.g. standard score of ≤ 2) and others scoring in the low average or above range (e.g. standard score ≥ 8). This suggests that the results may reflect individual variability in levels of motivation and co-operation during the assessment rather than a specific memory impairment.

The data suggest that different factors may have contributed to the poor performance of the Low group on the Names subtest of the Doors and People test. The significant impairment was found to result from the first set of names (which are designed to be easier than the second set). Poorer performance on the easier set of names suggests that these results do not reflect a difficulty with recognition per se. A deficit in recognition should result in poor performance on both sets of names (with the worst scores on the second set). Instead, perhaps the children in the Low group did not understand the requirements of the task. For example, the instructions do not explicitly say that the distractor names will have the same first name as the target name, and so successful recognition is reliant on paying attention to
the second name. Thus in the first set of names, the child may only have attempted to remember the first name. Learning the importance of the surname for successful recognition in the first set of names may then have aided the child’s performance on the second set.

In summary, therefore, it could be argued that the Low group showed no definite evidence of a recognition impairment beyond that expected from their IQ, and that poor motivation may have also affected their performance on some of the subtests. Although it is tempting to speculate on the implications of these results for the hypothesis of selective agent recognition deficit, such interpretation is inappropriate due to the lack of controls matched for the verbal ability of this group.

In conclusion, these results demonstrate good recognition skills in the High group and IQ appropriate recognition skills in the Low group. As all the recognition memory tasks in this section could be successfully completed on the basis of familiarity judgements, these findings seem consistent with there being no severe pathology affecting the parahippocampal region of the children with Autism.

### 3.3 Semantic Memory

#### 3.3.1 Introduction

Semantic memory (general knowledge about the world) has not been explicitly investigated in individuals with Autism. However, many verbal intelligence scales include subtests of semantic memory (such as general knowledge or vocabulary) and therefore many studies contain information regarding semantic memory in children with Autism.

It should be noted that inclusion of semantic memory subtests in verbal IQ assessments means that the two are inextricably linked. Almost by definition, therefore, children with Autism with an average verbal IQ have normal semantic memory, whilst children with
below average verbal IQ have poor semantic memory. There are many reports of individuals with Autism and normal (or even superior) intelligence (e.g. Minshew et al. 1992). Thus, at least some individuals with Autism have been reported to have normal semantic memory. Further, studies of subtest scores on the intelligence scale have found that scores on semantic memory subtests are on average higher than other subtests contributing to the verbal IQ score (e.g. Siegel et al. 1996).

This section investigates semantic memory in individuals with Autism, using the traditional verbal subtests from the intelligence scale and, additionally, a non-verbal test of semantic knowledge.

3.3.2 Methods

Subtests of the Wechsler Intelligence Scale [Wechsler 1991; Wechsler 1997]
Three subtests of the Wechsler Intelligence Scales were used as measures of semantic memory. All subtests were scored according to the manual, which provides norms for adults and children aged over 6. Standard scores are reported, which have a mean of 10 and standard deviation of 3.

Information: A series of orally presented questions that tap the child’s knowledge about common events, objects, places and people.

Similarities: A series of orally presented pairs of words for which the child explains the similarity of the everyday objects or concepts they represent.

Vocabulary: A series of words presented orally which the child defines.

Pyramids and Palmtrees Test [Howard and Patterson 1992]
This is a test of picture association which consists of triads of items where one item (presented at the top) has to be matched to one of two others. The child was asked to point to the picture that best ‘goes with’ the picture at the top. For example a pyramid has to be matched to a pine tree or a palm tree. The two choices are always semantic coordinates, whereas the item at the top is usually from a different category. Although not explicitly stated, the choice must be made on the basis of some property or association that is shared.
by the given item and the target. As the test involves only pictures, only concrete semantic representations are assessed. A total of 52 triads are presented and scores are reported as raw scores.

### 3.3.3 Results

Figure 3:4 shows the group scores on the three Wechsler subtests commonly regarded as measures of semantic knowledge. There was a significant group effect in all three subtests (see Table 3:2). In all three subtests, this was a reflection of the Low group performing significantly worse than the remaining two groups. It is not appropriate to examine correlations between these scores and verbal IQ (since the two are related). However, it is highly likely that the performance of the Low group would be accounted for by differences in verbal IQ.

Table 3:2 Statistical Results from Semantic Subtests of the Wechsler Intelligence Scales (ANOVA and Planned Comparisons)

<table>
<thead>
<tr>
<th>Subtest</th>
<th>F(2,41)</th>
<th>p</th>
<th>CvH</th>
<th>CvH df</th>
<th>CvH p</th>
<th>HvL</th>
<th>HvL df</th>
<th>HvL p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>26</td>
<td>&lt;0.001</td>
<td>0.1</td>
<td>41</td>
<td>0.9</td>
<td>6</td>
<td>41</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Similarities</td>
<td>7</td>
<td>0.002</td>
<td>-1</td>
<td>41</td>
<td>0.2</td>
<td>4</td>
<td>41</td>
<td>0.001</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>18</td>
<td>&lt;0.001</td>
<td>-0.07</td>
<td>41</td>
<td>0.9</td>
<td>5</td>
<td>41</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Figure 3:4 Results from WISC-III subtests (Information, Similarities, Vocabulary; Mean ± SEM)

Figure 3:5 shows the results from the Pyramids and Palmtrees test. There was a significant difference between the groups (ANOVA: F(2, 41) = 9 p = 0.001; Planned comparisons CvH: t = -0.2, df = 22.8, p = 0.8; HvL; t = 3, df = 18.3, p = 0.009). This reflects the Low group scoring significantly lower than either of the remaining two groups. The group effect was accounted for by differences in age and VIQ (ANCOVA F(2, 39) = 0.3 p = 0.7).

Figure 3:5 Results from Pyramids and Palmtrees Test (Mean ± SEM)
3.3.4 Discussion

These results demonstrate that the children in the High group had intact semantic memory. The Low group performed at a level consistent with their verbal IQ. These findings are in agreement with previous reports (e.g. Siegel et al. 1996). The implications of these results will be discussed in the general discussion, in the light of the remaining memory findings of the children in the High and Low groups.

3.4 Cued Recall And Recall Of Stimuli: Repeated Presentation

3.4.1 Introduction

Both cued recall and recall of stimuli following repeated presentation rely less on episodic or contextual memory than recall of stimuli presented only once (see above). Individuals with Autism have been shown to have good cued recall and spared recall of stimuli following repeated presentation. For example, in a study of cued recall, Bowler et al. demonstrated that individuals with Autism (CA 33, VIQ 99) perform similarly to controls on explicit word stem recall tasks [Bowler et al. 1997]. Boucher and Warrington (CA 9:11, VMA 5:5, NVMA 9) demonstrated that individuals with Autism are able to recall pictures they have seen if they are given either a semantic cue (e.g. something you sit on) or a phonetic cue (e.g. 'ch') [Boucher and Warrington 1976]. Boucher and Lewis demonstrated that children with Autism (CA 13:1, VMA 7:4, NVMA 11:3) are able to recall recent events if they are cued with contextual questions [Boucher and Lewis 1989].

Paradigms that use repeated presentation of stimuli to be recalled after a number of trials have generally found no gross impairment in the recall ability of individuals with Autism. For example Boucher and Warrington found no impairment in paired associate word learning, using a repeated presentation paradigm (CA 9:11, VMA 5:5, NVMA 9) [Boucher and Warrington 1976]. This finding was replicated by Minshew et al., 1993. Studies investigating word list recall after repeated presentation have also found no impairment.
In this section, the performance of the High and Low groups on cued recall and recall of stimuli following repeated presentation was assessed.

### 3.4.2 Methods

Subtests from two memory batteries were used in this section: the recall subtests of the Doors and People test (People and Shapes), and the subtests of dot location, word pairs and word list from the Children’s Memory Scale. Administration and marking of these subtests was as described below.

**Doors and People [Baddeley et al. 1994]**

**People**  The child was asked to learn the names of four people. Four cards were presented, each with a photograph and the person’s name and occupation written below it. Once the child had seen all four cards, he was asked for the doctor’s name, the postman’s name and so on. This procedure was repeated until the child recalled all four names correctly or for a maximum of three times. The child was awarded a mark for every part of a name recalled correctly (forename and surname) and a bonus mark if the full name was recalled.

**Shapes**  The child was shown 4 simple line drawings one at a time. Each depicted a cross and the child was required to copy it. After the last drawing had been completed and removed from view, the child was asked to draw all four from memory. Unless performance was perfect, the child was shown the four drawings again (for 3 seconds each) and then asked to draw them from memory. This was repeated again if performance was not perfect the second time. The child was awarded points for correct overall shape (e.g. long and thin), correct centre detail and correct detail of the ends of the cross. A total of three points was awarded for the recall of each cross.
Children’s Memory Scale [Cohen 1997]

*Dot Location*  In this subtest, the child was shown a stimulus item consisting of blue dots placed in various locations within a box. The stimulus item was then removed from view, and the child asked to place the response chips on the response grid in the same locations as the blue dots appeared in the stimulus item. After the three learning trials, a distraction trial with red dots was presented. For the immediate recall task, the child was asked to place the response chips on the response grid in the same locations that the blue dots appeared in the first stimulus item. After a filled delay (consisting of the remaining subtests of the CMS) of approximately 30 minutes, the child was asked to recall the locations of the blue dots. The child obtained two scores: dot learning score (made up of the three learning trials), and delayed recall of the dot locations.

*Word Pairs*  In this subtest the child was asked to remember a list of word pairs read aloud. The examiner then read the first word of each pair and asked the child to provide the second word from memory. This procedure was repeated for three trials and then the child was asked to provide both words of each pair from memory for the immediate recall task. In the delayed recall portion of this subtest, the child was asked to recall the word pairs learned during the learning trials of the word pairs subtest. The delayed recognition portion followed (see above). The child obtained two scores: word pair learning (consisting of the 3 cued trials) and a delayed recall score.

*Word Lists*  In this subtest, the child was asked to learn a list of words presented in four trials. For the first trial, the entire list was read out and the child was asked to recall as many of the words as possible in any order. On the second to fourth trials, the child was selectively reminded only of those words not recalled on the previous trial, and then asked to recall as many words as possible, including the words already recalled. After a distractor list (presented once) the child was asked to recall as many words as possible from the first list. In the delay condition of this subtest, the child was asked to recall the word list learnt during the learning trial of the word lists subtest. The delayed recognition condition of this subtest followed (see above). The child obtained two scores: Word List Learning (consisting of all 4 trials) and delayed recall.
3.4.3 Results

There was no significant group difference on the People recall subtest of the Doors and People Test (ANOVA: F(2, 41) = 2, p = 0.1) (see Figure 3:6). There was a significant group difference on the Shapes recall subtest (ANOVA: F(2, 41) = 5, p = 0.02: Planned comparisons: CvH: t = 0.9, df = 18.4, p = 0.4; HvL: t = 2, df = 13.4, p = 0.1). This reflects the Low group performing significantly worse than the controls. Scores were not accounted for by age or verbal IQ.

Figure 3:6 Results from Recall subtests of the Doors and People Test (Mean ± SEM)

There were no significant differences between the groups on either of the dot location scores. However there were significant differences between the groups on word pair learning and delayed recall, and on word list learning and delayed recall (see Table 3:3). The differences between the groups in word pair learning and delayed recall and in word list delayed recall were a reflection of the Low group performing significantly worse than the remaining two groups. The differences in the word pair subtests were accounted for by VIQ, whilst VIQ did not account for the difference in performance in the word list delayed recall subtest. The difference between the groups on word list learning was a reflection of the control group performing significantly better than both Autistic groups, and was not accounted for by VIQ (see Figure 3:7).
Table 3:3 Statistical Results from Cued and Repeated Presentation Verbal Subtests of the Children's Memory Scale (ANOVA and Planned Comparisons)

<table>
<thead>
<tr>
<th>Subtest</th>
<th>F</th>
<th>p</th>
<th>CvH t</th>
<th>CvH df</th>
<th>CvH p</th>
<th>HvL t</th>
<th>HvL df</th>
<th>HvL p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Pair Learning</td>
<td>8</td>
<td>0.001</td>
<td>-0.2</td>
<td>39</td>
<td>0.9</td>
<td>3</td>
<td>39</td>
<td>0.001</td>
</tr>
<tr>
<td>Word Pair Delay</td>
<td>14</td>
<td>&lt;0.001</td>
<td>0.3</td>
<td>41</td>
<td>0.8</td>
<td>4</td>
<td>41</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Word List Learning</td>
<td>6</td>
<td>0.007</td>
<td>2</td>
<td>41</td>
<td>0.04</td>
<td>1</td>
<td>41</td>
<td>0.2</td>
</tr>
<tr>
<td>Word List Delay</td>
<td>8</td>
<td>0.001</td>
<td>1</td>
<td>41</td>
<td>0.2</td>
<td>3</td>
<td>41</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Figure 3:7 CMS Subtests with Repetition: a) Dot Location learning; b) Dot Location delay; c) Word Pairs learning; d) Word Pairs delay; e) Word List learning; f) Word List delay (Mean ± SEM)

The results of the Word list learning were then further explored by examining the raw scores for each trial. Owing to the differing number of words presented to different age groups, only the scores for children aged between 9 and 16 were included (Control n = 16, High n = 12, Low n = 11; Max score = 14) (see Figure 3:8). The children in the High group
remembered fewer words on early trials than controls, but with repetition reaching a level similar to that of the control group.

Figure 3:8  Word List: Trial by Trial (Mean ± SEM)

1, 2, 3, 4 refer to learning trials. Immediate refers to the immediate recall trial (after the interference list) and Delay refers to the delay recall trial

3.4.4 Discussion

Results from the High Group

The results demonstrate that the children from the High group have good cued recall skills (word pair learning and people subtests) and good delayed recall of stimuli repeatedly presented (delayed recall of word pairs, word list and dot location). This is consistent with previous findings (e.g. Minshew et al. 1992; Minshew and Goldstein 1993; Buitelaar et al. 1999a).

The results from the word list learning trial demonstrate the although the High group had difficulty recalling the stimuli after single presentation of the stimuli, further repetition of the stimuli boosted their performance to control levels. The role of repetition in learning in Autism is further discussed in Section 3.8.
Not all previous studies of word list learning paradigms have found a deficit in recall during learning trials in Autism (e.g. Minshew et al. 1992; Minshew and Goldstein 1993; Buitelaar et al. 1999a). Other studies have found a deficit in recall during learning, but with performance on the initial trials being similar to controls and deficits showing in later trials (e.g. Bennetto et al. 1996). In the present data, there is no evidence of a decline with more repetitions. Differing intellectual abilities and/or ages of test cohorts and differing list lengths may in part explain these discrepancies.

Interestingly, there is no discrepancy between the High group and controls on either the shapes subtest of the Doors and People Test or the dot location learning subtest of the Children's Memory Scale. It might be expected that performance on these subtests and word list learning should be similar (since all reflect scores of learning trials over repeated presentation). There are, however, a number of important differences. The number of stimuli to be remembered was lower in the shapes and dot location subtests, and selective reminding was only used in the word list test. Such methodological differences may mean that the shapes and dot location learning subtests are insensitive to the subtle difficulties detected by the word list subtest. Another possible explanation is that the stimuli in the shapes and dot location are visual and meaningless, whilst the stimuli in the word list are verbal and meaningful.

Results from the Low group
The results from the Low group are inconsistent. Performance on dot location and names subtests suggest that the children have good cued recall and recall for repeated items. However, the children showed deficits in other subtests assessing these skills (i.e. shapes and word list subtests). Examination of the individual scores from the latter subtests once again sheds some insight into this apparent discrepancy. As found in the recognition subtests, there was a bimodal distribution of scores and there was no consistent profile of an individual performing poorly on all these subtests. Thus, these results are unlikely to reflect memory skills, but instead motivation and attention levels.
3.5 Recall Of Stimuli: Single Presentation

3.5.1 Introduction

Studies using single presentation paradigms (with their increased dependence on episodic memory) have found deficits in individuals with Autism. Bowler et al. (CA 33, VIQ 99) and Boucher (CA 14:2, VMA 5:9) demonstrated that subjects with Autism were impaired at recalling a list of words after a single presentation [Boucher 1978; Bowler et al. 1997]. Boucher and Lewis showed that children with Autism (CA 13:1, VMA 7:4, NVMA 11:3) were impaired at following instructions presented once, but not when given instructions with no memory load [Boucher and Lewis 1989]. Deficits in repeating novel actions of the examiner in children with Autism (CA < 7, IQ ~ 50) were also noted by DeMeyer et al. [1972].

Studies of recall of pictures presented have produced inconsistent results in children with Autism. For example, Boucher and Warrington found that children with Autism (CA 9:1, VMA 5:5, NVMA 9) were impaired at free recall of pictures presented once [Boucher and Warrington 1976]. Using a similar paradigm, Renner and colleagues found that higher functioning children (CA 10:2, VIQ 101, NVIQ 97) were able to recall pictures presented once as well as controls [Renner et al. 2000]. Factors which may contribute to these mixed findings include differing intellectual ability, varied proportions of agent and non-agent stimuli (see recognition section), and differing degrees of relatedness of the pictures. This latter possibility is raised by the finding that children with Autism benefit less from the stimuli being semantically related than controls do (e.g. Boucher 1978, see Section 3.6.4).

This section investigates recall of single presentation of verbal and visual material using recall of pictures and stories (see also previous section for discussion of recall of word lists after one presentation).
3.5.2 Methods

Two subtests from the Children’s Memory Scale were used.

*Family Pictures* In this subtest the child was shown a portrait of 6 family members and a dog. The child was then presented with four different scenes, each featuring three of the family members. Each scene was exposed for 10 seconds and the child was asked to remember as much as possible about each scene. After a 5 second delay, the child was shown the family portrait and a card that was identical to the scene just presented but with the family members missing. The child was asked to identify which family members were in the picture, where they were and what they were doing. After a filled delay the child was re-presented with the same portrait and blank scenes and asked to indicate who was in the picture, where they were and what they were doing. Children were awarded a point for correctly identified people, two points for their action and a point for identifying the quadrant of the picture they were in. Wrongly identified people (even with correct location and action) resulted in the entire response being given a score of 0.

*Stories* In this subtest the child listened to two stories read by the examiner. Immediately after hearing each story, the child was asked to retell it from memory exactly as read. After a filled delay, the child was asked to retell the two stories. The delayed recognition portion followed (see above). Points were awarded for correct segments (or close synonyms) as described in the manual.

3.5.3 Results

There were significant differences between the three groups on all the CMS subtests without repetition (see Figure 3:9 and Table 3:4). The differences in the Family pictures subtests were a reflection of the Control group scoring significantly higher than the remaining two groups, and were not accounted for by differences in VIQ. In contrast the differences in the stories subtests were a reflection of the Low group scoring significantly poorly and were accounted for by differences in VIQ (Stories Immediate: ANCOVA F(2, 40) = 0.8, p = 0.5; Stories Delay: ANCOVA F(2, 40) = 0.7, p = 0.5).
Table 3:4  Statistical Results from Single Presentation Recall Subtests of the Children's Memory Scale (ANOVA and Planned Comparisons)

<table>
<thead>
<tr>
<th>Subtest</th>
<th>F</th>
<th>p</th>
<th>CvH t</th>
<th>CvH df</th>
<th>CvH p</th>
<th>HvL t</th>
<th>HvL df</th>
<th>HvL p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>4</td>
<td>0.02</td>
<td>2</td>
<td>41</td>
<td>0.04</td>
<td>0.7</td>
<td>41</td>
<td>0.5</td>
</tr>
<tr>
<td>Immediate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family</td>
<td>5</td>
<td>0.01</td>
<td>3</td>
<td>41</td>
<td>0.01</td>
<td>0.2</td>
<td>41</td>
<td>0.8</td>
</tr>
<tr>
<td>Pictures Delay</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stories Immediate</td>
<td>4</td>
<td>0.02</td>
<td>0.8</td>
<td>41</td>
<td>0.4</td>
<td>2</td>
<td>41</td>
<td>0.5</td>
</tr>
<tr>
<td>Stories Delay</td>
<td>9</td>
<td>0.001</td>
<td>1</td>
<td>41</td>
<td>0.2</td>
<td>3</td>
<td>41</td>
<td>0.009</td>
</tr>
</tbody>
</table>
3.5.4 Discussion

Results from the High group
The results from the family pictures subtest demonstrate that the High group had difficulty recalling contextual information (who, what, where) from the pictures, despite the cue of the background. The stability of their scores across immediate and delay conditions suggests that any loss of information over time occurs at a normal rate. Given the animate, agent nature of the stimuli, it is difficult to state categorically that the High group had a generalised visual recall deficit when stimuli are presented once. Further work is required to assess if the deficit also occurs with non-animate stimuli.
Previous studies of picture recall have used black and white object drawings as stimuli [Boucher and Warrington 1976; Renner et al. 2000]. Studies of high functioning individuals with Autism have found no deficits in recall. However differences in stimuli (colour versus black and white, agent only versus mixed stimuli) and differences in recall procedure (free versus cued) make comparisons difficult.

The absence of a consistent deficit in story recall suggests that the High group has good verbal recall of related social stimuli presented once. Although this may suggest a verbal/visual discrepancy, a number of issues should be highlighted. Firstly, the predictable nature of the stories meant that recall of a general theme (girl witnesses boat sinking) would be sufficient for much of the story to be guessed correctly. Not only are the pictures less predictable, but the stories also contained more linked information. In some of the picture stimuli, all three individuals were completing apparently unrelated activities. Thus the two subtests are not necessarily equated for difficulty.

Results from the Low group

As mentioned previously, the lack of appropriate verbal IQ matched controls for the Low group limits interpretation of the results. However, it is interesting to note that when the verbal IQ discrepancy was accounted for, the Low group showed a similar performance profile on these subtests to the High group: deficits in recall of family pictures, and intact recall of stories, supporting the possibility that these deficits are characteristic of Autism.
3.6 Episodic (Contextual) Memory

3.6.1 Introduction

The previous sections have shown that the Autistic groups had an increased deficit in memory tasks that have greater reliance on episodic (or contextual) memory. In this section, further evidence is provided for impaired episodic memory in these children.

A number of studies have shown that individuals with Autism have difficulty with contextual memory. Boucher and colleagues showed that children with Autism (CA 13:2, VMA 7:2, NVMA 10:6) were significantly less able than controls to recall the context and order of recent activities in which they had participated. The same children, however, performed at levels similar to controls when asked questions about the events of the testing, which included contextual cues [Boucher 1981; Boucher and Lewis 1989]. In the more extensive of these two studies, children participated in four experiments over a period of 6-12 months. At least 1 month after completing the last experiment, the child was then asked to recall the activities previously completed with the experimenter [Boucher and Lewis 1989].

Other studies have shown impaired event memory in individuals with Autism. Jarrold and Russell asked the children in their study how they knew what was in a box (were they told or did they see) and who had placed a card into a grid (either the child, the experimenter or a doll manipulated by the experimenter or child). The children with Autism were severely impaired on this task in comparison to controls [Jarrold and Russell 1996]. Millward et al. further showed that children with Autism (CA 13:1, VMA 6:3) had more difficulties recalling events that they had personally experienced than events that they saw another child experience [Millward et al. 2000]. This last finding may be argued to be more a reflection of a lack of insight into self, rather than a memory deficit per se. However, the study has an important confound. Children's memories for events are enhanced by narration [Tessler and Nelson 1994], and it might be expected that control children are
more likely than children with Autism to provide a spontaneous narration of the events happening to them. Thus the children with Autism might have recalled events occurring to other children better due to the benefits of narration.

In this section, two measures of episodic memory are used, the Rivermead Behavioural Memory Test and the Sunderland Parental Memory Questionnaire.

3.6.2 Methods

The Rivermead Behavioural Memory Test [Wilson et al. 1991] was administered according to the manual, and age norms were used to calculate standardised scores. The test consists of a series of subtests, as detailed below. Raw scores from each subtest were converted into scaled scores, where 2 indicated normal performance, 1 indicated mild impairment and 0 indicated severe impairment. Scaled scores (for children aged 6+) were provided by the manual.

Remembering a name The child was shown a photographic portrait and told the first and second name of the person in the photograph. He was asked to repeat their name and then remember it. After a filled delay of over 20 minutes, the child was asked to recall the name. Partial credit was given for recall of the name following the prompt of the initial letter.

Belonging The child was asked for a belonging (such as a handkerchief, watch, etc.) and told that the examiner was going to hide the object. He was asked to remember where it was hidden and to ask for it when the examiner said that the test was over. After a filled delay of over 20 minutes, the child was told that the test was over. Credit was given for remembering the item and where it was hidden. Partial credit was given for remembering following a prompt.

Remembering an appointment The child was shown an alarm clock. The alarm was set for 20 minutes and the child was told to ask the examiner when the alarm rings when he would next be seeing the examiner. A practice was given. If, when the alarm rang, the child did not respond, he was prompted. Partial credit was given for remembering that he was supposed to do something.
Ten black and white line drawings were shown to the child, one a time, each for 5 seconds. After a brief, filled delay, the child was asked to pick out the drawings seen earlier from a set of 20 drawings.

The child was read a short story and asked to recall it immediately exactly as read. After a delay the child was asked to recall the story again.

Five faces were presented to the child, one at a time. The child was asked to remember them. After a brief delay, the child was shown 10 faces and asked to indicate which ones had been seen earlier.

The examiner traced a short route around the testing room. The route involved five locations. As the examiner completed the route, a verbal commentary was given, naming all the locations. The examiner took a message, and left it at stage 4 of the route. The child was required to follow the route immediately and after a delay of approximately 20 minutes.

The child was also scored on remembering to take the message and leaving it at the appropriate location. The child was prompted to take the message if he did not spontaneously pick it up. This was scored on both immediate and delayed recall of the route.

The child was asked 11 questions to determine orientation to date, time, place, awareness of age, birthdates, etc.

The Sunderland Parental Memory Questionnaire was used to assess aspects of real life everyday memory that are otherwise difficult to measure through standardised tests or through experimental measures [Sunderland et al. 1983]. Parents rated each of 28 items from A (has not occurred in the last 3 months) to I (occurs at least once a day). Scores of 1–9 were assigned to ratings A to I and these were summed for each child.

3.6.3 Results

Results from the Rivermead Behavioural Memory test (Figure 3:10) show that there was a significant difference between the three groups, with both the High and Low groups performing significantly worse than controls (ANOVA: F(2, 41) = 17, p < 0.001; Planned
Memory comparisons: CvH $t = 4$, df = 17.8, $p < 0.001$; HvL $t = 2$, df = 17.7, $p = 0.1$). These differences were not accounted for by VIQ (ANCOVA $F(2, 40) = 7$, $p = 0.003$) nor by sustained attention (ANCOVA $F(2, 39) = 11$, $p < 0.001$). This significant difference between the groups was also found when the analysis was restricted to just the boys’ data. The pattern of performance on the recognition of faces and pictures and story recall was similar to that reported for similar subtests in previous sections. The pattern of performance on the subtests can be seen in Figure 3:11.

Figure 3:10 Rivermead Behavioural Memory Test (Mean ± SEM)

![Figure 3:10](image)

Figure 3:11 Rivermead Behavioural Memory Subtests (Name, Belonging, Appointment, Picture recognition, Story immediate and delayed, Face recognition, Route Immediate and delay, message and orientation) (Mean ± SEM)

![Figure 3:11](image)
The results from the Sunderland Parental questionnaire confirm the results of the RBMT (see Figure 3:12). There was a significant difference between the three groups, with both Autistic groups scoring significantly higher (indicating greater impairment) than the Control group (ANOVA: $F(2, 38) = 12, p < 0.001$; Planned comparisons CvH: $t = -4, df = 13.8, p = 0.002$; HvL: $t = -0.7, df = 22.9, p = 0.5$). This difference was not accounted for by VIQ (ANCOVA: $F(2, 37) = 7, p = 0.002$).

Figure 3:12  The Sunderland Parental Memory Questionnaire (Mean ± SEM)

![Figure 3:12](image)

### 3.6.4 Discussion

The results from two well-established measures of episodic memory demonstrate that both Autistic groups are impaired on this aspect of memory. The Rivermead scores highlight prospective memory (i.e. memory that allows the realisation of delayed intentions) as particularly impaired.

No previous studies have investigated episodic memory in Autism using comparable paradigms. However these findings are consistent with studies investigating memory for events (e.g. Boucher 1981; Boucher and Lewis 1989; Jarrold and Russell 1996; Millward et al. 2000) (see Section 3.6.1). Gardiner and colleagues found a deficit in episodic memory in adults with Autism using a remember-know paradigm [Bowler et al. 2000].

Deficits in episodic memory in Autism may be related to other findings of abnormal use of context. Individuals with Autism do not appear to benefit from the context in which the
information is presented. For example Hermelin and O'Connor found that, compared to controls matched on digit span, children with Autism (CA 11:6, VMA 4:9) did not show the same enhanced memory for meaningful word phrases relative to unrelated words [Hermelin and O'Connor 1970]. A number of studies have suggested that individuals with Autism are comparatively more impaired at recall of related word lists compared to unrelated word lists (see Tager-Flusberg 1991a; Bowler et al. 1997 but see Beversdorf et al. 1998).

Echoing the animal work, as well as a failure to utilise relevant contextual information, individuals with Autism also show over-reliance on irrelevant contextual information. For example, children with Autism often fail to recognise familiar objects or people in a new context [Waterhouse 1987] and show disproportionate distress at changes in routine.

### 3.7 Novelty Preference

#### 3.7.1 Introduction

All of the previous sections have described paradigms that explicitly assess memory (the child is told to remember stimuli and later asked to recall them). However, as mentioned in the Introduction to this chapter, there are also methods of assessing explicit memory that do not require explicit instructions. One such method is visual paired comparison. This method assesses novelty preference and is described in this section.

It has been shown that, from a very young age, humans and animals spontaneously discriminate between novel and familiar stimuli [Fagan 1990; Bachevalier et al. 1993; McKee and Squire 1993; Pascalis and de Schonen 1994]. Visual paired comparison is designed to assess this spontaneous novelty preference. In this paradigm, the subject is shown a stimulus for familiarisation and then after a delay is shown the same stimulus paired with a novel one [Fagan 1970]. The key measure is the length of time spent fixating on each of the two stimuli. Children as young as 3-4 days old spend longer looking at
novel stimuli even after a two minute delay [Pascalis and de Schonen 1994]. Such discrimination is assumed to reflect recognition of the familiar stimulus.

Novelty preference is dependent on the integrity of the medial temporal lobes. Monkeys with bilateral lesions encompassing both the amygdala and hippocampal formation show impaired performance on visual paired comparison [Bachevalier et al. 1993]. More selective lesion studies in monkeys have shown that both the perirhinal cortex and the hippocampal formation are involved in novelty preference [Clark et al. 1996; Clark et al. 1997]. Human studies of individuals with selective hippocampal formation damage have also demonstrated deficits in novelty preference [Pascalis et al. 2000].

It should be noted that the results from visual paired comparison paradigms do not always correlate with overt measures of recognition (e.g. McKee and Squire 1993; Pascalis et al. 2000). Thus the paradigm may provide a measure of medial temporal lobe function that other recognition tasks do not.

In the light of the animal model of Autism proposed by Bachevalier [Bachevalier 1994], and the finding that neonatal amygdala and hippocampal formation lesions lead to impaired novelty preference [Bachevalier et al. 1993], the performance of individuals with Autism was investigated on this type of paradigm. In addition, the influence of contextual background manipulation was also investigated.

### 3.7.2 Methods

The Visual Paired Comparison Test (Preferential Looking) was used to examine recognition memory. The paradigm was based on that of Pascalis et al. [2000].

The child was first exposed to a visual stimulus and allowed to explore it visually during a familiarisation period. After a brief delay, during which the screen was blank, the child was shown the familiarised stimulus together with a new stimulus. The left/right position of the familiar/novel stimuli was counterbalanced across trials. The participant’s eye
movements were recorded on video so that the length of time spent gazing at each of the two pictures could be measured. The normal tendency is to look preferentially at the novel stimulus.

The effect of background context change between familiarisation and test was also investigated. In the Same Context condition, a coloured object was presented on the same coloured background during familiarisation and test. In the Different Context condition, the coloured background was changed between familiarisation and test. Two different delays between familiarisation and test were used, one of 0 seconds and one of 10 seconds.

The children were told that they were required to do nothing but watch the screen. They sat approximately 50 cm in front of a screen onto which the slides were back projected. On each trial, the sample stimulus was presented in the centre of the screen for 3 seconds and then removed. After the specified retention period (0, 10 seconds) the familiar stimulus and a new stimulus were projected side by side for 3 seconds. Three trials were used for each condition (total of 12 trials). Two random orderings of the trials were created and each was randomly assigned to half the participants in each group. The order of trials was counterbalanced across subjects.

### 3.7.3 Results

#### 3.7.3.1 Control Group

Results from the control group were first analysed separately to establish the validity of the paradigm. As Figure 3:13 shows, in all four conditions the children all looked longer at the novel item. Statistical analysis (using one-sample t tests with a test value of 50) confirmed that in the Same Context delay condition and the Different Context immediate condition this preference was statistically significant (Same-Delay: \( t = -3, df = 14, p = 0.006 \); Different-Immediate: \( t = -3, df = 15, p = 0.02 \)). However statistical significance for the
remaining two conditions was not reached (Same-Immediate: $t = -1$, $df = 15$, $p = 0.2$; Different-Delay: $t = -1$, $df = 15$, $p = 0.2$).

Figure 3:13 Percentage of time control children spent looking at novel stimulus (Mean ± SEM) (Line at 50% is % expected by chance)

![Figure 3:13](image)

**3.7.3.2 Autistic Groups Results**

In the light of the lack of significant novelty preferences in the Same Immediate and Different Delay condition in the control group, these results from the Autistic groups were not examined.

In the remaining two conditions, there was no significant overall group difference. As shown in Figure 3:14 and Figure 3:15, all three groups showed a preference for the novel stimuli. Statistical analysis (using a repeated measure ANOVA with within subject factor of condition (2 levels: Same delay and Different immediate) and between subject factor of group) revealed no effect of condition ($F(1, 35) = 0.4$, $p = 0.5$) or group ($F(2, 35) = 0.1$, $p = 0.9$) or group by condition interaction.
Figure 3:14 Percentage of time spent looking at novel item in Same-Delay condition (Mean ± SEM) (Line at 50% is % expected by chance)

Figure 3:15 Percentage of time spent looking at novel item in Different-Immediate condition (Mean ± SEM) (Line at 50% is % expected by chance)
3.7.4 Discussion

3.7.4.1 Control Group Results

The control group demonstrated a significant novelty preference in two of the four conditions: Same Context, delayed condition and Different Context, immediate condition. This is consistent with previous reports in the literature. Rather unexpectedly, however, the control children did not show a significant novelty (or familiarity) preference in the remaining two conditions. The reasons for this difference are unclear.

Novelty preference in the Same Context, immediate condition should in theory be the most robust condition: the child did not have to retain an image over a delay and the stimulus was presented in exactly the same context at test as during the familiarisation process. It is possible that since the task was very easy, the length of time at test (three seconds) might have been too long to be sensitive to the novelty preference. This explanation appears unlikely as pilot work with adults found a novelty preference using the same experimental set up.

The stimuli were not counterbalanced across subjects or conditions. Although the order of the trials was counterbalanced, the same stimulus was always the novel stimulus and the same stimuli pairs always made up each condition. The stimuli were those used by Pascalis et al., and so there was good reason to believe that the stimuli should elicit novelty preferences [Pascalis et al. 2000]. However, it is possible that stimuli-specific effects may have contributed to the lack of significant results. For example, the familiar stimuli in the Same Context, immediate condition may have been particularly interesting to the children. Such a hypothesis could be very simply assessed by displaying the pairs of stimuli to children without a familiarisation period and determining whether the amount of time looking at one of the stimuli is consistently higher than that for the other stimulus.
Two different orders of trial presentation were used and the trials were randomly ordered in each. Post-hoc examination of these orders showed that all three of the Same Context immediate condition trials occurred at the start of one of the orders. Although it is possible that the trials all occurring so early in the paradigm may have influenced the results, there was no significant effect of order on the results.

Familiarisation times have been shown to be important in inducing the novelty preference (at least in infants) [Rose et al. 1982; Slater 1995; Richards 1997]. Failure to encode the stimulus sufficiently may lead to the absence of a novelty preference, as the familiar stimulus is also to some extent unfamiliar. Familiarisation times may vary with age (at least in infants) (see Fagan 1990), and so it is possible that the broad age range of the children included in the control group may have masked the novelty preference shown by at least some of the children. However it is difficult to understand why such an effect should significantly affect only two of the conditions in the paradigm.

3.7.4.2 Autistic Group Results

In the two validated conditions, no significant group differences were detected. This suggests that the Autistic groups showed novelty preference in both conditions. Changes in context and delay appeared not to disrupt this preference significantly. This is the first study to investigate novelty preferences in children with Autism using the visual paired comparison task.

These results should not be taken to be inconsistent with the children with Autism having hippocampal formation abnormality, even though animals and humans with bilateral hippocampal damage have been shown to display no novelty preference [Bachevalier et al. 1995; Clark et al. 1996]. Monkeys with neonatal lesions of the medial temporal lobes show normal novelty preference in adulthood at short delays (~10 seconds) [Saunders et al. 1991; Bachevalier et al. 1999]. Thus, the results do not necessarily suggest that the children with Autism do not have hippocampal formation abnormality. Future studies of
novelty preference in children with Autism should explore longer delays between familiarisation and test (such as 2-3 minutes).

3.8 Discussion

3.8.1 Results From High Group

The overall results of the High group indicate preserved recognition, preserved semantic memory and preserved memory for stimuli presented several times, but difficulties remembering single presentations of unrelated verbal information and visual information, and difficulties with event or contextual memory. This pattern of strengths and weaknesses in memory function is consistent with the literature, and suggests abnormality in the hippocampal formation.

Two other groups of children with bilateral hippocampal atrophy have been reported as having such an explicit memory profile: children with developmental amnesia [Vargha-Khadem et al. 1997a] and extremely preterm children [Isaacs et al. 2000]. The scores of the children with Autism on the test of event or episodic memory fall between these two groups, with their episodic memory impairment being much less severe than that of the patients with developmental amnesia and slightly more severe than that of the preterm children. Interestingly, the pattern of impairment on the Rivermead differs in the preterm children and the High group. Although both groups had difficulty with aspects of prospective memory, the children with Autism, in contrast to the preterm group, had difficulty remembering the name but had no difficulty remembering the route around the room.

The results from the word list learning paradigm suggests that the memory impairment in the High group ameliorated with repetition of stimuli. Impairment in word list learning has been reported in the preterm children. In contrast to the results from the High group, the preterm children show no impairment on the first trial, but then perform significantly worse
than controls on later trials. However, differences in paradigm design should be noted (differing number of words and trials, selective reminding versus complete repetition of the stimuli, etc.). These differences prevent direct comparison of the scores. In addition, there is evidence that repetition improves recall in developmental amnesia. Baddeley et al. have shown that one individual with developmental amnesia (Jon) shows improved memory for video segments following repeated presentation [Baddeley et al. 2001].

As yet there are no published reports of the performance of individuals with developmental amnesia or the preterm children on the recall of family pictures. However, a deficit in such a task would be consistent with the reported memory profile of these children.

It should be noted that the children with developmental amnesia have a much more pervasive memory deficit affecting many tasks including recall of stories, word list recall and recall of word pairs [Gadian et al. 2000]. The difference in severity of memory impairment suggests that the nature of any hippocampal abnormality underlying the episodic memory impairment in the Autistic groups is of a different (and less devastating) nature.

3.8.2 Results From The Low Group

The results from the Low group need to be interpreted with caution due to the lack of verbal IQ matched controls. However, with this caveat, the results suggest that the Low group showed a memory profile similar to that of the High group: an episodic memory impairment with relatively preserved recognition and semantic memory. When VIQ differences were accounted for, the Low group was unimpaired on measures of semantic memory, story recall and visual recognition, but remained significantly impaired on measures of episodic memory (i.e. the Rivermead Behavioural Memory Test and Sunderland Parental Memory questionnaire).

In addition to the deficits found in the High group, the Low group was also impaired on delayed recall and recognition of the word list and word pairs, and on recall of the shapes.
These differences were not accounted for verbal IQ. As discussed in previous sections, the scores from these subtests suggest that these differences may not necessarily be reflecting memory deficit per se. Difficulties with motivation, co-operation and appreciation of task requirements may also have played a role. Further, even though analysis of covariance with verbal IQ does not account for the significantly impaired performance on these tasks, verbal IQ differences can not be ruled out as an important factor. Analysis of covariance can only account for linear relationships between the covariate verbal IQ and performance measure. There is no a priori reason to assume the relationship is linear.

In order to determine whether the results of the Low group are characteristic of low functioning children with Autism, or whether they are merely characteristic of low functioning children, it would be necessary to obtain results from a low functioning control group (although see Chapter 10 for a discussion of the problems with such comparisons). In the absence of this, further interpretation is difficult.

3.8.3 Summary

Selective episodic memory impairment was found in the High functioning group. This pattern was very similar to that found in the Low group when verbal IQ was accounted for. This leads to the obvious question: what is the relationship between the episodic memory impairment and Autism?

In answering this question, it is important to note that impaired episodic memory does not necessarily result in Autistic features. Children with developmental amnesia and preterm children are noted to have grossly normal social skills in the presence of episodic memory deficits [Vargha-Khadem et al. 1997a; Isaacs et al. 2000]. An impairment in episodic memory can not therefore be the single cause of the social difficulties in Autism.

The reverse possibility is that the symptoms of Autism are responsible for the episodic memory impairment. Researchers who believe that Autism is caused by a deficit in social attention (e.g. Pierce et al. 1997) might argue that episodic memory (unlike, say, semantic
memory) depends on attending to inherently social aspects of the world. Indeed, the Rivermead subtests on which the children with Autism most poorly were those which could be argued to have the heaviest ‘social component’. However, such an argument can not be used to explain the results of the word list learning (as word lists are not as inherently social).

A third alternative is that an episodic memory impairment is a contributing factor to the symptomatology of Autism. Perhaps this deficit when combined with, say, amygdala damage results in the syndrome of Autism. Distinguishing between these speculations is not possible from the current evidence. Such interpretation depends on refuting the remaining hypotheses (probably by finding counter examples such as a child with poor theory of mind and good episodic memory). Finding such evidence is notoriously difficult.
Quantitative cognitive and behavioural correlates of the amygdala are less well established than those of the hippocampal formation. For example, although social behaviour is widely acknowledged as being dependent on the amygdala (see Chapter 1), developing objective measures of social behaviour that truly reflect 'real-life' performance has proved difficult [Saver and Damasio 1991]. Extensive, complex behavioural observation checklists have been developed for use with monkeys (see e.g. Bachevalier et al. 2001), but these are unlikely to be appropriate for humans (as their environments are less controlled).

Face processing has been associated with amygdala function (see Chapter 1). In particular, many humans with amygdala damage display a profound impairment in recognition of emotional expression (e.g. Adolphs et al. 1994; Young et al. 1996; Broks et al. 1998). It is well established that processing of faces (and in particular facial expression) is abnormal in Autism, consistent with the notion of underlying amygdala damage (e.g. Hobson 1986a; Hobson et al. 1988; Tantam et al. 1989).

In the light of the wealth of research supporting an abnormality in face processing in Autism, an independent behavioural measure of the amygdala was investigated in this thesis, in an attempt to further characterise the nature of the abnormality. One recently developed measure is the emotional modulation of the startle response.

The startle response is a brain-stem-mediated motor response that occurs following the presentation of a sudden and intense stimulus. The vigour of the startle response varies systematically with the emotional state of the individual [Lang et al. 1990]. This emotional modulation of the startle response has been shown to be dependent on the amygdala (e.g. Rosen et al. 1996; Davis et al. 1999).
4.1 Introduction

Throughout the mammalian kingdom, presentation of a sudden, intense stimulus results in a startle response (see Lang et al. 1992). In humans, the fastest and most stable element of the startle response is the sudden closure of the eyelids [Anthony 1985]. This is the traditional experimental measure of the startle response in humans. The vigour of the startle response varies systematically with the affective status of the individual [Lang et al. 1990]. In humans the startle response is facilitated by unpleasant arousal and inhibited by pleasant arousal (e.g. Davis 1989a; Lang et al. 1990; Lang et al. 1992). Modulation of the individual’s emotional state is usually achieved with affect-arousing, visual stimuli and then an acoustic probe is used to elicit a startle response [Balaban and Taussig 1994]. However, modulation of the startle response can also occur when a subject recalls an emotional event or following chronic mood induction, and thus it is not dependent on the perceptual event itself [Vrana and Lang 1990; Cook et al. 1991].

In animals, the emotional state most widely used to investigate the emotional modulation of the startle response is fear. This is because there are quantifiable, objective physiological and behavioural changes in a fearful animal (e.g. freezing). Consistent with the findings in humans, animals show a fear-potentiation of the startle response (see Davis 1986). Further, although it is much harder to infer that an animal is in a hedonic state, animals anticipating food show an attenuation of the startle response [Armus et al. 1964].

Investigations into the neural basis of the startle response have revealed that it is controlled by two polysynaptic circuits, a direct reflex circuit and an additional modulating system. The direct reflex pathway for the acoustic startle response begins at the cochlear nucleus of the brain (activated by the startle noise), and projects via the lateral lemniscus to the reticular formation. The output pathway proceeds from the reticularis pontis caudalis nucleus, through the spinal neurons to the reflex effectors (for review see Davis 1989b).
Emotional Modulation of the Startle Response

The modulation circuit is less well characterised. However, animal studies have shown that a critical component of this circuit is the amygdala (and in particular the central nucleus). Amygdala lesions block the fear potentiation of the startle response, but have no effect on the magnitude of the baseline startle [Hitchcock and Davis 1986]. Similarly, partial kindling of the medial amygdala (inducing hyperexcitability in this region) results in an increase in the fear potentiation of the startle, with no effect on the baseline startle [Rosen et al. 1996]. Partial kindling of the lateral amygdala or the dorsal hippocampal formation does not affect the modulation of the startle response [Rosen and Davis 1988; Rosen et al. 1996].

Much of the characterisation of the neural basis of the modulation system has been carried out in animals. However, there is reason to believe that the same pathway is implicated across species. Humans show similar sensitivity of the responses to paradigm manipulations (such as anticipation of shock enhancing startle, temporal specificity of the response and sensitisation of the startle response by shock exposure [Grillon et al. 1993]). Drugs that have anxiolytic or anxiogenic properties in humans have been shown to inhibit or enhance (respectively) the potentiation of startle in animals in a fearful state [Davis 1979a; Davis 1979b; Davis et al. 1979]. Further, an individual with apparently selective amygdala damage was recently reported as failing to show the typical startle potentiation to aversive background stimuli [Angrilli et al. 1996].

The potential of the emotional modulation of the startle response paradigm is still being explored. There have been a few investigations using this paradigm to characterise clinical populations. For example, patients with phobia have been shown to have increased startle responses to images involving their fearful objects compared to other fearful images [Vrana et al. 1992]. Differential emotional modulation of the startle response has been observed in individuals with schizophrenia with high and low affective flattening. Those with high affective flattening showed greater modulation than those with low affective flattening [Schlenker et al. 1993]. Grillon et al. reported no difference in the emotional modulation of the acoustic startle response in individuals with panic disorder (although they did show increased startles at both baseline and under threat conditions) [Grillon et al. 1994].
The paradigm has also been used to demonstrate differential responses in the normal population. Subjects vulnerable to high anxiety showed reliable startle modification during imagery whilst those with low anxiety not [Cook et al. 1991].

Unfortunately, research into affective responses using the startle reflex in children and adolescents is still in its infancy, with few published studies. This dearth of studies can in part be explained by the fact that the stimuli used to potentiate the startle are not always appropriate for use with children and adolescents for ethical and practical reasons. Only highly arousing stimuli result in affective modulation [Cuthbert et al. 1996]. Less arousing unpleasant slides may be more appropriate, but may not induce the expected potentiation.

Studies that have investigated the emotional modulation of the startle response in children have produced mixed results. Balaban found that five month old infants show an augmented startle reflex to angry faces and an inhibited startle reflex to happy faces [Balaban 1995]. Nine month old children have been shown to have enhanced startle responses in a ‘stranger approach’ paradigm compared to baseline responses [Schmidt and Fox 1998].

Two studies have investigated the affective modulation in the startle response in school-aged children [Cook et al. 1995; McManis et al. 1995]. Neither study found significant modulation. However, Cook et al. noted that children prone to high fear showed smaller responses to fearful stimuli than to pleasant stimuli. A similar trend was noted in boys participating in McManis et al.’s study, but the girls showed the opposite trend. Adolescents, in contrast, have been shown to show fear potentiation of the startle response in experiments using darkness or aversive air puffs to induce emotional modulation [Grillon et al. 1999]. Different methodology and age ranges may be responsible for some of these conflicting results.

Despite the lack of understanding of the pattern of emotional modulation of the startle response in children, studies have demonstrated modulation differences between control children and children from clinical populations (e.g. Grillon et al. 1998).
The emotional modulation of the startle response in children with Autism was therefore investigated in this study. In the light of the medial temporal lobe hypothesis of Autism, it was predicted that the pattern of emotional modulation seen in the children with Autism would differ from that found in the control group. The paradigm chosen to elicit affective modulation involved presentation of visual stimuli. This was chosen for practical reasons (it is the most widely used paradigm in the literature, and additionally a balance between creating affective modulation and minimising the child’s distress was required).

Stimuli were deliberately selected to exclude facial stimuli. Children with Autism have been shown to be poor at recognition of facial expression. This impairment is therefore likely to confound any intended affective modulation in the paradigm (i.e. if the child does not recognise a fearful expression, such a face can not be assumed to engender a fearful emotional response).

4.2 Methods

Stimuli
Over 50 pictures were chosen from a variety of websites and rated as pleasant or unpleasant by 10 adults. From these, the 15 pictures rated as most pleasant and the 15 pictures rated as most unpleasant were selected. An additional 9 pictures were selected with varying ratings for use as filler stimuli in unprobed trials.

An acoustic startle probe (consisting of a 50 ms burst of white noise with instantaneous rise time) was presented binaurally over headphones. The intensity of the probe was chosen to be within a comfortable range for the children (at ~ 70 dB (Sound Pressure Level)).

Paradigm
Each picture was presented for 7 seconds and startle probes were presented at 1300 ms after slide picture onset. The pictures were presented on a Dell 1500 FP computer screen. Picture offset was followed by a blank screen.
The child was instructed that a series of slides would be presented and that each slide should be viewed for the whole time it was on the screen. In order to encourage the child to pay attention, he was told that he would be asked questions about the pictures at the end of the assessment. The child was told that occasional noises would be heard over the headphones which could be ignored.

After the startle probe series had been completed, the child was shown each picture again and asked to rate each picture as ‘Nice’ or ‘Scary’. In order to check comprehension of these concepts, the child was asked to give an example of something ‘Nice’ and ‘Scary’ on a previous assessment day. These ratings were used to produce individual categorisations of the stimuli (pleasant and unpleasant) for each child.

Data Recording
Unilateral right blink magnitude was measured by vertical electro-oculogram (vEOG) measurements using a pair of bipolar AgCl electrodes placed just above and below the orbit in a vertical line through the pupil. Sampling rate was 500 Hz and the data were recorded with a 50 Hz notch filter. Impedances were kept below 10 kΩ. Although it is more traditional to measure blink magnitude from EMG, when both EMG and vEOG were recorded they yielded highly similar results [Sugawara et al. 1994]. The raw vEOG signal was epoched and baseline corrected (interval –200 to 0 ms).

Data Analysis
The peak of the blink was defined as the point of maximum deflection before a return towards baseline that continued for 5 ms. Latency of response was defined as the latency of peak amplitude. When multiple blinks occurred, the response whose latency was closest to the mean latency for that condition was scored. Trials were deemed unscorable (and therefore rejected) if a blink was in progress at reflex stimulus onset or if the blink did not recover within the sampling period of 250 ms (return to at least 25% of peak amplitude).
4.3 Results

All the children were able to give appropriate examples of items or events that were scary or nice. Responses given included going on holiday, pets and chocolate (Nice), and spiders, the dark and horror films (Scary). There was no qualitative difference between the responses of the three groups.

The children’s individual ratings of the pictures were used to determine the categorisation of visual stimuli.

As Figure 4:1 shows, on average approximately half the 30 stimuli were rated as scary by the children. There was no group difference in the number of stimuli labelled as scary.

Figure 4:1  Mean number of pictures categorised as scary (Mean ± SEM)

Table 4:1  Mean number of trials with no blink response and mean number of rejected trials

<table>
<thead>
<tr>
<th>Group</th>
<th>No of trials with no response</th>
<th>No of trials rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>14.2 ± 2.1</td>
<td>1.7 ± 0.5</td>
</tr>
<tr>
<td>High</td>
<td>12.1 ± 2.7</td>
<td>2.9 ± 0.6</td>
</tr>
<tr>
<td>Low</td>
<td>8.3 ± 1.6</td>
<td>5.0 ± 0.9</td>
</tr>
</tbody>
</table>
As Table 4:1 shows, there was a high number of trials which did not elicit a response, but there was no significant difference in the number of no response trials between the groups (ANOVA: F(2, 39) = 2, p = 0.2). Trials with no blink response were also evenly distributed across the ‘Nice’ and ‘Scary’ categories. There were significantly more trials rejected in the Low group compared to the other two groups (ANOVA: F(2, 39) = 6, p = 0.005, Planned comparisons: CvH: t = -1, df = 39, p = 0.2; HvL: t = -2, df = 39, p = 0.04). This was not accounted for by verbal IQ.

Control Group:

In order to characterise the pattern of responses elicited by this paradigm in control children, prior to the group analysis the results from the control group were explored. Statistical analysis was carried out with a paired t-test.

Analysis of the results from the control group showed that there was a significant difference between the blink amplitudes to the two picture categories (t = -3, df = 15, p = 0.006). This was due to increased amplitude responses to pictures categorised as ‘Nice’. There was no significant difference between blink latencies to the two picture categories (Nice - Scary: Paired t test: t = 1, df = 15, p = 0.3) (see Figure 4:2 and Figure 4:3).

Group analysis

Statistical analysis of the group results was carried out using a repeated measure ANOVA design (within group factor: picture category).

**Latency**

There were no group or picture category differences on the mean latency of the blink response (Group: ANOVA: F(2, 34) = 1, p = 0.4; Picture category: ANOVA: F(1, 34) = 0.2, p = 0.7; Group x Picture category: ANOVA: F(2, 34) = 0.8, p = 0.4). Latency of responses was not correlated with verbal IQ, age or sex (see Figure 4:2).
Emotional Modulation of the Startle Response

Figure 4:2  Mean Latency of blink response by picture category (Mean ± SEM)

Amplitude
Analysis revealed an overall effect of picture category (ANOVA: F(1, 35) = 11, p = 0.003) and no effect of group or group by category interaction (see Figure 4:3).

Figure 4:3  Blink response according to individual child’s picture categorisation
4.4 Discussion

In this study, both Autistic groups showed the same pattern of emotional modulation of the startle response as the controls. To the extent that modulation of the startle response is regulated by the amygdala, the present results provided no evidence to suggest significant abnormality of this structure in the Autistic groups. However, there are a number of methodological and theoretical issues to address prior to concluding that the modulation of the startle response, and by implication, the function of the amygdala in children with Autism is normal.

Methodological Issues

The stimuli used in this study were chosen not to be as arousing as those typically used in adult paradigms (e.g. erotica or mutilated bodies) for ethical and practical reasons. As a result, it is possible that use of just 'moderately' arousing pictures rendered the paradigm less sensitive to detecting differences between the three groups.

Moreover, face stimuli were not included as stimuli for emotional modulation. Future studies should address whether face stimuli also produce the control pattern of modulation of startle. Such investigations could provide important insights into whether deficits in recognition of facial expression extend to implicit as well as explicit processing.

Studies of adults have often noted relative differences in startle size between men and women (e.g. Schupp et al. 1996). Although these sex differences are rarely statistically reliable [Bradley et al. 1999], one developmental study demonstrated significant differences between the emotional modulation of girls and boys aged 7-9 [McManis et al. 2001]. It is therefore possible that the absence of a group effect is due to a lack of sex matching between the control and Autistic groups. However, when only the boys’ data were included in the analysis, the results remained unchanged, suggesting that this is unlikely to explain the negative findings.
A further consideration in interpreting this study is the low response rate in all three groups. Although it is reported that up to 10% of subjects fail to show the startle response to even very intense stimuli [Omitz 1999], the response rate in this study fell below this level. This is likely to be due, at least in part, to the low decibel level of the startle probe (see Berg and Balaban 1999). This was chosen to minimise subjects failing to complete the paradigm due to discomfort or dislike of the probe. Maintenance of the child's co-operation was especially important in the light of the long protocol completed by participants in the investigations described in this thesis.

In this study, the child's own ratings were used to categorise pictures as 'Nice' and 'Scary'. This method was chosen as normed ratings of these stimuli were not available. Also, the procedure adopted has the advantage of allowing for individual variation in fears. For example, some children rated a picture of a wasp as 'Nice' whilst others rated it as 'Scary'. Such individual tailoring of categories maximises the possibility of detecting emotional modulation which might otherwise be masked by inclusion of non-scary stimuli in the scary category. It also prevents 'false' findings of abnormal response patterns merely due to any idiosyncratic fears or preferences of the children with Autism. This was desirable, as any such findings may not have been a reflection of the amygdala dependent modulation of the startle response.

Theoretical Issues
In addition to these methodological issues, there are a number of theoretical issues to address.

Firstly, the startle response only assesses one function of the amygdala. It is conceivable that whilst one function is intact, others are impaired. Future investigations should investigate other functional correlates of the amygdala (such as autonomic measures) (see Bechara et al. 1995).

There is some evidence to suggest that only some areas of the amygdala are involved in emotional modulation of the startle response. Stimulation studies in animals suggest that whilst central areas are involved, lateral areas of the amygdala are not [Rosen and Davis
Emotional Modulation of the Startle Response

1988; Rosen et al. 1996]. Therefore, the findings from this study do therefore not rule out the possibility of abnormality in the lateral amygdala in the children with Autism.

All of the studies that have implicated the amygdala in the emotional modulation of the startle response have been in adult animals and humans. It is possible that during development the amygdala is less critical to the emotional modulation of the startle response. Perhaps there are several pathways that initially support this modulation, and then gradually during development these pathways are pruned to the adult pattern. Investigations into the emotional modulation of the startle response in adults with Autism would address this question. It is also possible that neonatal lesions of the amygdala do not disrupt the emotional modulation of the startle response, even in adulthood. This would be consistent with the finding of normal fear responses in adult animals with neonatal medial temporal lesions compared to animals with medial temporal lesions sustained in adulthood [Nawla and Bachevalier 1991].

It is also possible that the reversed pattern of modulation found in children compared to adults obscures the detection of abnormality in the Autism groups. Why do the children in this study show the reverse pattern of modulation to adults? Trends in this direction have been noted in two previous studies of the startle response in school aged children [Cook et al. 1995; McManis et al. 1995] (although one of these only noted the trend in boys with girls showing the opposite trend). Neither of these studies used the children’s individual ratings of the stimuli, which may in part explain why the observed trends were not significant. In the current study, when the results from just the girls in the control group were analysed, the girls retained a trend towards fear-reduction startle. It is unclear why these results conflict with McManis et al.’s findings (although one possible explanation is the older age of the children in the present study).

It is, however, unclear why emotional modulation might show this inversion during childhood. One possible explanation is differing attentional resource allocation. It has previously been proposed that the modulation of the startle response in humans can reflect differing attentional processes [Bradley et al. 1993]. When more attention is paid to arousing stimuli, less attention is paid to the startle probe and the response to the latter is
therefore reduced. Perhaps the 'Scary' stimuli deemed appropriate for use with children are more arousing than the 'Nice' pictures chosen.

However, in conflict with this hypothesis, in a pilot study using these stimuli, adults demonstrated augmented startle responses to the 'Scary' pictures. Further, post hoc examination of the stimuli showed that pictures of the children's exemplars of items that were scary and nice had been included. This suggests that to the children (as well as to adults) the pictures were of a similar arousing nature regardless of the picture category. Additionally, all the pictures included in the stimuli set were similar (at least in description) to items rated as highly arousing in the pictures from the International Affective Picture System [Lang et al. 1995]. Future studies should address this issue more fully.

Perhaps the emotional stimuli are processed differently during development. However such a processing difference would need to account for findings of adult modulation patterns to facial stimuli in infants [Balaban 1995; Schmidt and Fox 1998] and to emotional contexts in adolescents [Grillon et al. 1999]. It remains to be established whether such an explanation is valid, and a number of issues regarding factors affecting differential processing (e.g. attention, variations in empathy, experience, etc.) remain.

In conclusion, in this study affective modulation of the startle response was found to be similar in controls and in the two Autistic groups. However, many questions were raised by the findings. This is perhaps inevitable given that the paradigm is relatively new and has not been used often with children. Replication of the study in adults and extension of stimulus conditions to include facial expressions are just two important aspects of future work. However, this study has demonstrated that the emotional modulation of the startle response is a practical tool for use in investigations in Autism. The non-verbal nature of the paradigm lends itself to the study of all individuals with Autism, regardless of their intellectual level.
Chapter 5  Executive Functions

This chapter describes investigations into the functions of the frontal lobes in Autism. There are a number of reasons to suspect that the frontal lobes (and in particular the orbitofrontal cortex) may be implicated in Autism. Firstly, the orbitofrontal cortex has extensive connections to the amygdala and the hippocampal formation (e.g. Barbas and De Olmos 1990; Barbas and Blatt 1995), and abnormal inputs from the medial temporal lobes have been shown to result in aberrant frontal lobe development [Bertolino et al. 1997; Hanlon and Sutherland 2000]. Secondly, damage to the orbitofrontal cortex results in abnormal social behaviour similar to that seen in Autism (e.g. Stone et al. 1998). Abnormalities in the structure and function of the frontal lobes have also been reported in Autism (e.g. Minshew et al. 1993; Carper and Courchesne 2000).

5.1 Introduction

The frontal lobe in humans (and other primates) is a large heterogeneous region that is widely connected with many cortical and subcortical structures. Different regions of the frontal lobes have different patterns of connections. One area of the frontal lobes with extensive connections to the medial temporal lobes is the orbitofrontal cortex [Porrino et al. 1981; Amaral and Price 1984; Barbas and De Olmos 1990; Morecraft et al. 1992; Barbas and Blatt 1995; Carmichael and Price 1995]. Normal development of the frontal lobes has been shown to be disrupted by lesions of the medial temporal lobe [Bertolino et al. 1997; Hanlon and Sutherland 2000]. In Autism, if the medial temporal lobes are abnormal, it is therefore likely that the development of the orbitofrontal cortex will also be affected. Whatever the initial pattern of abnormality in Autism (e.g. just medial temporal lobe, medial temporal lobe and orbitofrontal cortex, etc.) the orbitofrontal - medial temporal lobe connections are likely to confound normal development further. Other regions of the frontal lobes have few (if any) medial temporal lobe connections (for reviews see Barbas 1992; Barbas 2000). It may be speculated therefore, that these regions (even if affected in
the initial pattern of abnormality) may show comparatively preserved structure and function.

It is not only the anatomical connections of the orbitofrontal cortex that have led scientists to investigate this region in individuals with Autism. Patients with damage to this area show a pattern of behaviour that includes abnormal social behaviour. They appear able to analyse social situations correctly in the abstract, but have difficulty choosing appropriate forms of actions in real life and monitoring their own actions [Eslinger and Damasio 1985; Saver and Damasio 1991]. Their conversations tend to be perseverative and often inappropriate [Kaczmarek 1984; Mattson and Levin 1990]. Further, patients with bilateral orbitofrontal cortex damage have difficulty with Theory of Mind tasks [Stone et al. 1998] and understanding sarcasm [McDonald and Pearce 1996]. Functional imaging studies in normal adults have confirmed the role of the medial prefrontal cortex (including the orbitofrontal cortex) in ‘mentalising’ (e.g. Fletcher et al. 1995; Goel et al. 1995).

These observations have led many to investigate the integrity of the frontal lobes in Autism. Structural and functional abnormalities of the frontal lobes have been documented. Cortical abnormalities in the frontal lobes have been noted in some individuals with Autism at postmortem [Bailey et al. 1998]. Increased frontal lobe volume in individuals with Autism has also been reported [Carper and Courchesne 2000]. Other findings include decreased perfusion in the frontal lobes [George et al. 1992; Zilbovicius et al. 1995] and changes in metabolites involved in brain membrane synthesis and breakdown [Minshew et al. 1993]. A functional MRI study has revealed further abnormality: individuals with Autism participating in a theory of mind task did not display the control activation pattern in the medial frontal lobe [Happé et al. 1996].

It should be highlighted that selective abnormality of the orbitofrontal cortex in the frontal lobes of individuals with Autism has not yet been identified. Many of the above studies have used methods with too large a spatial scale to allow selective analysis (e.g. Minshew et al. 1993; Carper and Courchesne 2000). Others have demonstrated abnormality in medial regions but have not investigated the lateral regions (e.g. Happé et al. 1996).
Another approach investigating the integrity of the frontal lobes in Autism is to examine aspects of functioning reported to be abnormal in individuals with lesions to this region, e.g. executive functioning. 'Executive function' is an umbrella term describing the mental operations that allow one to guide behaviour by reference to mental models or future goals.

Executive function can be assessed using a variety of different paradigms. One example is the Wisconsin Card Sorting Test (WCST) [Heaton 1981]. The subject is presented with a set of cards that can be sorted by either colour, shape or number, and must deduce the correct sorting rule from the feedback given by the examiner who says whether each classification is right or wrong. Following 10 consecutive correct responses, the criterion for sorting is changed without warning the subject. Other tests of executive function include planning tasks (such as maze learning, Tower of London) and spatial working memory tasks.

It is well established that patients with frontal lobe damage show significant deficits in tests of executive function. They are impulsive, disinhibited, appear oblivious to the consequences of their actions and have difficulty learning from their mistakes [Bechara et al. 1994]. Similarly, studies of individuals with Autism have found impaired performance on a variety of executive function tasks, including WCST [Rumsey and Hamburger 1990; Ozonoff et al. 1991b], maze learning [Prior and Hoffmann 1990] and Tower of London [Hughes et al. 1994].

Failure on tests of executive function can not be taken as an indication of selective orbitofrontal cortex damage: lesions elsewhere in the frontal lobes also result in such deficits (e.g. Corcoran and Upton 1993; Upton and Corcoran 1995). Indeed tasks such as the Tower of London are thought to involve many areas of the frontal lobes (e.g. Rowe et al. 2001). However, error patterns on some tests of executive function have been shown to vary according to lesion location within the frontal lobe. For example perseverative errors on the WCST (repeatedly matching to an old criterion) are associated with dorsolateral prefrontal cortex damage, whereas failure to continue to match according to the criterion (i.e. failure to maintain set) is associated with orbitofrontal cortex damage [Stuss et al. 2000].
Individuals with orbitofrontal damage (in contrast to those with dorsolateral frontal lobe damage) have difficulty with extinguishing or reversing responses that have been previously rewarded [Rolls et al. 1994; Meunier et al. 1997]. After learning a stimulus-reward association, the reward contingencies are changed, and patients with orbitofrontal cortex lesions are unsuccessful in shifting their behaviour. This finding is consistent with functional imaging studies showing that the orbitofrontal cortex is involved in appraising the reward value of ongoing behaviour, especially under unpredictable circumstances [Elliott et al. 1997]. Additionally, in animals, electrophysiological recordings have shown that there are neurons in the orbitofrontal cortex that encode the reward value of visual stimuli used in learning and reversal tests [Thorpe et al. 1983; Rolls et al. 1996]. These neurons quickly change their firing rate to the visual stimuli depending on whether they are associated with reward.

In contrast, performance on spatial working memory and directed attention with resistance to interference tasks involves not the orbitofrontal cortex but areas elsewhere in the frontal lobes [Passingham 1985; Funahashi et al. 1989; Stuss 1991].

Although there is evidence of impairments in tasks associated with orbitofrontal cortex abnormality in children with Autism (e.g. reversal [Hughes et al. 1994]) the selective nature of such an abnormality has not been demonstrated, nor its association with pathology. Few studies have assessed frontal function comprehensively and there have been isolated reports of deficits consistent with dorsolateral prefrontal cortex damage (e.g. perseverative errors on WCST) [Rumsey and Hamburger 1990]. Further interpretation of these reports is confounded by the below average ability of the participants with Autism.

In the studies described in this chapter, the children's performance on tasks of executive function was assessed with the aim of delineating the nature of any deficit. Paradigms were selected to contrast performance on tasks associated with orbitofrontal function (reversal and WCST) [Stuss et al. 1983; Rolls et al. 1994; Nagahama et al. 1996; Stuss et al. 2000] with performance on tasks thought to be unaffected by orbitofrontal abnormality.
Executive Functions

(spatial working memory, directed attention with resistance to interference) [Stuss 1991; Vendrell et al. 1995; Owen et al. 1996a].

5.2 Methods

5.2.1 Tasks Sensitive To Orbitofrontal Cortex Abnormality

The Wisconsin Card Sorting Test (WCST) [Heaton 1981] measures conceptual problem solving abilities, including the ability to modify incorrect strategies flexibly and the ability to inhibit prepotent but incorrect responses. The child was given a series of cards and told “These are your cards. I want you to place the card underneath one of the four sample cards that you think it goes with best. I will tell you whether you are right or wrong. Your aim is to get as many right answers as possible”. The child was required to sort first by colour to criterion (10 correct consecutive responses), then shape and then number. Sorting criteria were changed without informing the child. The test stopped when 6 categories had been achieved or 128 cards were sorted.

Scoring noted the number of categories achieved, the number of correct responses (not including those in a run of 10 correct responses) and the class of the errors. The errors were classified as perservative (if the card was sorted according to the preceding criterion), unique (if the card didn’t match the sample card on either colour, form or number) and non-perservative errors (all remaining errors). A high number of correct responses is thought to be a reflection of a failure to maintain set.

Stimulus-Reward Associations

Reversal This test was based on the paradigm described in Rolls et al. [1994]. In this test the child learnt to touch one of two simple patterns that appeared one at a time on a touch screen. The stimuli were highly discernible coloured fractal images. The child gained one point for touching the correct pattern and one point for not touching the incorrect pattern. A point was lost for touching the incorrect pattern and another for not touching the correct pattern.
Patterns remained on the screen for seven seconds if not touched and disappeared immediately when touched (regardless of whether the response was correct). Once the pattern disappeared, it was replaced with a message providing feedback. Running totals of the child’s score (which could fall below zero) were displayed on the screen. Correct and incorrect responses were also signalled by different sounds.

The task was self-paced. Each child was asked to gain as many points as possible. Once a criterion of 9 correct responses out of the preceding 10 trials had been reached, the rule reversed: i.e. the relation between the patterns and the consequences of touching or not touching them was reversed without warning. Testing continued for 30 trials after the first reversal. If performance after the reversal reached criterion, then further reversals occurred whenever the criterion was reached again (up to a maximum of 3).

Performance was scored for the number of reversals achieved and the number of the last error trial on the first reversal (i.e. prior to the second reversal).

Extinction This test (also based on that described in Rolls et al. 1994) used two more highly discriminable fractal images as stimuli and was very similar to the reversal test. However, after criterion had been reached, the rule changed, and it became without warning incorrect to touch either pattern. Points were then only won by refraining from touching both of the patterns and were lost by touching either of them. The extinction test was run after the reversal test, with a gap of at least five minutes between tests.

Performance was measured by the number of the last error trial in extinction.

5.2.2 Tasks Insensitive To Orbitofrontal Cortex Damage

Spatial Working Memory
This test was based on a paradigm described in [Morris et al. 1988; Owen et al. 1990; Owen et al. 1992; Owen et al. 1993]. The child was required to ‘search through’ a number of boxes presented on the screen by clicking on each one which resulted in it
Executive Functions

‘opening’ to reveal what was inside. The object of the task was to collect ‘yellow tokens’ hidden inside the boxes. At any one time, there was a single token hidden inside one of the boxes and the child was required to search until this was found, at which point the next token was hidden. The key instruction was that once a yellow token had been found within a particular box, then that box was not used again to hide a token.

The child could search the boxes in any order. The number of empty boxes searched (excluding errors) before a token was found was randomly determined by the computer. After four practice trials with two boxes, there were four test trials with each of three, four, six and eight boxes.

Two types of search error were possible. First, a child could return to open a box in which a yellow token had already been found during the same trial (a ‘between-search’ trial). Secondly a child could return to a box which had already been opened and shown to be empty earlier in the same search sequence (a ‘within-search’ trial). The task was scored according to the number of ‘between’ and ‘within’ search errors at each level of difficulty.

The Stroop Colour and Word Test [Golden 1978]

This test was made up of three conditions: words to be read, colours to be named and interference. In each condition the child was asked to perform the task as quickly as possible for 45 seconds.

*Words to be read:* The child read aloud words arranged in columns. The words were colour names (blue, green and red) written in black ink. *Colours to be named:* The child named the colour of ‘XXXX’ stimuli printed in either blue, green, or red presented on a white sheet. *Interference:* The child named the colour (blue, green or red) of each word (blue, green or red) written in an incongruent ink colour.

For all conditions the experimenter noted the number of items completed. Normative data (with age correction) were used (as described in the manual) to calculate an interference score. The interference score compares the number of words completed in the first two trials with the number completed in the final trial. It therefore reflects not the child’s
reading or naming speed, but the change in this speed that occurred in the interference trial (i.e. the disruption caused by incongruent stimuli). The mean value of the interference score is 0; scores above 0 indicate high resistance to interference.

5.3 Results

5.3.1 Tasks Sensitive To Orbitofrontal Cortex Abnormality

Figure 5:1 shows the results from the Wisconsin Card Sorting Test. There was a significant difference between the groups on the number of categories achieved (ANOVA: F(2, 38) = 9, p = 0.001; Planned comparisons: CvH: t = 3, df = 38, p = 0.01; HvL: t = 2, df = 38, p = 0.1). This was a reflection of both Autistic groups attaining significantly fewer categories than the controls. This difference was not accounted for by verbal IQ differences.

There was no significant difference in the percentage of errors that were perseverative or unique (Perseverative: ANOVA: F(2, 38) = 1, p = 0.4; Unique: ANOVA: F(2, 38) = 0.8, p = 0.5). There was a significant difference between the groups in the number of correct responses (ANOVA: F(2, 38) = 5, p = 0.01; Planned comparisons: CvH: t = -3, df = 38, p = 0.004; HvL: t = 0.9, df = 38, p = 0.4). This was a reflection of both Autistic groups obtaining significantly more correct responses than controls and was not accounted for by differences in verbal IQ or sustained attention. The number of correct responses significantly correlated with the children’s performance on the Rivermead Behavioural Memory Test (see Chapter 3) (r = 0.4, p = 0.02).
Figure 5:1 Wisconsin Card Sorting Test  
(a) Number of categories; (b) Number of correct responses (Mean ± SEM)

![Graph showing the number of categories and correct responses for Control, High, and Low groups.]

Figure 5:2 shows the results from the stimulus-reward reversal test. There was no significant difference between the means of the groups on either measure (Number of reversals: ANOVA: F(2, 40) = 1, p = 0.2; Last error trial: ANOVA: F(2, 40) = 1, p = 0.4).

Figure 5:2 Reversal Results  
(a) Number of reversals achieved; (b) Last error trial on first reversal (Mean ± SEM)

![Graph showing the number of reversals and last error on reversal for Control, High, and Low groups.]

133
Figure 5:3 shows the results from the extinction test. The data were not normally distributed, so non-parametric statistical assessment was carried out. There was no significant difference between the means of the groups ($K-W \chi^2 = 4, df = 2, p = 0.1$).

5.3.2 Tasks Insensitive To Orbitofrontal Cortex Abnormality

Spatial Working Memory

Figure 5:4 shows the results from the spatial working memory task. The number of within-search errors increased significantly with search set size ($F(1.3, 37) = 22, p < 0.001$ (with Greenhouse-Geisser correction)). There was no significant difference between the means of the groups on the number of within-search errors (ANOVA: $F(2, 40) = 0.5, p = 0.6$).

The number of between-search errors increased significantly with search set size ($F(1.3, 37) = 150, p < 0.001$ (with Greenhouse-Geisser correction)). There was a significant difference between the means of the groups on the number of between-search errors (ANOVA: $F(2, 40) = 3, p = 0.04$: Planned comparisons: CvH: $p = 0.6$; HvL: $p = 0.06$). This is a reflection of the Low group scoring significantly more between-search errors than the control group. This difference was not correlated with age, sex or VIQ.
The Stroop Colour and Word Test
There were no significant differences between the means of the groups on the Stroop (ANOVA: $F(2, 40) = 1, p = 0.4$).

### 5.4 Discussion

The Autistic groups were found to achieve significantly fewer categories on WCST than controls. Although such a result is consistent with orbitofrontal cortex abnormality, similar performances have been noted following damage to other regions of the frontal lobe and other areas of the brain [Corcoran and Upton 1993; Upton and Corcoran 1995; Stuss et al. 2000]. Error pattern analysis in the two Autistic groups demonstrated no significant increase in perservative errors, but a significant increase in correct responses. This pattern reflects a loss of set (the child was able to match the cards according to the current criterion, but could not sustain this consistently for 10 consecutive cards for the category to be achieved).

Several studies have reported poor performance of individuals with Autism on the WCST (e.g. Prior and Hoffmann 1990; Rumsey and Hamburger 1990; Bennetto et al. 1996). The nature of the errors has not been reported in some of these studies (e.g. Rumsey and Hamburger 1990), making interpretation of the results difficult. Bennetto et al. reported...
no significant difference in set maintenance between individuals with Autism and their matched controls [Bennetto et al. 1996]. This discrepancy may be due to the lower intellectual ability of the subjects in Bennetto’s study (mean VIQ = 85). Individuals with learning disability may show increased set loss compared to average ability controls. This issue could be addressed by comparing the Low group (in this study) with ability matched controls. Further research is therefore needed.

Findings in the literature relating to the perseverative errors on the WCST in Autism are conflicting. One study reported increased levels of perseverative errors [Bennetto et al. 1996]. However, Hughes et al. report that perseverative error rates in a computerised task similar to WCST were not different in subjects with Autism from the rates of controls matched for learning disability [Hughes et al. 1994].

Failure to maintain set has been associated with orbitofrontal cortex abnormality. Stuss et al. reported increased set loss in patients with orbitofrontal cortex lesions [Stuss et al. 2000] and in frontal leucotomy patients [Stuss et al. 1983]. Nagahama et al. found in a PET study that controlling for maintenance of set reduced orbitofrontal cortex activation [Nagahama et al. 1996].

Several possible cognitive reasons for set loss have been proposed [Stuss et al. 2000]. Firstly, failure to suppress responses to irrelevant (but salient) stimuli may be responsible. However, the results from the Stroop test suggest that, at least to some extent, this skill is intact in individuals with Autism. Poor sustained attention would also result in set loss. However, there was no significant correlation between sustained attention scores and the number of correct responses on WCST. A third alternative is that the children had difficulty remembering which category they were currently sorting to. Consistent with this, there was a significant correlation between the Rivermead scores and the number of correct responses.

The performance of the three groups did not differ significantly on the two other measures aimed at assessing orbitofrontal cortex function (rule reversal and extinction). This test has been shown to be sensitive to gross neuropathology in the orbitofrontal cortex in adults
Executive Functions

[Rolls et al. 1994] and children (unpublished data from Vargha-Khadem’s laboratory). The intact performance of the children with Autism on the rule reversal and extinction may arguably suggest that poor orbitofrontal function can not explain the WCST results. However the complexity of WCST, and hence its sensitivity, is far greater than that of the computer tasks. Consistent with this, Hughes et al. found that Autistic children were able to complete the early stages of a computerised rule shifting paradigm (comparable to the rule reversal task administered in this task) [Hughes et al. 1994]. However, as the task became more complex (requiring both intra and extra dimensional shifts), the children with Autism were impaired.

The spatial working memory results suggested that the High group were similar to the control group on these measures. This lack of difference is unlikely to be due to a floor effect: neither group performed near chance levels [Owen et al. 1996b]. Further, the scores of both groups are comparable to other reports on children’s performance on this task [Luciana and Nelson 1998]. Imaging studies suggest that the dorsolateral prefrontal cortex is important for successful performance on this task [Owen et al. 1996a]. Animal studies have demonstrated that whilst dorsolateral prefrontal lesions affect spatial working memory, orbitofrontal lesions do not [Bachevalier and Mishkin 1986]. As such, the results are consistent with relatively preserved dorsolateral prefrontal cortex function in the High group.

The Low group showed a selective impairment in the spatial working memory task. They returned to a box where they had already found a token significantly more often than the control group (between-search errors). However, the Low group did not return to a box already found to be empty within a trial more often than controls (within-search errors). It is important that these results are interpreted with caution: the lack of verbal ability matched controls prevents demonstration that the deficit is associated with Autism and not learning disability. Additionally, it should be noted that the distribution of the between-search errors on this task was bimodal (for similar pattern see Chapter 3). Two of the children completed the task with a mean of less than 4 errors, whilst the remainder of the children made at least 8 errors.
Spatial working memory has not been reported before in Autism. Bennetto et al. investigated verbal working memory (e.g. Bennetto et al. 1996) and found deficits. However the tasks used in this study required the children to remember a list of sentences, and to remember the number of yellow dots on a series of presented cards. Both of these tasks can be argued to be assessing contextual memory rather than working memory. The children did not have to manipulate the stimuli on-line in any manner.

All three groups performed similarly on the Stroop task (consistent with previous reports, e.g. Blair et al. 2001). Although the Stroop is widely recognised as a test of frontal lobe function, understanding of the more specific areas involved is lacking. Regression analysis in a lesion study suggested that the lateral frontal cortex plays an important role [Vendrell et al. 1995]. Importantly for this study, patients with orbitofrontal cortex damage have been reported to perform similarly to controls on the Stroop task [Stuss 1991].

In summary, therefore, the results from this study are consistent with the hypothesis of selective abnormality of the orbitofrontal cortex in children with Autism. The possibility that this abnormality is associated with abnormality in the medial temporal lobes is conceivable in the light of the pattern of medial temporal and frontal connections (see introduction). However, a number of issues should be highlighted.

Firstly, deficits in executive functions have been found in patients with pathology outside the frontal lobes (e.g. Corcoran and Upton 1993; Upton and Corcoran 1995). Secondly, much of the evidence for localisation of function within the frontal lobes comes from adult lesion studies. It is possible that developmental abnormality of the frontal lobes may result in reorganisation of function, leading to different localisation of function. However, although developmental frontal lobe damage appears to result in a more aberrant social profile than similar damage in adulthood [Dolan 1999; Anderson et al. 2000], there have been no reports of functional localisation in these children. Functional imaging is likely to be an invaluable tool in addressing this question.

There is an additional factor that should be considered when interpreting the results of this study. It has been suggested that individuals with Autism perform better on computer tasks
than on the equivalent pen and paper tasks [Ozonoff 1995]. Although the pattern of deficit found in this chapter is not directly related to the mode of presentation, it remains possible that, for example, a manual spatial working memory task may reveal a deficit not detected by the computer version, or that a computerised version of WCST might show different results. Further research is needed to investigate this possibility and to attempt to determine what factors create any differences (such as requirement to interact with examiner). This issue will be returned to in the final chapter of this thesis.

Finally it should be repeated that whilst the inter-connectivity of the orbitofrontal cortex and medial temporal lobes suggests that any abnormality in either region would be compounded by aberrant projections, this does not preclude the possibility of areas outside the orbitofrontal cortex in the frontal lobes also being abnormal. Firstly, the initial pattern of pathology may include these regions. Secondly, the orbitofrontal cortex is extensively connected to the remainder of the frontal lobes (for reviews see Pandya et al. 1988; Barbas 1992), so abnormal orbitofrontal connections could affect the rest of the frontal lobes. Additionally, the medial temporal lobe is involved in functional circuits that include other regions of the frontal lobes (e.g. hippocampal formation and dorsolateral frontal cortex in spatial working memory (e.g. Aggleton et al. 1986; Owen et al. 1996a; Olton and Papas 2001)). Abnormality in these (indirect) connections may induce abnormality (albeit to a lesser extent) in these frontal lobe regions.

It is pertinent to address in this chapter the claim that an executive function disorder associated with frontal lobe abnormality underlies the syndrome of Autism (e.g. Hughes et al. 1994). There are a number of reasons to believe that this explanation of Autism is inadequate. Deficits in executive function have been noted in other neurodevelopmental psychiatric disorders including ADHD [Chelune et al. 1986], conduct disorder [Lueger and Gill 1990] and Tourette's syndrome [Incagnoli and Kane 1981]. Deficits in executive function therefore do not necessarily result in social impairment. Further, although early frontal lobe abnormality does results in aberrant social behaviour, affected children are not necessarily Autistic [Anderson et al. 2000]. This issue will be re-examined in Chapter 10.
Part II  Neuropathological Investigations

In Part I the neuropsychological profile of the children with Autism was extensively investigated. In Part II the focus turns to brain structure. A variety of different MR sequences were used to obtain complementary information about the children’s brains. These included clinical multi-slice MR scans, a 3-dimensional scan and scans providing additional focal information about the amygdala and hippocampal formation (called T2 maps).

Analysis of the MR scans was undertaken using a number of different methods. The clinical scans were examined by a neuroradiologist and the T2 maps were analysed using traditional analysis of variance statistics. The 3-dimensional scans were used to measure the volumes of the amygdala and hippocampal formation. In addition, these scans were analysed with voxel-based morphometry, a whole brain automated analysis technique.

Chapter 7 contains the results of the analyses seeking to characterise structural abnormalities in the groups of children with Autism. Prior to this, Chapter 6 describes the development of a new voxel-based morphometry (VBM) application which exclusively searches for bilateral abnormalities. Appendix D contains a series of technical studies which seek to validate different aspects of VBM methodology.

II.1. Voxel-based Morphometry

Voxel-based morphometry was developed to characterise cerebral grey and white matter differences in structural MRI scans. In contrast to methods that frame the search in terms of regions of interest, voxel-based morphometry can detect structural differences with uniform sensitivity throughout the brain. Voxel-based morphometry is essentially a technique that compares images of grey matter (or white matter) obtained from segmented MRI images. This comparison uses statistical parametric mapping to identify, and make inferences about, regionally-specific differences.
Before statistical analysis can be carried out, the 3D datasets are first normalised to place the data into a common stereotactic frame. The data are then segmented into grey matter and white matter using prior probability maps. The data are then smoothed using a Gaussian kernel. This step not only helps to ensure that the assumptions of the statistical model are met, but also sensitises the analysis to a specific spatial scale (see Figure II:1).

Figure II:1  Preprocessing Steps of Voxel-based Morphometry

Following completion of these pre-processing sets, the data are then analysed. The analysis carries out a Student’s t test at each volume element (voxel). Due to the large number of voxels in the brain, it is necessary to apply a correction to the significance values (to avoid false positives). The correction used is based on Gaussian Random Field theory and permits interpretation of significant results over the entire brain (corrected p values), or over a small volume (small volume correction (SVC) p values).

The results are commonly presented in two formats: the glass brain images and mean images. The glass brain images are shown in Figure II:2 and show all voxels significant above a given threshold (such as corrected p < 0.05). Alternatively, the significant voxels can be superimposed on the mean normalised images. This latter presentation facilitates identification of the location of the peaks (see Figure II:3).
Figure II:2  Examples of VBM results displayed as glass brain images. Each image shows all voxels significant above a given threshold, collapsed over one axis.

Figure II:3  Example of VBM results displayed superimposed on mean images
Chapter 6  

Detecting Bilateral Abnormalities With Voxel-Based Morphometry

6.1 Introduction

In this chapter, a technique is described that detects bilateral structural abnormalities using voxel-based morphometry and new results from Gaussian Random Field theory that enable inferences about conjoint changes in homologous brain regions. Below is first a review of voxel-based morphometry (VBM), a discussion of the importance of bilateral changes and then an introduction to the use of VBM for detection of bilateral effects using conjunction analyses. This chapter is based on the published paper: Salmond et al. [2000].

6.1.1 Voxel-Based Morphometry

Voxel-based morphometry has been used to look at structural abnormalities in a variety of patient populations. Wright et al. first used it to examine schizophrenic patients. They found correlations between syndrome scores and grey and white matter densities [Wright et al. 1995]. Vargha-Khadem et al. applied the technique to the KE family [Vargha-Khadem et al. 1998]. Half of the members of this family have an inherited speech and language disorder and voxel-based morphometry highlighted abnormalities in regions that included the caudate nuclei bilaterally. Gadian et al. have employed voxel-based morphometry to investigate children with perinatal hypoxic-ischaemic damage and found bilateral hippocampal abnormalities consistent with the atrophy seen on volumetric measurements [Gadian et al. 2000]. Sowell et al. have used voxel-based morphometry to explore neuroanatomical changes that occur during development [Sowell et al. 1999]. In many instances systematic differences in macroscopic grey matter anatomy were revealed that were not evident on conventional neuroradiological examination.
6.1.2 Unilateral And Bilateral Pathologies

Unilateral pathology in adults usually leads to severe, selective, high-level deficits such as apraxia, alexia or agnosia with double or even triple dissociations [Shallice 1988; Warrington and Warrington 1990]. However in children the loss of function following unilateral damage tends to be relative rather than absolute, leading to a diagnosis of dyspraxia, dyslexia or dysnosia [Vargha-Khadem et al. 1994]. Unilateral damage does not result in mutism or aphasia in childhood [Basser 1962]. In addition, unilateral damage is usually associated with a more general or non-specific decline in cognitive ability. Speech and language are often spared at the expense of other abilities provided that the damage is sustained in childhood [Vargha-Khadem and Mishkin 1997].

The differential outcome of unilateral pathology, depending on whether it is developmental or acquired in adulthood, may reflect the greater plasticity of the young brain, which in turn may facilitate reorganisation and compensation [Teuber 1975; Goldman 1979]. It should be noted that the site and size of the lesion plays a critical role, with some systems showing greater resilience than others [Teuber 1975]. Instances of both intra- and inter- hemispheric reorganisation have been noted [Satz et al. 1990].

The cognitive outcome of children with bilateral pathology is much poorer. There have been two case reports of children with bilateral perisylvian damage who have failed to acquire intelligible speech [Laudau and Kleffner 1957; Vargha-Khadem et al. 1985]. Members of the KE family suffer from a severe verbal dyspraxia and have been shown to have bilateral caudate abnormalities [Vargha-Khadem et al. 1998]. A further example comes from children who have suffered perinatal hypoxic-ischaemic injury. These children have a severe episodic memory impairment with relatively preserved semantic memory associated with bilateral hippocampal atrophy [Gadian et al. 2000].

These results have led to the conclusion that rescue of function, after brain injury in childhood is only possible if substrates for that function are preserved and operational in at least one hemisphere [Vargha-Khadem and Mishkin 1997]. Bilateral pathology should therefore be suspected in any patient with a selective impairment of cognitive function of a
Detecting Bilateral Abnormalities with Voxel-based Morphometry

developmental nature. However, it should be noted that unilateral lesions could result in similar cognitive deficits in patients who may have reduced neuronal plasticity due to late onset of pathology or because of the specificity of the site of lesion.

In addition to neurodevelopmental syndromes, there are other groups of patients who should also be suspected of having bilateral damage. For example patients with temporal lobe epilepsy may have bilateral abnormalities. Incisa della Rocchetta et al. have shown that even patients with well lateralised temporal lobe epilepsy may have bilateral pathology [Incisa-della-Rocchetta et al. 1995].

The functional importance of bilateral pathology in developmental disorders motivated the analytic developments described in this section. Currently voxel-based morphometry does not test explicitly for bilateral abnormalities. One problem is that unless a symmetric template is used, there is no guarantee that homologous regions are symmetrically positioned following non-linear spatial normalisation. Also, although statistical parametric mapping (SPM) gives a p value for any regionally specific effect on the right and a p value for any homologous effect on the left, there is no measure for the chance probability of a conjoint left and right (i.e. bilateral) effect. Since the latter’s chance probability is expected to be lower than either of the two unilateral probabilities, current approaches underestimate the significance of bilateral effects *per se* (i.e. a conjunction of homologous effects)

### 6.1.3 Voxel-Based Morphometry For Bilateral Effects

Here a new method is described, using SPM99 with modified algorithms, that searches explicitly for bilateral abnormalities. In order to examine 3D data sets for bilateral effects, a conjunction analysis is used that relies on new results from Gaussian field theory (see Worsley 1999 and Friston et al. 1999). Briefly, a conjunction analysis compares the results of two (or more) SPMs to find regions that are significant in both. To do this it creates a new statistical parametric map which contains the least significant t value from all the SPMs entered in the conjunction. After thresholding, this conjunction SPM can be
regarded as the intersection of the component SPMs to assess, in the present context, the conjoint expression of grey matter density changes, relative to controls, in both hemispheres. The two component SPMs are obtained by analysing flipped (right to left) and unflipped data. P values for maxima in the conjunction SPM are then corrected for the volume analysed.

6.2 Methods

6.2.1 Patients

Five patients (mean age 12.4; 4 males; 1 female) who had developmental amnesia associated with early hypoxic-ischaemic episodes were chosen for the test sample. All these patients have been shown to have bilateral hippocampal atrophy using volumetric methods in addition to voxel-based morphometry (for further details see Gadian et al. 2000). Eight controls (mean age 13.9; 3 males; 5 females) were also selected.

6.2.2 MRI Data Acquisition

All subjects were scanned on a 1.5 T Siemens Vision scanner, using a T1 weighted 3D MPRAGE sequence [Mugler and Brookeman 1990] with the following parameters: TR 10 ms, TE 4 ms and TI 200 ms; flip angle 12°; matrix size 256 x 256; field of view 250 mm; partition thickness 1.25 mm; 128 sagittal partitions in the third dimension; acquisition time 8.5 mins; no gap.

6.2.3 Data Analysis

The 3D data sets were analysed in SPM99 (Wellcome Department of Cognitive Neurology, London, UK). Each scan was normalised to a symmetric template to reduce structural asymmetries and more closely co-localise homologous regions. As SPM weights the fitting with a mask image of ones where there is brain, and zeros otherwise, this weighting image was made symmetric (by averaging with itself flipped in the transverse plane). The
algorithm was otherwise as described in Friston et al. [1995a] and Ashburner and Friston [1999a]. The data were normalised by global grey matter. This is a fundamental component that accounts for any differences among subjects that are simply due to differences in brain size.

The images were then segmented using a symmetric probability template (created by averaging with itself flipped) using the Bayesian algorithm described in [Ashburner and Friston 1997]. This produced continuous probability maps where the values correspond to the posterior probability that the voxel belonged to the grey matter partition. These grey matter images were then duplicated and flipped in the transverse plane. Flipping the image merely consists of a transposition across the sagittal plane defined as x = 0 in Talairach coordinates. The grey matter images were smoothed with a 4 mm isotropic Gaussian kernel. This smoothing renders the voxel values an index of the amount of grey matter per unit volume under the smoothing kernel. The term “grey matter density” is generally used to refer to this probabilistic measure. 4mm was chosen as the smoothing parameter as this corresponds roughly to the cross-sectional dimensions of the hippocampal formation and, by the matched filter theorem, sensitised the analysis to differences at this spatial scale.

The following statistical analyses were then carried out using SPM99. Four group-specific effects were modelled in the design matrix as shown below.

1: Controls (n = 8)
2: Patients (n = 5)
3: Flipped Controls (n = 8)
4: Flipped Patients (n = 5)

All differences of interest were assessed with contrasts of the group effects. A contrast refers to a linear compound of parameter estimates where the compound is defined by a vector of contrast weights. For example (1 −1 0 0...) would test the null hypothesis that the parameters associated with the first two columns of the design matrix are the same. The design matrix is simply a collection of explanatory variables that may explain variance in
the response variable. In our case, the explanatory variables are group membership and the response variable is grey matter density. Inferences about contrasts are made using the standard parametric statistics (in our case the t statistic, which is the contrast divided by its standard error). A conjunction is simply a significant effect expressed jointly over two or more contrasts. The following contrasts were tested:

**Standard Contrast**

The data were first assessed using a conventional single contrast (i.e. 1 -1 0 0). This tests for relative decreases in grey matter density in the patients relative to controls.

**Average Contrast**

The data were then examined for symmetric abnormalities by looking at the contrast (1 -1 1 -1). This creates a statistical parametric map that looks for significant differences between the controls and patients averaged over both hemispheres.

**Conjunction Analysis**

The data were finally examined for symmetric abnormalities using conjunction analysis, testing for the conjoint expression of bilateral effects. The two contrasts entered into the conjunction were 1 -1 0 0 and 0 0 1 -1.

For all three analyses proportional scaling to a grand mean of 100 was used to adjust for global grey matter differences.

### 6.3 Results

Inferences (i.e. p values) are restricted to the intensity statistics (i.e. height of maxima in the ensuing SPMs), as the corresponding Gaussian field results for spatial extent of regional effects have not yet been determined for conjunction SPMs.

Figure 6:1 shows the glass brain images of the three analyses and demonstrates that the results of the standard analysis (a) are, as expected, not symmetric whereas the SPMs
corresponding to the average analysis (b) and the conjunction analysis are symmetric (c). Also from this Figure it is clear that the conjunction analysis is more selective than the average analysis.

Figure 6:1  a: Standard Contrast; b: Average Contrast; c: Conjunction Analysis (Thresholded at uncorrected p = 0.001)

Figure 6:2 shows the position and extent of the grey matter abnormalities in the patients identified by all three analyses in the same coronal slice. The images are superimposed on the average of the normalised scans of the patients and the controls. From Figure 6:2 it can be seen that all three contrasts detect bilateral hippocampal effects. However the extent of the abnormality on the left is smaller with the standard contrast (a) than with the average contrast (b). The extent of the abnormality detected in the conjunction analysis (c) is smaller than with both the previous contrasts because the statistical acceptance criteria are more stringent (i.e. the abnormality needs to be present on the right and on the left in the conjunction analysis).
The t values of the average contrast appear much higher than the t values of the conjunction contrast (see Table 6:1). This is simply a reflection of the fact that a different t statistic is being used as the basis of inference. In the average analysis the t statistic is a single t value testing for the average effect. In the conjunction analysis the t value is the minimum of the two t values for each of the contrasts (1 -1 0 0 and 0 0 1 -1). This new t value has a different distribution under the null hypothesis and hence the p values differ. As a result of this, the conjunction analysis has smaller t values than the average contrast, but it yields much more significant p values.
Figure 6:3 demonstrates the differences between the corrected probability in the three different contrasts. All points that were significant in the conjunction analysis at the corrected level of $p = 0.05$ are shown. As can be seen, all the voxels in the standard analysis were less significant than in either of the two other analyses. In turn the average analysis results were less significant than the conjunction analysis at all points (there is no entry for some points as they had no corresponding maximum).

Figure 6:3 Comparison of Probability Values (Dotted line represents significance level - $p = 0.05$, All the points that are significant in the conjunction analysis at the corrected significance level of $p = 0.05$. Points 1, 3, 5, 6 are bilateral abnormalities and 2, 4 are abnormalities in the midline). See Table 6:1.

Further results of the conjunction analysis are shown in Figure 6:4. Regions that were significant at the correct 0.05 level were hippocampal (a and b), in the putamen (c) or on the sagittal midline (see below).
Detecting Bilateral Abnormalities with Voxel-based Morphometry

Figure 6:4  Conjunction Analysis Results (uncorrected p value = 0.01, left is left) a: bilateral hippocampal formation abnormality ±15 -36 0; b: bilateral hippocampal formation abnormality ±32 -28 -12; c: bilateral putamen abnormality ±34 3 -4,

6.4 Discussion

These results demonstrate that the conjunction analysis provides an improvement over the conventional analyses for the detection of bilateral abnormalities. The conjunction analysis has more specificity and sensitivity by virtue of employing two null hypotheses instead of one. Essentially one null hypothesis (no unilateral changes) constrains the assessment of the other (no homologous change in the other hemisphere), rendering the correction for multiple comparisons much less severe.

The conjunction analysis is therefore able to use the prior information (that bilateral abnormalities are suspected) to increase the power. This technique is clearly limited by the ability of the spatial normalisation to remove anatomical asymmetries. In regions that show great asymmetry, some bilateral deficits may be less easy to detect if that asymmetry cannot be adequately removed by spatial normalisation.
Detecting Bilateral Abnormalities with Voxel-based Morphometry

The hippocampal abnormality fits with both volumetric and standard SPM analyses and the putamen abnormality has been noted using a standard SPM approach (see Gadian et al. 2000).

All the midline abnormalities are likely to be associated with the smoothness in the data and should be discounted. This is because the data at the midline are the same in the flipped and unflipped scans, so clearly there is a higher probability of the voxels from both flipped and unflipped exceeding the threshold (i.e. a conjunction). Conjunction analyses require the contrasts to test for independent (or at least orthogonal) effects. This is violated at the midline. Only those voxels on the midline that are significant with the standard contrast should be reported. In short the conjunction analysis shows good specificity for bilateral abnormalities, although it is prone to false positives in the midline.

It is not proposed that the average contrast be used to detect bilateral abnormalities. This is because, if there were a severe unilateral effect, it is possible that this might be sufficient for the average of the flipped and unflipped differences to give rise to a significant bilateral abnormality. There is no danger of this happening in the conjunction analysis because both sides have to show an effect. In conditions such as those reviewed in the introduction, it is appropriate to hypothesise bilateral abnormalities and in these cases the use of a conjunction analysis as described should be a more powerful approach.

In this example, the data were smoothed to 4mm to sensitise the analysis to structures of a comparable size to the hippocampal formation, according to the matched filter theorem. The choice of smoothing kernel may impact on statistical validity: the general linear model assumes that after fitting, the residuals are normally distributed. The segmented images contain values that are mostly either zero (and slightly above), or one (and slightly below). Therefore, without any smoothing, the residuals are far from normally distributed. With more smoothing, the values are more normally distributed (by the Central Limit Theorem), and so the parametric models become more valid. Demonstration that 4mm smoothing is sufficient to prevent the violations of the assumptions of the statistical model is detailed in Appendix D. An alternative method is being developed by Taylor et al. [1998b] (re-
Detecting Bilateral Abnormalities with Voxel-based Morphometry

weighted logistical regression) which addresses this issue. In summary, by the Central Limit Theorem, the residuals will be normally distributed after spatial convolution.

An additional concern may appear to be that large areas of grey matter may render the spatial correlations among the residuals greater than those at the cortical sheet. This however is not a problem for the analysis since inferences made in SPM99 are valid in the context of non-stationary and non-isotropic smoothness. It should also be noted that only voxels surviving a grey matter threshold (0.8) are included in the analysis. This precludes analysing areas with near-zero variance (e.g. extra-cranial regions).

The general approach outlined in this section can be extended to functional characterisations. Bilateral structural lesions are not necessary to produce a severe developmental impairment. An example of this is children with Landau-Kleffner syndrome [Laudau and Kleffner 1957]. After having acquired normal speech, these children become aphasic as a result of bilateral interference with activity in the perisylvian areas by seizures emanating, in some cases, from the auditory cortex of one temporal lobe [Morrell et al. 1990; Pateau et al. 1991]. A further example of a functional bilateral abnormality is a child, Alex, with Sturge-Weber syndrome affecting the left hemisphere: he suffered from seizures and remained mute with a mental age of 3 or 4 until he was 8 and a half. At this point following a left hemispherectomy he became seizure free. Several months later he began to speak, eventually achieving a language level of an 11 year old [Vargha-Khadem et al. 1997b]. It is assumed that these cases result from the functional interference or inhibition of one hemisphere by the other. In such cases it may be appropriate to look for bilateral functional abnormalities. The conjunction analysis and flipping techniques described in this section may be suitable for this purpose.
Table 6:1  Location of all points significant in the conjunction analysis at the corrected level of $p = 0.05$ (excluding two midline points). Coordinates give the location of maxima in the Conjunction Contrast. Corresponding maxima for the other two contrasts were within 2 to 3mm. (Hpcpal is Hippocampal)

<table>
<thead>
<tr>
<th>Label on Figure</th>
<th>Region</th>
<th>x in mm</th>
<th>y in mm</th>
<th>z in mm</th>
<th>Standard Contrast t value</th>
<th>Average Contrast t value</th>
<th>Conjunction Contrast t value</th>
<th>Standard Contrast Corrected p value</th>
<th>Average Contrast Corrected p value</th>
<th>Conjunction Contrast Corrected p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1L</td>
<td>Hpcpal formation (Left)</td>
<td>-26</td>
<td>-38</td>
<td>-2</td>
<td>6.53</td>
<td>8.31</td>
<td>5.26</td>
<td>0.243</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>1R</td>
<td>Hpcpal formation (Right)</td>
<td>26</td>
<td>-38</td>
<td>-2</td>
<td>6.43</td>
<td>8.31</td>
<td>5.26</td>
<td>0.281</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>3L</td>
<td>Hpcpal formation (Left)</td>
<td>-15</td>
<td>-36</td>
<td>0</td>
<td>4.71</td>
<td>7.31</td>
<td>4.71</td>
<td>1.000</td>
<td>0.064</td>
<td>0.003</td>
</tr>
<tr>
<td>3R</td>
<td>Hpcpal formation (Right)</td>
<td>15</td>
<td>-36</td>
<td>0</td>
<td>5.72</td>
<td>7.31</td>
<td>4.71</td>
<td>0.721</td>
<td>0.064</td>
<td>0.003</td>
</tr>
<tr>
<td>5L</td>
<td>Hpcpal formation (Left)</td>
<td>-32</td>
<td>-28</td>
<td>-12</td>
<td>5.61</td>
<td>6.10</td>
<td>4.07</td>
<td>0.791</td>
<td>0.461</td>
<td>0.037</td>
</tr>
<tr>
<td>5R</td>
<td>Hpcpal formation (Right)</td>
<td>32</td>
<td>-28</td>
<td>-12</td>
<td>4.67</td>
<td>6.10</td>
<td>4.07</td>
<td>1.000</td>
<td>0.461</td>
<td>0.037</td>
</tr>
<tr>
<td>6L</td>
<td>Putamen (Left)</td>
<td>-34</td>
<td>3</td>
<td>-4</td>
<td>4.28</td>
<td>5.90</td>
<td>4.06</td>
<td>1.000</td>
<td>0.596</td>
<td>0.039</td>
</tr>
<tr>
<td>6R</td>
<td>Putamen (Right)</td>
<td>34</td>
<td>3</td>
<td>-4</td>
<td>4.18</td>
<td>5.90</td>
<td>4.06</td>
<td>1.000</td>
<td>0.596</td>
<td>0.039</td>
</tr>
</tbody>
</table>
Chapter 7  Neuropathology Of Autism

In this chapter, a description is given of experiments aiming to characterise the neuropathology of the High and Low groups. Voxel-based morphometry [Ashburner and Friston 1999a] was used to obtain a comprehensive characterisation of the entire brain, whilst a combination of other techniques (including volumetric measurements [Watson et al. 1992], and T2 maps of the amygdala and the hippocampal formation [Van Paesschen et al. 1996; Van Paesschen et al. 1997] was used to investigate the integrity of the medial temporal lobes in more detail.

7.1 Introduction

As discussed at length in the first chapter, there is abundant evidence to suggest that the medial temporal lobes play an important role in the neuropathology of Autism. However there are a number of reasons why abnormality in this area is unlikely to be the only abnormality found in individuals with Autism. Firstly, the environmental and genetic processes which induce medial temporal lobe abnormalities may also impact on the development of other areas. Secondly, development is an interdependent dynamic process, and so abnormality in the medial temporal lobes is likely to affect the development of other areas, such as the frontal lobes [Bertolino et al. 1997; Hanlon and Sutherland 2000]. It has also been argued that the complex nature of the symptomatology of Autism is suggestive of the involvement of several brain areas. These abnormalities may contribute to the behavioural profile of Autism. The location of these abnormalities may also provide clues as to the timing of the pathological process, as different sites or systems in the brain are likely to have distinct time periods when they are particularly sensitive to environmental or genetic processes or aberrant inputs.
Neuropathology of Autism

The chronic and selective developmental nature of the impairments characteristic of Autism is indicative of another important aspect of its associated neuropathology. In contrast to unilateral damage suffered in adulthood, comparable unilateral damage resulting from early onset neurodevelopmental or acquired pathology rarely leads to chronic and selective deficits in cognitive and behavioural functioning [Shallice 1988; Warrington and Warrington 1990; Vargha-Khadem et al. 1994]. The absence of such deficits is believed to be due to the increased plasticity of the immature brain which supports functional reorganisation of tissue in the homologous regions contralateral to the damaged side [Teuber 1975; Goldman 1979]. It is only when early damage to the neural substrates subserving a cognitive or behavioural system is bilateral that profound and chronic impairments are documented. Therefore, in Autism, where a genetic and a neurodevelopmental basis of the disorder is strongly implicated, the putative neuronal abnormality is suspected to be bilateral (see Chapter 6 for more details).

Having argued that there is likely to be bilateral abnormality in the brains of individuals with Autism outside the medial temporal lobes, the question arises as to where these regions are. Review of the literature reveals two other areas of the brain which, like the medial temporal lobes, have been implicated in Autism by a range of techniques: the frontal lobes and the cerebellum.

Before reviewing the evidence for the involvement of these areas, it must first be considered if abnormality in all three regions is plausible from a developmental perspective. There are a number of lines of evidence suggesting that this may be the case. Firstly, there are extensive connections between frontal areas (especially the orbitofrontal cortex) and the medial temporal lobes (especially the amygdala) [Porrino et al. 1981; Amaral and Price 1984; Barbas and De Olmos 1990; Morecraft, et al. 1992; Barbas and Blatt 1995; Carmichael and Price 1995] between the cerebellum and the medial temporal lobes [Heath and Harper 1974; Heath et al. 1978]; and between the cerebellum and the frontal lobes [Sasaki et al. 1979; Middleton and Strick 1994; Schmahmann and Pandya 1997]. Regardless of whether the abnormalities have a common aetiology or result from one region influencing another, the reciprocal connections between these regions would
have a sustained detrimental influence on development. Secondly, neonatal damage to the medial temporal lobes in animals has been shown to affect frontal lobe development [Bertolino et al. 1997; Hanlon and Sutherland 2000]. Thirdly, several neurodevelopmental events co-occur in these three regions. For example, neurogenesis in the hippocampal formation and amygdala occurs at the same time in the first trimester as the migration of the Purkinje cells to the posterior vermis in the cerebellum [Bayer and Altman 1987]. Furthermore, cell migration in the dentate gyrus and migration of the granule cells of the cerebellum occur within the same time frame [Rakic 1971; Abraham et al. 1999]. Finally, in a recent MR study, a correlation was found between the degree of abnormality in the frontal lobes and the degree of abnormality in the cerebellum in individuals with Autism [Carper and Courchesne 2000].

In the next three sections, the evidence for the involvement of all three areas in the neuropathology of Autism is summarised.

### 7.1.1 Evidence For Medial Temporal Lobe Abnormality

Post mortem studies of individuals with Autism have revealed abnormalities in the medial temporal lobes. Reports have included findings of increased cell density and abnormally small cells in the amygdala, hippocampal formation and entorhinal cortex [Bauman and Kemper 1985; Raymond et al. 1989; Bauman 1991; Bauman and Kemper 1994; Bailey, et al. 1998], decreased dendritic extents in the CA1 and CA4 regions of the hippocampal formation [Raymond et al. 1996], decreased hippocampal formation volume [Bauman and Kemper 1988] and neurofibrillary tangles in the entorhinal and perirhinal cortex [Hof et al. 1991].

A few imaging studies have quantitatively investigated the volumes of the medial temporal lobe structures in Autism [Saitoh et al. 1995; Piven et al. 1998]. Most studies, which have adjusted for total brain volume, have reported abnormalities in the amygdala and hippocampal formation (e.g. Aylward et al. 1999; Howard et al. 2000, but see Piven, et al. 1998). Whilst the hippocampal formation has been reported as abnormally small, the
Neuropathology of Autism

The amygdala has been found to be abnormally large in some studies and abnormally small in others [Aylward et al. 1999; Howard et al. 2000]. Different inclusion criteria, age and verbal intelligence of subjects may at least in part explain these conflicting results.

Other abnormalities have been reported in the medial temporal lobes using different techniques, including voxel-based morphometry, spectroscopy, regional cerebral blood flow (rCBF) and functional magnetic imaging (fMRI). Abell et al. reported abnormally high grey matter density in the left amygdala in a voxel-based morphometry study [Abell et al. 1999]. MRS studies have suggested neuronal loss or damage in the amygdala-hippocampal region [Otsuka et al. 1999]. Abnormal rCBF at rest has also been reported in the right medial temporal lobe [Ohnishi et al. 2000]. In a recent imaging study, Baron-Cohen et al. reported that a group of individuals with Autism showed the same brain activation pattern as controls in an fMRI task, with the exception of the amygdala, which was only activated in the controls [Baron-Cohen et al. 1999].

Although these studies are consistent with structural and functional abnormalities in the medial temporal lobes, it should be noted that attempts at replication of these findings is warranted. The known heterogeneity of Autism requires that such replication studies be carried out to validate the generalisation of the findings beyond a specific experimental group. The unilateral nature of many of the findings also raises questions regarding the role of these abnormalities in the pathology of Autism, as unilateral developmental abnormality rarely results in profound, selective deficits (see above and Chapter 6).

Another important line of evidence supporting the role of the medial temporal lobes in Autism comes from studies of behaviour in animals and humans with medial temporal lobe damage. Bachevalier and colleagues have demonstrated that bilateral ablations of the medial temporal lobes in neonatal monkeys result in a pronounced and sustained constellation of behavioural abnormalities reminiscent of Autism [Bachevalier 1994]. By the age of 6 months, the monkeys spend less time in social interaction, actively avoid social contacts, develop locomotor stereotypies (significantly more than control monkeys) and self-directed behaviours, have poor facial and body expression and display little eye contact (see Chapter 1 for more information). Difficulties interpreting facial expressions have been
reported in groups of individuals with amygdala damage (e.g. Adolphs et al. 1994; Adolphs et al. 1995; Young et al. 1995a), as well as in children with Autism (e.g. Hobson, et al. 1988; Teunisse and Gelder 2001). Selective memory deficits for events (episodic memory) have been reported in individuals with bilateral developmental hippocampal damage [Vargha-Khadem et al. 1997a; Gadian et al. 2000]. Although episodic memory has not been systematically investigated in Autism, the pattern of memory impairment reported in the literature is consistent with a selective deficit in episodic memory (see Chapter 3).

7.1.2 Evidence For Frontal Lobe Abnormality

Abnormalities in the frontal lobes in individuals with Autism are less well documented than those in the medial temporal lobes. However, a number of different functional imaging studies have indicated abnormality in this region. For example, decreased perfusion in the frontal lobes has been reported by several authors [George et al. 1992; Zilbovicius et al. 1995; Carper and Courchesne 2000]. Minshew et al. reported changes in metabolites involved in brain membrane synthesis and breakdown in the frontal lobes [Minshew et al. 1993]. Autistic individuals participating in a theory of mind fMRI task did not show the control activation pattern in the medial frontal lobe [Happé et al. 1996].

Structural abnormalities in the frontal lobes of individuals with Autism have also been reported. Cortical abnormalities have been found in the frontal lobe of some individuals with Autism at postmortem [Bailey et al. 1998]. Some individuals with Autism have been found to have increased total volume of the frontal lobe [Carper and Courchesne 2000]. The similarities in the behavioural profiles of individuals with frontal lobe damage and those with Autism have led many to hypothesise abnormality in this area. Similarities include difficulties with social skills (as illustrated on tasks using Theory of Mind, faux pas and sarcasm paradigms) [McDonald and Pearce 1996; Stone et al. 1998] and poor executive skills (such as rule reversal and perseveration) (e.g. Hughes et al. 1994) (for further details see Chapter 5).
7.1.3 Evidence For Cerebellar Abnormality

There is also convergent evidence of cerebellar abnormality from imaging and post mortem studies. At post-mortem, reduced number of Purkinje cells in the cerebellum have been found [Bauman and Kemper 1985]. Abnormalities reported in the cerebellum in imaging studies include both hypoplasia and hyperplasia of the vermal lobules 6 and 7 [Courchesne, et al. 1988; Courchesne et al. 1994a]. Additionally MRS studies have shown decreased NAA concentration (reflecting neuronal loss or damage [Gadian 1995]) in the cerebellum of individuals with Autism [Otsuka et al.1999].

Little is known about the functional role of the cerebellum in behaviour (although a role in shifting attention has been hypothesised [Courchesne et al. 1994b]). Unlike the cerebral cortex, different areas of the cerebellum have not yet been associated with functional specialisation. However cerebellar abnormality is associated with deficits in motor coordination [Verleger et al. 1999; Kakizawa et al. 2000] and poor fine and gross motor skills are frequently reported in individuals with Autism (e.g. Mawson et al. 1985; Tantam 1988a; Cox 1991).

7.1.4 Summary

There is evidence for brain abnormality in three different regions of the brains of individuals with Autism: the medial temporal lobes, the frontal lobes and the cerebellum. In the following sections, a number of structural magnetic resonance (MR) studies are described. Two studies investigate the structural integrity of the entire brain (clinical neuroradiological assessment and voxel-based morphometry), whilst the reminder focus on structures in the medial temporal lobes (volumetrics and T2 maps of the hippocampal formation and the amygdala). A general discussion of all the results follows the description of the individual studies (see section 7.5).
7.2 Clinical Analysis Of MRI Scans

7.2.1 Introduction

It is generally agreed that the brains of children with Autism are grossly normal. Although gross structural abnormalities were initially reported in Autism, these findings have since been traced back to other associated genetic, infectious or neurological conditions [Minshew and Dombrowski 1997]. Clinical neuroradiological assessment provides a qualitative assessment of brain abnormalities that can be seen on visual inspection of the images.

7.2.2 Methods

7.2.2.1 MRI Acquisition

All subjects were scanned unsedated on a 1.5 T Siemens Vision scanner. A series of scans were collected, resulting in a total scanning time of ~ 30 minutes (see Table 7:1). As part of this series of scans, coronal and axial T2-weighted scans were collected.

Table 7:1 MR Protocol Details

<table>
<thead>
<tr>
<th>Scan</th>
<th>TR (ms)</th>
<th>TE (ms)</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2 Axial</td>
<td>3458</td>
<td>96</td>
<td>19 slices, 5mm thick</td>
</tr>
<tr>
<td>T2 Coronal</td>
<td>3458</td>
<td>96</td>
<td>19 slices, 5mm thick</td>
</tr>
<tr>
<td>3D FLASH</td>
<td>16.8</td>
<td>5.7</td>
<td>Flip = 12°, Voxel size 0.8 x 0.8 x 1 mm</td>
</tr>
<tr>
<td>Hippocampal T2 Map</td>
<td>2400</td>
<td>22-262</td>
<td>1 slice, 5mm thick</td>
</tr>
<tr>
<td>Amygdala T2 Map</td>
<td>2400</td>
<td>22-262</td>
<td>1 slice, 5mm thick</td>
</tr>
</tbody>
</table>
7.2.2.2 Data Analysis

The images were reviewed by an experienced paediatric neuroradiologist, blind to the group membership of the children, who recorded presence or absence of abnormality on visual inspection of the images. Particular attention was paid to the amygdala, hippocampal formation, orbitofrontal cortex and cerebellum. Any ventricular asymmetry was also noted.

7.2.3 Results

As shown in Table 7:2, the majority of the children showed no abnormality on neuroradiological assessment. Only one child (who was in the High group) was reported to have abnormality in the hippocampal formation. No abnormalities in the amygdala or orbitofrontal cortex were noted. Abnormalities in ventricular asymmetry and the cerebellum were found in both the Autistic groups and the control group, with approximately the same frequency.

<table>
<thead>
<tr>
<th>Number of children showing abnormality</th>
<th>Control</th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hippocampal Formation</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Amygdala</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Orbitofrontal cortex</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Ventricular Asymmetry</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
7.2.4 Discussion

This section demonstrated that, on visual inspection of MRI scans, there were no obvious
differences among the three groups. This is consistent with previous reports of children
with Autism with no known aetiology, such as encephalitis or rubella (e.g. Minshew and
Dombrowski 1997). However, this does not preclude the possibility of more subtle
differences that are not visible on neuroradiological examination. In the following sections,
more sensitive methods are used to further investigate the integrity of the brains of the
children with Autism.

7.3 Voxel-Based Morphometry

7.3.1 Introduction

Voxel-based morphometry (VBM) is a technique that characterises cerebral grey and white
matter differences in structural MRI scans. VBM has uniform sensitivity throughout the
entire brain. It essentially compares images of grey matter (or white matter, obtained from
segmented MR images), using statistical parametric maps to identify and make inferences
about regionally specific differences.

As described in Chapter 6, it is possible to restrict the analysis in VBM to exclusively
search for bilateral abnormalities. This method is applied here, to search for bilateral
abnormality throughout the entire brain.
7.3.2 Methods

7.3.2.1 MRI Acquisition

A T1 weighted 3D FLASH sequence was used as detailed in Table 7:1. This sequence gives good grey-white matter differentiation.

7.3.2.2 Data Analysis

The 3D data sets were analysed in SPM99 (Wellcome Department of Cognitive Neurology, London, UK) in accordance with the method described in Salmond et al. [2000]. Briefly each scan was normalised to a symmetric template [Friston et al. 1995b; Ashburner and Friston 1999b]. The data were normalised by global grey matter. This is a fundamental component that accounts for any differences among subjects that are simply due to differences in brain size. The images were then symmetrically segmented using a modified version of the Bayesian algorithm described in Ashburner et al. [1997]. This produced continuous probability maps where the values correspond to the posterior probability that the voxel belonged to the grey matter partition. The grey matter images were then duplicated and one copy of each scan was flipped in the transverse plane. The grey matter images were then smoothed with 12 mm and 4 mm isotropic Gaussian kernels. This smoothing renders the voxel values an index of the amount of grey matter per unit volume under the smoothing kernel. The term “grey matter density” is generally used to refer to this probabilistic measure. 4mm and 12mm were chosen as the smoothing parameters as these correspond roughly to the cross sectional dimensions of the hippocampal formation and amygdala, respectively and, by the matched filter theorem, sensitised the analysis to differences at these spatial scales. The data were then examined for symmetric abnormalities using conjunction analysis (see Chapter 6). Age and sex were included as covariates. Similar analyses were carried out with the white matter images.
Inferences from statistical parametric maps were made at two different threshold levels. Firstly, the significance values corrected for multiple comparisons across the entire brain were used to report abnormalities throughout the entire brain. Secondly, incorporating our prior central hypothesis of bilateral medial temporal lobe abnormality, a small volume correction was also used. This small volume correction was made using a sphere (radius 15 mm, centred at \(-27, -15, -18\)) (see Figure 7:1). This small volume was chosen to be highly conservative. Not only does it encompass the amygdala and much of the hippocampal formation, its total volume (~13,500 mm\(^3\)) is far greater than the volume of these structures. All peaks that survived correction for multiple comparisons (either at total brain or small volume level (labelled SVC)) are reported, provided the x co-ordinate of the peak was more than 2 times the resolution of the statistical parametric map (to avoid false positives in the sagittal midline, see Salmond et al. [2000] and Chapter 6).

The statistical parametric maps were superimposed on the mean normalised image of the group data in order to aid anatomical location. Identification of the anatomical location of the areas of abnormality was carried out with reference to Duvernoy’s atlas [Durvernoy 1991]. All figures show the statistical parametric maps superimposed on the mean normalised image of the group data, at a threshold of uncorrected \(p < 0.005\). The cross hairs on the figures indicate the location of the maximal peak. Figures are displayed in neurological convention (left hemisphere is on the left).

In order to determine the common areas of abnormality and significant differences in the areas of abnormality between the High and Low groups, a further conjunction analysis was carried out. Areas of interest in this analysis were only areas where either or both groups were significantly different from controls (in order to detect pathological differences and similarities only). In order to restrict the analysis to only these areas, control groups were included in the analysis. The conjunction analysis requires subject groups to be orthogonal and so two different (non-overlapping) control groups were used. Twelve controls for the High group were selected from the main pool of control children who participated in this study, matching as closely as possible for age and verbal IQ. The controls for the Low group consisted of the remaining controls from this main pool and an additional 7 children who had taken part in previous studies at Great Ormond Street Hospital. Selection was
carried out in this way to ensure that the controls for the High group were as closely matched as possible. Such close matching for the Low group would not have been possible due to the low verbal IQ of this group. Unfortunately (as can be seen from the average VIQ in Table 7:3), the additional controls used in the Low group analysis had above average verbal IQ. This led to a substantial verbal skills discrepancy between the Low group and their controls. Additional analyses were carried out with the Low group including verbal IQ as a covariate in an attempt to account for this discrepancy.

For consistency, the comparisons of High versus Controls and of Low versus Controls were also carried out with these control groups.

Table 7:3 Characteristics of Control Groups for VBM

<table>
<thead>
<tr>
<th></th>
<th>Mean Age</th>
<th>Sex</th>
<th>VIQ</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls for</td>
<td>12</td>
<td>5 male; 9</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>High group</td>
<td></td>
<td>female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls for</td>
<td>11</td>
<td>7 male; 3</td>
<td>111</td>
<td>(includes 7 children not</td>
</tr>
<tr>
<td>Low group</td>
<td></td>
<td>female</td>
<td></td>
<td>assessed behaviourally)</td>
</tr>
</tbody>
</table>

Figure 7:1 Centre of Small Volume Correction
7.3.3 Results

7.3.3.1 Results From The Comparison Of The High Group Versus Controls

There were no significant decreases in grey matter density in the High group compared to the controls at either level of smoothing. Neither were there any significant differences in white matter density (either increases or decreases) in the High group compared to the controls at either level of smoothing.

Significant increases in grey matter density were found in the High group at both levels of smoothing (4mm and 12mm). These increases were found in the amygdala and hippocampal formation, the orbitofrontal cortex, the superior temporal gyrus and the cerebellum (see Figure 7:2 - Figure 7:7 and Table 7:4 below).

Increases in Grey Matter Density in High group compared to controls

Figure 7:2  ±27 –18 –18 Hippocampal Formation (4mm, SVC)
Figure 7:3  ±20 -10 -24 Amygdala (+ rostral hippocampal formation) (12mm, SVC)

Figure 7:4  ±24 21 -18 Orbitofrontal cortex (12mm)

Figure 7:5  ±52 18 -21: Rostral Superior Temporal Gyrus (12mm)
Figure 7:6 ±24 –74 –45 Cerebellum (Inferior region (including VIIB, VIII A and Crux II)) (4mm)

Figure 7:7 ± 28 –66 –22 Cerebellum (Superior, including areas VI, V, IV) (4mm)
Table 7:4 Increases in Grey Matter density in High group compared to Controls (* SVC)

<table>
<thead>
<tr>
<th>Coordinates</th>
<th>Location</th>
<th>Z score</th>
<th>Corrected p value</th>
<th>Smoothing</th>
<th>Figure Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>±24 -74 -45</td>
<td>Cerebellum (Inferior, VIIB, VIIB, Crux II)</td>
<td>5.88</td>
<td>0.001</td>
<td>4mm</td>
<td>Figure 7:6</td>
</tr>
<tr>
<td>±28 -66 -22</td>
<td>Cerebellum (Superior, VI, V, IV)</td>
<td>5.49</td>
<td>0.010</td>
<td>4mm</td>
<td>Figure 7:7</td>
</tr>
<tr>
<td>±27 18 18</td>
<td>Hippocampal Formation</td>
<td>4.35*</td>
<td>0.045*</td>
<td>4mm</td>
<td>Figure 7:2</td>
</tr>
<tr>
<td>±52 18 21</td>
<td>Rostral Superior Temporal Gyrus</td>
<td>3.45</td>
<td>0.021</td>
<td>12mm</td>
<td>Figure 7:5</td>
</tr>
<tr>
<td>±24 21 18</td>
<td>Orbitofrontal Cortex</td>
<td>3.41</td>
<td>0.025</td>
<td>12mm</td>
<td>Figure 7:4</td>
</tr>
<tr>
<td>±20 10 24</td>
<td>Amygdala</td>
<td>3.58*</td>
<td>0.042*</td>
<td>12mm</td>
<td>Figure 7:3</td>
</tr>
</tbody>
</table>

As the control group was not matched to the High group on sex, analyses were run comparing just the males from the High group (n = 11) with just the males in the control group (n = 5), and then with just the females in the control group (n = 9). Z scores at the six peaks were compared (rather than the p values). As Table 7:5 shows, even though reducing the number of individuals in a comparison reduces the power of the analysis, the significance of the peaks in the two comparisons are very similar. This suggests that the results are not due to failure to match the control group to the High group on sex.
Table 7:5  The effect of sex on the VBM results

<table>
<thead>
<tr>
<th>Coordinates</th>
<th>Location</th>
<th>Z score Girls (n = 9)</th>
<th>Z score Boys (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>±24 -74 -45</td>
<td>Cerebellum</td>
<td>2.84</td>
<td>3.9</td>
</tr>
<tr>
<td>±28 -66 -22</td>
<td>Cerebellum</td>
<td>3.27</td>
<td>3.04</td>
</tr>
<tr>
<td>±27 -18 -18</td>
<td>Hippocampal Formation</td>
<td>3.97</td>
<td>3.85</td>
</tr>
<tr>
<td>±52 18 -21</td>
<td>Rostral Superior</td>
<td>3.01</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>Temporal Gyrus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>±24 21 -18</td>
<td>Orbitofrontal cortex</td>
<td>2.23</td>
<td>3.76</td>
</tr>
<tr>
<td>±20 -10 -24</td>
<td>Amygdala</td>
<td>&lt;1.92</td>
<td>&lt;1.92</td>
</tr>
</tbody>
</table>

In order to determine the homogeneity of the neuropathology of the children in the High group, individual analyses were also carried out. The methods used were identical to those described in Section 7.3.2.2, comparing each child in the High group against the entire control group. Interpretation of such unbalanced designs is complex, as these analyses are prone to false positives (see Appendix D) and do not possess the statistical power of group comparisons. The latter point can be addressed by using a lenient statistical threshold of uncorrected $p < 0.001$ and acknowledging that absence of abnormality in a particular region is not evidence of normality. In order to address the problem of false positives, the results from the statistical parametric maps were not interpreted as described in Section 7.3.2.2. Instead, presence or absence of abnormality was noted in each of the 5 regions where group abnormality was found: hippocampal formation, amygdala, orbitofrontal cortex, superior temporal gyrus and cerebellum.

The results from the individual voxel-based morphometric analyses are shown in Table 7:6. These are summarised in Table 7:7.
Table 7:6  Results from individual VBM analyses Y indicates presence of abnormality at threshold uncorrected $p < 0.001$, N indicates no significant abnormality at threshold.

<table>
<thead>
<tr>
<th>Name</th>
<th>Hippocampal formation</th>
<th>Amygdala</th>
<th>Orbitofrontal cortex</th>
<th>Superior Temporal Gyrus</th>
<th>Cerebellum</th>
</tr>
</thead>
<tbody>
<tr>
<td>AACA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>ASRO</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>BETR</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>CRWA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>JOCO</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>JOWE</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>JONA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>KERE</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>LEDA</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>MAWA</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>RONI</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>RYWI</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>TICH</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>TIWI</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 7:7  Summary of Group Variation According to VBM analyses

<table>
<thead>
<tr>
<th>Area</th>
<th>Number of individuals showing abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hippocampal formation</td>
<td>7</td>
</tr>
<tr>
<td>Amygdala</td>
<td>7</td>
</tr>
<tr>
<td>Orbitofrontal Cortex</td>
<td>13</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>10</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>11</td>
</tr>
</tbody>
</table>
Correlations with Cognitive and Behavioural Skills

Many cognitive and behavioural skills have been associated with the integrity of different areas of the brain (e.g. memory and the medial temporal lobes, see Chapter 3). Voxel-based morphometry can be used to determine areas of the brain that are correlated with indices of performance. In the light of the results from Part I of the thesis, correlations were therefore investigated in Motor-ABC (Motor co-ordination); Rivermead Behavioural Memory Test (episodic memory) and Number of categories and number of correct responses in the Wisconsin Card Sorting Test (executive function). Additionally, correlations with the children’s scores on the Autistic Behaviour Checklist, Children’s Obsessions and Compulsions Inventory and the Children’s Communication Checklist were also examined.

The methods described in Section 7.3.2.2 were used. Instead of inter-group comparisons, intra-group correlations were entered into the design matrix, with age and sex as covariates of no interest. Significant results were determined by using a corrected statistical threshold for whole brain (as described in Section 7.3.2.2) and a small volume correction of 5mm radius around the coordinates of the peaks of abnormality detected in grey matter density (see Table 7:5).

A significant correlation was found in the amygdala with the Autism Behaviour Checklist (p = 0.01 SVC). No other significant correlations were found using voxel-based morphometry with any of the measures of performance.

Individual voxel-based morphometry analyses of the children in the High group showed differing patterns of significant areas of abnormality (see above). Performance indices associated with the functions of these areas were examined to determine if there appeared to be a similar pattern of abnormality. Statistical analysis was not used, due to the small sample sizes. As can be seen in Table 7:8 and Table 7:9, the children with significant abnormality detected in individual VBM analyses showed a similar behavioural and cognitive profile to those children in the High group who did not show significant abnormality.
### Table 7.8 Performance Scores according to presence of detected abnormality on individual VBM

<table>
<thead>
<tr>
<th>Area of abnormality</th>
<th>Range of scores of children in High group with significant detected abnormality in this region</th>
<th>Range of scores of children in High group with no significant detected abnormality in this region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hippocampal Formation (Rivermead)</td>
<td>13-19</td>
<td>13-22</td>
</tr>
<tr>
<td>Orbitofrontal Cortex (Number of correct responses)</td>
<td>6-50</td>
<td>44 (*)</td>
</tr>
<tr>
<td>Cerebellum (M-ABC)</td>
<td>3-84</td>
<td>0-53</td>
</tr>
</tbody>
</table>

* n = 1

### Table 7.9 Symptom severity according to presence of detected abnormality on individual VBM

<table>
<thead>
<tr>
<th>Area of abnormality</th>
<th>Range of scores on Autistic Behaviour Checklist of children in High group with significant detected abnormality in this region</th>
<th>Range of scores on Autistic Behaviour Checklist of children in High group with no significant detected abnormality in this region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hippocampal Formation</td>
<td>10-76</td>
<td>35-94</td>
</tr>
<tr>
<td>Orbitofrontal Cortex</td>
<td>10-94</td>
<td>35 (*)</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>10-94</td>
<td>46-74</td>
</tr>
</tbody>
</table>

* n = 1
7.3.3.2 Results From The Comparison Of The Low Group Versus Controls

The results from the Low group comparison showed more widespread abnormality. Decreased grey matter density was found in the amygdala, parahippocampal region, superior precentral gyrus, operculum bank of the subcentral gyrus and the fusiform gyrus (see Figure 7:8 - Figure 7:12 below and Table 7:10).

Low: Decreases in Grey Matter Density

Figure 7:8 ±16 −4 −20 Amygdala (12mm) (SVC)

Figure 7:9 ±30 −26 −27 Parahippocampal region (4mm) (SVC)
Figure 7:10  ±44 -2 48 Superior precentral gyrus (12mm)

Figure 7:11  ±42 -10 14 Operculum bank of the subcentral gyrus (12mm)

Figure 7:12  ±39 -34 -28 Fusiform gyrus (12mm)
Table 7:10 Decreases in Grey Matter density in Low group compared to Controls (* SVC)

<table>
<thead>
<tr>
<th>Coordinates</th>
<th>Location</th>
<th>Z score</th>
<th>Corrected p value</th>
<th>Smoothing</th>
<th>Figure Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>±39 -34 -38</td>
<td>Fusiform gyrus</td>
<td>5.65</td>
<td>0.005</td>
<td>4mm</td>
<td>N/A</td>
</tr>
<tr>
<td>±30 -26 -27</td>
<td>Parahippocampal region</td>
<td>4.46*</td>
<td>0.032*</td>
<td>4mm</td>
<td>Figure 7:9</td>
</tr>
<tr>
<td>±44 -2 48</td>
<td>Superior precentral gyrus</td>
<td>5.72</td>
<td>0.001</td>
<td>12mm</td>
<td>Figure 7:10</td>
</tr>
<tr>
<td>±39 -34 -28</td>
<td>Fusiform gyrus</td>
<td>5.47</td>
<td>0.002</td>
<td>12mm</td>
<td>Figure 7:12</td>
</tr>
<tr>
<td>±42 -10 14</td>
<td>Operculum bank of the subcentral gyrus</td>
<td>4.82</td>
<td>0.041</td>
<td>12mm</td>
<td>Figure 7:11</td>
</tr>
<tr>
<td>±16 -4 -20</td>
<td>Amygdala</td>
<td>3.89*</td>
<td>0.018*</td>
<td>12mm</td>
<td>Figure 7:8</td>
</tr>
</tbody>
</table>

The Low group comparison also revealed increases in grey matter density in the orbitofrontal cortex (see Figure 7:13 and Table 7:11).

Low: Increases in Grey Matter

Figure 7:13 ±26 32 -14 Posterior orbitofrontal cortex (12mm)
Abnormalities were also detected in the Low group in the white matter density comparisons. Increased white matter density was found in the corona radiata at the level of the caudal putamen and in the superior precentral gyrus. No decreases in white matter density were found in the Low group compared to controls (see Figure 7:14 and Figure 7:15 below and Table 7:12).

**Low: Increases in White Matter**

**Figure 7:14** -24 –32 27 Corona radiata at the level of the caudal putamen (4mm)
7.3.3.2.1 Results from the Comparison of the Low Group with Controls accounting for verbal IQ

In order to attempt to account for the significant difference in VIQ between the control group and the Low group, the conjunction analysis was repeated using VIQ as a linear and quadratic covariate of no interest.

**Grey Matter** Accounting for VIQ did not alter the decreases found in the grey matter density in the fusiform gyrus and the amygdala (12mm). However, no other grey matter differences found in the initial analysis were replicated in this analysis. Additionally,
significant decreases in grey matter differences were found in the OFC and the superior cerebellum (4mm) and the superior temporal gyrus (12mm).

**White Matter** Accounting for VIQ did not alter the increases found in white matter density in the corona radiata (4mm). However, the increases in the superior precentral gyrus were not replicated in this analysis.

### 7.3.4 High And Low Group Comparisons

The conjunction analysis between the High and Low groups detected common areas of increased grey matter density in the posterior orbitofrontal cortex, the rostral superior temporal gyrus, the cerebellum and the claustral insula (see Figure 7:16 to Figure 7:20 below and Table 7:13).

**High versus Low Similarities: Increased Grey Matter Density**

**Figure 7:16** Posterior orbitofrontal cortex (±27 22 −16) (12mm)
Figure 7:17  Rostral superior temporal gyrus (±45 10 -26) (12mm)

Figure 7:18  Cerebellum (Inferior including VIIIB) (± 21 -48 -51) (4mm)

Figure 7:19  Cerebellum (Inferior, including VIIB, VIIIA, VIIIB) (±30 -54 -51) (12mm)
Figure 7:20  Claustral Insula ±36 0 –9 (4mm)

Table 7:13  Increases in Grey Matter Density in High and Low Groups

<table>
<thead>
<tr>
<th>Coordinates</th>
<th>Location</th>
<th>Z score</th>
<th>Corrected p value</th>
<th>Smoothing</th>
<th>Figure Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>±21 –48 –51</td>
<td>Cerebellum (Inferior, VIIIIB)</td>
<td>6.03</td>
<td>0.001</td>
<td>4mm</td>
<td>Figure 7:18</td>
</tr>
<tr>
<td>±36 0 –9</td>
<td>Claustral Insula (Inferior, VIIIIB)</td>
<td>5.44</td>
<td>0.017</td>
<td>4mm</td>
<td>Figure 7:20</td>
</tr>
<tr>
<td>±27 22 –16</td>
<td>Posterior OFC</td>
<td>5.96</td>
<td>&lt;0.001</td>
<td>12mm</td>
<td>Figure 7:16</td>
</tr>
<tr>
<td>±30 –54 –51</td>
<td>Cerebellum (Inferior, VIIB, VIIIIA/B)</td>
<td>5.47</td>
<td>0.003</td>
<td>12mm</td>
<td>Figure 7:19</td>
</tr>
<tr>
<td>±45 10 –26</td>
<td>Rostral superior temporal gyrus</td>
<td>5.01</td>
<td>0.025</td>
<td>12mm</td>
<td>Figure 7:17</td>
</tr>
</tbody>
</table>

Significant differences in the pattern of grey matter abnormality were detected between the High and Low groups. There were significant interactions in the hippocampal formation, the amygdala, the rostral superior temporal gyrus, the cerebellum and the postcentral gyrus (see Figure 7:21 to Figure 7:25 and Table 7:14).
Interaction of High and Low groups

Figure 7:21  Hippocampal Formation (±24 -15 -18) (4mm)

Figure 7:22  Amygdala (±18 -6 -21) (12mm) (SVC)

Figure 7:23  Rostral Superior Temporal Gyrus (±45 24 -24) (12mm)
Figure 7:24  Cerebellum (Superior including VI) (±28 –63 –24) (4mm)

Figure 7:25  Postcentral gyrus (somatosensory motor cortex) (±40 –30 66) (12mm)
Table 7:14 Grey Matter Density Differences between High and Low Groups

<table>
<thead>
<tr>
<th>Coordinates</th>
<th>Location</th>
<th>Z score</th>
<th>Corrected p value</th>
<th>Smoothing</th>
<th>Figure Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>±28 -63 -24</td>
<td>Cerebellum (Superior, VI)</td>
<td>6.00</td>
<td>0.000</td>
<td>4mm</td>
<td>Figure 7:24</td>
</tr>
<tr>
<td>±24 -15 -18</td>
<td>Hippocampal Formation</td>
<td>5.38</td>
<td>0.013</td>
<td>4mm</td>
<td>Figure 7:21</td>
</tr>
<tr>
<td>±20 -9 -21</td>
<td>Amygdala</td>
<td>4.56*</td>
<td>0.018*</td>
<td>4mm</td>
<td>N/A</td>
</tr>
<tr>
<td>±45 24 -24</td>
<td>Rostral superior temporal gyrus</td>
<td>5.79</td>
<td>&lt;0.000</td>
<td>12mm</td>
<td>Figure 7:23</td>
</tr>
<tr>
<td>±40 -30 66</td>
<td>Postcentral gyrus</td>
<td>4.83</td>
<td>&lt;0.000</td>
<td>12mm</td>
<td>Figure 7:25</td>
</tr>
<tr>
<td>±18 -6 -21</td>
<td>Amygdala</td>
<td>4.42*</td>
<td>0.002*</td>
<td>12mm</td>
<td>Figure 7:22</td>
</tr>
</tbody>
</table>

7.3.4.1.1 Results from High and Low comparison accounting for verbal IQ

Accounting for VIQ, the significant similarities between the High and Low groups were increased grey matter density in the inferior cerebellum (4mm, 12mm), claustral insula (4mm), OFC (12mm) and the superior temporal gyrus (12mm).

Accounting for VIQ, the significant differences between the High and Low groups were in the superior cerebellum, hippocampal formation and amygdala (4mm), the superior temporal gyrus and amygdala (12mm)

Interpretation of the significant interactions found between the High and Low group requires inspection of the Z scores at these locations in the appropriate comparisons (see Table 7:15). These scores show that in the cerebellum, hippocampal formation and superior temporal gyrus there were decreases (some merely trends, others reaching corrected significance) in grey matter density in the Low group and increases (some trends, others reaching significance) in grey matter density in the High group, resulting in the
significant interaction term. The interaction in the postcentral gyrus results from a highly
significant abnormality in the Low group, compared to little or no trend of abnormality in
the High group in this region.

Table 7:15  Z scores in High and Low comparisons at coordinates of significant
interaction in conjunction analyses

<table>
<thead>
<tr>
<th>Location</th>
<th>Coordinates</th>
<th>Smoothing</th>
<th>Z Score High (Inc in GM)</th>
<th>Z Score Low (Dec in GM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum</td>
<td>±28 -63 -24</td>
<td>4 mm</td>
<td>3.24</td>
<td>1.8</td>
</tr>
<tr>
<td>Hippocampal Formation</td>
<td>±24 -15 -18</td>
<td>4 mm</td>
<td>2.37</td>
<td>2.18</td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td>±40 -30 66</td>
<td>12 mm</td>
<td>&lt;1.3</td>
<td>3.38</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>±45 24 -24</td>
<td>12 mm</td>
<td>2.79</td>
<td>3.14</td>
</tr>
</tbody>
</table>
Table 7:16 Summary of Grey Matter Density Results (↓ significant decreases compared to controls; ↑ significant increases compared to controls; # significant interaction; blank = no significant difference)

<table>
<thead>
<tr>
<th>Region</th>
<th>High</th>
<th>Low</th>
<th>High and Low</th>
<th>Low w VIQ</th>
<th>High and Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum (Inferior)</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td></td>
<td>↑</td>
</tr>
<tr>
<td>Cerebellum (Superior)</td>
<td>↑</td>
<td></td>
<td>#</td>
<td>↓</td>
<td>#</td>
</tr>
<tr>
<td>Hippocampal</td>
<td>↑</td>
<td></td>
<td>#</td>
<td></td>
<td>#</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>↑</td>
<td>↑#</td>
<td>↓</td>
<td>↑#</td>
<td></td>
</tr>
<tr>
<td>Posterior OFC</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>Amygdala</td>
<td>↑</td>
<td>↓</td>
<td>#</td>
<td>↓</td>
<td>#</td>
</tr>
<tr>
<td>Fusiform gyrus</td>
<td>↓</td>
<td></td>
<td></td>
<td>↓</td>
<td></td>
</tr>
<tr>
<td>ParaHCal region</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior precentral gyrus</td>
<td>↓</td>
<td></td>
<td>#</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opercular bank of subcentral gyrus</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td></td>
<td></td>
<td>#</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claustral Insula</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td>↑</td>
</tr>
</tbody>
</table>
7.3.5 Discussion

7.3.5.1 Results From The High Group

The results from the VBM analysis of the High group indicate bilateral abnormality in the amygdala, hippocampal formation, the orbitofrontal cortex, the superior temporal gyrus and the cerebellum. All these areas were found to have increased grey matter density. More typically in developmental disorders, decreases in grey matter density (assumed to reflect atrophy) are found. However this study suggests, perhaps counter-intuitively, that increases in grey matter density can be associated with Autism. Increased grey matter density may reflect a lack of appropriate levels of programmed cell death or a failure of neurons to migrate appropriately.

There is only one other report of VBM in Autism in the literature: a study of 15 high functioning adults with Autism [Abell et al. 1999]. This study found abnormalities in the cerebellum (bilaterally) and the left amygdala, as well as the middle and inferior temporal gyri, left inferior frontal gyrus, left occipitotemporal junction and the paracingulate sulcus. Although these findings seem to be in conflict with the present results from the High group, a number of methodological differences must be noted. Abell and colleagues did not use the bilateral method (which was not available at the time), thereby reducing the sensitivity of the analysis to bilateral abnormalities. They also did not use non-linear spatial transformations (which is known to improve the accuracy of the colocalisation of structures (see Appendix D)). Further, statistical inferences were made at an uncorrected threshold, increasing the likelihood of Type I errors.

Medial temporal lobes

The findings of bilateral increases in grey matter density in the amygdala and hippocampal formation are consistent with reports of abnormalities in these areas using other techniques. Post mortem studies have found bilateral increases in cell density in the amygdala and the hippocampal formation (e.g. Bauman and Kemper 1994) and abnormal volumes of these
structures have been reported (e.g. Aylward et al. 1999; Howard et al. 2000). It should be emphasised, however, that grey matter density is not a direct measure of cell density. Grey matter density is an index of signal intensity on MR scans after normalisation, and signal intensity is affected by many factors, including neuronal cell density. Grey matter density should be considered a metric that is sensitised to the volume of structures relative to the surrounding structures.

Orbitofrontal Cortex

Abnormality in the orbitofrontal cortex is consistent with behavioural patterns of children with Autism. For example, difficulties in set-shifting are characteristic of individuals with Autism and individuals with orbitofrontal damage (see Chapter 5). Happe et al. found an area of the medial frontal cortex that was functionally inactive in individuals with Autism participating in a PET study compared to controls [Happe et al. 1996]. The location of the abnormality found in the High group of the present study was more lateral and inferior than the area reported in this study. This apparent contradiction between the two data sets, raises a number of possible explanations. Perhaps there are at least two distinct, independent areas of abnormality in the frontal lobes, or maybe the structural abnormality is interfering with the functioning of cortex remote from the implicated site (such as that found to be inactive in Happe et al.’s study).

A number of other studies have found abnormalities in the frontal lobes using techniques with a much larger spatial scale. Abnormalities at these larger spatial scales (such as those reported by Minshew et al. 1993; Carper and Courchesne 2000) are not inconsistent with the current findings.

There have been a few reports of abnormalities of the rostral superior temporal gyrus in individuals with Autism. Structural abnormalities in the superior temporal gyrus were noted in a case study of twins [Kates et al. 1998] and Ohnishi and colleagues reported bilaterally abnormal rCBF in the superior temporal gyrus [Ohnishi et al. 2000]. Abnormality in the rostral superior temporal gyrus in Autism is consistent with studies suggesting the involvement of this region in social processing. Damage to the superior temporal gyrus impairs face perception [Campbell et al. 1990]. Single cell recordings
demonstrate the involvement of the superior temporal gyrus in detection of gaze [Perrett et al. 1985], whilst making mentalistic inferences from the eyes activates this area in functional imaging studies in normal subjects [Baron-Cohen et al. 1999]. However, it should be noted that normal activation in the superior temporal gyrus was found in individuals with Autism participating in the latter study.

A further role that abnormality in the superior temporal gyrus might play in the syndrome of Autism is suggested by a series of functional imaging studies investigating language processing. These have suggested a role for the superior temporal gyrus (bilaterally) in the processing of signals with dynamic pitch variation (such as speech and music) [Scott et al. 2000]. The left superior temporal gyrus seems to have a specialised role in the prelexical processing of phonetic cues and features and in their sequence processing. Perhaps abnormality in the superior temporal gyrus in individuals with Autism may result in the flat intonation and abnormal prosody that many of the affected individuals show.

Cerebellum
The findings of abnormalities in the cerebellum in the High group are consistent with many other studies. There is, however, little agreement in the literature as to which areas of the cerebellum are abnormal and the nature of these abnormalities. Some reports suggest selective abnormality of vermal lobules VI and VII (e.g. Courchesne et al. 1988) whilst others indicate diffuse abnormality throughout the cerebellum (e.g. Arin et al. 1991). In this study, abnormalities were found in lobules IV – VIII. Factors which may account for these inconsistent findings include incidence of epilepsy (Purkinje cell loss is a complication of epilepsy [Honavar and Meldrum 1997]), different subject inclusion criteria and different levels of verbal skills, etc. (see below). Courchesne has argued that there are two subgroups of individuals with Autism, some showing hypoplasia of the cerebellum and others hyperplasia of the cerebellum [Courchesne et al. 1994a]. Whilst it is not clear whether these subgroups can be distinguished behaviourally, the variation in the cerebellar pathology may suggest differing aetiology.

The relationship between the cerebellar abnormalities and the cognitive and behavioural profile of Autism is not straightforward. Children with cerebellar dysfunction (or agenesis)
occurring early in life often present with developmental delay, poor motor co-ordination and mental retardation, but intact social skills [Rubinstein and Freeman 1940; Jervis 1950; Macchi and Bentivoglio 1977]. The cause and specificity of the cerebellar abnormalities in these children is varied. There is, however, some limited evidence of cerebellar involvement in social cognition. Following lesion resection, some children with posterior fossa tumours have been reported to show Autistic-like symptoms [Riva and Giorgi 2000]. However, these symptoms appear to ameliorate with time (at least in some cases), suggesting that the trauma of surgery and/or treatment may also play a role in the social impairments.

There is ample evidence that the cerebellum is involved in motor co-ordination (for reviews see Houk et al. 1996; Thach 1996). Abnormal development in the cerebellum leads to poor motor co-ordination. For example, in mice, interruption of synapse elimination in the cerebellum leads to mild but persistent loss of motor co-ordination [Kakizawa et al. 2000]. The abnormalities in the cerebellum found in children with Autism may therefore play a role in the increased incidence of clumsiness in these children (see Chapter 2).

The individual VBM analyses found that most children in the High group showed evidence of abnormality in the orbitofrontal cortex, superior temporal gyrus and the cerebellum. Fewer individuals showed evidence of medial temporal lobe abnormality. This might suggest that abnormality in the medial temporal lobe plays a less important role in the neuropathology of Autism than the other areas of abnormality. However, it is important to emphasise that the analyses do not demonstrate normality in these regions (see methods).

The correlation between the scores on the Autism Behaviour Checklist and grey matter density in the amygdala highlights the importance of the amygdala in the neuropathology of Autism. The absence of correlations with other cognitive indices should not be interpreted as evidence against links between the detected neuropathology and the cognitive and behavioural impairments. A number of factors could account for the lack of significant correlation. There may have been insufficient variation in scores and/or grey matter density to identify a linear relation between the two. Further, the VBM correlation analyses only examine linear relations: yet there is no a priori reason to suppose such a direct relationship
between the two variables. There have been no reports of significant correlations between grey matter density and any cognitive performance indices in the literature. Perhaps grey matter density is too indirect a measure of pathology (compared to, say, cell density, etc.) for correlations with cognitive performance to be detectable.

No differences were found in the cognitive and behavioural patterns in the High group according to detection of significant VBM abnormality in individual cases. This is consistent with the caveat above, regarding the lack of power of individual VBM analyses. Absence of detected abnormality cannot be taken as an indication of normality.

7.3.5.2 Results From Low Group

The results from the Low group indicate more widespread abnormality throughout the brain (in both white and grey matter) compared to the High group. Abnormalities were found in the fusiform gyrus, subcentral gyrus, precentral gyrus, amygdala and orbitofrontal cortex.

Interpretation of these findings is confounded by the significant VIQ difference between the Low group and their controls. Thus differences may be attributable to either Autism or to learning disability. Although it is tempting to speculate on the significance of abnormalities in these areas (such as the fusiform gyrus which is known to be involved in face processing (e.g. Rossion et al. 2000)), such analysis is premature.

To aid interpretation of the present results, VIQ was included as a covariate in a second analysis, in an attempt to account for VIQ differences. The results, however, should still be interpreted with caution as there is no reason to believe that the relationship between VIQ and grey (or white) matter is completely described by linear or quadratic terms. Indeed, the relationship is likely to be highly complex and interdependent.

When VIQ was included as a covariate, the abnormalities in the subcentral and precentral gyri were no longer found. Although the crude nature of the correction demands cautious interpretation, this is consistent with these abnormalities being associated with low VIQ
and not Autism per se. In contrast, the abnormalities in the amygdala, orbitofrontal cortex and fusiform gyrus survive the crude correction, consistent with the possibility that these areas are associated more with Autism than with verbal IQ deficits. It should be noted that, although it appears that correcting for VIQ resulted in a direction change in abnormality in the orbitofrontal cortex (from increase to decrease), the location and spatial scale of the abnormality within the orbitofrontal cortex also changed substantially (from -26 32 -14 (12mm) to -12 42 -24 (4mm)).

7.3.5.3 Results From The Comparison Of The High Group Versus The Low Group

Although interpretation of the Low group results must necessarily be cautious (see above), comparisons between the High and Low groups highlight a number of important points.

Firstly, the different loci of abnormality in the High and Low groups when verbal IQ is not accounted for, and the reduction of these differences when verbal IQ is accounted for, emphasise the importance of matching the control group on verbal IQ (see Table 7:16). Perhaps some of the difficulty replicating results in groups of children with Autism may be attributable to the lack of appropriate controls. However the difficulties of finding an appropriate control group for children with below average IQ should be acknowledged. Do some ‘normal’ children have below average IQ, and if so, what is the relationship between low IQ and brain structure? Although the neural underpinnings of IQ are not understood, it seems likely that learning disabilities are associated with neural abnormalities. Until such differences are well characterised, interpretation of studies using normal children as controls is complex.

Although accounting for verbal IQ rendered the patterns of abnormality in the two groups more similar, significant differences (identified in the interaction analysis) were found in the hippocampal formation, the amygdala, superior temporal gyrus and the cerebellum. As revealed in Table 7:15, the interaction in all these areas was due to increases in grey matter density in the High group and decreases in grey matter density in the Low group. Although
these differences may be a results of inadequate correction for verbal IQ, there have been reports of both increased and decreased volumes of the amygdala [Aylward et al. 1999; Howard et al. 2000] and the cerebellum (e.g. Courchesne et al. 1994a). Post mortem studies have found variation in the pattern of neuropathology with the individual’s IQ [Bauman 1996]. These differences may therefore reflect differing pathological aetiologies in the two groups or variation in the consequences of the same aetiology. The current data do not permit resolution of the cause of the difference.

The conjunction of the Low and High groups reveals additional abnormality in the claustral insula in both groups. Whilst structural abnormality of the insula has not been reported in Autism, functional abnormalities have been reported [Ohnishi et al. 2000]. There are many connections between the insula and the paralimbic systems [Mufson and Mesulam 1981; Mufson et al. 1981; Mesulam and Mufson 1982; Morecraft et al. 1982; Augustine 1985]. Abnormality in the insula may therefore be a ‘knock-on’ effect of abnormalities in the medial temporal lobes.

### 7.4 Volumetrics

#### 7.4.1 Introduction

One method of analysis of 3D MR data sets is VBM. Another more traditional method is to measure the volumes of discrete structures by tracing the boundaries by hand. This technique is used here.

MRI permits accurate representation of in vivo brain anatomy. Precise knowledge of morphological changes in cerebral structures is important for diagnosis of neurological disorders and for understanding of the underlying pathophysiology [Bilir et al. 1998]. Volumetric changes in brain structures have been reported in a number of neurological and psychiatric disorders such as temporal lobe epilepsy [Cendes et al. 1993], schizophrenia (e.g. Zipursky et al. 1994) and bipolar disorder [Altshuler et al. 2000], as well as
developmental disorders such as orofacial dyspraxia [Vargha-Khadem et al. 1998], developmental amnesia [Vargha-Khadem et al. 1997a] and paediatric generalised anxiety disorder [De Bellis et al. 2000]. Although atrophy is most commonly reported in patient populations, abnormally large structures have also been associated with psychiatric disorders (e.g. Altshuler et al. 2000; Tebartz Van Elst et al. 2000).

Recent investigations have demonstrated the reliability and sensitivity of volume measurements of the hippocampal formation and amygdala (e.g. Watson et al. 1992; Bartzokis et al. 1998; Honeycutt et al. 1998; Convit et al. 1999). Volumetric measures have also been shown to correlate with neuropsychological measures (such as memory) (e.g. Baxendale et al. 1998). As discussed in Section 7.1, a number of studies have reported volumetric abnormalities in the hippocampal formation and the amygdala of individuals with Autism (e.g. Saitoh et al. 1995; Aylward et al. 1999; Howard et al. 2000; Pierce et al. 2001). Volumetric measures of these structures were also undertaken in this study.

7.4.2 Methods

7.4.2.1 MRI Acquisition

The same 3D FLASH data sets as described in Section 7.3 were used for volumetric assessment.

7.4.2.2 Data Analysis

The data sets were reformatted into 1mm thick contiguous slices in a tilted coronal plane that was perpendicular to the long axis of the hippocampal formation. Tracings were carried out in MEDx (Version 3.30, Sensor Systems) and the volumes were calculated by summing the hippocampal cross sectional areas and multiplying by the distance between the slices (Cavalieri's principle [Gundersen and Jensen 1987; Cook et al. 1992]).
Hippocampal boundaries were defined as described previously [Watson et al. 1992], with the modification that the posterior boundary was extended to include the full length of the hippocampal formation (rather than stopping when the fornix is seen in full profile).

Following the demarcation of the hippocampal formation boundary, the boundaries of the amygdala were delineated in accordance with the method described in Watson et al. [1992]. It should be noted that this method undoubtedly excludes small amounts of the medial and central nuclei (at the superior border) and includes small amounts of the entorhinal cortex. Although this is undesirable, it is generally felt acceptable in order to maintain reproducibility of the boundaries.

Intracranial volume (ICV) was measured on the sagittal unreformatted FLASH dataset measuring every tenth slice, as described previously [Van Paesschen et al. 1997].

Statistical analysis was carried out using a repeated measure ANOVA (within subject factor of Side (2 levels) and between subject factor of group) with ICV as a covariate.

The mean, mean difference and reproducibility coefficient (RC) between the first and second measurements were calculated. The RC is 2 standard deviations of the mean difference between two sampling strategies and has been suggested to be the best method to assess agreement between two measurements [Bland and Altman 1986].

7.4.3 Results

7.4.3.1 Reliability Studies

The results are shown in Table 7:17. The mean difference between measurements was small for both the amygdala and the hippocampal formation and within 5% of the mean amygdala volume and 4% of the mean hippocampal volume. The RC for the right amygdala volume was 150 mm$^3$. This means that 95% of the re-measurements would be
within 150 mm$^3$ of the first measurement. The remaining RC values can be similarly interpreted.

Table 7:17 Percentage mean difference and RC for first and second measurement of Amygdala and Hippocampal formation volumes. % Mean difference is the mean difference between the two measurements expressed as a percentage of the mean volume of the structure. RC is 2 standard deviations of the mean difference between two sampling strategies (mm$^3$) (see Bland and Altman 1986)

<table>
<thead>
<tr>
<th>Structure</th>
<th>% Mean difference (Left)</th>
<th>% Mean difference (Right)</th>
<th>RC (Left) (mm$^3$)</th>
<th>RC (Right) (mm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amygdala</td>
<td>4.4</td>
<td>5.0</td>
<td>162</td>
<td>150</td>
</tr>
<tr>
<td>Hippocampal Formation</td>
<td>3.7</td>
<td>3.5</td>
<td>252</td>
<td>167</td>
</tr>
</tbody>
</table>

7.4.3.2 Volumetric Results

The results are shown in Figure 7:26, Figure 7:27 and Figure 7:28. The volumes of the hippocampal formation in the two Autistic groups were larger than in the control group bilaterally. The volumes of the amygdala in the two Autistic groups showed slight variations compared to controls, with the High group mean being slightly larger bilaterally, whilst the Low group mean was slightly lower. Both Autistic groups had larger intracranial volumes (ICVs) compared to controls.

Statistical analysis (including ICV as a covariate) revealed a significant difference between the mean hippocampal volumes of the groups ($F(2, 34) = 6, p = 0.005$), and no interaction of group and side. There was a significant main effect of side ($F(1, 34) = 6, p = 0.02$) which reflected an increased volume in the right hippocampal formation compared to the left. Planned comparisons revealed that the group effect was due to the Low group having significantly larger hippocampi than the High group ($CvH \ p = 1.0; \ HvL \ p = 0.004$).
Inclusion of sex, age or verbal IQ as covariates in the analysis did not account for the significant effect. However analysis of only the boys did reduce the significance of the group effect to a trend ($p = 0.07$).

Statistical analysis revealed no significant difference between the mean amygdala volumes of the groups ($F(2, 34) = 0.6$, $p = 0.6$), and no main effect of side or interaction of group and side. Inclusion of sex as a covariate in the analysis or analysis of only the boys did not alter these results.

**Figure 7:26 Volumes of the Hippocampal Formation (Mean ± SEM)**

![Graph showing volumes of the hippocampal formation with mean ± SEM for control, high, and low groups.](image)

**Figure 7:27 Volumes of the Amygdala (Mean ± SEM)**

![Graph showing volumes of the amygdala with mean ± SEM for control, high, and low groups.](image)
Intragroup correlations were carried out to investigate the relationship between the episodic memory impairment noted in Chapter 3 and the volume of the hippocampal formation. Partial correlations were used, with ICV controlled for. Only the left hippocampal volume in the High group was significantly correlated with the index of episodic memory (coeff = -0.7; p = 0.006) (Rivermead Behavioural Memory Test).

7.4.4 Discussion

These results show that the hippocampal volume of the children in the Low group was significantly different from that of the controls. In comparison, no significant difference between the Autistic groups and controls was observed in the amygdala volumes. The size of brain structures has been found to be determined by the number and size of neurons and glial cells, packing density, vascularity and matrix composition of the structure [Giedd et al. 1996a]. The results are therefore consistent with abnormalities in the cellular pattern of the hippocampal formation in the children in the Low group.

The lack of significant difference in the volumes of the amygdala in the Autistic groups should not be taken to imply that the amygdala were normal. The method of measuring the volume had a 5% accuracy, and therefore differences in volume of the order of this
accuracy would not have been detectable. Further, the dependence of the amygdala volume measures on the gyral patterns of the individual (which is chosen to aid reproducibility, see Watson et al. 1992) is likely to mean that the amount of amygdala actually included in the measure varies between individuals. It is therefore possible that with a more sensitive method differences would have been observed.

In this section, intracranial volume was used as a covariate to correct for total brain volume. Although the relative merits of using absolute size or correcting for brain volume are a source of considerable debate [Arndt et al. 1991], both approaches have potential utility in elucidating structure-function relationships in the brain. The method of correction used here (covariation) was chosen to avoid the reduction in reliability inherent in other methods (such as forming ratios) (see Arndt et al. 1991 for further discussion of this issue).

Although there is no evidence to suggest that verbal IQ is associated with hippocampal volume, it should be noted that the increase in hippocampal volume in the Low group may at least in part be due to the discrepancy in verbal IQ between the groups. Interpretation of the significantly larger hippocampal volumes in the Low group is also complicated by the lack of sex matching. Sex differences in intracranial volume, hippocampal formation volume and amygdala volume have been noted. Boys have larger intracranial volumes [Giedd et al. 1996a]. During childhood, selective increases in volume of the amygdala have only been observed in boys, whilst girls show selective increases in volume of the hippocampal formation [Giedd et al. 1996b]. Although inclusion of sex as a covariate did not alter the significance of the results, analysis of only the boys reduced the effect to a trend. The importance of this latter finding is difficult to establish, as the power of the analysis is reduced by the reduction in the number of children. Thus whilst there is evidence to suggest that the change in volume of the hippocampal formation is not related to sex, further studies are required to address this issue more fully.

Previous studies investigating the volume of the hippocampal formation in individuals with Autism have either reported decreases in volumes or no significant differences (e.g. Saitoh et al. 1995; Piven et al. 1998; Aylward et al. 1999; Howard et al. 2000; Saitoh et al. 2001). Whilst some of these studies have used more crude measurements of volume (such
as Saitoh et al. 1995), other studies have used very similar methods (e.g. Piven et al. 1998; Aylward et al. 1999). Factors that might account for the differing findings in this study include sex matching (see above), differing age ranges of subjects and differing levels of intellectual ability. This latter suggestion may be particularly relevant in the light of the non-significant difference between the High group and the controls in this study. Many previous studies have investigated individuals with average range verbal IQ, making their results more comparable to the High group. An additional factor is that many of these studies have not carried out volumetric analysis blind (e.g. Howard et al. 2000), stating that it was possible to distinguish between the controls scans and scans of individuals with Autism with the naked eye. This may be due to movement artefacts being more prominent in the scans of individuals with Autism in their study.

Several studies of Autism have investigated the volume of the amygdala. Their findings were inconsistent. Howard and colleagues reported that the individuals with Autism had increased amygdala volume, whilst others reported significant decreases in volume [Aylward et al. 1999; Howard et al. 2000; Pierce et al. 2001]. The finding in this study therefore seems to fall between the two, and it is interesting to note that whilst the High group show a trend for increased volume, the Low group show a trend for decreased volume. Further studies are required to establish the presence and nature of any abnormality in amygdala volume in individuals with Autism.

7.5 **T2 Maps**

7.5.1 **Introduction**

For this final MR study, another quantitative measure was used to assess the tissue integrity of the amygdala and hippocampal formation. T2 relaxometry provides a quantitative way of detecting abnormalities that are more conventionally evaluated by visual inspection of T2-weighted images. T2 relaxation times have been shown to detect lesions or abnormalities that may not be identified in standard clinical imaging (e.g. Van Paesschen et al. 1996).
7.5.2 Methods

7.5.2.1 MRI Acquisition

Hippocampal T2 (HCT2) and Amygdala T2 (AT2) maps were obtained using a 16 echo sequence as previously described (HC Van Paesschen et al. 1997; Amygdala Van Paesschen et al. 1996) (see Table 7:1). The HCT2 map was oriented in a tilted coronal plane along the anterior border of the brainstem, perpendicular to and at the level of the body of the hippocampal formation. The AT2 was oriented in a tilted axial plane parallel to and above the long axis of the hippocampal formation.

7.5.2.2 Data Analysis

Both HCT2 and AT2 were measured by placing the largest possible circle as a region of interest (ROI) within the hippocampal formation and amygdala (respectively) while avoiding boundaries where partial volume effects within CSF might occur. HCT2 and AT2 values are expressed in milliseconds (ms).

7.5.3 Results

There were no significant differences between the AT2 and HCT2 values of any of the three groups (AT2 (Left): F(2, 25) = 1, p = 0.4; AT2 (Right): F(2, 26) = 2, p = 0.2; HCT2 (Left): F(2, 30) = 0.2, p = 0.8; HCT2 (Right): F(2, 30) = 2, p = 0.2). Figure 7:29 and Figure 7:30 show the range of results obtained.
7.5.4 Discussion

These results demonstrate no group difference in the $T_2$ maps in the amygdala and hippocampal formation. This suggests that the tissue present in these structures is within normal limits as characterised by the $T_2$ relaxation time. It should be emphasised that this does not imply that these structures are normal. $T_2$ maps have been shown to be sensitive
to some types of neuropathology, but may not detect all abnormalities. For example, cases of hippocampal abnormality and normal HCT2 have been reported in the literature [Van Paesschen et al. 1997; Gadian et al. 2000]. It should also be noted that a number of children (especially from the Low group) are not represented in these data.

7.6 Discussion

VBM analysis indicates areas of bilateral abnormality in the High group in the amygdala, hippocampal formation, cerebellum, rostral superior temporal gyrus and the orbitofrontal cortex. These areas are highly interconnected. In particular the rostral superior temporal gyrus, orbitofrontal cortex and cerebellum are all connected with the medial temporal lobes [Barbas and De Olmos 1990; Barbas and Blatt 1995]. Abnormalities were also found in the Low group using VBM. These abnormalities were more widespread and may at least in part be related to the discrepancy in verbal IQ between the Low group and the controls.

Volumetric analysis failed to detect significant volume changes in the amygdala or the hippocampal formation in the High group. However, this is not necessarily incompatible with the VBM results. Firstly, the error inherent in the volumetric analyses is likely to prevent detection of small changes in volume. In the light of this possibility it is interesting to note that the trends seen in the volumes of these structures in the High group versus the controls are consistent with the increased grey matter density observed in the VBM results. Unlike VBM, volumetric analyses rely on identification of structural boundaries by the experimenter. Particularly in the case of the amygdala, the definition of the boundaries is according to gyral patterns, which may vary independently of the amygdala volume. Volumetric analyses also reflect the global volume of a structure. Thus, volume changes of a particular region of a structure, such as a nucleus of the amygdala, or the anterior portion of the hippocampal formation, may only be detected by VBM (see Maguire et al. 2000a). New methods such as tensor based morphometry and deformation field morphometry may be better suited to providing corroborative evidence to support the VBM results (see Thompson and Toga 1999).
Neuropathology of Autism

For the Low group, no difference in amygdala volume was found, although the trend is once again consistent with the decrease in grey matter density found in the VBM results. The results of the volumes of the hippocampal formation are more difficult to reconcile with the VBM analyses. No significant abnormality in the hippocampal formation is detected in the Low VBM analysis, and, when compared to the High group in the VBM conjunction analysis, the Low group was found to have a slight non-significant decrease in grey matter density (see Table 7:15). In contrast, the Low group is found on volumetric analysis to have significantly larger hippocampal volumes than the High group. These results may suggest that whilst there is a focal region of non-significant decrease in volume in the hippocampal formation, there is an increase in the overall volume. This might not be detected by VBM if it occurs along the length of the body of the hippocampal formation, so that no cluster of voxels reach significance. It should also be noted that some of the control children included in the VBM analysis of the Low groups were not included in the volumetric analysis. The control group used in the Low group VBM analysis were matched more closely on sex: this may potentially explain the seemingly discrepant results.

As argued above, the normal T2 maps found in both Autistic groups are not incompatible with findings of abnormality in these regions using other techniques. In particular, the VBM results indicating abnormality were based on T1-weighted scans. T2 relaxation times and signal intensity on T1 scans can reflect different processes [Gadian 1995].

From the results of this chapter, it is not possible to determine the developmental progression of abnormality. Perhaps the pattern of abnormality reflects only one area primarily being abnormal and the abnormal outputs from this area affecting development in other regions. Alternatively, an environmental or genetic process may directly affect development of all identified areas of abnormality, with aberrant interconnections compounding the disturbance. Addressing this issue requires neuropathological studies of the brains of individuals with Autism at early stages in development. At present diagnosis rarely occurs before the age of 3 and so such studies are currently impossible.

These findings relate to children aged between 8 – 18 years. During this time many areas of the brain are still developing [Huttenlocher 1979; Huttenlocher et al. 1982; Easter et al.]

206
This raises the question of how applicable these findings are to adults with Autism. It is possible that some of the abnormalities detected may be a reflection of lagging development. For example, it might be hypothesised that increased grey matter density is associated with inefficient cell pruning. Differences due to such delayed development may reduce with age and eventually resolve, with the adult brain of individuals with Autism "catching up". Consistent with this, Saitoh et al. recently reported differences in the volume of the dentate gyrus and CA4 being most significant in the youngest age range they examined (2-4 years) [Saitoh et al. 2001]. Only by investigating the neuropathological profile of adults with Autism can this issue be directly addressed.
Part III  Event Related Potentials

Thus far, this thesis has documented the brain structure and behaviour of individuals with high functioning Autism. In this part, the brain function of these children is investigated using event related potentials (ERPs).

ERPs consist of a series of transient perturbations of the spontaneous electroencephalogram that are time locked to some externally definable event such as the onset of a stimulus. They provide a partial record of the neural events underlying stimulus processing. Although ERPs are usually swamped by the ongoing electroencephalic activity, signal averaging allows extraction of an average estimate of the ERP.

ERPs are recorded from an electrode montage placed on the scalp (see Figure III:1). They are characterised in terms of their polarity (positive or negative), their latency (in order of ms), their magnitude (or amplitude) (typically of the order of \( \mu V \)) and their topographic distribution. A sample ERP is shown in Figure III:2. Many of these characteristics may show developmental trends, such as decreasing latencies, changing topography or altered magnitude (see Kurtzberg et al. 1984).
Figure III:1  A typical electrode montage. F, C and P refer to frontal, central and parietal electrodes, whilst z, 3, and 4 refer to the midline, left and right electrodes, respectively, placed according to 10-20 system.
Figure III:2  A typical ERP component (the P300: a positive peak at 300 ms). Time is plotted in ms along the x axis (in 0.2 sec intervals). Stimulus onset occurs at t = 0.0 ms. Voltage (in μV) is plotted on the y axis (negative is up). The scale of the y axis is indicated by the bar. The amplitude (A) is inferred from the deviation from the baseline, the latency (L) is defined as the time at which the maximal amplitude is obtained. The topography of the response can be inferred by comparing the responses at different electrodes.

ERPs have exceptional temporal resolution compared to many other non-invasive neural measures and so provide information about the time courses of cognitive processes. However, the spatial resolution of ERPs is limited, since the signal measured at the scalp is likely to be a composite of neural activity at a variety of depths and locations in the brain. These multiple generators can be modelled, but such models do not have unique solutions. Anatomical location of the neural generators of the ERP therefore relies on inferences from studies of patients with lesions and intracranial recordings carried out for pre-surgical work up. ERPs are therefore highly complementary to MR techniques which provide exceptional spatial resolution but poor temporal resolution.

Two paradigms are investigated in the following two chapters. Each was selected to address the integrity of ERPs thought to be modulated in the medial temporal lobes. ERPs
are usually generated by multiple neural sources and this precludes any definitive conclusion that abnormalities in an ERP are due to selective abnormality in one of the neural generators. Instead, evidence of abnormality is sought which would be consistent with medial temporal lobe abnormality.
Chapter 8  P300

In this chapter we investigate the functional integrity of two ERP components elicited by the auditory 'oddball' task. This task consists of a string of frequent stimuli combined with rare deviant (target) stimuli and unexpected novel stimuli (e.g. Opitz et al. 1999). The subject is instructed to pay attention to the target stimuli (either counting or indicating when targets are presented).

Two of the ERPs evoked by this paradigm are the novelty P300 and the target P3b. As the names suggest, the first is associated with the novel stimuli whilst the latter component is associated with detection of target stimuli.

In Section 8.1 these two components and their neural generators will be discussed. Following characterisation of the components generated in this paradigm, first in adult controls (Section 8.2) and then in child controls (Section 8.3), Section 8.4 describes the group comparisons between control children and children with Autism.

8.1 Introduction

8.1.1 Novelty P300

A positive wave at a latency of around 300 ms (known as the novelty P300) is elicited by an unpredictable novel change in an ongoing repetitive series of stimuli. This occurs even under circumstances where the stimuli are defined as task irrelevant and ostensibly are not attended [Squires et al. 1975; Opitz et al. 1999]. The novelty P300 amplitude has been shown to decrease with novel event occurrence repetition, regardless of whether a single novel stimulus [Knight 1984], repeated novel stimuli [Kazmerski and Friedman 1995; Cycowicz et al. 1996] or many unique stimuli [Courchesne 1978a; Friedman and Simpson 1994] are used. It has been suggested that it reflects a cerebral component of the orienting response [Ritter et al. 1968; Roth 1973; Roth and Kopell 1973].
Recognisable novel auditory sounds elicit a novelty P300 with a central maximum (e.g. Friedman and Simpson 1994). This is broadly developmentally invariant. Indeed, there is little evidence of significant developmental change in the novelty P300, although slight decreases in latency with age and increasing magnitude during childhood before returning to adult levels have been reported [Cycowicz et al. 1996]. Despite these slight differences, both adult and child novelty P300 are modulated by similar factors (such as a decrease in amplitude with repetition of the novelty stimuli) [Cycowicz et al. 1996].

EEG and MEG scalp recordings have shown that the novelty P300 reflects activity of a widespread neuronal circuit, including frontal, parietal and medial temporal lobe structures [Mecklinger and Ullsperger 1995; Alho et al. 1998]. Intracranial recordings and studies of patients with focal lesions have further highlighted the particular importance of the hippocampal formation, the temporal lobes and the dorsolateral prefrontal cortex [Knight 1984; Baudena et al. 1995; Halgren et al. 1995a; Halgren et al. 1995b; Knight 1996]. For example, Knight showed that patients with hippocampal damage (including damage to parahippocampal regions) had widespread reductions of the novelty P300, with decrements most pronounced over the prefrontal regions [Knight 1996].

This evidence suggests that if the medial temporal lobe region is implicated in the neuropathology of Autism, the novelty P300 may be abnormal in these individuals. There is behavioural evidence to support the possibility of abnormalities in the novelty P300 in individuals with Autism. In particular, they have been noted to have an aberrant orienting response. For example, at times children with Autism appear oblivious to unexpected novel sounds and give no overt indication that they are startled or that they are curious as to the source of the noise [Courchesne 1987].

The novelty P300 has not been extensively investigated in Autism, though Courchesne and colleagues (and others) have reported that unexpected novel stimuli evoke an abnormally small P300 response in individuals with Autism compared to the response to expected non-novel background stimuli [Courchesne et al. 1984; Courchesne et al. 1985; Kemner et al. 1996].
Interpretation of these results should be conservative, however, as the Autistic group was not matched on intelligence to the control group.

8.1.2 Task Relevant P3b

An enhanced positive component with a latency of around 300 ms (P3b) is elicited by a specific 'anticipated' target signal that occurs unpredictably within a series of non-target stimuli and demands a special cognitive or motor response [Squires et al. 1975]. This modulation occurs with a wide variety of different target signals in any modality (or even by omitted stimuli) [Hillyard et al. 1976; Ritter et al. 1976] and only occurs when active attention is paid towards the targets (e.g. Picton and Hillyard 1974).

The P3b has a widely distributed scalp topography with a maximum amplitude over the centro-parietal regions (e.g. Courchesne et al. 1975; Squires et al. 1975). This topography shows little developmental variance. The latency of the P3b decreases with age, whilst the amplitude shows a peak in adolescence before decreasing to adult levels [Courchesne 1978b; Taylor 1988; Travis 1998]. These changes are usually accompanied by a decrease in reaction time with age [Courchesne 1978b].

Identification of the P3b is complicated by an earlier P300 component that may also occur in response to the unpredictable, target stimuli [Squires et al. 1975]. Similar to the novelty P300, this component (usually termed P3a) is probability dependent and task independent. P3a and P3b can be dissociated, however, by their different topography (P3a is usually maximal in frontal areas) and the slightly later latency of P3b.

The neural sources of P3b, like the novelty P300, are likely to include a distributed network of areas. Implicated areas include the medial temporal lobes, temporo-parietal junction, parietal cortex and dorsolateral prefrontal cortex [Knight et al. 1989; Smith et al. 1990; Knight 1991]. Source localisation studies has suggested temporal cortex (especially the superior temporal gyrus) as well as subcortical generators [Alho et al. 1998; Mecklinger et al. 1998]. Intracranial recordings have implicated the amygdala and the hippocampal
formation in the neural generation of the P3b [McCarthy et al. 1982; McCarthy et al. 1989; Paller et al. 1992]. Abnormality in these regions reduces the P3b depth potential [Meador et al. 1987; Wood et al. 1988; Puce et al. 1989].

However several studies of individuals with temporal lobectomies have found intact P3b scalp potentials [Wood et al. 1982; Johnson and Fedio 1986; Johnson 1988; Johnson 1989; Knight 1996]. Other studies investigating the P3b in individuals with medial temporal damage (including some with bilateral medial temporal lobe damage) have found the P3b response is within normal limits [Rugg et al. 1991; Polich and Squire 1993]. Although these lesion studies may suggest that the amygdala and the hippocampal formation may not be critically involved in the generation and propagation of the P3b to the scalp, more extensive electrode recordings have demonstrated that the P3b was disrupted by temporal lobectomies at anterior and mid-temporal electrodes [Nishitani et al. 1999].

The P3b has been more extensively investigated in individuals with Autism than the novelty P300. However, findings regarding the amplitude of the P3b have been inconsistent. Some researchers report significantly smaller amplitude than controls [Lelord et al. 1973; Novick et al. 1979; Martineau et al. 1980; Novick et al. 1980; Courchesne et al. 1984; Courchesne et al. 1985; Oades et al. 1988 but see Kemner et al. 1995]. Interpretation of these findings, however, is confounded by the discrepancy in intelligence between the controls and children with Autism, differences in behavioural performance and inconsistent replications (see Kemner et al. 1995).

Using the ERP components described here as functional indicators of the integrity of the medial temporal lobes, we therefore pursue the hypothesis of dysfunction of this region in children with Autism.
8.2 Adults

8.2.1 Introduction

In order to demonstrate the validity of the experimental set-up and paradigm in evoking the novelty P300 and P3b, this section describes a pilot study of 10 adults (aged 20 – 40).

8.2.2 Methods

Subjects

10 adults (aged 20-40; 4 males) were recruited. None had any history of neurological difficulties, hearing impairments, psychiatric illness or head injury.

Stimuli

The auditory stimuli were comprised of

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Stimulus Description</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>Low tones (700 Hz, 200 ms)</td>
<td>Standards</td>
</tr>
<tr>
<td>10%</td>
<td>High tones (1000 Hz, 200 ms)</td>
<td>Targets</td>
</tr>
<tr>
<td>10%</td>
<td>Novel environmental sounds</td>
<td>Novels</td>
</tr>
</tbody>
</table>

The novels and targets were randomly interspersed with the standard tones. Stimuli were presented binaurally over headphones at a rate of 1 per second. A total of 880 stimuli were presented. Each novel sound was presented four times, to permit examination of amplitude variation of the novelty P300 with repetition.

Experimental Set-Up

The subject sat in an easy chair and was asked to indicate presentation of target stimuli by clicking on a computer mouse button. The subject was asked to look at a computer screen during the recording and watch a slowly rotating cross in the centre of the screen. A practice run was carried out to ensure the task was understood. Breaks were given approximately every 5 minutes.
Behavioural measures

Behavioural measures were made by button presses and were recorded simultaneously with the ERPs (reaction time to targets, number of correctly detected targets and the number of false alarms (responses to non-target stimuli)).

ERP recording

The electrode arrangement used to measure the ERPs was a 13 electrode montage (comprising A1, A2, F3, F4, Fz, C3, C4, Cz, P3, P4, Pz, Reference (nose) and Ground (forehead)) according to the International 10-20 system, and two bipolar electrodes attached above and below the right eye (vEOG) and the outer canthus of the left and right eyes (hEOG) to monitor eye movement.

The recording system had gain of 75000, A/D conversion rate (or sampling rate) of 500 Hz, a band pass of 0.01 – 100 Hz and a notch filter at 50 Hz was used. Resistance at each electrode was kept below 10 kΩ. Silver chloride electrodes were used.

Data Analysis

An ocular artefact reduction algorithm (Scan 4.1, NeuroScan Labs) was used to remove the artefacts introduced into the ERPs by blinking. Any trials containing peak-to-peak deflections exceeding ±100 µV after this were excluded from the analysis. Artefact rejection led to an average of 6% of trials being rejected. A minimum of 10 trials was required for inclusion of average waveform in subsequent analysis. Band pass filtering offline was carried out using settings of 0.1 – 100 Hz. ERPs time-locked to the target, standard and novel stimuli were computed for each subject at all recording sites, with epochs extending from 200 ms before stimulus onset until 1400 ms thereafter. The average voltages in the 200 ms preceding the stimulus presentation served as a baseline.

The Novelty P300 was measured as the mean voltage in the 200-350 ms interval and the P3b was measured as the mean voltage in the 250-400 ms interval. The peak latency was
P300 defined as the time point of the maximal positive (Novel P300 and P3b) deflection within this time interval.

For statistical analysis, repeated measures ANOVA were used with the Greenhouse-Geisser correction for analyses with more than 1 degree of freedom. The design (for all ERP analyses unless otherwise specified) was within subject factors of electrode row (3 levels, Frontal, Central, Parietal), electrode column (3 levels, left, central, right) and stimulus type (novel versus standard for novel P300 and target versus standard for P3b).

ERP waveforms are displayed in figures below. These represent the average of the mean traces from a specified stimulus type (e.g. novel) from each individual at each of the 9 electrode positions (F3, etc.). The vertical axis is voltage (with negative up) and the horizontal axis is time in ms (-200 – 900 ms).

8.2.3 Results

8.2.3.1 Behavioural Measures

The behavioural measures indicated that all the subjects had completed the task appropriately (mean target detection = 94.1 %, mean number of false alarms = 3.2).

8.2.3.2 ERP Responses To Novel Stimuli

The ERP waveforms evoked by the novel and standard stimuli for the adult control group are displayed in Figure 8:1. The novelty P300 showed a central maximum, with little hemispheric asymmetry, with a latency of ~ 300 ms. A repeated measures ANOVA revealed a main effect of stimulus type (F(1, 9) = 40, p < 0.001) and a column by stimulus type interaction (F(1.8, 16.3) = 13, p < 0.001). This reflects the midline maximum of the novel P300 response. There was no significant row by stimulus type interaction, reflecting minimal anterior-posterior variation in response (see Figure 8:1 and Table 8:1).
Table 8:1  Mean voltage of adult novel responses at midline electrodes (Mean ± SEM; Voltages in µV, Time in ms)

<table>
<thead>
<tr>
<th></th>
<th>Fz</th>
<th>Cz</th>
<th>Pz</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel P300</td>
<td>5.7 (0.9)</td>
<td>8.2 (1.2)</td>
<td>7.6 (1.5)</td>
<td>287 (16)s</td>
</tr>
<tr>
<td>Std</td>
<td>-0.0 (0.5)</td>
<td>0.0 (0.7)</td>
<td>0.5 (0.8)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 8:1  ERP waveforms evoked by the novel and standard stimuli for adult control group.

8 µvolts. Horizontal axis - 200 - 900 ms, ticks represent 200 ms intervals. Dashed line is response to standard stimuli, solid line is response to novel stimuli. Negative is up

When the novel responses were average by presentation number (first presentation, second, etc.) the expected decrease in amplitude of the P3 was observed.
8.2.3.3 ERP Responses To Target Stimuli

The ERP waveforms evoked by the target and standard stimuli for the adult control group are displayed in Figure 8:2. The P3b showed a parietal maximum, with a latency of approx. 350 ms. A repeated measures ANOVA revealed a main effect of stimulus type ($F(1, 9) = 22, p = 0.001$) and a row by stimulus type interaction ($F(1.3, 12.1) = 5, p = 0.04$). This reflects the parietal maximum of the target P300 response. There was no significant column by stimulus type interaction, reflecting minimal left-right variation in response (see Table 8:2 and Figure 8:2).

Figure 8:2 ERP waveforms evoked by the target and standard stimuli for adult control group.
Table 8:2 Mean voltage of adult target responses at midline electrodes compared to mean voltage in same time window of standard response (Mean ± SEM; Voltages in µV, Time in ms)

<table>
<thead>
<tr>
<th></th>
<th>Fz</th>
<th>Cz</th>
<th>Pz</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target P300</td>
<td>2.2 (0.6)</td>
<td>2.9 (0.9)</td>
<td>4.5 (1.0)</td>
<td>348 (13)</td>
</tr>
<tr>
<td>Std</td>
<td>-0.2 (0.5)</td>
<td>-0.2 (0.6)</td>
<td>0.2 (0.6)</td>
<td></td>
</tr>
</tbody>
</table>

8.2.4 Discussion

The ERP results from the adult controls are consistent with results from other laboratories (e.g. Courchesne et al. 1975; Squires et al. 1975; Friedman and Simpson 1994; Opitz et al. 1999). The target P3b occurs later than the novelty P300, and both show the expected latency and order of magnitude. Further the novelty P300 shows sensitivity to repetition as previously reported (e.g. Kazmerski and Friedman 1995).

8.3 Control Children

8.3.1 Introduction

In the previous section, the paradigm and experimental set-up were validated. In this section the response pattern of the control group of children is characterised. In the light of known developmental changes in the ERP responses, the results from this section characterise the nature of the average response of the child control group to further validate the experimental set-up.
8.3.2 Methods

Subjects:
The children in this section are the same control children as those described in Chapter 3.

Stimuli
Stimuli were as described in Section 8.2, with the exception that 440 stimuli were presented.

The experimental set-up, behavioural measures, ERP recording, and data analysis methods were as described in Section 8.2. Latency measures were obtained from Pz. Artefact rejection led to the rejection of 38% of trials and one individual was excluded from further analysis as the number of trials in the average was less than 10.

Statistical analysis
In addition to the statistical analysis described in Section 8.2, correlations between all performance measures and age and sex were carried out.

8.3.3 Results

No correlations between performance measures and age or sex were observed.

8.3.3.1 Behavioural Measures

The behavioural measures indicated that all the subjects had completed the task appropriately (mean target detection = 90.0 %, mean number of false alarms = 13.1).

8.3.3.2 ERP Responses To Novel Stimuli

The ERP waveforms evoked by the novel and standard stimuli for the child control group are displayed in Figure 8:3. The novelty P300 response had a parietal maximum and
occurred with a mean latency of ~270 ms. As can be seen from Figure 8:3 (especially at Cz), there was an earlier positive peak superimposed on the novelty P300.

A repeated measures ANOVA revealed a main effect of stimulus type ($F(1, 16) = 13, p = 0.002$) and a row by stimulus type interaction ($F(1.2, 18.8) = 8, p = 0.01$). This reflects the parietal maximum of the novel P300 response. There was no significant column by stimulus type interaction, reflecting the hemispheric symmetric nature of the response.

Figure 8:3 ERP waveforms evoked by the novel and standard stimuli for child control group.

16 μvolts. Horizontal axis -200 - 900 ms, ticks represent 200 ms intervals. Dashed line is response to standard stimuli, solid line is response to novel stimuli. Negative is up.
Table 8:3  Mean voltage of child control novel responses at midline electrodes (Mean ± SEM; Voltages in μV, Time in ms)

<table>
<thead>
<tr>
<th></th>
<th>Fz</th>
<th>Cz</th>
<th>Pz</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel P300</td>
<td>1.1 (2.0)</td>
<td>4.1 (2.1)</td>
<td>6.3 (2.0)</td>
<td>269 (10)</td>
</tr>
<tr>
<td>Std</td>
<td>-2.6 (0.8)</td>
<td>-2.9 (0.8)</td>
<td>-2.1 (1.2)</td>
<td></td>
</tr>
</tbody>
</table>

When the novel responses were averaged by presentation number (first presentation, second, etc.) the expected decrease in amplitude was observed (see Figure 8:9).

8.3.3.3  ERP Responses To Target Stimuli

The ERP waveforms evoked by the target and standard stimuli for the child control group are displayed in Figure 8:4 and Table 8:3. The target P3b response showed a parietal maximum and a latency of around ~ 330 ms. As can be seen from Figure 8:4 (especially at Cz), there was an earlier positive peak superimposed on the novelty P300. A repeated measures ANOVA revealed a main effect of stimulus type (F(1, 16) = 10, p = 0.007), a row by stimulus type interaction (F(1.4, 21.9) = 5, p = 0.02) and a row by column by stimulus type interaction (F(2.6, 41.9) = 3, p = 0.4). This reflects the parietal maximum of the target response, with a left sided bias.
Figure 8:4   ERP waveforms evoked by the target and standard stimuli for child control group.

Table 8:4   Mean voltage of child target responses at midline electrodes compared to mean voltage in same time window to standard (Mean ± SEM; Voltages in µV, Time in ms)

<table>
<thead>
<tr>
<th></th>
<th>Fz</th>
<th>Cz</th>
<th>Pz</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target P300</td>
<td>4.4 (1.8)</td>
<td>5.4 (2.3)</td>
<td>7.2 (1.9)</td>
<td>331 (10)</td>
</tr>
<tr>
<td>Std</td>
<td>-1.1 (0.9)</td>
<td>-1.9 (0.7)</td>
<td>-2.2 (1.1)</td>
<td></td>
</tr>
</tbody>
</table>
8.3.4 Discussion

The ERP results from the child control group show the expected pattern. Both the target P3b and novelty P300 showed parietal maxima with the latency of the novelty P300 being earlier than the target P3b (compare to Courchesne 1978b; Travis 1998). Rather surprisingly, however, there was an earlier positive component superimposed on the mean traces of both the target and novel responses. On inspection of the individual data, this seemed to be due to the responses of the younger children. There have been no reports of such a component in the literature and it is unclear why such a peak should occur. However, in the visual domain, Friedman also reported multiple positive components in children that were not present in adults [Friedman et al. 1978].

Consistent with the literature, both the adult and child groups showed similar amplitude novel P300 responses, with a central maximum in the adult group and a more posterior parietal maximum in the children. The latency of the responses were similar in the adult and child control group. Decreasing amplitude of the novelty P300 with repetition of novelty stimuli was observed in the children as previously documented by Cycowicz et al. [1996].

The target P3b showed a parietal maximum in both the adult and child controls. This is consistent with the literature. The amplitude of the response was larger in the children than in the adults. It has been demonstrated that the magnitude of the target P3b increases with age in childhood before dropping back to adult levels [Courchesne 1978b; Taylor 1988; Travis 1998].
8.4 Children With Autism

8.4.1 Introduction

In this final section, the children with Autism completed the experiment. The data from the previous section (the child control group) were used as controls for these results.

8.4.2 Methods

Subjects
The children with Autism were as described in Chapter 2.

Methods were as described in Section 8.3 with the exception of the statistical analysis. Artefact rejection led to the rejection of 38% of trials in the Control group, 29% in the High group and 48% in the Low group. One individual (in the control group) was excluded from further analysis as the number of trials in the average was less than 10.

Statistical Analysis
Topographical differences were assessed using McCarthy and Wood's scaling factor [McCarthy and Wood 1985]. This scaling permits interpretation of topographical differences in terms of different neural sources. Statistical analysis was carried out using repeated measures ANOVA with within subject factors of row (3 levels) and column (3 levels) and a between subject factor of group. Amplitude differences were assessed at the electrode with maximal response in the controls (Pz). One way ANOVA was used as described in Chapter 2. This method of analysis is as recommended in Picton et al. [2000]. Analyses were only completed with the target P3b and novel P300 responses. Standard responses were not analysed (since the medial temporal lobes are not implicated in the neural generation of these responses).
8.4.3 Results

8.4.3.1 Behavioural Measures

As Figure 8:5 shows, there was a significant difference between the three groups on the number of correctly identified targets and number of false alarms (Target: ANOVA: F(2, 40) = 5; p = 0.02: Planned comparisons: CvH: t = 0.2, p = 0.9, HvL: t = 2, p = 0.06; False Alarms: ANOVA: F(2, 40) = 9, p = 0.001: Planned comparisons: CvH: t = -0.3, p = 0.7, HvL: t = -3, p = 0.001). In both cases, this reflects the Low group performing poorly and was accounted for by differences in verbal IQ.

Figure 8:5 a) Percentage of targets correctly identified; b) False alarms (Mean ± SEM)

As Figure 8:6 shows, there was no significant difference between the three groups on reaction time to the target stimuli. All three groups had an average response time of ~500 ms.
8.4.3.2 ERP Responses To Novel Stimuli

The ERP waveforms evoked by novel stimuli for the child control group and the High and Low Autistic groups are displayed in Figure 8:7. The latency of the novelty P300 was similar in all three groups. This was confirmed by statistical analysis (ANOVA: F(2, 40) = 2, p = 0.1). Examination of the topography of the novelty P300 across all three groups shows a similar anterior-posterior gradient of increasing magnitude to the posterior (parietal) electrodes. Statistical analysis of the scaled data confirms that the topographical distribution of the response was not significantly different across the groups (there was no significant group effect, or group by row or column interaction). The amplitude of the response in the Low group seems to be larger than those of the other two groups (especially posteriorly). However this was not statistically significant (ANOVA: F(2, 42) = 0.9, p = 0.4).
Figure 8:7 ERP waveforms evoked by the novel stimuli for all 3 groups (Solid line = Controls; Dashed line = High group, Dotted line = Low group).

16 μvolts. Horizontal axis -200 - 900 ms, ticks represent 200 ms intervals. Dashed line is response to standard stimuli, solid line is response to target stimuli. Negative is up.

Figure 8:8 Anterior – Posterior Topographical differences in response to novel stimuli (Negative is up)
Figure 8:9 shows how the magnitude of the novelty P300 varies with repeated presentation of the novel stimuli. Whilst the controls showed decreased amplitude with repetition, the High group showed no change in amplitude and the Low group actually showed an increasing response with repetition. Unfortunately, the low number of trials averaged in each individual render statistical analysis of these trends inappropriate.

Figure 8:9 Effect of repeated presentation of novel sounds (Mean ± SEM) Negative is down

Table 8:5 Mean amplitude and latency of novel P300 responses to novel stimuli (Mean ± SEM) (Voltages in μV, Time in ms)

<table>
<thead>
<tr>
<th></th>
<th>Control Novel</th>
<th>High Novel</th>
<th>Low Novel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fz</td>
<td>1.1 (2.0)</td>
<td>4.4 (1.7)</td>
<td>4.0 (2.0)</td>
</tr>
<tr>
<td>Cz</td>
<td>4.1 (2.1)</td>
<td>8.5 (2.5)</td>
<td>8.2 (2.0)</td>
</tr>
<tr>
<td>Pz</td>
<td>6.3 (2.0)</td>
<td>8.2 (2.1)</td>
<td>11.1 (2.5)</td>
</tr>
<tr>
<td>Latency</td>
<td>268 (10)</td>
<td>278 (12)</td>
<td>298 (9)</td>
</tr>
</tbody>
</table>
8.4.3.3 ERP Responses To Target Stimuli

The ERP waveforms evoked by target stimuli for the control group and the High and Low Autistic groups are displayed in Figure 8:11. The latency of the target P3b response across the groups was similar. This was confirmed by statistical analysis ($F(2,40) = 2, p = 0.1$). The topographical distribution of the response was also similar between the three groups with an increasing response along the anterior-posterior axis (see Figure 8:10). This was confirmed by statistical analysis of the scaled response (no group, or group by row or column interaction). The Low group appear to show a smaller response to the targets than the controls (see Pz in Figure 8:10 and Figure 8:11). However, this trend was not statistically significant ($F(2, 41) = 0.7, p = 0.5$) (see Table 8:6).

Figure 8:10 Mean Voltage of target P3b on midline electrodes (Mean ± SEM) Negative is up

![Graph showing ERP responses for control, High, and Low groups]
Figure 8.11  ERP waveforms evoked by the target for all three groups. (Solid line = Controls; Dashed line = High group, Dotted line = Low group).

Table 8.6  Mean amplitude and latency of responses to target stimuli (Mean ± SEM)
(Voltages in μV, Time in ms)

<table>
<thead>
<tr>
<th></th>
<th>Control Target</th>
<th>High Target</th>
<th>Low Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fz</td>
<td>4.4 (1.8)</td>
<td>2.9 (1.3)</td>
<td>0.25 (1.6)</td>
</tr>
<tr>
<td>Cz</td>
<td>5.4 (2.3)</td>
<td>4.3 (1.7)</td>
<td>0.7 (1.8)</td>
</tr>
<tr>
<td>Pz</td>
<td>7.2 (1.9)</td>
<td>7.4 (1.8)</td>
<td>4.1 (1.7)</td>
</tr>
<tr>
<td>Latency</td>
<td>331 (10)</td>
<td>322 (11)</td>
<td>317 (11)</td>
</tr>
</tbody>
</table>
8.4.4 Discussion

Novelty P300

All three child groups showed a similar magnitude, distribution and latency of the novelty P300. This was a posterior parietal maximum response. However, there was a non-significant trend for the Low group to show a larger amplitude novelty P300. This appeared to be related, at least in part, to the Low group not showing the expected decrease in amplitude with repeated novelty stimuli presentation.

Although previous studies of the novelty P300 in children with Autism have found latency and distributional responses similar to those of controls [Courchesne et al. 1984; Courchesne et al. 1985], the similarity in the magnitude of the novelty P300 found in this study is in contrast to previous reports. Courchesne and colleagues reported abnormally small novelty P300 in two different studies [Courchesne et al. 1984; Courchesne et al. 1985], a finding further replicated by Kemner and colleagues [1995]. There are a number of possible explanations for this difference. These include ERP methodological differences and subject selection differences.

Methodological Differences

One important difference between the current study and previous work is the inclusion of repeated novel stimuli in this paradigm. In other studies, each novel sound was only presented once. The apparent lack of repetition-associated decrease noted in the Low group would therefore not have been detected in other studies. The results from the first presentation of the novel stimuli are therefore more comparable to these studies. However, as Figure 8:9 shows, the novelty P300 to the first presentation of the novel stimuli in the Low group was of a similar magnitude to the controls. This is in contrast to the reduced magnitude response noted by the Courchesne and Kemner studies.

The nature of the novel stimuli used in previous studies is another potential source of variation. Non-environmental sounds (such as digitally synthesised nonsense sounds) have been included in the novel stimuli in previous studies. The current study only used environmental sounds. There is some evidence to suggest that environmental novel sounds
are differentially processed in adult controls, although this probably begins after the novelty P300 [Mecklinger et al. 1997; Opitz et al. 1999].

Previous studies have quantified their results by measuring the peak voltage within a given time window, rather than obtaining a mean voltage over this time window. Peak measurements are particularly susceptible to noise and slight changes in latency of the component. However such a difference is unlikely to account for the discrepant results, as the results of this study remain unchanged when analysed by peak voltages (data not shown).

It should also be noted that the Low group had a significantly higher rate of false alarms compared to the remaining two groups. Many of these responses were to novel stimuli. It is therefore possible that the children in the Low group were treating the novel stimuli, at least to some extent, as targets. It is possible that inclusion in the analysis of trials with incorrect behavioural responses is therefore masking abnormalities in the waveforms of the Low group in correct trials. Incorrect trials were not excluded from the analysis as the reduction in the number of trials would have reduced the number of individuals with sufficient trials to an unacceptable level. However, it appears unlikely that the false alarms to novels can explain the results of the study. If the novels were being processed as targets, the magnitude of the responses to the novels and targets should be similar in the Low group. This was not the case. Additionally, previous studies have not excluded incorrect trials. Nevertheless, future studies should include more trials and address the issue of false alarms.

**Subject Selection Differences**

In previous studies, subjects have not been of average intelligence, nor have there been intelligence-matched controls. This could explain the lack of replication in the High group. However, this can not explain the discrepancy between the results from the Low group and previous reports. No significant difference in amplitude was noted in the Low group and, indeed, the trend appeared to be increased amplitude in the Low group.
Previous studies have not excluded children with epilepsy from participation and some have not used the DSM-IV diagnostic criteria. The age range of the children studies has also varied (13 – 25 years : [Courchesne et al. 1984]; 13 – 25 years : [Courchesne et al. 1985]; mean age 10 years: [Vendrell et al. 1995]) compared to this study. However the age range included in this study spans these ranges, making developmental differences an unlikely explanation.

Target P3b
There was no significant difference in the latency, amplitude or distribution of the ERP responses to the target stimuli between the groups. The findings that the latency and topography of the target P3b were not significantly different from controls is in agreement with previous studies. However, although some have also found no significant amplitude differences compared to controls [Kemner et al. 1995; Kemner et al. 1999], others have reported that the amplitude of the response is significantly smaller [Novick et al. 1979; Novick et al. 1980; Courchesne et al. 1984; Courchesne et al. 1985; Verbaten et al. 1991]. There does not seem to be any experimentally consistent factor that varies with the differing results. For example, both significant and non-significant amplitude differences have been found using control children unmatched on VIQ and using differing latencies to define P3b (e.g. Courchesne et al. 1985; Kemner et al. 1995).

Such inconsistent results have also been found for the visual target P3b. Whilst some have found that visual P3bs show a pattern of abnormality similar to that of auditory P3bs in Autism [Novick et al. 1979; Verbaten et al. 1991], others have reported intact visual P3bs [Courchesne et al. 1985].

The factors underlying these variations are poorly understood. It is interesting to note that the reduced amplitude P3bs have also been reported in other psychiatric populations (including schizophrenia [Frodl-Bauch et al. 1999] and ADHD [Rowe et al. 2001]). The findings of abnormal P3bs in other studies may therefore be due to the inclusion of children with co-morbid disorders. A further possibility is that the responses of the children are unstable and variable. In support of this, Courchesne reported that inspection of the ERP traces trial by trial reveals that the children with abnormally small P3b responses show
occasional normal amplitude P3b responses. He suggests that the systems that generate the P3b are capable of response, but perhaps they are usually prevented from responding by some sort of abnormal neural interference [Courchesne 1987]. A simpler explanation is that the children's attention to the task is inconsistent. The target P3b response only occurs during active attention [Picton and Hillyard 1974], so it is plausible that the smaller response is a reflection of the child failing to attend. Although this explanation is consistent with differences in behavioural response accuracy between the control and Autistic groups reported in some studies (e.g. Verbaten et al. 1991), others have found no difference in behavioural responses, yet reduced P3b amplitudes (e.g. Courchesne et al. 1985).

In summary, these results show that the novelty P300 and target P3b in children with Autism were not significantly different from controls. These results do not preclude, however, any abnormality in the regions associated with the generation and propagation of these potentials. It is possible that selective abnormalities within these regions might not affect the neural generation of the P300 responses. Indeed, lesions to the medial temporal lobe have been found to not disrupt the target P3b response [Knight 1996]. Additionally, abnormality early in development in these areas may lead to reorganisation of the neural generators of the responses, resulting in no residual abnormality. This last suggestion could be addressed by examination of the responses of children with known developmental medial temporal lobe abnormality (such as individuals with developmental amnesia [Vargha-Khadem et al. 1997a]). These children have bilateral hippocampal atrophy often following very early insults (such as hypoxia at birth). These results also raise an interesting question regarding novelty responses in children with Autism. Future studies should examine more closely the novelty P300 and its reduction following stimulus repetition in children with Autism. A failure to respond differentially to novel stimuli presented previously may offer an insight into some of the symptomatology associated with Autism (such as insistence on sameness and difficulties of integration in context).
Chapter 9  N400

In this chapter another ERP component is investigated. The N400 is a negative ERP deflection that occurs in response to anomalous stimuli. Its generators are thought to be near the amygdala. Hence, in this chapter, the integrity of the N400 in Autism is investigated.

9.1 Introduction

The N400 was first described in an experiment comparing ERPs to words that either formed anomalous or predictable completions to sentences [Kutas and Hillyard 1980]. ERPs to anomalous sentence endings contain a centro-parietal distributed negative wave (the N400) which is not present in the ERPs to predictable sentence endings. In the auditory domain, the N400 begins about 200 ms after stimulus onset and peaks between 250-500 ms [Picton and Stuss 1984; Rugg et al. 1986] and tends to show little hemispheric asymmetry.

Since its original description, it has been demonstrated that the N400 can also be evoked in a number of different experimental paradigms which contrast semantic congruence. Larger N400s have been noted to presentation of a word following an unrelated (prime) word compared to presentation of a word following a semantically related prime [Bentin et al. 1985a; Bentin et al. 1985b; Rugg 1985]. These studies have attempted to elucidate the functional significance of the N400. The amplitude of the N400 appears to vary primarily as a function of a word’s relation to the ongoing semantic context and thus to its ease of integration (see Kutas and Van Petten 1994).

The N400 effect has been demonstrated in children as young as five years of age [Holcomb et al. 1992; Byrne et al. 1999]. The magnitude of the N400 is inversely related to age [Holcomb et al. 1992].

The scalp distribution of the N400 is complex and is likely to represent temporally overlapping activity from different neural generators. The broad midline central
distribution is suggestive of contributions from deep neural generators [Kutas et al. 1988a; Kutas et al. 1988b]. Although the location of these generators is difficult to ascertain, there is convergent evidence of the importance of the medial temporal lobes in the generation of the N400.

The N400 is largely attenuated following unilateral anterior temporal lobectomy, suggesting that the temporal lobe plays a role in the generation of the N400 [Smith and Halgren 1988; Smith and Halgren 1989]. Consistent with this, intra-cranial recordings have reported large negative potentials with a peak latency near 400 ms elicited in the anterior medial temporal lobes by anomalous sentence-ending words bilaterally (termed AMTL-400) [McCarthy et al. 1995; Guillem et al. 1995]. The AMTL-400 and the scalp recorded N400 respond to the same manipulations (e.g. semantic priming) [Smith et al. 1986; McCarthy et al. 1995; Nobre and McCarthy 1995] and it is therefore highly likely that the AMTL-400 contributes to the neural generation of the N400. Further, intracranial recordings have noted characteristics of local voltage gradients that implicate the amygdala, areas anterior to the hippocampal formation, the fusiform and parahippocampal gyri in the neural generation of the AMTL-400 [Smith et al. 1986; Nobre et al. 1994; McCarthy et al. 1995]. Other areas of the brain have also been implicated in the neural generation of the N400, including posterior temporal lobe and prefrontal regions [Guillem et al. 1995].

The N400 has not been investigated in individuals with Autism. However, there is behavioural evidence that individuals with Autism have difficulty integrating contextual information (see Chapter 1). Indeed, one of the leading psychological theories of Autism suggests that Autism is characterised by a specific imbalance in integration of information at different levels. In other words, individuals with Autism do not limit their interpretation of information by the context in which stimuli are presented. Frith argues that this weak central coherence could explain parsimoniously the cognitive profile of Autism [Frith 1989; Happé 1994; Happé 1999].

Abnormality of the N400 in Autism would be consistent with medial temporal lobe abnormality. It was therefore decided to investigate the integrity of the N400 in children with Autism. A word paradigm was chosen to minimise the confounds of syntax, context
and the heavy load on semantic working memory which complicate interpretation of sentence completion paradigms [Weisbrod et al. 1999]. A behavioural task was included, with the subject asked to indicate if the words were semantically related. Use of a task provides behavioural evidence of attention to the stimuli and the task was designed to encourage relational processing as this has been shown to augment the N400 effect [Kutas and Hillyard 1989].

The experimental paradigm was also designed to address a second question. One hypothesis of Autism is that the children have a specific impairment of the 'social' brain (i.e. abnormalities are restricted only to social tasks, stimuli, etc.) (e.g. Baron-Cohen et al. 1985b; Baron-Cohen et al. 1999). Word pairs were therefore categorised as 'social' and 'non-social' pairs. The results were then investigated for evidence of differential processing of the social and non-social stimuli in Autism.

9.2 Adults

9.2.1 Introduction

As in Chapter 8, the experimental set-up and paradigm were first validated in a pilot study of 10 adults. Replication of previous reports of the N400 was sought. In addition, this adult pilot study was used to characterise any differences in processing of social pairs compared to non-social pairs.

9.2.2 Methods

Subjects

10 adults (as described in Chapter 8) completed the paradigm.

Stimuli

The stimuli used were spoken words. Words were chosen from the MRC psycholinguistic database (www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm) and only words with an age
of acquisition below 9 years were selected. The words were then rated on a scale of 1 to 5 as socio-emotional or non-socio-emotional by 15 adults. Words with a score 2 or below were then paired together to create related or unrelated social pairs. Words with a score 4 or above were similarly paired to create related or unrelated non-social pairs. Word pairs were then rated again by 15 adults (some of whom took part in the first rating) on a scale of 1 to 5 as related or unrelated.

From this, 50 word pairs were designated as related social word pairs, 50 word pairs as unrelated social word pairs, 50 word pairs as related non-social word pairs and 50 word pairs as unrelated non-social word pairs. For example: scout - grave was a unrelated social word pair, shriek - scream was a related social word pair, sweep - rake was a related non-social word pair and juice - mud was a unrelated non-social word pair. The first word of each pair was called a 'prime' and the second word a target. No word was used twice.

The duration of the spoken words varied from 461 ms to 915 ms. There was no significant difference in duration between related and unrelated primes and targets (ANOVA: F(3, 396) = 2, p = 0.2). There was no significant difference between the related and unrelated primes and targets on the Kucera-Francis frequency measure (ANOVA: (F(3, 396) = 0.6, p = 0.6). There was a significant difference between the unrelated and related word pairs on the age of acquisition measure (ANOVA: F(3, 396) = 6, p < 0.001). This was a reflection of the related word pairs having a significantly lower mean age of acquisition compared to the unrelated word pairs (see Table 9:1).
Table 9:1 Characteristics of word stimuli (Mean ± SEM)

<table>
<thead>
<tr>
<th>Word type</th>
<th>Mean duration of word (ms)</th>
<th>Mean frequency*</th>
<th>K-F</th>
<th>Mean Age of acquisition**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Related prime</td>
<td>713 (11)</td>
<td>62 (9)</td>
<td>319 (8)</td>
<td></td>
</tr>
<tr>
<td>Related target</td>
<td>703 (11)</td>
<td>62 (12)</td>
<td>324 (8)</td>
<td></td>
</tr>
<tr>
<td>Unrelated prime</td>
<td>735 (9)</td>
<td>47 (7)</td>
<td>361 (8)</td>
<td></td>
</tr>
<tr>
<td>Unrelated target</td>
<td>722 (9)</td>
<td>66 (12)</td>
<td>350 (8)</td>
<td></td>
</tr>
</tbody>
</table>

* Frequency as described in Kucera and Francis [1967] (Numbers represent frequency of occurrence of words in Brown Corpus (a collection of 500 text samples of a variety of genre))

** Scale ranged from 100-700, from 100 (age 0-2 years) to 700 (13 years and older). Intermediate points were identified with 2 year age bands (i.e. 100 times those quoted in Gilhooly and Logie [1980]).

The words were recorded in a random order, spoken by a female. The recordings were digitised. Prior to the experiment, each word was edited for a precise time of onset so as to permit synchronisation with ERP recording and behavioural reaction. During the experiment the digital representations of the prime and target stimuli were output through a digital-to-analogue converter and played to the subject over headphones.

The ERP recordings were triggered by the start of the word. Word pairs were presented in a random order. Total time for the paradigm was approximately 20 minutes. The interstimulus interval was 1500 ms and the intertrial interval was determined by the subject’s response (minimum time 1500 ms).

Experimental Set-up
The subject sat in an easy chair during the recording. Breaks were given approximately every 5 minutes. A practice run was carried out, to ensure the task was understood.
Behavioural measures
The subject heard a pair of words and was asked to answer the question “Do these words go together?” by pressing one of two mouse buttons with their preferred hand (in all cases this was the right hand). The button assigned to Yes was counterbalanced across the groups.

ERP Recording
This was completed as in Chapter 8.

Data Analysis
This was completed as in Chapter 8, with the exception of off-line filtering. This was done using bandpass settings of 0.01 – 30 Hz. Artefact rejection led to an average of 3% trials being rejected.

Mean voltage values were determined for the time period 350-700 ms (previously shown to be an appropriate response window for the N400 in children aged over 8 [Byrne et al. 1999]). Latencies were not measured, as the broad and slow nature of the N400 effect means that determination of this factor is of limited value.

Statistical Analysis
For statistical analysis, repeated measures ANOVA were conducted, with the Greenhouse-Geisser correction used whenever the degrees of freedom were greater than 1. As the N400 effect is centro-parietal, analysis was restricted to the centro-parietal electrodes (see Byrne et al. 1999; Weisbrod et al. 1999). The design, unless otherwise stated, had within subject factors of electrode row (2 levels), electrode column (3 levels) and stimulus type (related or unrelated targets).

As in Chapter 8, ERP waveforms are displayed in figures. These represent the average of the mean traces from a specified stimulus type (e.g. Related target) from each individual at each of the 9 electrode positions (F3, etc.). The vertical axis is voltage (with negative up) and the horizontal axis is time in ms (-200 – 900 ms).
9.2.3 Results

9.2.3.1 Behavioural Measures

The adult behavioural results demonstrated that the task was completed appropriately. The mean reaction time was 1.4 s (related pairs) and 1.7 s (unrelated pairs). Subjects correctly categorised 90.3% of the word pairs. They categorised correctly significantly fewer of the non-social word pairs than the social word pairs (Mean non-social pairs = 87%; mean social pairs = 94%; Student's t test: t = 3, df = 7, p = 0.01).

9.2.3.2 Event Related Potentials

The ERP waveforms evoked by the related and unrelated targets for the adult control group are displayed in Figure 9:2. The N400 showed a maximum with differences in ERP responses between related and unrelated targets, beginning very early post stimulus onset. It showed little hemispheric asymmetry. Statistical analysis revealed no overall effect of stimulus type (F(1, 9) = 1, p = 0.2), but a row by stimulus interaction (F(1,9) = 7, p = 0.03). No stimulus by column interaction was noted (F(1.6, 14.1) = 0.3, p = 0.8). This reflects the parietal distribution of the N400 effect.

Figure 9:1 Anterior-Posterior Variations in response to related and unrelated target words in adult control group (Negative is up)
Figure 9:2  ERP waveforms evoked by the related and unrelated targets for adult control group.

---

Social versus Non-social Results
As Figure 9:3 shows, no differences were found between the processing of social and non-social pairs. Statistical analysis confirmed this (ANOVA: Rptd measures: (within subject effects row (2), column (3), stimulus (2), social (2)) Effect of social (F(1, 9) = 0.009, p = 0.9)
9.2.4 Discussion

Differential neural responses were found to unrelated targets compared to related targets. The size and distribution of the N400 effect is similar to that in previous reports (e.g. see Holcomb and Neville 1990). Although it may appear surprising that the responses to related and unrelated targets diverge so early (especially as the average word duration was ~ 720 ms), divergence at 200 ms has been reported in the literature (e.g. Holcomb and Neville 1991). The onset latency of the N400 to semantic anomalies has been shown to vary with the discrepancy point (the point at which the acoustic information first diverges from expectation) (Van Patten and colleagues (unpublished data quoted in Kutas and Iragui 1998)). In this study the difference appears even earlier than 200 ms. This may be associated with stimuli specific effects (such as word presentation, discrepancy point, nature of word, etc. since primes and targets were not counterbalanced) or the early age of acquisition of the words.

The inclusion of a behavioural task to ensure the subject attended to the stimuli yields a confounding factor. A decision following presentation of the target word generates late
ERP components that may overlap in timing with N400 and thus distort measurements of its timing, amplitude and topography. There was an overall significant difference between the reaction times to the related and unrelated pairs. It might, therefore, be argued that the N400 effect might be merely a reflection of the potentials induced in order for the motor response to occur. However, if this was the case, the N400 effect should show asymmetries (since responses were always carried out using the right hand). Further, the average length of reaction time was ~1.5 seconds, long after the N400 effect is observed (beginning at ~200 ms).

It should be acknowledged that the preceding words in the paradigm might act as primes for the target, not just the intended prime word. However, it has been found that statistically significant semantic priming effects only occur to consecutive words [Dannenbring and Briand 1982]. Thus whilst the target might affect the ERP associated with the following prime, it does not affect the response to the following target.

### 9.3 Control Children

#### 9.3.1 Introduction

In the previous section, the paradigm and experimental set-up were validated. In this section the response pattern of the control group of children is characterised. In the light of known developmental changes in the ERP responses, the results from this section characterise the nature of the average response of the child control group to further validate the experimental set-up.
9.3.2 Methods

Subjects:
The control children in this section are the same as those described in Chapter 2.

Stimuli, Experimental Set-up, Behavioural measures, ERP Recording, Data analysis and Statistical analysis
All are as described in Section 9.2. Rejected trials made up 37% of trials and no individuals were excluded from further analysis due to number of trials in average < 10.

9.3.3 Results

9.3.3.1 Behavioural Measures

The behavioural responses indicated that the children were able to complete the tasks (% correct = 81.9, reaction time: Related pairs 1.7 s; Unrelated pairs 2.1 s). The children obtained significantly more correct answers on the social word pairs than the non-social word pairs (t = 3, p = 0.004).

9.3.3.2 Event Related Potentials

The ERP waveforms evoked by the related and unrelated targets for the child control group are displayed in Figure 9:5. The N400 was seen most clearly in the central electrodes. It had a centro-parietal distribution with a slight left-sided bias. This pattern was confirmed by a repeated measures ANOVA. This revealed a borderline significant effect of stimulus type (F(1, 16) = 4, p = 0.05) and no interaction of stimulus by column or row.
Figure 9:4  Anterior-Posterior Responses along midline in child control group (Negative is up)

Figure 9:5  ERP waveforms evoked by the related and unrelated targets for child control group.

-8 μvolts. Horizontal axis -200 - 800 ms, ticks represent 200 ms intervals. Solid line is response to related target, dashed line is response to unrelated target. Negative is up. Stimulus onset at t = 0.
Social Nature of Response

Figure 9:6 shows the mean difference in the N400 response to related and unrelated words sorted by the social nature of the pairs. Although there was no significant difference, there was a trend for a decreased N400 effect to the non-social pairs.

Figure 9:6 Difference in magnitude of N400 effect in control children to related and unrelated target words, averaged by the social nature of the pair.

9.3.4 Discussion

The ERP results show that the children in the control group had differential neural responses to the related and unrelated targets. This difference took the form of a centro-parietal increased negativity to unrelated targets compared to related targets, peaking between 400 – 500 ms. The distribution, amplitude and time course were consistent with previous reports of the N400 (e.g. Holcomb and Neville 1990; Byrne et al. 1999).

The size of the N400 effect was larger in the control children than in the adult controls. This is consistent with a previous study of individuals aged 5 – 26. All age groups showed an N400 effect to an anomalous sentence paradigm and the effect was inversely correlated with age [Holcomb et al. 1992]. One possible explanation for this pattern is that the unrelated word pairs have some associations for adults, whilst the children’s more limited life experience means the words are more truly unassociated.
The social - non-social comparison showed, as in the adult data, no significant differences in responses. However the graph does suggest that the N400 effect tended to be larger to social pairs than to non-social word pairs. This result may be linked to the reduced accuracy of responses in the non-social pairs versus the social pairs. The related non-social pairs may also have less strong semantic associations which would reduce the N400 effect. However this effect is perhaps unlikely as the adults showed similar N400 effects for social and non-social pairs despite significant differences in the accuracy of the Behavioural response.

Alternatively the results may reflect a true effect that may have been significant if a larger number of pairs were used. It is interesting to note that Byrne and colleagues found an increased N400 effect in normal children to pairs of words made up of vocabulary with an age of acquisition below their chronological age compared to those made up of vocabulary acquired more recently [Byrne et al. 1999]. The social words did not differ from the non-social words in mean frequency, but unavoidably did have an earlier age of acquisition. As suggested by Byrne, this may therefore reflect a change in processing strategy as a function of the level of challenge [Byrne et al. 1995; Byrne et al. 1999]

9.4 Children with Autism

9.4.1 Introduction

In this final section, the children with Autism completed the experiment. The data from the previous section (the child control group) were used as controls for these results.
9.4.2 Methods

Subjects
The children with Autism were as described in Chapter 2.

Methods were as described in Section 8.3 with the exception of the statistical analysis. Rejected trials made up 37% of trials in the Control group, 31% in the High group and 36% in the Low group. No individuals were excluded from further analysis due to number of trials in average < 10.

Statistical Analysis
Topographical differences were assessed using McCarthy and Wood’s scaling factor [McCarthy and Wood 1985]. This scaling permits interpretation of topographical differences in terms of different neural sources. Statistical analysis was carried out using repeated measures ANOVA with within subject factors of row (2 levels) and column (3 levels) and a between subject factor of group. Analyses were completed with the difference wave (related – unrelated) since the modulation, not the absolute values of the response, was the critical variable. Amplitude differences were assessed at the maximal response electrode in the control group (Cz). One way ANOVA was used as described in Chapter 2. This method of analysis is recommended in Picton et al. [2000]. Correlations between performance measures and sex and age were examined.

9.4.3 Results

9.4.3.1 Behavioural Measures

There were two behavioural measures: the children’s percentage of correct semantic judgements and additional behavioural responses. The latter was a count of the number of extra button presses the children made. Figure 9:7 shows that there were significant group differences on both behavioural measures (Semantic judgements: ANOVA: F(2, 39) = 21, p
< 0.001: Planned comparisons CvH: t = 0.8, df = 39, p = 0.4; HvL: t = 5, df = 39, p < 0.001; Additional behavioural responses: ANOVA: F(2, 39) = 9, p < 0.001: Planned comparisons CvH: t = -0.4, df = 39, p = 0.7, HvL: t = -4, df = 39, p = 0.001). Both these differences reflect a significant difference between the Low group and the other two groups. Only the additional behavioural responses were accounted for by including verbal IQ and age as covariates. All three groups scored above chance with the correct semantic judgements. Further, there was no indication of perseverative responding in the Autistic groups. Inspection of the responses revealed no unusually long sequences of one answer (e.g. ‘Yes’) to whether the words were related and, as Figure 9:8 shows, the children with Autism responded ‘Yes’ to the question of semantic congruency with a frequency similar to that of the controls.

Figure 9:7 Behavioural results from N400 paradigm a) % Correct semantic judgements; b) Additional behavioural responses (Mean ± SEM)

Figure 9:8 Percentage of responses ‘Yes’ to semantic congruency question (Mean ± SEM)
The reaction times of the behavioural responses are shown in Figure 9:9.

Figure 9:9  
Reaction Times for related and unrelated word pairs (Mean ± SEM)

When the behavioural responses were divided into social and non-social pairs, the Autistic groups showed increased accuracy for social pairs compared to non-social pairs. This pattern mirrored that of the controls (see Figure 9:10) (Rptd Measures (2 levels: social, non-social) F(1, 37) = 0.5, p = 0.6).

Figure 9:10  
Percentage of correct behavioural responses divided according to the social nature of the word pairs. (Mean ± SEM).
9.4.3.2 Event Related Potentials

The difference waveforms between the traces evoked by related and unrelated targets (related – unrelated) for the three different groups are displayed in Figure 9:12. The results demonstrate that the High group showed an N400 effect similar to that of the control group at the central and parietal electrodes with a central maximum. The High group showed a reduced N400 effect at the frontal electrodes. The N400 effect in the High group appeared to be more prolonged, with a late ‘peak’ at 600 ms. The Low group appeared to show little N400 effect across all electrodes. Statistical analysis revealed no significant differences between the groups at Cz (F(2, 41) = 2, p = 0.2). The topographical analysis of the scaled data revealed no significant effects of group (F(2, 39) = 1, p = 0.2) or interactions of group by column or row.

Figure 9:11 Anterior-Posterior variation of N4 effect in the three groups (Mean ± SEM)
Figure 9:12 Difference ERP waveforms (related – unrelated) for all three groups (Solid line = Controls; Dashed line = High group, Dotted line = Low group).

Social Nature of Response

Further analysis of the results from subsets of all three groups demonstrated that there was no significant difference in the amplitude of the N400 effect to social and non-social pairs (Repeated ANOVA: $F(1, 29) = 0.03$, $p = 0.9$). The topography of the response (according to the scaled data) showed no difference between the groups.
9.4.4 Discussion

The results show that there were no significant group differences in the magnitude or topography of the N400 effect, despite apparent differences that can be seen on visual inspection of the waveforms. The significance of the results may be masked by a number of factors. Only 9 electrodes were used in this study: it is plausible that more extensive electrode recording may have detected significant abnormalities. Secondly, a source of variance in this study is the wide age range of the subjects. There is some evidence to suggest the N400 changes with age [Holcomb et al. 1992], and so it is possible that use of a smaller age range may increase sensitivity.

On visual inspection, it appears that there is little if any N400 effect in the Low group. It is possible that with a higher signal-to-noise ratio (obtained by an increased number of trials, subjects, electrodes, etc.) this effect may become significant. Interpretation of this difference is confounded by a lack of verbal-IQ matched controls and the significantly poorer performance of the Low group on behavioural measures. The apparent lack of an
N400 effect may merely be suggestive that children in the Low group were not processing the word pairs in a manner similar to that of the controls. One possibility is that the words were unfamiliar (although this is unlikely as all the children in this study had a reading age of 8 years or above). Another possibility is that the children in the Low group were failing to attend to the paradigm. There is little evidence to support this, as the behavioural responses were not perseverative and the children scored above chance. Further the N400 effect is not dependent on behavioural responses [Bentin et al. 1993]

It was recently established that there is a negative ERP deflection around 400 ms after presentation of a novel stimulus [Mecklinger et al. 1997]. The neural generators of this so-called novelty N400 and the relation of this to the N400 described here are not yet known. However, consistent with the apparent lack of the N400 effect in the word paradigm, the Low group also show little (or no) N400 effect in the novel ERPs (see Chapter 8).

The N400 has not been studied in Autism. However difficulties with contextual integration are the basis of a popular psychological theory of Autism (weak central coherence). In addition, children with high functioning Autism have been described as having difficulty with metaphors and other complex aspects of language.

Results from the control group indicate that normally developing children do not show differential processing of social and non-social stimuli as evidenced by the N400 effect. This result was mirrored in the children with Autism. This evidence is consistent with the hypothesis that children with Autism process both social and non-social stimuli similarly to controls. However, as the stimuli were all verbal, it could be argued that both the social and non-social pairs were inherently social (since they represent communicative intent) and so a selective deficit in social stimuli processing would not be detected in this study.

In conclusion, no significant abnormality in the N400 effect in children with Autism was found. Further studies are required to address the significance of the abnormalities noted in the Low group on visual inspection of the waveforms.


Chapter 10  General Discussion and Future Studies

This chapter begins by first summarising the results from the neuropsychological (Part I), neuroimaging (Part II) and electrophysiological (Part III) studies, and then relating these to the findings that have been reported in the literature. A discussion of the generalisability of these findings to the wider population of individuals with Autism then follows. After consideration of the implications of the findings for the hypothesis of medial temporal lobe involvement in Autism, the final section discusses the question of directions of future work.

10.1 Summary of Findings

The aim of this thesis was to provide convergent evidence for medial temporal lobe abnormality in children with Autism. In pursuit of this goal, three groups of children were studied: a control group, a group of high functioning children with Autism (i.e. verbal IQ ≥ 85), and a group of low functioning children with Autism (i.e. verbal IQ < 85). The children in the Autistic groups had been diagnosed with high functioning Autism or Asperger’s syndrome by an appropriate clinician. The results from Chapter 2 confirmed the nature of their difficulties.

The main comparison of interest in this thesis was between the High group and the controls. These groups were matched on verbal IQ and age, meaning that any significant difference between them could be reasonably attributed to the diagnoses of Autism in the High group. A second comparison of interest was between the High and Low groups. Although interpretation of the results from the Low group versus the High group was confounded by their verbal IQ discrepancy, comparisons revealed important differences. The Low group, consistent with their low verbal IQ, showed a greater degree of impairment on many of the psychological assessments (e.g. memory) than the High group. In addition voxel-based morphometry revealed a more extensive pattern of brain abnormality in these children, only some of which appeared related to their low verbal IQ.
In the following sections the results from the High group will be discussed according to implicated brain regions. This will be followed by a discussion of the results from the Low group.

10.1.1 Medial temporal lobes

The structure and function of the medial temporal lobes in Autism were extensively investigated in this thesis. In Part I (Neuropsychology), a cognitive correlate of hippocampal function (episodic memory) and a behavioural correlate of the amygdala (startle modification) were assessed. In a comprehensive memory evaluation, the High group showed good recognition and semantic memory with a selective impairment in episodic memory. The children showed particular difficulties with the subtests of the Rivermead that involved prospective memory (such as remembering to ask a question when an alarm rang). No previous studies have used comparable paradigms to investigate episodic memory in children with Autism. However the results are consistent with reports of impaired event memory [Boucher 1981; Boucher and Lewis 1989; Millward et al. 2000] and good recognition [Ameli et al. 1988; Bennetto et al. 1996; Brian and Bryson 1996; Farrant et al. 1998] in children with high functioning Autism.

This memory profile has also been described in two patient populations with bilateral hippocampal abnormality of developmental origin: individuals with developmental amnesia [Vargha-Khadem et al. 1997a; Gadian et al. 2000] and those born preterm [Isaacs et al. 2000]. The severity of the impairment found in the High group appears to fall between the severe deficit in developmental amnesia and the comparatively mild deficit in preterm children.

One result from the memory battery that might conflict with an interpretation of hippocampal abnormality is the finding of apparently normal performance on the visual paired comparison task. Both the animal and human literature suggest that the visual paired comparison paradigm is sensitive to hippocampal abnormality [Bachevalier et al. 1993; Clark et al. 1996; Pascalis et al. 2000]. However there is evidence from the monkey
Discussion and Future Directions

literature to suggest that lesions of the hippocampal formation during early development do not affect performance on the visual paired comparison task at short delays [Bachevalier et al. 1999]. The performance of individuals with perinatal hippocampal abnormality on this type of task has not yet been investigated. Overall, it appears that the pattern of impaired and spared memory abilities of children with Autism is consistent with hippocampal abnormality.

As to the behavioural correlate of amygdala function, the results from the emotional modification of the startle response suggest that the children with Autism show emotional modulation of the startle similar to the controls. This negative finding indicates that amygdala function as assessed by the modulation of the startle response may not be grossly abnormal in individuals with Autism. However as discussed in Chapter 4, the paradigm may not assess the function of all the nuclei in the amygdala and may be only sufficiently sensitive to detect gross deficiencies in modulation. Furthermore, the plasticity of the developing brain may lead to reorganisation of the neural pathways that subserve such an evolutionary important response.

In Part II, the structure of the amygdala and hippocampal formation was investigated using a variety of different MR techniques. One technique, voxel-based morphometry, revealed bilateral abnormality in the amygdala and the hippocampal formation. Abnormality in the amygdala was found to correlate with a measure of Autistic symptomatology. Another measure, volumetric analysis, failed to detect significant abnormality in the hippocampal formation and the amygdala in the High group. A significant correlation was found between the left hippocampal volume and the Rivermead scores in the High group. The volumetric analyses were not automated and relied on visual identification of the boundaries of the hippocampal formation and the amygdala. As discussed in Chapter 7, this introduces an inherent error (particularly when gyral patterns are used as markers as in the amygdala boundaries) and therefore reduces sensitivity. This may explain the conflicting results.

T2 maps revealed no significant abnormality in the children with Autism in either the hippocampal formation or the amygdala. However T2 values within the normal range do
not preclude abnormality in these regions (see Van Paesschen et al. 1997). In summary therefore, significant abnormality in the amygdala and hippocampal formation was detected in the children with Autism, using voxel-based morphometry. Volumetric analyses and T2 maps failed to further characterise this abnormality.

Only one other study has used voxel-based morphometry to investigate the neuropathology of Autism, and this study did not use the bilateral method. Despite this, abnormality in the amygdala was detected [Abell et al. 1999]. Previous volumetric studies of the amygdala and hippocampal formation have reported mixed results. Some have found increases in volume [Howard et al. 2000], others decreases [Aylward et al. 1999; Pierce et al. 2001; Saitoh et al. 2001] and in others no significant abnormality was detected [Saitoh et al. 1995; Piven et al. 1998]. As discussed in Chapter 7, interpretation of these results is confounded by a number of factors including non-blind assessment, differences in age ranges and intellectual ability.

In Part III, event related potentials were used to attempt to find evidence of abnormality consistent with medial temporal abnormality. No significant group differences were detected, providing no corroborative evidence of pathology in the medial temporal lobes in children with Autism. The results conflict with some previous reports of the target and novel P300 components in Autism (e.g. Courchesne et al. 1984; Lincoln et al. 1993; Kemner et al. 1995). However, replication of these studies has not been consistent and confounds such as a lack of verbal IQ matched controls complicate interpretation.

The results from these three parts are therefore consistent with the hypothesis that individuals with Autism having subtle abnormalities of the hippocampal formation and the amygdala. Not all aspects of the structure and functions of these areas were affected, but in the hippocampal formation, convergent evidence of both structural and functional abnormality was obtained.
10.1.2 Frontal Lobes

The integrity of the functions of the frontal lobes in Autism was investigated in Chapter 5. Tests were selected to assess functions dependent on different aspects of the frontal lobes. The children in the High group were found to perform similarly to the controls on a test of spatial working memory and the Stroop colour word test (both thought to be independent of the orbitofrontal cortex). In contrast, the children from the High group were impaired at the Wisconsin Card Sorting Test. They achieved significantly fewer categories which was found to be related to a failure to sustain correct matching of the cards. This pattern of performance has been reported to be associated with orbitofrontal cortex function [Stuss et al. 1983; Nagahama et al. 1996; Stuss, et al. 2000]. The subtlety of this deficit is highlighted by the children's good performance on rule reversal and extinction tasks, which has previously been found to be disrupted by gross orbitofrontal cortex lesions [Rolls et al. 1994].

Voxel-based morphometry of the whole brain found focal areas of abnormality in the orbitofrontal cortex bilaterally. No significant abnormalities were found in the dorsolateral areas of the frontal lobes (although it should be emphasised that absence of abnormality does not imply presence of normality). These results therefore provide convergent evidence of both behavioural and structural brain abnormalities associated with the orbitofrontal area of the frontal lobes.

What, if any, is the relationship between the medial temporal lobe abnormalities in Autism and the orbitofrontal cortex abnormality? Although this study has not demonstrated a link between the two regions, there is anatomical and behavioural evidence to support this. Firstly, behaviourally, there was a correlation between the Wisconsin set maintenance scores and the Rivermead scores. This suggests that children who performed poorly on the Rivermead also had difficulty maintaining set on the Wisconsin Card Sorting Test. Secondly those areas of the frontal lobes with few, if any, direct connections with the medial temporal lobes (such as the dorsolateral prefrontal cortex) [Barbas 1992; Barbas 2000] have not been found to be significantly abnormal. In contrast, there are extensive
connections between the orbitofrontal cortex and the medial temporal lobes [Barbas and De Olmos 1990; Barbas and Blatt 1995]. Structural abnormality of either area during development is likely to affect both structural and functional development of these regions. For example neonatal lesions to the medial temporal lobes have been shown to affect frontal lobe development [Bertolino et al. 1997; Hanlon and Sutherland 2000]. Additionally if both areas are (potentially independently) abnormal, development of both areas will be further confounded by aberrant interconnections.

Whilst there is evidence (e.g. from voxel-based morphometry) to support the role of the hippocampal formation in the deficit in episodic memory in Autism, the finding of abnormality in the frontal lobes raises another possibility. Functional imaging studies have implicated the frontal lobes in episodic memory in humans (e.g. Lepage et al. 2000) and lesions to the frontal lobes in monkeys lead to memory impairments [Bachevalier and Mishkin 1986; Meunier et al. 1997]. It is therefore possible that the abnormality in the frontal lobes is also contributing to the episodic memory deficit described in this thesis.

Such a possibility can not be ruled out, especially in the light of the nature of the abnormality in the hippocampal formation. In the Autistic group, an increase in grey matter density was found whilst the ‘reverse pattern’ (i.e. a decrease) has been found in previous studies of children with selective memory impairments [Gadian et al. 2000; Isaacs et al. 2000]. Such a contrast is suggestive of different pathological processes in development, which may not necessarily result in the same cognitive profile. Perhaps increased grey matter density in the hippocampal formation only results in an episodic memory impairment when combined with other abnormalities (such as in the orbitofrontal cortex).

Can hippocampal abnormality be contributing to the loss of set found in the Wisconsin Card Sorting Test? Abnormalities outside the frontal lobes have been found to affect WCST performance [Corcoran and Upton 1993; Upton and Corcoran 1995]. For example, Corcoran et al. reported that patients with hippocampal abnormality showed an increase in perseverative errors [Corcoran and Upton 1993]. However, it is unlikely that the hippocampal abnormality was the dominant factor affecting the children’s performance on WCST in that increased perseverative errors were not seen in the children with Autism.
However, the performance of children with bilateral hippocampal abnormality of developmental origin on the WCST has not been reported, so it would be premature to rule out the involvement of the hippocampal abnormality in the children’s performance.

10.1.3 Cerebellum

In Chapter 2, the children in the High group were shown to score poorly on a parental assessment of motor co-ordination. The cerebellum has been repeatedly implicated in motor co-ordination [Verleger et al. 1999; Kakizawa et al. 2000]. In Chapter 7, cerebellar abnormalities were found in the voxel-based morphometry analysis. This provides convergent evidence of cerebellar abnormality in the children with Autism.

What, if any, is the relationship between the cerebellar abnormality and the orbitofrontal cortex and medial temporal lobe abnormalities? The role of the cerebellum in cognition is poorly understood, so it is difficult to link cerebellar abnormalities to the cognitive patterns found in the children with Autism. However, the cerebellum does have connections with both orbitofrontal cortex [Sasaki et al. 1979; Middleton and Strick 1994; Schmahmann and Pandya 1997] and medial temporal lobes [Heath and Harper 1974; Heath et al. 1978] and neural development of these regions co-occurs (see Chapter 7). Thus abnormality in all three areas is consistent with the dynamic and interdependent nature of developmental processes. In the absence of more direct evidence of behavioural impairment indicative of cerebellar dysfunction, however, it is premature to speculate further on the role of the cerebellar abnormalities in the Autistic population.

10.1.4 Low Group Results

Interpretation of many of the results from the Low group is complicated by the lack of a verbal-IQ matched control group. Determining which deficits or abnormalities are associated with low verbal IQ and which are due to Autism per se is difficult. Whilst a more appropriate control group might appear to address this issue, it should be noted that
identifying a suitable control group for a below average IQ group is fraught with difficulty. Is the ideal control group one matched on chronological age and verbal IQ; or chronological age and performance IQ; or normally developing children matched on mental age? All of these possibilities have drawbacks: for example, does a ‘normal’ child with a below average verbal IQ have normal brain structure? If the groups are matched for mental age, does the difference in chronological age (and therefore brain development) confound any results suggesting neural abnormality? Interpretation of the results from a group such as the Low group is therefore highly complex regardless of the control group.

Nonetheless, comparisons of the results with the High group can provide important insights into the syndrome of Autism. Firstly, abnormalities or deficits present only in the Low group and not the High group provide one line of evidence. These differences may reflect associated, but non-essential features of Autism. This could aid identification of the ‘core’ deficits and abnormalities in Autism. Cognitive differences may reveal difficulties that the children in the High group are able to overcome due to their higher intellectual ability. Alternatively the differences may point towards different aetiologies leading to Autism with and without intellectual impairment. Similarly any abnormalities or deficits in the High group not found in the Low group may represent non-essential associated features or point to differing aetiologies. Findings of abnormality in both groups may highlight ‘core’ features strongly associated with Autism.

Comparisons between the High and Low groups also provide another insight. The literature on Autism contains many conflicting findings and inconsistencies, reflecting an imprecise understanding of the neurobiology of Autism. The differences between the High and Low groups may shed some light on some of the causes of such varied results.

As mentioned earlier, the Low group performed significantly worse than the High group on many of the neuropsychological assessments. Many of these differences are likely to be associated with the discrepancy in verbal IQ between the two groups and were accounted for by inclusion of verbal IQ as a covariate. Deficits in these domains (such as attention, receptive language, recognition and semantic memory) therefore are likely to be non-essential, associated impairments of Autism in low functioning individuals. Indeed at least
Discussion and Future Directions

some of these may be more associated with low verbal ability than Autism. These results shed light on a number of the discrepancies in the literature regarding the performance of children with Autism in these areas.

Other differences between the High and Low group could not be accounted for by a linear relationship between verbal IQ and performance. For example the Low group scored significantly worse than the High group on the motor co-ordination checklist, and the Children’s Communication Checklist. This may reflect good compensation of the difficulties by the children in the High group, a genuine difference in the nature of the impairment in the High and Low groups or a bias in parental reporting associated with the lower intellectual abilities of the Low group. More direct measures and a longitudinal study are required prior to deciding between these several alternatives.

None of the neuropsychological results revealed impaired performance in the High group but relatively intact performance in the Low group. Indeed the weight of the cognitive profile found in the High group is strengthened by the Low group results. The pattern of impairment in the Low group included all the deficits in the High group.

Significant differences in neuropathology between the High and Low group were found, even with attempts to account for verbal IQ differences. The areas implicated were the hippocampal formation, amygdala, rostral superior temporal gyrus and cerebellum. Of particular interest is the difference in the amygdala. In both groups, significant abnormality was found in the amygdala: in the High group this was an increase in grey matter density, whilst in the Low group this was a decrease. This suggests that whilst abnormality in the amygdala is associated with Autism, the nature of such an abnormality may vary. This may be indicative of differing pathological processes in the two groups. However, this could be a single cause with differing consequences dependent on environment, a common cause occurring at different neurodevelopmental times, or differing aetiology. The other differences in neuropathology may also be a reflection of these processes. This finding may explain previously discrepant results (e.g. Aylward et al. 1999; Howard et al. 2000).

It is interesting to note that such a dichotomy in pathology has previously been reported in
the cerebellum by Courchesne. He found hypoplasia of the cerebellum in some individuals and hyperplasia in others [Courchesne et al. 1994a].

The conjunction analysis found common areas of significant abnormality in the orbitofrontal cortex, superior temporal gyrus and cerebellum, as well as in the insula. This highlights the importance of these areas in the neuropathology of Autism and suggests that even if differing pathological processes are involved in the medial temporal lobe, there are similarities in other areas. In the light of such consistent results across a spectrum of ability, these areas may be good candidates for investigations searching for diagnostic neuropathology.

10.2 Generalisation of the Results to the Autistic Population

One very important question regarding the results from this thesis is the extent to which the findings can be generalised beyond the specific children in this study to the Autistic population. There are two main issues to address in answering this question: sex and age. These will be discussed in turn.

10.2.1 Sex

Due to a sex bias in the children from local schools who volunteered to act as controls, there was a large proportion of females in the control group. The reverse bias was present in the two Autistic groups. To what extent may this failure to match on sex restrict the generalisability of the results?

Unfortunately it is difficult to fully assess the effects of the failure to match on the basis of sex. Two types of analyses could provide some insights: including sex as a covariate in the analysis or restricting the analysis to include just boys. Both these techniques have limitations however. Inclusion of sex as a covariate may not accurately account for any differences. Restricting the analysis to include only boys reduces the power of the
statistical comparison (as it reduces the number of subjects). If the results remain significant when the analysis is restricted to the boys, however, this indicates that sex differences are not responsible for the observed difference. Results that are significant when sex is included as a covariate may also suggest that sex is not an important factor, even if the boys-only analysis result is only a trend.

It is more difficult to determine the impact of sex on non-significant results. A lack of sex matching may reduce the sensitivity of a comparison. However inclusion of only the data from the boys also reduces the sensitivity (since it reduces the power), so it is not possible to rule out a role of sex in these results. It should therefore be particularly emphasised that absence of significant abnormality does not imply normality.

When children are matched on intelligence, there is very little evidence to support sex differences in memory, motor co-ordination and executive function [Chelune and Baer 1986; Magalhaes et al. 1989; Ionescu 2000]. Further, when analyses were carried out with just the boys in the groups, the results remained significant. For example, the episodic memory deficit detected by the Rivermead Behavioural Memory Test remained significant when only the boys were included in the analysis.

There have been mixed results from studies investigating the influence of sex on emotional modulation of the startle response [Cook et al. 1995; Bradley et al. 1999; McManis et al. 2001]. However, in this study, even when only the boys' scores were analysed, or when sex was included as a covariate, the results remained unchanged. This suggests that a lack of sex matching can not account for the findings in these sections.

The results from the voxel-based morphometry were also not due to a failure to match on sex. As shown in Chapter 7, the results were the same when only the boys were included in the analysis. This is important in the light of sex differences found in the brain [Giedd et al. 1996b; Giedd et al. 1999]. The structural abnormalities in the brain are therefore unlikely to be accounted for poor sex matching.
Finally, the investigations of the ERP components found no significant abnormality in the Autistic groups. This was not altered by the inclusion of sex as a covariate, or restricting the sample to boys alone. Further, there is little evidence to suggest that ERP components are sensitive to sex differences in development [Martin et al. 1988].

In summary, therefore, there is little evidence to suggest that the significant abnormalities found in this thesis are attributable to sex differences between the control group and the Autistic groups. Furthermore, no evidence was found to suggest that the absence of significant differences on behavioural measures or on indices of neuropathology was associated with sex differences. However, difficulties in assessing such a possibility mean that this can not be ruled out on the basis of these data.

10.2.2 Age

All the participants in this study were children or adolescents. How applicable are the results to adults with Autism? It is possible that at least some of the results are a reflection of delayed development and that in time the abnormalities and deficits may resolve. The pattern of resolution would provide information about the ‘core’ deficits of Autism (i.e. those that do not resolve with time) and associated deficits. A definitive answer to this question is dependent on a replication of this study including adults with high functioning Autism. However, there are a number of lines of evidence to suggest that the abnormalities are unlikely to completely resolve.

Episodic memory skills improve with age [Fivush 1997] and so the deficit observed in the children with Autism could be a reflection of developmental delay. Such a delay would need to be relatively selective, as semantic memory and recognition memory appear to be preserved. This might be plausible as competence at recognition and non-episodic memory appears earlier in development than episodic memory skills (see Bauer 1997). The measure of episodic memory used in this thesis was standardised for age, so if resolution was likely to occur it might be expected that the biggest discrepancies in episodic memory would be at the younger ages studied. However this was not the case. Further Gardiner et al. have
reported episodic memory deficits in adults with Autism [Bowler et al. 2000] suggesting that the abnormality does not resolve with age.

Performance on the Wisconsin Card Sorting Test also improves with age [Chelune and Baer 1986] and so similarly the deficit observed in the High and Low groups could again be attributed to a developmental delay. However once again the nature of the delay would have to be very selective, as perseverative errors, unique errors and other errors were not significantly different from controls. Further, the pattern of development suggests that set loss shows only a minor developmental improvement, making delay an unlikely explanation [Chelune and Baer 1986].

Although deficits in motor co-ordination could be attributed to developmental delay, adults with Autism have also been reported to be clumsy and poorly co-ordinated [Tantam 1988b; Tantam 1988c]. Whilst the severity of the deficit is likely to reduce at least to some extent with age (and therefore practice), it seems that abnormality in co-ordination is likely to be a finding applicable to individuals with Autism of all ages.

The neuropathological findings of abnormality in the hippocampal formation, amygdala, orbitofrontal cortex, superior temporal gyrus and cerebellum may be associated with delayed development. The increased grey matter density in the amygdala (for example) in individuals with Autism may be a reflection of delayed (or interrupted or maladaptive) development. Cell death is an important mechanism in neural development (see Bredesen 2000); a reduced rate of cell death could be responsible for the noted abnormalities.

Although brain development is most dynamic in utero, modification of brain structure and function continues throughout childhood and adolescence [Giedd 1999]. For example, temporal lobe grey matter follows a non-linear developmental course with maximal size not being reached in controls until aged 16 years [Giedd et al. 1999]. Recent MR studies have suggested that the hippocampal formation and the amygdala continue to grow throughout adolescence [Giedd et al. 1996b]. The frontal lobes have a particularly prolonged maturation, with many connections developing during childhood [Huttenlocher 1979]. Whilst abnormalities in the cerebellum, amygdala and hippocampal formation have been
reported in adults with Autism [Abell et al. 1999; Howard et al. 2000; Saitoh et al. 2001], no previous studies have used the bilateral voxel-based morphometry method described in this thesis. The applicability of the findings to adults with Autism must therefore become a subject of future research.

Finally it should be noted that the findings of bilateral abnormality of the hippocampal formation and the orbitofrontal cortex do not support the hypothesis that abnormalities in functions associated with these areas will resolve in time. A lack of eloquent brain tissue in homologous regions is thought to prevent rescue of function (see Chapter 6). Further, delayed development rarely resolves completely. Often performance reaches a plateau and fails to improve beyond this level (see discussion in Vargha-Khadem et al. 1997b).

10.3 The Medial Temporal Lobe Hypothesis of Autism

The results from this study have characterised structural and functional deficits that are associated with specific areas in the medial temporal lobes in individuals with Autism, consistent with the hypothesis of medial temporal lobe dysfunction [Bachevalier 1994]. Other abnormalities were detected in areas with multiple connections to the medial temporal lobe. In light of these results, what is the status of the medial temporal lobe hypothesis?

The usefulness of a hypothesis regarding the neuropathological basis of Autism is determined by its universality, its primacy and its specificity. In addition any viable hypothesis should be compatible with established features of the disorder (such as increased incidence of Autism in males and the cognitive profiles leading to psychological theories). These issues will be addressed in turn.
10.3.1 Universality

Structural and functional abnormalities were detected in both the High and Low groups. Both groups showed an impairment in episodic memory and abnormality in the hippocampal formation (although the latter finding involved the use of different techniques for the two groups). Structural abnormality of the amygdala was also detected in both groups, although this was reflected as an increase in grey matter density in the High group and a decrease in the Low group. Further a correlation was found between the abnormality in the amygdala as detected by VBM and the children’s scores on the Autistic Behaviour Checklist.

Whilst it appears that both groups show evidence of medial temporal lobe abnormality, it has yet to be determined whether every individual shows the same pattern of abnormality. As shown in Chapter 3, at least one child in the High group obtained a perfect score on the Rivermead Behavioural Memory Test. Although this does not preclude some subtle functional deficit of the hippocampal formation that might be revealed by a more sensitive test, it questions the assumption of universality. Similarly the neuropathology results from both volumetric analysis and individual voxel-based morphometry suggest that not all individuals had detectable abnormalities compared to the controls. As before, these results, particularly those based on the voxel-based morphometry, may be simply a reflection of reduced sensitivity. Nonetheless, the results suggest that abnormalities of the medial temporal lobes (at least at the scale detected in this study) may not be universal in Autism.

10.3.2 Primacy

Are the medial temporal lobes the primary site of abnormality? This study can not provide direct evidence for or against such a hypothesis. However a number of observations may provide some insight.
Firstly, abnormalities were detected outside the medial temporal lobe. However, given the interdependent and dynamic nature of development, if the medial temporal lobes were the primary site of abnormality, then it might be expected that other areas (especially those with extensive connections to the medial temporal structures such as the orbitofrontal cortex) would develop abnormally due to aberrant inputs from the medial temporal lobes.

Secondly, if the medial temporal lobes were the primary site of abnormality, it might be expected that these areas would show more significant abnormality than other areas of the brain. However, in the individual voxel-based morphometric analyses, the only consistent abnormality found in all the individuals in the High group was in the orbitofrontal cortex. Whilst the reduced power of these analyses limits the sensitivity, this may indicate that the orbitofrontal cortex is the primary site of abnormality and the medial temporal lobe abnormality is a consequence of this.

10.3.3 Specificity

Does developmental abnormality of the medial temporal lobes always result in Autism? The results of this study can only address this question in a very limited manner. The cognitive impairments found in the Autistic groups in this study in isolation are certainly not specific to Autism. Children with developmental amnesia due to hippocampal atrophy have severe episodic memory deficits [Gadian et al. 2000], and children with frontal lobe damage show deficits on the Wisconsin Card Sorting Test [Vargha-Khadem et al. 2000]. It is possible that the combination of the two deficits (i.e. impairment in episodic memory and the difficulty with maintenance of set) may be specific to Autism. Studies of other children with developmental disorders are required to address this question.

Abnormality in the hippocampal formation, amygdala, orbitofrontal cortex, superior temporal gyrus and cerebellum in isolation are unlikely to be specific to Autism. For example, isolated developmental abnormality of the hippocampal formation appears to result in a different behavioural and cognitive profile, namely developmental amnesia [Vargha-Khadem et al. 1997a]. It is possible that the co-occurrence of neuropathology in a
number of other regions is specific to Autism, but once again further studies of other
developmental disorders are required to address this issue adequately.

A number of other studies have now reported children with developmental medial temporal
lobe abnormality who do not display Autistic symptoms (e.g. Maurer 1992; Isaacs et al.
2000). Indeed Maurer described a child with extensive congenital bilateral lesions of the
medial temporal lobes who was not Autistic. Do these cases invalidate the medial temporal
lobe hypothesis of Autism?

Rather than invalidating the hypothesis, these cases perhaps lead to a refinement of the
hypothesis. The preterm children and children with developmental amnesia show
decreased hippocampal formation volume and no apparent abnormality of the amygdala. In
comparison, the results from this study suggest increased hippocampal volumes combined
with amygdala abnormality. The damage to the medial temporal lobes in Maurer’s case
was thought to result from a stroke in utero. The timing of the stroke may have been
different from the timing of the abnormality in Autism, which due to the strong genetic
indicators may find its origins during embryonic development. This would lead to different
consequential or ‘knock-on’ abnormality from that found in congenital or perinatally
acquired disorders. Additionally, the syndrome of Autism may only be observed when the
abnormality of the medial temporal lobes is restricted to the amygdala and hippocampal
formation (for example) and not following more extensive abnormality.

10.3.4 Increased incidence of Autism in Males

Can the theory of the medial temporal lobes explain the higher occurrence of Autism in
boys (typical ratios are 3:1 or 4:1 males : females) [Volkmar et al. 1993]? Sex differences
in the incidence of Autism potentially point to very important information regarding the
underlying causes of Autism.

Sexual dimorphism in the medial temporal lobes has been noted. Hormone receptors in the
medial temporal lobe are distributed non-uniformly, with the amygdala having a
predominance of androgen receptors and the hippocampal formation a predominance of oestrogen receptors [Morse et al. 1986; Clark et al. 1988; Sholl and Kim 1989]. It is not implausible that abnormality in these areas may have sex specific consequences and thus contribute to sex bias observed in Autism. Male and female rats show a different magnitude of response to neonatal limbic damage. Specifically, males show a greater reduction than females in volume of forebrain structures and a greater increase in lateral ventricle volumes after neonatal damage [Hanlon and Sutherland 2000].

It should also be noted that the increased incidence of Autism in males is not restricted to this disorder alone. Males are more frequently affected than females by almost all behavioural disorders, mental handicaps and learning problems [Rutter et al. 1970; Eme 1979; Finucci and Childs 1981]. It is possible that male development may therefore have greater vulnerability or sensitivity to environmental or genetic influences involved in neurodevelopmental disorders. As such it may not be necessary for a neuropathological theory of Autism to be able to completely explain the genetic bias.

10.3.5 Cognitive Profiles and Psychological Theories of Autism

This thesis attempted to explain Autism in terms of a neurobiological theory (namely the medial temporal lobe theory of Autism). In addition to other neurobiological explanations, a number of psychological explanations have been hypothesised in the literature. These argue that a single cognitive deficit is responsible for the symptomatology of Autism.

Perhaps the most widely known psychological hypothesis of Autism is the Theory of Mind. Theory of Mind is a generic term that refers to the ability to think about a range of mental states (intentions, desires, thoughts, beliefs, dreams, pretence, etc.) [Baron-Cohen et al. 1994]. Individuals with Autism have been shown to be impaired on a range of Theory of Mind tasks [Baron-Cohen et al. 1985b; Loveland 1991; Critchley et al. 2000]. It has been argued that this deficit underlies their social and communicative difficulties.
Whilst the Theory of Mind hypothesis has increased our understanding of the nature of the social deficits characteristic of Autism, the specificity, universality and primacy of the hypothesis has come into question. Firstly a number of patient populations with non-autistic disorders have been shown to perform poorly on Theory of Mind tasks, including deaf individuals, and individuals with mental retardation, ADHD or developmental language disorders [Benson et al. 1993; Shields et al. 1996; Buitelaar et al. 1999b]. Secondly, a proportion of individuals with Autism succeed on Theory of Mind tasks. Whilst increasing the complexity of the tasks reduces this proportion, some of these individuals are capable of showing some evidence of mentalising outside the laboratory [Frith et al. 1994].

If deficits in Theory of Mind are responsible for the symptomatology of Autism, abnormal social behaviours should start to emerge in children with Autism at the age when normally developing children start to develop Theory of Mind skills. However, this hypothesis is not supported (at least with some children with Autism). Klin et al. found deficits in very basic and early emerging socially adaptive behaviours (such as reaching for a familiar person) [Klin et al. 1992]. Also the Theory of Mind hypothesis provides no adequate explanation for some of the non-social characteristics of Autism (such as stereotypies, insistence on sameness and executive function deficits [Hughes et al. 1994]).

Integrating the Theory of Mind hypothesis into the findings of this thesis is far from straightforward. The neuroanatomical basis of Theory of Mind is unknown. Imaging studies have implicated the anterior cingulate, medial prefrontal cortex and the superior temporal gyrus in the performance of Theory of Mind tasks [Fletcher et al. 1995; Happé et al. 1996], with individuals with Autism failing to show medial prefrontal activation when engaged in the task. The neuropathology of Autism as found in this study does not include this area of the frontal lobes.

The cognitive profile of deficits in episodic memory would be consistent with the predicted impairment in Theory of Mind. Children develop episodic memory as they develop a Theory of Mind [Perner 1990]. It has been argued that deficits in Theory of Mind result in
a lack of experiential awareness and therefore in deficits in episodic memory [Tager-Flusberg 1991b].

A second, more recent psychological account of Autism is weak central coherence. A characteristic of normal information processing appears to be the tendency to draw together diverse information to construct higher-level meaning in context. Frith refers to this as 'central coherence'. It has been argued that disturbance of central coherence would parsimoniously explain the troughs and peaks of cognition in Autism [Frith 1989; Happé 1999]. Consistent with this hypothesis, a number of studies report unusual attention to detail and difficulty with global processing (e.g. Shah and Frith 1983). Although originally suggested to account for Theory of Mind deficits, weak central coherence is now argued to be a second cognitive characteristic of Autism [Happé 1994].

Yet the neurobiological underpinning of weak central coherence is also unknown, and consequently reconciliation between this theory and the findings of this thesis is difficult. However, weak central coherence suggests that the children with Autism have difficulty utilising context appropriately. The hippocampal formation has been associated with appropriate use of context, not just in terms of episodic memory. Animals with hippocampal lesions have abnormally high sensitivity to irrelevant contexts and abnormally low sensitivity to relevant contexts [Winocur and Olds 1978; Winocur and Gilbert 1984; Winocur et al. 1987; Good and Honey 1991]. It is plausible therefore that medial temporal lobe abnormality may be associated with some of the behavioural characteristics associated with weak central coherence.

A third psychological theory of Autism is that deficits in executive function account for the syndrome of Autism (e.g. Hughes et al. 1994). There are a number of reasons to believe that this explanation of Autism is inadequate. As discussed in Chapter 5, deficits in executive function have been noted in other neurodevelopmental psychiatric disorders including ADHD [Chelune et al. 1986], conduct disorder [Lueger and Gill 1990] and Tourette's syndrome [Incagnoli and Kane 1981]. Deficits in executive function therefore do not necessarily result in social impairment. Lesions to the frontal lobes in childhood have been found to result in severe deficits in executive function. Whilst these children do
have difficulties with appropriate social behaviour, they are of a qualitatively different nature. Unlike children with Autism, children with developmental frontal lobe pathology appear to understand that others have feelings and simply have no desire to alter their behaviour to accommodate these feelings [Anderson et al. 2000].

10.4 Future Studies and Directions

In order to extend the findings of this thesis to the general population of individuals with Autism, a similar study should be undertaken with adults. The sensitivity of this study was limited by a lack of sex matching and the broad range of ages included. Future studies should match on sex and might be restricted to a more selective age range to minimise the confounding factor of broad performance scores in control groups due to developmental differences. Additionally future studies should consider use of an additional control group matched on verbal ability to individuals with below average verbal IQ. However the difficulties of selecting an appropriate group have already been discussed (see above).

The results from many of the experimental tasks included in this study raise many more questions than they answer. For example, the emotional modulation of the startle reflex is an elegant paradigm that offers the potential for an objective method suitable for children of all abilities. Studies of the modulation of the startle response in normally developing children are needed to help interpret results, and future studies of Autism should address the use of different stimuli and experimental paradigms. Similarly, the paradoxical ERP results (in particular, the apparent increase in the ‘novelty’ P300 response to repeated novels) suggest that pursuit of this area may be important in the search for objective measures of the children’s difficulties and differences.

In addition to replications of the paradigms used in this study, a number of other techniques may offer valuable insights into the neurobiology of Autism. Functional imaging offers a unique insight into the correlates of neural functioning and has already been used in a number of studies of individuals with Autism performing Theory of Mind tasks (e.g. Happé et al. 1996; Baron-Cohen et al. 1999). Use of paradigms aimed at assessing the function
Discussion and Future Directions

of the medial temporal lobes (such as episodic memory (e.g. Maguire et al. 2000b) or facial expression (e.g. Whalen et al. 1998) would further clarify the nature of any dysfunction.

Whilst this study demonstrated a relatively selective episodic memory deficit in individuals with Autism using the Rivermead Behavioural Memory Test, the nature of such a deficit should be further explored, perhaps using a remember-know paradigm. Use of such a paradigm with children may prove difficult however, as pilot studies at Great Ormond Street Hospital have suggested that children may find understanding the distinction between remember and know difficult to understand. Familiarity and recollection have recently been shown to have unique ERP ‘signatures’ [Duzel et al. 2001]. This method may prove a more feasible way of exploring the memory deficit further. In addition, in the light of the evidence that memory for emotional material is impaired in individuals with amygdala damage, it would be interesting to explore whether children with Autism show a more severe deficit for emotional material than for non-emotional material.

In this thesis, in the light of the extensive literature on the subject, it was assumed that the children with Autism would show deficits in the processing of facial expressions. Such deficits are widely reported in the literature, so such an assumption is appropriate. However, future studies addressing the abnormality of the amygdala should include such an assessment to gather together cognitive evidence confirming amygdala dysfunction.

Future neuropsychological assessments should also include an extensive battery of executive tasks, firstly to establish how selective the deficit really is and then to characterise further the nature of the difficulty related to the orbitofrontal cortex abnormality. Additionally, inclusion of a gambling task such as that described by Damasio et al. (see Bechara et al. 1996) might provide a useful insight into the similarities and differences between the individuals with Autism and those with large ventromedial lesions. Furthermore, behavioural correlates of abnormality in the superior temporal gyrus should be investigated (e.g. detection of gaze [Campbell et al. 1990] and pitch processing [Scott et al. 2000]). This could provide further convergent evidence of abnormality in this region.
Further, a more direct and less crude method of assessing motor co-ordination should be included. Asking relatives to fill in questionnaires should reflect fairly accurately the difficulties the individuals have, but the scales may mean different things to different people and such assessments are rarely totally objective.

The focus of this study led to extensive MR investigations of the amygdala and hippocampal formation. However the abnormalities in the orbitofrontal cortex, superior temporal gyrus, and cerebellum have not been comprehensively investigated. Volumetric measures of the cerebellum and the orbitofrontal cortex are possible (e.g. Courchesne, et al. 1988; Crespo-Facorro et al. 1999). Both these techniques should be used to further explore the abnormalities found in this study. In particular volumetric analysis of the cerebellum would enable comparisons with the work of Courchesne et al.

The motivation for the research in this thesis was to assess the validity of the medial temporal lobe hypothesis of Autism. However, the results suggest that the neuropathology associated with Autism affects multiple neural systems. Attempting to relate Autism to one neuroanatomical site may be therefore incorrect. Indeed, it may be that it is the coincident abnormalities in the hippocampal formation, amygdala, orbitofrontal cortex, superior temporal gyrus and the cerebellum that is unique to this syndrome. In the light of these results, further investigation into this possibility is warranted.
Appendix A  Diagnostic Criteria: DSM-IV

A.1 Autistic Disorder [DSM-IV 1994]

The child must meet all three (A-C) of the criteria below.

A:  A total of six or more items from (1), (2), and (3), with at least two from (1) and one each from (2) and (3):

1. Qualitative impairment in social interaction as manifested by at least two of the following:
   a) marked impairment in the use of multiple non-verbal behaviours such as eye to eye gaze, facial expression, body postures, and gestures to regulate social interaction
   b) failure to develop peer relationships appropriate to developmental level
   c) a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people. For example by a lack of showing, bringing or pointing out objects of interest
   d) lack of social or emotional reciprocity

2. Qualitative impairments in communication as manifested by at least one of the following:
   a) delay in, or total lack of the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)
   b) in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
   c) stereotyped and repetitive use of language or idiosyncratic language
   d) lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level

3. Restricted repetitive and stereotyped patterns of behaviour, interests and activities, as manifested by at least one of the following:
Appendix A: Diagnostic Criteria

a) encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
b) apparently inflexible adherence to specific non-functional routines or rituals
c) stereotyped and repetitive motor mannerisms (e.g. hand or finger flapping or twisting, or complex whole body movements)
d) persistent preoccupation with parts of objects

B: Delays or abnormal functioning in at least one of the following areas, with onset prior to age 3 years, with onset prior to three years of age:

(1) Social interaction
(2) Language as used in social communication
(3) Symbolic or imaginative play

C: The disturbance is not better accounted for by Rett’s Disorder or Child Disintegrative Disorder

A.2 Asperger’s Syndrome [DSM-IV 1994]

The presence of a marked abnormal or impaired development in social interaction and communication and a markedly restricted repertoire of activity and interests.

Places autistic disorders in the category pervasive developmental disorders (which are disorders characterised by severe impairments in more than one area of development).

A Qualitative impairment in social interaction, as manifested by at least two of the following:
1. marked impairment in the use of multiple non-verbal behaviours such as eye to eye gaze, facial expression, body postures and gestures to regulate social interaction
2. failure to develop peer relationships appropriate to developmental level
3. a lack of spontaneous seeking to share enjoyment, interests or achievements with other people (e.g. by a lack of showing, bringing or pointing out objects of interest to other people)
4. lack of social or emotional reciprocity
Appendix A: Diagnostic Criteria

B Restricted repetitive and stereotyped patterns of behaviour, interests and activities, as manifested by at least one of the following:
1. encompassing preoccupation with one or more stereotypes and restricted patterns of interest that is abnormal either in intensity or focus
2. apparently inflexible adherence to specific, non-functional routines or rituals
3. stereotyped and repetitive motor mannerisms (e.g. hand or finger flapping or twisting or complex whole body movements)
4. persistent preoccupation with parts or objects

C The disturbance causes clinically significant impairment in social, occupational or other important areas of functioning.

D There is no clinically significant general delay in language (e.g. single words used by age 2 years, communicative phrases used by 3 years)

E There is no clinically significant delay in cognitive development or in the development of age-appropriate self-help skills, adaptive behaviour (except in social interaction) and curiosity about the environment in childhood

F Criteria not met for another specific Pervasive Developmental Disorder or Schizophrenia
### Appendix B  Missing Data

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Child (Group)</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selective Attn</td>
<td>STPH (L)</td>
<td>Failed to score enough on motor control</td>
</tr>
<tr>
<td></td>
<td>SOFA (L)</td>
<td>Failed to understand instructions</td>
</tr>
<tr>
<td>Divided Attn</td>
<td>CRWA (H)</td>
<td>Failed to follow instructions</td>
</tr>
<tr>
<td></td>
<td>SOFA (L)</td>
<td>Failed to understand instructions</td>
</tr>
<tr>
<td>CCC</td>
<td>AMCR (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>GESC (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td>ABC</td>
<td>AMCR (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>GESC (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>TICH (H)</td>
<td>Incomplete</td>
</tr>
<tr>
<td>MABC</td>
<td>AMCR (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>GESC (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>BETR (H)</td>
<td>Parents felt unable to judge</td>
</tr>
<tr>
<td></td>
<td>JONA (H)</td>
<td>Parents felt unable to judge</td>
</tr>
<tr>
<td></td>
<td>CLDA (L)</td>
<td>Parents felt unable to judge</td>
</tr>
<tr>
<td>ChOCI</td>
<td>AMCR (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>GESC (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>CLDA (L)</td>
<td>Parents felt unable to judge</td>
</tr>
<tr>
<td>Sunderland Questionnaire</td>
<td>AMCR (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>GESC (C)</td>
<td>Not returned by parents</td>
</tr>
<tr>
<td></td>
<td>JONA (H)</td>
<td>Parents felt unable to judge</td>
</tr>
<tr>
<td>WCST</td>
<td>JOWE (H)</td>
<td>Became distressed</td>
</tr>
<tr>
<td></td>
<td>CLDA (L)</td>
<td>Threw cards over floor</td>
</tr>
<tr>
<td></td>
<td>STPH (L)</td>
<td>Became distressed</td>
</tr>
<tr>
<td>Stroop</td>
<td>BHTE (C)</td>
<td>Mildly colour blind</td>
</tr>
<tr>
<td>OFC Reversal</td>
<td>SETH (L)</td>
<td>Computer error</td>
</tr>
<tr>
<td>OFC Extinction</td>
<td>BHTE (C)</td>
<td>Experimenter ill</td>
</tr>
<tr>
<td>Spatial working memory</td>
<td>BHTE (C)</td>
<td>Experimenter ill</td>
</tr>
</tbody>
</table>
### Appendix B: Missing Data

<table>
<thead>
<tr>
<th>Test Type</th>
<th>Subject(s)</th>
<th>Reason for Missing Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERP (N4)</td>
<td>ALMI (C)</td>
<td>Did not wish to participate</td>
</tr>
<tr>
<td></td>
<td>SOFA (L)</td>
<td>Distressed</td>
</tr>
<tr>
<td>ERP (P300)</td>
<td>ALMI (C)</td>
<td>Did not wish to participate</td>
</tr>
<tr>
<td>Dot location (Immediate and delay)</td>
<td>AMCR (C)</td>
<td>Not part of adult WMS</td>
</tr>
<tr>
<td></td>
<td>JONA (H)</td>
<td>Not part of adult WMS</td>
</tr>
<tr>
<td></td>
<td>STPH (L)</td>
<td>Administration error</td>
</tr>
<tr>
<td>Word pair leaning</td>
<td>AMCR (C)</td>
<td>Not part of adult WMS</td>
</tr>
<tr>
<td></td>
<td>JONA (H)</td>
<td>Not part of adult WMS</td>
</tr>
<tr>
<td>Names (D &amp; P)</td>
<td>CHAU (L)</td>
<td>Unable to read names with confidence</td>
</tr>
<tr>
<td></td>
<td>SAAS (L)</td>
<td>Unable to read names with confidence</td>
</tr>
<tr>
<td></td>
<td>TOWA (L)</td>
<td>Unable to read names with confidence</td>
</tr>
<tr>
<td>All MR protocol</td>
<td>ALMI (C)</td>
<td>Did not wish to participate</td>
</tr>
<tr>
<td></td>
<td>GESC (C)</td>
<td>Unable to attend last afternoon</td>
</tr>
<tr>
<td></td>
<td>ROHA (L)</td>
<td>Shoulder breadth too large for scanner</td>
</tr>
<tr>
<td></td>
<td>SAAS (L)</td>
<td>Unable to tolerate scanner environment</td>
</tr>
<tr>
<td></td>
<td>SOFA (L)</td>
<td>Unable to tolerate scanner environment</td>
</tr>
<tr>
<td>T2 Hippocampal Maps</td>
<td>ALSU (C)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>LEAL (C)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>VILO (C)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>CRWA (H)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>MAWA (H)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>RONI (H)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>CHCO (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>CLDA (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>SARU (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>T2 Amygdala Maps</td>
<td>ALSU (C)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>JEWI (C)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td></td>
<td>LEAL (C)</td>
<td>CSF artefact prevented measurement (on right)</td>
</tr>
<tr>
<td></td>
<td>PHWH (C)</td>
<td>Artefact through amygdala</td>
</tr>
<tr>
<td></td>
<td>CRWA (H)</td>
<td>Images too blurred</td>
</tr>
</tbody>
</table>

287
<table>
<thead>
<tr>
<th>Category</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAWA (H)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>RONI (H)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>CHCO (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>CHAU (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>CLDA (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>FRSN (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>SARU (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>STPH (L)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>Volumes ALSU (C)</td>
<td>Images too blurred</td>
</tr>
<tr>
<td>ERP (P300) ALMI (C)</td>
<td>Did not wish to participate</td>
</tr>
<tr>
<td>ERP (N400) ALMI (C)</td>
<td>Did not wish to participate</td>
</tr>
<tr>
<td>SOFA (L)</td>
<td>Computer malfunction</td>
</tr>
<tr>
<td>Startle ALMI (C)</td>
<td>Did not wish to participate</td>
</tr>
<tr>
<td>SOFA (L)</td>
<td>Computer malfunction</td>
</tr>
<tr>
<td>Magnitude of Startle AMCR (C)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>AACA (H)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>JOWE (H)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>ROHE (H)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>Latency of Startle AMCR (C)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>LEAL (C)</td>
<td>No non-zero responses to one category</td>
</tr>
<tr>
<td>AACA (H)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>JOWE (H)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>ROHE (H)</td>
<td>No valid responses to ≥ 1 category</td>
</tr>
<tr>
<td>Number of pictures 'Scary' TICH (H)</td>
<td>Score sheet mislaid</td>
</tr>
<tr>
<td>Picture Metaphor (RHBT) ALSU (C)</td>
<td>Administration error</td>
</tr>
<tr>
<td>Inference (RHBT) SOFA (L)</td>
<td>Distressed</td>
</tr>
<tr>
<td>Joke (RHBT) SOFA (L)</td>
<td>Distressed</td>
</tr>
<tr>
<td>Visual Paired Comparison DIKI (C)</td>
<td>Administration error (on Same Delay cond’n)</td>
</tr>
<tr>
<td></td>
<td>Contrast on video too low</td>
</tr>
<tr>
<td></td>
<td>JEWI (C)</td>
</tr>
<tr>
<td></td>
<td>Contrast on video too low</td>
</tr>
</tbody>
</table>
Appendix B: Missing Data

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHWH (C)</td>
<td>Contrast on video too low</td>
</tr>
<tr>
<td>JOCO (H)</td>
<td>Contrast on video too low</td>
</tr>
<tr>
<td>MAWA (H)</td>
<td>Contrast on video too low</td>
</tr>
<tr>
<td>CLDA (L)</td>
<td>Contrast on video too low</td>
</tr>
</tbody>
</table>
Appendix C  Statistical Results

Since a large number of statistical tests are performed in this thesis, significant results may capitalise on chance and the overall probability of a Type I error may exceed 5%. However the primary hypotheses of this thesis predict a specific pattern of results across a range of tasks, which includes both impaired and intact performances. Setting the acceptable alpha too low would reduce the power of detecting a group difference on the tasks for which intact performance is predicted. To lower the probability of capitalising on chance and to reduce the number of statistical comparisons, specific analyses were planned a priori.

Statistical analyses (except where otherwise stated) were carried out using an analysis of variance model with appropriate checks for the normal distribution of the data. Planned comparisons were carried out using a Student’s t test comparing the Control group with the High group and the High group with the Low group. The comparison of the Control group versus the Low group was not carried out, as this is dependent on the results of the previous two comparisons. The homogeneity of variance was measured using Levene’s test, and when violated, results are reported without assumptions for equal variance. This correction results in non-integral degrees of freedom.

When appropriate, correlations of potentially confounding factors (such as age, verbal IQ, etc.) were investigated and if significant, analyses of covariance were carried out. Statistical analysis when normality assumptions were violated was carried out using the Kruskal-Wallis test.

All statistics are quoted to the nearest integer or one significant figure (with the exception of non integral degrees of freedom, which are quoted to one decimal place to indicate use of correction).

It should be emphasised that statistical analysis techniques used in voxel-based morphometry are unique to this method and details are given in Chapters 6 and 7.
<table>
<thead>
<tr>
<th>Chapter Test</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>CvHt</th>
<th>CvHp</th>
<th>HvLt</th>
<th>HvLp</th>
<th>VIQ corr r</th>
<th>VIQ p</th>
<th>Age corr r</th>
<th>Age p</th>
<th>ANCOVA F</th>
<th>ANCOVA df</th>
<th>ANCOVA p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 ABC</td>
<td>34</td>
<td>2</td>
<td>0.38</td>
<td>-0.001</td>
<td>&lt;0.001</td>
<td>-2</td>
<td>0.1</td>
<td>-0.5</td>
<td>&lt;0.01</td>
<td>-0.04</td>
<td>ns</td>
<td>22</td>
<td>2, 37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2 BAS</td>
<td>8</td>
<td>2</td>
<td>0.41</td>
<td>0.001</td>
<td>1</td>
<td>0.2</td>
<td>2</td>
<td>0.02</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.4</td>
<td>-0.01</td>
<td>1, 39</td>
<td>0.3</td>
</tr>
<tr>
<td>2 BPVS</td>
<td>10</td>
<td>2</td>
<td>0.41</td>
<td>0.001</td>
<td>-0.009</td>
<td>0.9</td>
<td>4</td>
<td>&lt;0.001</td>
<td>0.8</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>0.2, 40</td>
<td>0.9</td>
</tr>
<tr>
<td>2 CCC</td>
<td>58</td>
<td>2</td>
<td>0.39</td>
<td>0.001</td>
<td>-0.001</td>
<td>-2</td>
<td>0.03</td>
<td>-0.5</td>
<td>&lt;0.01</td>
<td>0.001</td>
<td>ns</td>
<td>37</td>
<td>2, 38</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2 ChOCI Compulsions</td>
<td>14</td>
<td>2</td>
<td>0.38</td>
<td>0.001</td>
<td>-0.001</td>
<td>-4</td>
<td>0.3</td>
<td>-0.4</td>
<td>&lt;0.05</td>
<td>0.09</td>
<td>ns</td>
<td>9</td>
<td>2, 37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2 ChOCI Obsessions</td>
<td>8</td>
<td>2</td>
<td>0.38</td>
<td>0.001</td>
<td>-0.001</td>
<td>-4</td>
<td>0.3</td>
<td>-0.2</td>
<td>&lt;0.05</td>
<td>0.2</td>
<td>ns</td>
<td>9</td>
<td>2, 37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2 Dewey (correct norm)</td>
<td>0.7</td>
<td>2</td>
<td>0.39</td>
<td>0.5</td>
<td>0.4</td>
<td>0.7</td>
<td>0.8</td>
<td>0.4</td>
<td>ns</td>
<td>0.8</td>
<td>ns</td>
<td>2</td>
<td>2, 38</td>
<td>0.4</td>
</tr>
<tr>
<td>2 Dewey (correct viol)</td>
<td>6</td>
<td>2</td>
<td>0.39</td>
<td>0.004</td>
<td>1</td>
<td>0.3</td>
<td>2</td>
<td>0.07</td>
<td>0.4</td>
<td>&lt;0.01</td>
<td>0.3</td>
<td>ns</td>
<td>2, 38</td>
<td>0.4</td>
</tr>
<tr>
<td>2 Divided attention (TEACh)</td>
<td>13</td>
<td>2</td>
<td>0.39</td>
<td>0.001</td>
<td>2</td>
<td>0.06</td>
<td>3</td>
<td>0.004</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>-0.2</td>
<td>ns</td>
<td>0.2, 38</td>
<td>0.3</td>
</tr>
<tr>
<td>2 EFT (Number correct)</td>
<td>7</td>
<td>2</td>
<td>0.41</td>
<td>0.003</td>
<td>0.2</td>
<td>0.9</td>
<td>2</td>
<td>0.004</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.2</td>
<td>ns</td>
<td>0.8, 40</td>
<td>0.4</td>
</tr>
<tr>
<td>2 EFT (Time)</td>
<td>2</td>
<td>2</td>
<td>0.39</td>
<td>0.004</td>
<td>0.1</td>
<td>-2</td>
<td>0.1</td>
<td>-0.6</td>
<td>0.5</td>
<td>0.004</td>
<td>0.007</td>
<td>0.3</td>
<td>2, 38</td>
<td>0.4</td>
</tr>
<tr>
<td>2 Inference (RHB)</td>
<td>11</td>
<td>2</td>
<td>0.40</td>
<td>0.001</td>
<td>2</td>
<td>0.07</td>
<td>3</td>
<td>0.007</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.2</td>
<td>ns</td>
<td>3, 38</td>
<td>0.06</td>
</tr>
<tr>
<td>2 Joke (RHB)</td>
<td>7</td>
<td>2</td>
<td>0.40</td>
<td>0.001</td>
<td>0.002</td>
<td>0.08</td>
<td>0.9</td>
<td>0.002</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.1, 38</td>
<td>0.9</td>
</tr>
<tr>
<td>2 M-ABC</td>
<td>20</td>
<td>2</td>
<td>0.39</td>
<td>0.001</td>
<td>0.004</td>
<td>-4</td>
<td>0.002</td>
<td>-2</td>
<td>0.04</td>
<td>0.004</td>
<td>0.007</td>
<td>0.3</td>
<td>0.4, 39</td>
<td>0.7</td>
</tr>
<tr>
<td>2 Picture Metaphor (RHB)</td>
<td>4</td>
<td>2</td>
<td>0.40</td>
<td>0.002</td>
<td>0.02</td>
<td>-0.6</td>
<td>0.6</td>
<td>3</td>
<td>0.001</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.09</td>
<td>0.4, 39</td>
<td>0.7</td>
</tr>
<tr>
<td>2 Selective attention (TEACh)</td>
<td>2</td>
<td>2</td>
<td>0.39</td>
<td>0.2</td>
<td>0.5</td>
<td>0.6</td>
<td>1.2</td>
<td>0.2</td>
<td>0.7</td>
<td>0.01</td>
<td>0.5</td>
<td>0.2</td>
<td>0.4, 01</td>
<td>0.1</td>
</tr>
<tr>
<td>2 Sustained attention (TEACh)</td>
<td>6</td>
<td>2</td>
<td>0.41</td>
<td>0.008</td>
<td>3</td>
<td>0.005</td>
<td>-0.3</td>
<td>0.8</td>
<td>0.9</td>
<td>0.002</td>
<td>0.002</td>
<td>0.5</td>
<td>0.4, 01</td>
<td>0.1</td>
</tr>
<tr>
<td>2 Written Metaphor (RHB)</td>
<td>8</td>
<td>2</td>
<td>0.41</td>
<td>0.002</td>
<td>1</td>
<td>0.2</td>
<td>3</td>
<td>0.02</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>1, 39</td>
<td>0.3</td>
</tr>
<tr>
<td>3 Doors</td>
<td>4</td>
<td>2</td>
<td>0.41</td>
<td>0.002</td>
<td>2</td>
<td>0.1</td>
<td>1</td>
<td>0.3</td>
<td>0.4</td>
<td>&lt;0.005</td>
<td>0.4</td>
<td>&lt;0.01</td>
<td>2, 39</td>
<td>0.1</td>
</tr>
<tr>
<td>3 Dot location delay</td>
<td>0.5</td>
<td>2</td>
<td>0.40</td>
<td>0.6</td>
<td>-0.3</td>
<td>0.8</td>
<td>1</td>
<td>0.3</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>3 Dot location learning</td>
<td>2</td>
<td>2</td>
<td>0.40</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.8</td>
<td>0.8</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>3 Face Del</td>
<td>2</td>
<td>2</td>
<td>0.40</td>
<td>0.1</td>
<td>0.2</td>
<td>0.9</td>
<td>0.4</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>3 Face Imm</td>
<td>2</td>
<td>2</td>
<td>0.40</td>
<td>0.1</td>
<td>0.2</td>
<td>0.8</td>
<td>0.8</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>3 Family Pictures Del</td>
<td>5</td>
<td>2</td>
<td>0.41</td>
<td>0.001</td>
<td>3</td>
<td>0.01</td>
<td>0.2</td>
<td>0.8</td>
<td>0.2</td>
<td>ns</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>3 Family Pictures Imm</td>
<td>4</td>
<td>2</td>
<td>0.41</td>
<td>0.002</td>
<td>2</td>
<td>0.04</td>
<td>0.7</td>
<td>0.5</td>
<td>0.1</td>
<td>ns</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>3 Information</td>
<td>26</td>
<td>2</td>
<td>0.41</td>
<td>&lt;0.001</td>
<td>0.1</td>
<td>0.9</td>
<td>6</td>
<td>&lt;0.001</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Chapter Test</td>
<td>F</td>
<td>df</td>
<td>p</td>
<td>CvH t</td>
<td>CvH p</td>
<td>HvL t</td>
<td>HvL p</td>
<td>VIQ corr r</td>
<td>VIQ p</td>
<td>Age corr r</td>
<td>Age p</td>
<td>ANCOVA F</td>
<td>ANCOVA df</td>
<td>ANCOVA p</td>
</tr>
<tr>
<td>--------------</td>
<td>---</td>
<td>----</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------------</td>
<td>------</td>
<td>------------</td>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>3 Names</td>
<td>72,41</td>
<td>0.003</td>
<td>0.2</td>
<td>0.9</td>
<td>3</td>
<td>0.003</td>
<td>0.3</td>
<td>ns</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>52,37</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Names set 1</td>
<td>102,38</td>
<td>&lt;0.001</td>
<td>2</td>
<td>0.1</td>
<td>2</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Names set 2</td>
<td>12,38</td>
<td>0.3</td>
<td>-0.8</td>
<td>0.5</td>
<td>2</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 People</td>
<td>22,41</td>
<td>0.1</td>
<td>0.3</td>
<td>0.7</td>
<td>2</td>
<td>0.1</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Pyramids and Palmtrees</td>
<td>92,41</td>
<td>0.001</td>
<td>-0.2</td>
<td>0.8</td>
<td>3</td>
<td>0.009</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>0.4</td>
<td>&lt;0.05</td>
<td>0.3</td>
<td>2,39</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>3 Rivermead</td>
<td>172,41</td>
<td>&lt;0.001</td>
<td>4</td>
<td>&lt;0.001</td>
<td>2</td>
<td>0.1</td>
<td>0.6</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>72,40</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Shape</td>
<td>52,41</td>
<td>0.02</td>
<td>0.9</td>
<td>0.4</td>
<td>2</td>
<td>0.1</td>
<td>0.3</td>
<td>ns</td>
<td>0.2</td>
<td>ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Similarieties</td>
<td>72,41</td>
<td>0.002</td>
<td>-1</td>
<td>0.2</td>
<td>4</td>
<td>0.001</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Stories Delay</td>
<td>92,41</td>
<td>0.001</td>
<td>1</td>
<td>0.2</td>
<td>3</td>
<td>0.009</td>
<td>0.7</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>0.72,40</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Stories Immed</td>
<td>42,41</td>
<td>0.02</td>
<td>0.8</td>
<td>0.4</td>
<td>2</td>
<td>0.5</td>
<td>0.6</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>0.062,40</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Story Del Recog</td>
<td>62,41</td>
<td>0.004</td>
<td>0.9</td>
<td>0.4</td>
<td>3</td>
<td>0.01</td>
<td>0.6</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>0.22,40</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Sunderland</td>
<td>122,38</td>
<td>&lt;0.001</td>
<td>-4</td>
<td>0.002</td>
<td>-0.7</td>
<td>0.5</td>
<td>-0.4</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>72,37</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Vocabulary</td>
<td>182,41</td>
<td>&lt;0.001</td>
<td>-0.07</td>
<td>0.9</td>
<td>5</td>
<td>&lt;0.001</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Word List Del Recog</td>
<td>82,41</td>
<td>0.001</td>
<td>0.2</td>
<td>0.9</td>
<td>3</td>
<td>0.002</td>
<td>0.4</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>52,40</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Word list delay</td>
<td>82,41</td>
<td>0.001</td>
<td>1</td>
<td>0.2</td>
<td>3</td>
<td>0.01</td>
<td>0.4</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>42,40</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Word list learning</td>
<td>62,41</td>
<td>0.007</td>
<td>2</td>
<td>0.04</td>
<td>1</td>
<td>0.2</td>
<td>0.3</td>
<td>&lt;0.05</td>
<td>na</td>
<td>na</td>
<td>32,40</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Word pair delay</td>
<td>142,41</td>
<td>&lt;0.001</td>
<td>0.3</td>
<td>0.8</td>
<td>4</td>
<td>&lt;0.001</td>
<td>0.6</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>22,40</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Word pair learning</td>
<td>82,39</td>
<td>0.001</td>
<td>-0.2</td>
<td>0.9</td>
<td>3</td>
<td>0.001</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>22,38</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Word Pair Recog</td>
<td>82,41</td>
<td>0.002</td>
<td>0.9</td>
<td>0.4</td>
<td>3</td>
<td>0.02</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>na</td>
<td>na</td>
<td>12,40</td>
<td>0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Nill responses</td>
<td>22,39</td>
<td>0.2</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 No of pictures scary</td>
<td>012,38</td>
<td>0.9</td>
<td>0.4</td>
<td>0.7</td>
<td>0.02</td>
<td>1</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Rejected trials</td>
<td>62,39</td>
<td>0.005</td>
<td>-1</td>
<td>0.2</td>
<td>-2</td>
<td>0.04</td>
<td>-0.3</td>
<td>ns</td>
<td>-0.1</td>
<td>ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Categories (WCST)</td>
<td>92,38</td>
<td>0.001</td>
<td>3</td>
<td>0.01</td>
<td>2</td>
<td>0.1</td>
<td>0.4</td>
<td>&lt;0.01</td>
<td>0.1</td>
<td>ns</td>
<td>42,37</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Correct (WCST)</td>
<td>52,38</td>
<td>0.01</td>
<td>-3</td>
<td>0.004</td>
<td>0.9</td>
<td>0.4</td>
<td>-0.05</td>
<td>ns</td>
<td>-0.1</td>
<td>ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Last error (Reversals)</td>
<td>12,40</td>
<td>0.4</td>
<td>-0.2</td>
<td>0.8</td>
<td>-1</td>
<td>0.3</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix C: Statistical Results

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Test</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>CvH t</th>
<th>CvH p</th>
<th>HvL t</th>
<th>HvL p</th>
<th>VIQ corr r</th>
<th>VIQ p</th>
<th>Age corr r</th>
<th>Age p</th>
<th>ANCOVA F</th>
<th>ANCOVA df</th>
<th>ANCOVA p</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Number of reversals (Reversals)</td>
<td>1</td>
<td>2,40</td>
<td>0.2</td>
<td>0.05</td>
<td>0.7</td>
<td>1</td>
<td>0.2</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>5</td>
<td>Preservative errors (WCST)</td>
<td>1</td>
<td>2,38</td>
<td>0.4</td>
<td>-1</td>
<td>0.3</td>
<td>-0.2</td>
<td>0.8</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>5</td>
<td>Stroop</td>
<td>1</td>
<td>2,40</td>
<td>0.4</td>
<td>1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.7</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>5</td>
<td>Unique errors (WCST)</td>
<td>0.8</td>
<td>2,38</td>
<td>0.5</td>
<td>-0.4</td>
<td>0.7</td>
<td>1</td>
<td>0.2</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>7</td>
<td>AT2 Left</td>
<td>1</td>
<td>2,25</td>
<td>0.4</td>
<td>-0.3</td>
<td>0.8</td>
<td>-1.2</td>
<td>0.2</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>7</td>
<td>AT2 Right</td>
<td>2</td>
<td>2,36</td>
<td>0.2</td>
<td>-1.1</td>
<td>0.2</td>
<td>-0.8</td>
<td>0.4</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>7</td>
<td>HCT2 Left</td>
<td>0.2</td>
<td>2,30</td>
<td>0.8</td>
<td>-0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.6</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>7</td>
<td>HCT2 Right</td>
<td>2</td>
<td>2,30</td>
<td>0.2</td>
<td>-2</td>
<td>0.06</td>
<td>0.9</td>
<td>0.4</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>8</td>
<td>False alarms</td>
<td>9</td>
<td>2,40</td>
<td>0.001</td>
<td>-0.3</td>
<td>0.7</td>
<td>-3</td>
<td>0.001</td>
<td>-0.5</td>
<td>&lt;0.001</td>
<td>-0.2</td>
<td>ns</td>
<td>0.08</td>
<td>2,39</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Latency of Novel</td>
<td>2</td>
<td>2,42</td>
<td>0.1</td>
<td>-0.7</td>
<td>0.5</td>
<td>0.2</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>8</td>
<td>Latency of Target</td>
<td>0.5</td>
<td>2,41</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.8</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>8</td>
<td>Novel amplitude</td>
<td>0.9</td>
<td>2,42</td>
<td>0.4</td>
<td>-0.7</td>
<td>0.5</td>
<td>-0.7</td>
<td>0.5</td>
<td>-0.2</td>
<td>ns</td>
<td>-0.03</td>
<td>ns</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>8</td>
<td>Reaction time</td>
<td>1</td>
<td>2,40</td>
<td>0.4</td>
<td>1</td>
<td>0.1</td>
<td>0.7</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>8</td>
<td>Target amplitude</td>
<td>0.7</td>
<td>2,41</td>
<td>0.5</td>
<td>-1.1</td>
<td>0.2</td>
<td>0.8</td>
<td>0.4</td>
<td>-0.01</td>
<td>ns</td>
<td>0.1</td>
<td>ns</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>8</td>
<td>Targets identified</td>
<td>5</td>
<td>2,40</td>
<td>0.02</td>
<td>0.2</td>
<td>0.9</td>
<td>2</td>
<td>0.06</td>
<td>0.5</td>
<td>0.001</td>
<td>0.3</td>
<td>&lt;0.05</td>
<td>0.02</td>
<td>2,38</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>% correct</td>
<td>21</td>
<td>2,39</td>
<td>&lt;0.001</td>
<td>0.8</td>
<td>0.4</td>
<td>5</td>
<td>&lt;0.001</td>
<td>0.6</td>
<td>&lt;0.001</td>
<td>0.4</td>
<td>&lt;0.05</td>
<td>4.2</td>
<td>37</td>
<td>0.03</td>
</tr>
<tr>
<td>9</td>
<td>Additional behavioural response</td>
<td>5</td>
<td>2,39</td>
<td>&lt;0.001</td>
<td>0.4</td>
<td>0.7</td>
<td>-4</td>
<td>0.001</td>
<td>-0.4</td>
<td>0.003</td>
<td>-0.3</td>
<td>&lt;0.05</td>
<td>2.2</td>
<td>37</td>
<td>0.1</td>
</tr>
<tr>
<td>9</td>
<td>Age of Acquisition</td>
<td>6</td>
<td>3,396</td>
<td>&lt;0.001</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>9</td>
<td>Duration of word</td>
<td>2</td>
<td>3,396</td>
<td>0.2</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>9</td>
<td>N400 Cz difference</td>
<td>2</td>
<td>2,39</td>
<td>0.2</td>
<td>-0.6</td>
<td>0.6</td>
<td>2</td>
<td>0.06</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
</tbody>
</table>

293
## Appendix C: Statistical Results

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Test</th>
<th>Stats Test</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VIQ-PIQ (High)</td>
<td>Paired t test</td>
<td>1.1</td>
<td>13</td>
<td>0.3</td>
</tr>
<tr>
<td>1</td>
<td>VIQ-PIQ (Low)</td>
<td>Paired t test</td>
<td>-1.1</td>
<td>11</td>
<td>0.3</td>
</tr>
<tr>
<td>1</td>
<td>Basic reading - Inference</td>
<td>Pearson's correlation</td>
<td>0.5</td>
<td>-</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>Rivermead - sustained attention</td>
<td>Pearson's correlation</td>
<td>0.6</td>
<td>-</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>Rivermead - ABC</td>
<td>Pearson's correlation</td>
<td>-0.5</td>
<td>-</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>Visual paired comparison Controls SI</td>
<td>One sample t test (50)</td>
<td>-1</td>
<td>15</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>Visual paired comparison Controls SD</td>
<td>One sample t test (50)</td>
<td>-3</td>
<td>14</td>
<td>0.006</td>
</tr>
<tr>
<td>3</td>
<td>Visual paired comparison Controls DI</td>
<td>One sample t test (50)</td>
<td>-3</td>
<td>15</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>Visual paired comparison Controls DD</td>
<td>One sample t test (50)</td>
<td>-1</td>
<td>15</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>VPC: Group comparison</td>
<td>Rpt Measures ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>VPC order effect in control group SI</td>
<td>ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>VPC order effect in control group SD</td>
<td>ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>VPC order effect in control group DI</td>
<td>ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>VPC order effect in control group DD</td>
<td>ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Control Magnitude</td>
<td>Paired t test</td>
<td>-3</td>
<td>15</td>
<td>0.006</td>
</tr>
<tr>
<td>4</td>
<td>Control Latency</td>
<td>Paired t test</td>
<td>-1</td>
<td>15</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>Latency blink</td>
<td>Rptd measures ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Magnitude blink</td>
<td>Rptd measures ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Latency - VIQ</td>
<td>Pearson's correlation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Latency - Age</td>
<td>Pearson's correlation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

294
### Appendix C: Statistical Results

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Test</th>
<th>Stats Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Latency - Sex</td>
<td>Pearson's correlation</td>
</tr>
<tr>
<td>4</td>
<td>Magnitude - VIQ</td>
<td>Pearson's correlation</td>
</tr>
<tr>
<td>4</td>
<td>Magnitude - Age</td>
<td>Pearson's correlation</td>
</tr>
<tr>
<td>4</td>
<td>Magnitude - Sex</td>
<td>Pearson's correlation</td>
</tr>
<tr>
<td>5</td>
<td>Correct (WCST) - sustained attention</td>
<td>Pearson's correlation</td>
</tr>
<tr>
<td>5</td>
<td>Within error (Spatial Working memory)</td>
<td>Rptd measures ANOVA:</td>
</tr>
<tr>
<td>5</td>
<td>Between error (Spatial Working memory)</td>
<td>Rptd measures ANOVA:</td>
</tr>
<tr>
<td>5</td>
<td>Last error (Extinction)</td>
<td>Kruskal Wallis</td>
</tr>
<tr>
<td>7</td>
<td>HC Volumes</td>
<td>Rptd measures ANCOVA (covariate - ICV)</td>
</tr>
<tr>
<td>7</td>
<td>Amygdala Volumes</td>
<td>Rptd measures ANCOVA (covariate - ICV)</td>
</tr>
<tr>
<td>7</td>
<td>Amygdala Volumes boys only</td>
<td>Rptd measures ANCOVA (covariate - ICV)</td>
</tr>
<tr>
<td>7</td>
<td>Rivermead and HF volume correlations</td>
<td>Partial correlation with ICV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix C: Statistical Results

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Test</th>
<th>Stats Test</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 Child: Novel</td>
<td>Rpt ANOVA</td>
<td>Row</td>
<td>36</td>
<td>1,3,53.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x Grp</td>
<td>0.9</td>
<td>2.7, 53.8</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column</td>
<td>3</td>
<td>1.7, 69.7</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column x Grp</td>
<td>1</td>
<td>3.5, 69.7</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x Column</td>
<td>1</td>
<td>2.8, 112.4</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x Column x Grp</td>
<td>2</td>
<td>5.6, 112.4</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group</td>
<td>0.9</td>
<td>2, 40</td>
<td>0.4</td>
</tr>
<tr>
<td>8 Child: Target</td>
<td>Rpt ANOVA</td>
<td>Row</td>
<td>19</td>
<td>1.7, 65.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x Grp</td>
<td>1</td>
<td>3.4, 65.6</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column</td>
<td>1</td>
<td>3.6, 60.6</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column x Grp</td>
<td>1</td>
<td>3.1, 60.6</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x Column</td>
<td>9</td>
<td>3.4, 128.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x Column x Grp</td>
<td>2</td>
<td>3.4, 128.2</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group</td>
<td>0.5</td>
<td>2, 39</td>
<td>0.6</td>
</tr>
<tr>
<td>9 Adults % correct responses (social-non)</td>
<td>Paired t test</td>
<td>t = 3, df = 7, p = 0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 N400 Adults</td>
<td>Rpt ANOVA</td>
<td>Stimulus</td>
<td>1</td>
<td>1, 9</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stimulus x Row</td>
<td>7</td>
<td>1, 9</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stimulus x Column</td>
<td>0.3</td>
<td>1.6, 14.1</td>
<td>0.8</td>
</tr>
<tr>
<td>9 N400 Adults Social - Non-social</td>
<td>Rpt ANOVA</td>
<td>Row</td>
<td>25</td>
<td>1, 9</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column</td>
<td>4</td>
<td>1.7, 15.1</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Related</td>
<td>1</td>
<td>1, 9</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Social</td>
<td>0.009</td>
<td>1, 9</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x column</td>
<td>3</td>
<td>1.7, 15.0</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x rel</td>
<td>7</td>
<td>1, 9</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x social</td>
<td>0.02</td>
<td>1, 9</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column x rel</td>
<td>0.3</td>
<td>1.5, 13.9</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column x social</td>
<td>3</td>
<td>1.8, 16.0</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel x Social</td>
<td>2</td>
<td>1, 9</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C x Row x Rel</td>
<td>0.4</td>
<td>1.4, 12.7</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x Rel x soci</td>
<td>0.003</td>
<td>1, 9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column x rel x social</td>
<td>2</td>
<td>1.8, 16.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Column x row x social</td>
<td>0.07</td>
<td>1.7, 15.7</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Row x column x related x social</td>
<td>2</td>
<td>1.5, 13.7</td>
<td>0.1</td>
</tr>
</tbody>
</table>
## Appendix C: Statistical Results

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Test</th>
<th>Stats Test</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Control children % correct responses</td>
<td>Paired t test</td>
<td>t = 3, df = 16, p = 0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Control Children N400</td>
<td>Rpt ANOVA</td>
<td>F</td>
<td>df</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td>Stimulus</td>
<td>4</td>
<td>1, 16</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stimulus x Row</td>
<td>0.9</td>
<td>1, 16</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stimulus x Column</td>
<td>0.1</td>
<td>1.7, 27.3</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row ANOVA</td>
<td>F</td>
<td>df</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row</td>
<td>50</td>
<td>1, 10</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column</td>
<td>0.2</td>
<td>1.5, 14.6</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Related</td>
<td>7</td>
<td>1, 10</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social</td>
<td>2</td>
<td>1, 10</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row x column</td>
<td>1</td>
<td>1.8, 18.0</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row x rel</td>
<td>0.05</td>
<td>1, 10</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row x social</td>
<td>1</td>
<td>1, 10</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column x rel</td>
<td>0.3</td>
<td>1.2, 11.7</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column x social</td>
<td>2</td>
<td>1.3, 13.4</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rel x Social</td>
<td>1</td>
<td>1, 10</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C x Row x Rel</td>
<td>1</td>
<td>1.5, 14.6</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row x Rel x soci</td>
<td>0.05</td>
<td>1, 10</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column x rel x social</td>
<td>0.2</td>
<td>1.2, 12.3</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column x row x social</td>
<td>0.6</td>
<td>1.5, 14.6</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row x column x related x social</td>
<td>2</td>
<td>1.7, 17.4</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Children % correct</td>
<td>Rpt ANOVA</td>
<td>F(1, 37) = 0.5, p = 0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>RMS CZ Difference N400</td>
<td>Rpt ANOVA</td>
<td>F</td>
<td>df</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td>Row</td>
<td>0.6</td>
<td>1, 39</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column</td>
<td>2.7</td>
<td>1.8, 68.5</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group</td>
<td>1</td>
<td>2, 39</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Row x Group</td>
<td>2</td>
<td>2, 39</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column x Group</td>
<td>0.4</td>
<td>3.5, 68.5</td>
<td>0.8</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix C: Statistical Results

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Test</th>
<th>Stats Test</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 RMS CZ Difference Social-Non-social</td>
<td>Rpt ANOVA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Row</td>
<td></td>
<td></td>
<td>0.5</td>
<td>1,29</td>
<td>0.5</td>
</tr>
<tr>
<td>Column</td>
<td></td>
<td></td>
<td>1.5</td>
<td>43.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Social</td>
<td></td>
<td></td>
<td>0.01</td>
<td>1,29</td>
<td>0.9</td>
</tr>
<tr>
<td>Row x column</td>
<td></td>
<td></td>
<td>0.1</td>
<td>1,9, 55.7</td>
<td>0.9</td>
</tr>
<tr>
<td>Row x social</td>
<td></td>
<td></td>
<td>0.03</td>
<td>1,29</td>
<td>0.9</td>
</tr>
<tr>
<td>Column x social</td>
<td></td>
<td></td>
<td>3.1</td>
<td>42.7</td>
<td>0.1</td>
</tr>
<tr>
<td>Column x row x social</td>
<td></td>
<td></td>
<td>0.5</td>
<td>1.7, 49.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Row x Group</td>
<td></td>
<td></td>
<td>0.5</td>
<td>2,29</td>
<td>0.6</td>
</tr>
<tr>
<td>Column x Group</td>
<td></td>
<td></td>
<td>0.8</td>
<td>3.0, 43.4</td>
<td>0.5</td>
</tr>
<tr>
<td>Social x Group</td>
<td></td>
<td></td>
<td>0.4</td>
<td>2,29</td>
<td>0.7</td>
</tr>
<tr>
<td>Row x column x Group</td>
<td></td>
<td></td>
<td>0.7</td>
<td>3.8, 55.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Row x social x Group</td>
<td></td>
<td></td>
<td>0.4</td>
<td>2,29</td>
<td>0.7</td>
</tr>
<tr>
<td>Column x social x Group</td>
<td></td>
<td></td>
<td>2.3</td>
<td>3.0, 42.7</td>
<td>0.08</td>
</tr>
<tr>
<td>Column x row x social x Group</td>
<td></td>
<td></td>
<td>0.6</td>
<td>3.4, 49.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
<td>3.2</td>
<td>2,29</td>
<td>0.06</td>
</tr>
</tbody>
</table>
Appendix D Voxel-based Morphometry

D.1 Precision of Spatial Normalisation in Voxel-Based Morphometry

The first processing step in voxel-based morphometry is spatial normalisation. The purpose of this step is to place all the data into a common stereotactic space. It permits comparison across subjects of homologous regions. The accuracy of this normalisation therefore limits the precision of the anatomical localisation of any abnormality. In this section, the normalisation accuracy was investigated.

D.1.1 Introduction

Voxel-based morphometry depends on spatially normalising all the images into the same stereotactic space, extracting the grey (or white) matter from the normalised images, smoothing and finally performing a statistical analysis to localise and make inferences about group differences. The output of the method is a statistical parametric map (SPM) showing regions where grey (or white) matter density differs significantly among groups.

Normalisation is required to ensure that homologous regions are compared across subjects. It is not intended to be an exact process (since this would remove all anatomical differences among the groups, rendering the differences in grey matter segments negligible). The normalisation employed should only remove anatomical differences down to a specified spatial scale so that grey matter structures are registered but their size differences are preserved. The accuracy of colocalisation of anatomical structures across subjects is a critical factor in ensuring that the results can be interpreted in terms of grey matter changes per se. The aim of the work described here was to investigate the precision of the normalisation procedure.
Appendix D: Voxel-based Morphometry

D.1.1.1 Spatial Normalisation

Normalisation is essentially the process of warping MRI images from different subjects into a standard space as defined by a template image [Ashburner and Friston 1999b]. One way of normalising data involves identifying certain brain ‘landmarks’ by inspection and then manipulating the images so that these landmarks are brought into register. However, to overcome the inherent subjectivity of this method, voxel-based morphometry uses a non-label based technique that is fully automated.

Non-label based normalisation techniques minimise some index of the difference between the source image and the template image [Ashburner and Friston 1999b; Ashburner and Friston 1999c]. SPM99 minimises the sum of the squared differences between the image to be normalised and the template whilst maximising the prior probability of the transformation (also known as regularisation). The maximum a posteriori solution is found iteratively: the algorithm starts with an initial parameter estimate and searches from there. The algorithm stops when criterion is achieved (the weighted sum of square differences no longer decreases or after a finite number of iterations).

Normalisation can be divided into two components: affine (or linear) and transformations. Affine transformations are generally carried out first and account for differences in position, orientation and overall brain size. Then non-linear normalisation is used to account for low spatial frequency global variability in head shape. In SPM99, the non-linear transformations are restricted to linear combinations of 3-dimensional discrete cosine functions [Ashburner and Friston 1999c].

The normalisation is therefore shaped by a number of constraints: hard constraints (such as the number of discrete cosine functions employed) and soft constraints (such as the degree of regularisation). The construct validity of the normalisation procedure has been addressed by Ashburner and Friston [1999a]. These investigations focussed on ensuring that the normalisation procedure minimises various measures of template and data differences (such as membrane energy). In the present work, face validity is assessed in terms of the precision of the normalisation procedure, by examining the variation in the co-ordinate location of various anatomical landmarks across a group of individuals. The effect
Appendix D: Voxel-based Morphometry

of changing the constraints on the non-linear transformations as well as the dependency of the results on the template used is investigated.

D.1.2 Changing the number of basis functions (hard constraints)

D.1.2.1 Materials and Methods

All subjects (20 children, mean age 13 years; 11 males; 9 females with no known neurological or psychiatric history) were scanned, unsedated, on a 1.5 T Siemens Vision scanner, using a T1 weighted 3D MPRAGE sequence as described in Chapter 6.

The 3D data sets were analysed in SPM99 (Wellcome Department of Cognitive Neurology, London, UK). Each scan was normalised [Friston et al. 1995b; Ashburner and Friston 1999b] to the T1 template supplied with SPM99. This template is constructed from 152 T1 weighted scans from the Montreal Neurological Institute (supplied by Alan Evans, Montreal Neurological Institute, Canada (ICBM, NIH P-20 project, Principal Investigator John Mazziotta)). Normalisation was performed using different numbers of non-linear basis functions in three orthogonal directions to give 4 levels of constraint:

Level 1  No non-linear transformations
Level 2  4 x 5 x 4 non-linear basis functions
Level 3  7 x 8 x 7 non-linear basis functions
Level 4  10 x 11 x 10 non-linear basis functions

All other normalisation parameters were constant across the groups (medium regularisation, 16 iterations). To assess face validity, 8 anatomical landmarks were identified as described in Table D:1. These comprised “bounding box”-like landmarks and specific landmarks in the medial temporal region (see Figure D:1). The investigations focussed on the medial temporal precision given the central hypothesis of the thesis (the role of the medial temporal lobes in Autism).
The coordinates of these landmarks were measured on each normalised scan, with the assessor blinded to the level of constraint. The mean position for each anatomical landmark was calculated within each level and the standard deviation about this position was calculated. Measurements were carried out twice on Level 3 to assess test-retest reliability. The precision data (i.e. the standard deviation of the 8 landmark locations) were analysed in SPSS using repeated measures and Pearson’s correlation coefficients.

**D.1.2.2 Results and Discussion**

The reliability of anatomical landmark identification was confirmed by Pearson’s correlation coefficients for all 8 landmarks. All correlation coefficients were significant at $p < 0.05$ level. Figure D.2 shows that the four landmarks in the plane $z = 0$ were colocalised to within ~1 mm ($< 1$ voxel) regardless of how many non-linear basis functions were used. This is a reflection of the fact that these ‘basic’ aspects of colocalisation are well accommodated by linear transformations. The four landmarks in the coronal plane through the pons intersection were significantly less well localised than these basic measures ($F(3.7, 71.1) = 35; p < 0.001$). In addition, from Figure D:2, it can be seen that 4 x 5 x 4 non-linear basis functions gives better colocalisation (significantly smaller standard deviation from mean anatomical location) for all four medial temporal landmarks.
Figure D:2 Standard deviation from mean position of landmarks using medium regularisation Group 1: No non-linear transformations; Group 2: 4 x 5 x 4 non-linear transformations; Group 3: 7 x 8 x 7 non-linear transformations; Group 4: 10 x 11 x 10 non-linear transformations.

These results suggest that a smaller number of non-linear basis functions can be "better" than a larger number. This may appear counter-intuitive, since more basis functions endow the transformation with more degrees of freedom, which in turn should make the normalisation more accurate. However this intensity matching is not spatial matching and does not necessarily give the most spatially precise registration. In other words, it is possible that whilst the algorithm minimises the difference between the template and the image, the amount of regularisation fails to prevent anatomically unlikely transformations, i.e. the algorithm prioritises reducing the difference between the template and the image at the expense of anatomical validity. This possibility was investigated in the next analyses in which evidence of an interaction between the basis functions and regularisation was sought.

D.1.3 Changing the degrees of regularisation (soft constraints)

Regularisation refers to the inclusion of a penalty term for unlikely warps (based on specified priors) that enter into the minimisation. This log likelihood term penalises rough, quickly changing warps. Operationally this implies penalising the use of high spatial frequency basis functions more than lower frequencies (see Ashburner and Friston, 1999b).
D.1.3.1 Materials and Methods

The 3D data sets (as described above) were normalised with two different sets of non-linear basis functions with high regularisation to form two additional normalisation levels:

Level 2A 4 x 5 x 4 non-linear basis functions  
Level 3A 7 x 8 x 7 non-linear basis functions

The data were otherwise normalised exactly as described above. The coordinates of landmarks were measured and analysed as above.

D.1.3.2 Results and Discussion

As Figure D:3 shows, the normalisation precision was unchanged by increasing regularisation with 7 x 8 x 7 non-linear basis functions. With 4 x 5 x 4 non-linear basis functions, the accuracy was reduced by increasing the regularisation. These results suggest that reducing the importance of minimising the template and image difference relative to the probability of the anatomical warping interacts with the number of basis functions in terms of face validity. Regularising the larger basis set has little effect. However, regularising the smaller (optimal for these 8 landmarks) set overly constrains the warp and reduces the precision. These findings suggest that having a greater number of basis functions does not necessarily improve normalisation accuracy and the hard constraints offered by basis functions can actually lead to better spatial co-registration.
Figure D:3  Standard deviation from mean position of landmarks using varying levels of regularisation. Group 2: 4 x 5 x 4 non-linear transformations, Group 3: 7 x 8 x 7 non-linear transformations.

It should be noted that whilst 4 x 5 x 4 non-linear basis functions gave the most accurate normalisation in the range tested, it may not be the optimal number for all landmarks. However since 4 x 5 x 4 non-linear basis functions colocalised all eight landmarks to within ~1.5 mm and this is the size of the voxels in the normalised images, it is likely that this accuracy is adequate for most applications.

D.1.4 Changing the template

The normalisation procedures of VBM are intended to provide good anatomical colocalisation in relation to the template the data are normalised to. The more basis functions used in the non-linear normalisation, the greater the possibility that the data are over-fitted, rendering the results sensitive to the choice of template. In this section the template dependent nature of the colocalisation accuracy was investigated. Having established registration accuracy in the medial temporal lobe to within 3 mm using the landmark approach, the impact of changing templates on the anatomical precision of VBM per se was examined. This was done using a test sample comprising of patients with
bilateral hippocampal atrophy. To extend the analysis of the previous sections, the effect of

template on landmark colocalisation using $7 \times 8 \times 7$ basis functions was also assessed.

**D.1.4.1 Materials and methods**

5 patients (mean age 12.4; 4 males, 1 female) who had developmental amnesia associated
with early hypoxic-ischaemic episodes comprised the test sample (as described in Chapter
6). All these patients have been shown to have bilateral hippocampal atrophy using
volumetric methods in addition to VBM [Gadian et al. 2000]. Eight controls (mean age
13.9; 3 males; 5 females) were selected for comparison. All subjects were scanned
unsedated on a 1.5 T Siemens Vision scanner, using a T1 weighted 3D MPRAGE sequence
as described before.

The 3D data sets were analysed in SPM99 (Wellcome Department of Cognitive Neurology,
London, UK). Normalisation was carried out using two different templates: the MNI
template (supplied with SPM99, 152 adult T1 scans) and an in-house template (GOS
template: 27 children scanned on the same scanner with the same image acquisition
sequence as the subjects used in this analysis; mean age 14). Normalisation was carried out
using two different sets of non-linear functions: $4 \times 5 \times 4$ and $7 \times 8 \times 7$ with medium
regularisation. The data were otherwise normalised as described above. The data were
then segmented [Ashburner and Friston 1997], smoothed with a Gaussian isotropic kernel
of FWHM 4mm and analysed to look for decreases in grey matter density in the patients
with amnesia versus the controls.

The extent and location of the abnormalities within the hippocampal formation were then
compared between the two templates at both $4 \times 5 \times 4$ and $7 \times 8 \times 7$ non-linear basis
function levels.

In addition, the scans of the 20 controls used in previous sections were normalised to the
GOS template with $7 \times 8 \times 7$ non-linear basis functions to determine whether the template
chosen affected the colocalisation precision.
**D.1.4.2 Results and Discussion**

Figure D:4 shows the results of the analysis (with $7 \times 8 \times 7$ non-linear basis functions) of the left hippocampal formation. The extent of the abnormality appeared greater using the GOS template relative to the MNI template analysis. In contrast, Figure D:5 shows that using $4 \times 5 \times 4$ non-linear basis functions results in similar extent and location abnormalities within the left hippocampal formation (this pattern is also present in the right hippocampal formation (not shown)).

Figure D:4 Significant areas of decreased grey matter in the amnesic children versus controls in left hippocampal formation ($7 \times 8 \times 7$ non-linear basis functions). Results are superimposed on the mean normalised images and thresholded at uncorrected $p < 0.001$. a) GOS template b) T1 template

![Figure D:4](image1)

Figure D:5 Significant areas of decreased grey matter in the amnesic children versus controls in left hippocampal formation ($4 \times 5 \times 4$ non-linear basis functions). Results are superimposed on the mean normalised images and thresholded at uncorrected $p < 0.001$. a) GOS template b) T1 template

![Figure D:5](image2)
Figure D:4 and Figure D:5 suggest that when 4 x 5 x 4 non-linear basis functions are used, the results of the analyses are minimally dependent on the template used. This suggests that use of the default template even when comparing children is suitable when using the optimal number of non-linear basis functions, at least in the medial temporal lobe. However, with larger basis function sets, the template does appear to affect the results of the analysis. Investigators using larger number of basis functions need to consider which template might be most appropriate for their study and allow for the possibility of overfitting.

Template dependencies such as those seen in Figures 4a and b remain when the smoothing was raised to 8 mm from 4 mm FWHM (data not shown).

There was no significant difference between the normalisation precision of any of the 8 landmarks regardless of which template was used.

It is important to note that the precision of the normalisation, as reflected in the landmark colocalisation, does not depend on the template. This is to be expected as long as the warps required to match an image to the template are not too “unlikely”. This finding moderates enthusiasm for “custom” templates unless one is dealing with very different brains (e.g. patients with gross pathology, etc.).

**D.1.4.3 Conclusion**

In conclusion, our results suggest that when optimum normalisation parameters are used, anatomical landmarks (at least in the medial temporal lobe) are colocalised to within a standard deviation of about 1 mm. When suboptimal parameters are used this standard deviation can increase up to 3mm. Interestingly the optimal parameters are those that give a rather constrained normalisation and are not those that optimise the intensity matching at the expense of using ‘unlikely’ warps. The face validity of the normalisation was not improved by the use of “custom” templates, even for children.
Restricting the number of basis functions, which produced empirically better results, effectively meant that warps above a certain frequency were penalised infinitely. This could imply that the form of regularisation used by the warping algorithm may need to impose much higher penalties against high frequency deformations. Currently the model minimises the membrane energy of the warps, which is based on the sum of squares of the first derivatives of the deformations. It is possible that the regularisation should minimise the sum of squares of a higher derivative. Regularisation based on higher order derivatives has the effect of increasing the penalty against higher frequency deformations relative to those of lower frequencies. These forms may turn out to be better models of the true distribution of warps likely to be needed for registering brain images.
D.2 Distributional assumptions in Voxel-Based Morphometry

Voxel-based morphometry (VBM) uses the general linear model to construct parametric statistical tests. In order for these statistics to be valid, a small number of assumptions must hold. A key assumption is that the model's error terms are normally distributed. This is usually ensured through the Central Limit Theorem by smoothing the data. However, there is increasing interest in using minimal smoothing (in order to sensitise the analysis to regional differences at a small spatial scale). The validity of such analyses was therefore investigated.

D.2.1 Introduction

As described earlier, voxel-based morphometry involves spatially normalising all the images into the same stereotactic space, extracting the grey matter (or white matter) from the normalised images, smoothing and finally performing a statistical analysis to localise and make inferences about group differences. The output of the method is a statistical parametric map showing regions where grey matter density differs significantly among groups.

D.2.1.1 Normality of Residuals

The parametric statistical tests are carried out within the framework of the general linear model. In order for these tests to be valid, the errors must be normally distributed. Prior to smoothing, the segmented images may have a highly non-normal density function, with most voxels having a value close to the extremes of the range of 0-1. This is because the voxel values in the segments correspond to the probability that the voxel is grey matter. The distribution of errors about any group mean will show a similar non-normal distribution. However by appealing to the Central Limit Theorem, it is generally assumed that the errors are rendered normally distributed by spatial smoothing.

An important question in this context is "what circumstances would result in deviations from normality, sufficient to render the tests invalid?". One potential situation is the
comparison of a single subject and a group. In parametric statistics it is assumed that the
group difference (more formally a contrast of parameter estimates) is normally distributed.
Dividing this contrast by an estimate of its standard deviation gives the t-statistic.
Generally this difference will be well-behaved because the group mean represents an
average over many observations and will have a normal distribution by the central limit
theorem. However when one of the groups has only one subject the difference may be
highly non-normal and the distribution of the ensuing statistic will not conform to
parametric assumptions.

In short, for most designs, inferences are quite robust to violations of normality, but there
are some (e.g. unbalanced) designs which may be less robust. In other words there may be
an interaction between the degree of non-normality and experimental design that renders
the tests invalid.

Recent applications of voxel-based morphometry have included analyses with small
smoothing kernels (e.g. 4mm Gadian et al. 2000) and comparisons of an individual versus
a group of controls (e.g. Woermann et al. 1999a; Woermann et al. 1999b). Smaller
Gaussian kernels are used to sensitise the analysis to a spatial scale equivalent to the
structure of interest (e.g. the hippocampal formation). Investigations into the
neuropathology of single cases are particularly important in clinical diagnosis and the field
of clinical neuropsychology, where individual cognitive and behavioural profiles prevent
the formation of a homogeneous clinical group. The validity of the parametric statistical
tests in this context is investigated in this work.

D.2.2 Materials and Methods

MRI data acquisition
All subjects (20 children, mean age 13 years; 11 males; 9 female, with no known
neurological or psychiatric history) were scanned on a 1.5 T Siemens Vision scanner, using
a T1 weighted 3D MPRAGE sequence as described in Chapter 6.
Appendix D: Voxel-based Morphometry

**Pre-processing Steps**

The 3D data sets were analysed in SPM99 (Wellcome Department of Cognitive Neurology, London, UK).

Each scan was spatially normalised [Friston et al. 1995b; Ashburner and Friston 1999c]. The images were then segmented using the Bayesian algorithm described in [Ashburner and Friston 1997]. This produced continuous probability maps where the values correspond to the posterior probability that a voxel belongs to the grey matter partition. The grey matter images were smoothed with 12 mm, 8 mm, 4 mm and 0 mm isotropic Gaussian kernels. This smoothing renders the voxel values an index of the amount of grey matter per unit volume under the smoothing kernel. The term “grey matter density” is generally used to refer to this measure.

**D.2.2.1 Effects of Smoothing on Normality of Residuals**

*QQ Plots:* It is not possible to prove that data are normally distributed. However it is possible to determine if data are significantly non-normal. One method is the QQ plot. A QQ plot is a plot of the sample quantile versus the sample quantile that would be expected if the data were normally distributed. For normally distributed data, the QQ plot of the data should be a straight line. A significant deviation from a straight line can be identified by computing the correlation coefficient of the plot (as described by Johnson and Wichern, 1998). To test for normality the correlation coefficient is calculated and a look up table is used to reject the null hypothesis. If the correlation coefficient falls below a particular value, given a certain sample size, non-normality can be inferred. For more information see [Ashburner and Friston 1999a].

12 sets of 10 scans were selected randomly, from the 20 possible scans, for three levels of smoothness (8mm, 4mm, 0mm). Correlation coefficients from a QQ plot were computed over all voxels where the mean intensity over all the images was greater than 0.05 (adimensional units of probability). Voxels of low mean intensity were excluded as they would not be included in a conventional SPM analysis. The QQ plots were calculated...
using the residuals of a model that accounted for the confounding effects of age, sex and total amount of grey matter in each volume.

Histograms of the QQ correlation coefficients were computed over all voxels. This procedure was repeated but replacing the data with simulated Gaussian noise for comparison. In addition the proportion of voxels where the correlation coefficient fell below the tabulated value (indicating residuals that were not normally distributed at $p < 0.05$) was also calculated. Under the null hypothesis or normally distributed errors, this proportion would be expected to be 0.05.

Transforming the smoothed, segmented data with a 'logit' transform prior to performing statistical tests may render the errors more normally distributed. This is because every voxel in the smoothed image segment has a value between 0 and 1. In order to assess the improvement the logit transform makes, the QQ analyses were calculated with and without the logit transform.

Rate of False Positives: Whilst QQ analyses can determine whether the data are not normally distributed, they can not demonstrate the influences that any non-normality may have on subsequent statistical inference. One way of assessing this crudely is to look at the false positive rate. The rate of false positives was assessed by randomly assigning the 20 children into 2 groups (each of size 10). Confounding factors of age, sex and total amount of grey matter were included in the model. This was repeated a total of 10 times at three levels of smoothing (8mm, 4mm, 0mm). Significant increases and decreases in grey matter density were assessed, resulting in a total of 20 SPMs of the t-statistic, for each smoothing level. The number of analyses with one or more false positives (at $p = 0.05$ corrected) was assessed.

Assuming false positive SPMs are encountered like “rare events”, the Poisson distribution is used to compare the probability of obtaining the observed number of SPMs with one or more maxima at a corrected level of significance. This probability assumes the tests are exact and normality has not been violated. Although these p values do not establish that
Appendix D: Voxel-based Morphometry

VBM is valid they do allow us to say that the tests are invalid if the p value falls below a critical threshold (i.e. $p = 0.05$).

D.2.2.2 Effects of Experimental Design on Robustness

The data from 17 of the 20 children were used in this investigation. One child was randomly chosen and compared against the remaining 16 children. This was repeated a total of 10 times at all three levels of smoothing. Confounding factors of age, sex and total amount of grey matter were included in the model. Significant increases and decreases in grey matter density were assessed, resulting in a total of 20 SPMs of the t-statistic, at three smoothing levels (12mm, 8mm, 4mm). The number of SPMs with one or more false positives (at $p = 0.05$ corrected) was assessed, and the probability of getting this number or more was computed as above with reference to the Poisson distribution.

D.2.3 Results

Inferences (i.e. p values) are restricted to the intensity statistics (i.e. height of maxima in the ensuing SPMs), as inferences based on the spatial extent are invalid for voxel-based morphometry. This is because volumes are measured in voxels in SPM99 whereas they should be measured in RESELs when the spatial smoothness is not stationary (see Ashburner and Friston 1999a).

D.2.3.1 Effects of Smoothing on Normality of Residuals

The effect of smoothing on balanced designs is shown in Figure D:6. It demonstrates that the proportion of voxels violating an assumption of normality (based on QQ plots) is below the expected limit ($\leq 0.05$) for smoothing kernels of 4mm and 8mm. The proportion under no smoothing (0 mm) is higher than expected for both the logit transformed data and the raw data. The logit transform does reduce the proportion at all smoothing levels, suggesting any excess is indeed due to violations of the assumption of normality. The proportion of non-normal voxels was less than expected even for the simulated data. This suggests the test for non-normality based on the correlation coefficient is not exact.
Figure D:6 QQ Plots: The proportion of data points significantly violating the assumptions of normality at 8 mm, 4 mm, 0 mm. The dotted line is the expected proportion of data points.

Figure D:7 demonstrates that there were no analyses with one or more false positives at \( p \leq 0.05 \) (corrected), at any of the three smoothing levels (8 mm, 4 mm, and 0 mm).

Figure D:7 Number of analyses with one or more false positives at 12 mm, 8 mm, and 4 mm: Balanced design * indicate significantly more analyses meeting criteria than would be expected by chance. Hard line indicates level expected by chance.
D.2.3.2 Effects of Experimental Design on Robustness

The effect of unbalanced designs on robustness is shown in Figure D:8. It demonstrates that the number of SPMs with one or more false positives at $p \leq 0.05$ (corrected) decreases with smoothing. The false positive rate was within the expected range for 12mm ($p = 0.4$). There were significantly more analyses with false positives than would be expected by chance at 4mm and 8mm (8mm: $p = 0.01$; 4mm: $p = 0.00001$). Figure 3 also demonstrates that no false positives were found in those that looked at increases in grey matter in the individual versus the group. As expected the number of false positives was greater for the unbalanced design (1 vs 16) relative to the balanced design (10 vs 10).

Figure D:8 Number of analyses with one or more false positives at 12mm, 8mm and 4mm: Unbalanced design. 16>1 refers to the contrast examining decreases in the individual versus the group, whilst 16<1 refers to that examining increases in the individual versus the group. * indicate significantly more analyses meeting criteria than would be expected by chance. Hard line indicates level expected by chance.
D.2.4 Discussion

D.2.4.1 Effects of Smoothing on Normality of Residuals

When using balanced group comparisons, smoothing at 4mm appears to be sufficient to ensure that any non-normality has little effect both as assessed using the QQ plots and the false positive rate. This is consistent with, and extends the results of, Ashburner and Friston who found that 12mm was sufficient. The non-normality detected by the QQ plots for unsmoothed data does not markedly inflate the rate of false positives. While this conclusion should be moderated by the limited number of SPMs used, it is not necessarily surprising. Despite the lack of spatial smoothing, the group analysis is rendered robust by averaging over subjects. By the central limit theorem this renders the contrasts normally distributed.

D.2.4.2 Effects of Experimental Design on Robustness

When comparing a single subject to a group, the false positive rates were only within the expected range for smoothing kernels of 12mm, there was no evidence that these analyses are invalid. Reducing the smoothing kernel to 8mm or 4mm does render the analyses significantly prone to false positives, suggesting that the tests are no longer robust. The implicit interaction between group size and smoothing on false positives is to be expected: (as discussed in introduction) group size (i.e. the design) may influence the robustness to violations of normality at low levels of smoothing. Reducing the smoothing kernel size to 4mm when investigating an individual’s neuropathology should therefore be avoided.

Increases versus decreases in Grey Matter

An interesting observation was that the false positive rate appears to be systematically higher in the tests for decreases in grey matter in an individual versus a group, compared to increases. This suggests a ‘skew’ in the distribution of differences between a single subject and a group. In other words the probability that a small region contains grey matter is skewed toward high values. This is intuitively sensible, even after smoothing, because the probability of a region being void of grey matter is much greater than the probability that it
is all grey matter. This is because grey matter conforms to a sheet or manifold embedded in a volume of non-grey matter.

In conclusion, our results indicate that non-normality in the error terms can be an issue in VBM. However, provided the data are smoothed with a 12mm FWHM kernel, non-normality is sufficiently attenuated to render the tests valid in all situations. An important caveat, however, is that unbalanced designs appear to be less robust to violations of normality: a significant number of false positives arise at a smoothing of 4mm and 8mm when comparing a single subject to a group. This is in contrast to the observation that conventional group comparisons appear to be robust, remaining valid even with no smoothing.
Appendix D: Voxel-based Morphometry

D.3 Independence of Residuals

A second assumption made in voxel-based morphometry is that the residuals are independent. This is usually automatically ensured by each individual subject appearing only once in the design matrix. However when using the bilateral method described in Chapter 6, this is not the case. Independence of the residuals, in this special case, was therefore investigated.

D.3.1 Introduction

Traditionally voxel-based morphometry has been used to search for unilateral abnormalities in grey matter density (see for example Wright et al. 1995; Abell et al. 1999; Gadian et al. 2000). Recently a new application for voxel-based morphometry has been developed, which searches exclusively for bilateral anatomical differences. This method is of particular relevance for investigations in developmental neuropsychology, where profound selective cognitive impairments are usually associated with bilateral damage (see Chapter 6).

In this method, each subject’s scan appears twice in the design matrix (once unflipped and once flipped in the transverse plane). The independence of the residuals is therefore investigated using the rate of false positives as an indication of violation of this assumption.

D.3.2 Materials and Methods

MRI data acquisition

The scans used in this analysis were identical to those described above.
D.3.2.1 Pre-processing Steps

Bilateral Analysis: The raw data were normalised in accordance with the bilateral method described in Chapter 6. Briefly the scans were normalised to a symmetric template, segmented symmetrically and duplicated. One copy of each scan was then flipped in the transverse plane. The grey matter segments were smoothed with 12mm, 8mm, 4mm, and 0mm isotropic Gaussian kernels as in the unilateral analyses.

D.3.2.2 Data Analysis

Effects of Smoothing on Normality of Residuals

Rate of False Positives: The rate of false positives was assessed by randomly assigning the 20 children into 2 groups (each of size 10). Confounding factors of age, sex and total amount of grey matter were included in the model. This was repeated a total of 10 times at three levels of smoothing (8mm, 4mm, 0mm). Significant increases and decreases in grey matter density were assessed, resulting in a total of 20 SPMs of the t-statistic, for each smoothing level. The number of analyses with one or more false positives (at \( p = 0.05 \) corrected) was assessed. Only significant differences away from the midline were assessed (since the data along \( x = 0 \) is not independent between the flipped and unflipped images, see Salmond et al., 2000).

Assuming false positives SPMs are encountered like “rare events” we used the Poisson distribution to compare the probability of obtaining the observed number of SPMs with one or more maxima at a corrected level of significance. This probability assumes the tests are exact and independence has not been violated. Although these \( p \) values do not establish that VBM is valid they do allow us to say that the tests are invalid if the \( p \) value falls below a critical threshold (i.e. \( p = 0.05 \)).
Effects of Experimental Design on Robustness

17 children from the 20 data sets were selected. One child was randomly chosen and compared against the remaining 16 children. This was repeated a total of 10 times at all three levels of smoothing. Confounding factors of age, sex and total amount of grey matter were included in the model. Significant increases and decreases in grey matter density were assessed, resulting in a total of 20 SPMs of the t-statistic, at three smoothing levels (12mm, 8mm, 4mm). The number of SPMs with one or more false positives (at $p = 0.05$ corrected) was assessed and the probability of getting this number or more was computed as above with reference to the Poisson distribution.

D.3.3 Results

D.3.3.1 Effects of Smoothing on Normality of Residuals

Figure D:9 demonstrates that there were no analyses with one or more false positives at $p \leq 0.05$ (corrected), at 8 mm and only one at 4mm and 0mm using the bilateral method.

Figure D:9 Number of analyses with one or more false positives at 12mm, 8mm, and 4mm: Balanced design; Bilateral method * indicate significantly more analyses meeting criteria than would be expected by chance. Hard line indicates level expected by chance.
D.3.3.2 Effects of Experimental Design on Robustness

Figure D:10 demonstrates that the number of SPMs with one or more false positives at p ≤ 0.05 (corrected) when using the bilateral method. In contradistinction to the conventional (unilateral) analyses, there were significantly more false positives SPMs than would be expected by chance at all three smoothing levels: 12 mm, 8mm and 4mm (12mm: 16>1: p < 0.001; 16<1: p < 0.001; 8 mm 16>1: p < 0.001; 4mm: 16>1: p < 0.001 ). Figure D:10 also demonstrates that similar to the unilateral method, the bilateral method is also more susceptible to false positives when examining decreases in grey matter in the individual versus the group than when investigating increases in grey matter.

Figure D:10  Number of analyses with one or more false positives at 12mm, 8mm and 4mm: Unbalanced design; Bilateral method. 16>1 refers to the contrast examining decreases in the individual versus the group, whilst 16<1 refers to that examining increases in the individual versus the group. * indicate significantly more analyses meeting criteria than would be expected by chance. Hard line indicates level expected by chance.

As the pre-processing steps on the bilateral method are slightly different from those of the unilateral method (see Materials and Methods), additional investigations were carried out to determine whether the pre-processing steps or the search for bilateral abnormalities per se are responsible for the increase in false positive rates seen between the unilateral and
bilateral methods. These investigations demonstrated that the pre-processing steps are not the cause of the increased false positive rates.

D.3.4 Discussion

Using the bilateral method in group comparisons appears to be valid, suggesting dependence among errors is not a problem. However, the surprising results of this study suggest that use of the bilateral method to investigate the neuropathology is apparently invalid at all levels of smoothing investigated (12mm, 8mm, 4mm), in single subjects (relative to a group). However this is not the case. The unilateral simulations suggest that any normality violation cannot explain readily the profound increase in positive results using the bilateral analysis. The explanation is not that any assumption has been violated but that the tests are correctly rejecting the null hypothesis. The reason the null hypothesis is wrong lies in the difference between the unilateral and bilateral analyses. The bilateral analysis uses a conjunction analysis, where under the null hypothesis the error terms from the two hemispheres are independent. In this instance the error terms are simply endogenous subject-specific variations in anatomy which are likely to show a degree of bilateral symmetry. In short a sufficient explanation for the elevated results in Figure D:10 is that idiosyncratic differences among subjects are more likely to be expressed bilaterally rather than not. This means bilateral differences could be real and unique to that subject. Whether they are construed as "pathological" or not is beyond the scope of inference. Idiosyncratic differences are averaged out in the group analysis and the positive rate falls to that predicted by the null hypothesis (see Figure D:9– bilateral results).
## Table D.1 Anatomical landmarks location

<table>
<thead>
<tr>
<th>Label</th>
<th>Category</th>
<th>Z co-ordinate</th>
<th>Localisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>Basic Landmarks</td>
<td>In plane z = 0</td>
<td>Maximal y value</td>
</tr>
<tr>
<td>Back</td>
<td>Basic Landmarks</td>
<td>In plane z = 0</td>
<td>Minimal y value</td>
</tr>
<tr>
<td>Left</td>
<td>Basic Landmarks</td>
<td>In plane z = 0</td>
<td>Minimal x value</td>
</tr>
<tr>
<td>Right</td>
<td>Basic Landmarks</td>
<td>In plane z = 0</td>
<td>Maximal x value</td>
</tr>
<tr>
<td>Left white matter</td>
<td>Medial Temporal Landmarks</td>
<td>In coronal plane through superior pons-brainstem join (chosen in the sagittal plane x = 0)</td>
<td>Most medial white matter on lateral border of left hippocampal formation</td>
</tr>
<tr>
<td>Right white matter</td>
<td>Medial Temporal Landmarks</td>
<td>In coronal plane through superior pons-brainstem join (chosen in the sagittal plane x = 0)</td>
<td>Most medial white matter on lateral border of right hippocampal formation</td>
</tr>
<tr>
<td>Left pons</td>
<td>Medial Temporal Landmarks</td>
<td>In coronal plane through superior pons-brainstem join (chosen in the sagittal plane x = 0)</td>
<td>Widest point of the pons brainstem on left</td>
</tr>
<tr>
<td>Right pons</td>
<td>Medial Temporal Landmarks</td>
<td>In coronal plane through superior pons-brainstem join (chosen in the sagittal plane x = 0)</td>
<td>Widest point of the pons brainstem on right</td>
</tr>
</tbody>
</table>
## Appendix E  Abbreviations and Definitions

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>Autism Behavioural Checklist</td>
</tr>
<tr>
<td>ANCOVA</td>
<td>Analysis of covariance</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>AS</td>
<td>Asperger’s syndrome</td>
</tr>
<tr>
<td>AT2</td>
<td>T2 map of Amygdala</td>
</tr>
<tr>
<td>CCC</td>
<td>Children’s Communication Checklist</td>
</tr>
<tr>
<td>ChOCI</td>
<td>Children’s Obsessions and Compulsions Inventory</td>
</tr>
<tr>
<td>CMS</td>
<td>Children’s Memory Scale</td>
</tr>
<tr>
<td>CvH</td>
<td>Controls versus High</td>
</tr>
<tr>
<td>Df</td>
<td>Degrees of freedom</td>
</tr>
<tr>
<td>ERPs</td>
<td>Event Related Potentials</td>
</tr>
<tr>
<td>FLASH</td>
<td>Fast low angle shot (MR sequence)</td>
</tr>
<tr>
<td>HCT2</td>
<td>T2 map of Hippocampal formation</td>
</tr>
<tr>
<td>HFA</td>
<td>High functioning Autism</td>
</tr>
<tr>
<td>High</td>
<td>Group of participants in study with Autism and VIQ &gt;85</td>
</tr>
<tr>
<td>Hippocampal formation</td>
<td>The dentate gyrus, areas CA 1-3 and the subiculum(^1)</td>
</tr>
<tr>
<td>HvL</td>
<td>High versus Low</td>
</tr>
<tr>
<td>ICV</td>
<td>Intracranial volume</td>
</tr>
<tr>
<td>Low</td>
<td>Group of participants in study with Autism and VIQ &lt;85</td>
</tr>
<tr>
<td>M-ABC</td>
<td>Movement Assessment Battery for Children</td>
</tr>
<tr>
<td>MNI</td>
<td>Montreal Neurological Institute</td>
</tr>
<tr>
<td>MPRAGE</td>
<td>Magnetisation-prepared rapid acquisition gradient echo (MR sequence)</td>
</tr>
<tr>
<td>MR</td>
<td>Magnetic Resonance</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>MRS</td>
<td>Magnetic Resonance Spectroscopy</td>
</tr>
<tr>
<td>NVIQ</td>
<td>Non-verbal Intelligence Quotient</td>
</tr>
<tr>
<td>NVMA</td>
<td>Non-verbal Mental Age</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>Parahippocampal region</td>
<td>Presubiculum, parasubiculum, entorhinal and perirhinal and</td>
</tr>
<tr>
<td></td>
<td>parahippocampal cortices&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>PET</td>
<td>Positron Emission Tomography</td>
</tr>
<tr>
<td>PIQ</td>
<td>Performance Intelligence Quotient</td>
</tr>
<tr>
<td>RC</td>
<td>Reproducibility coefficient</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard Error Of The Mean</td>
</tr>
<tr>
<td>SPM</td>
<td>Statistical Parametric Mapping</td>
</tr>
<tr>
<td>SVC</td>
<td>Small volume correction</td>
</tr>
<tr>
<td>TE</td>
<td>Echo time (MR parameter)</td>
</tr>
<tr>
<td>TEA-Ch</td>
<td>Test Of Everyday Attention For Children</td>
</tr>
<tr>
<td>TR</td>
<td>Repetition time (MR parameter)</td>
</tr>
<tr>
<td>VBM</td>
<td>Voxel-based Morphometry</td>
</tr>
<tr>
<td>VIQ</td>
<td>Verbal Intelligence Quotient</td>
</tr>
<tr>
<td>VMA</td>
<td>Verbal Mental Age</td>
</tr>
<tr>
<td>WCST</td>
<td>Wisconsin Card Sorting Test</td>
</tr>
</tbody>
</table>

<sup>1</sup> See [Scharfman et al. 2000]
Appendix F References


Adolphs R, Cahill L, Schul R, Babinsky R. Impaired declarative memory for emotional material following bilateral amygdala damage in humans. Learning and Memory 1997; 4: 291-300.


Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. I. Acquisition of social knowledge is related to the prefrontal cortex. J Neurol 2000; 247: 72-73.


Bechara A, Tranel D, Damasio H, Damasio AR. Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. Cerebral Cortex 1996; 6: 215-225.


Buitelaar JK, van der Wees M, Swaab BH, van der Gaag RJ. Verbal memory and Performance IQ predict theory of mind and emotion recognition ability in children with...


References


Davis M. Diazepam and flurazepam: Effects on conditioned fear as measured with the potentiated startle paradigm. Psychopharmacology 1979a; 62: 1-7.

Davis M. Morphine and naloxone: Effects on conditioned fear as measured with the potentiated startle paradigm. Journal of Pharmacology 1979b; 54: 341-347.


Davis M, Redmond DEJ, Baraban JM. Noradrenergic agonists and antagonists: Effects on conditioned fear as measured by the potentiated startle paradigm. Psychopharmacology 1979; 65: 111-118.


DSM-IV. 299.00 Autistic Disorder DSM. 1994: 66-78.


References


References


References


Heberlein AS, Adolphs R, Tranel D, Kemmerer D, Anderson S, Damasio AR Impaired attribution of social meanings to abstract dynamic geometric patterns following damage to the amygdala Society for Neuroscience Abstracts 1998; 24: 1176.


Hitchcock J, Davis M. Lesions of the amygdala, but not of the cerebellum or red nucleus, block conditioned fear as measured with the potentiated startle paradigm. Behav.Neurosci. 1986; 100: 11-22.


Ionescu MD. Sex differences in memory estimates for pictures and words. Psychological Reports 2000; 87: 315-322.


Liebenauer LL, Slotnick BM. Social organisation and aggression in a group of olfactory bulbectomised male mice. Physiol Behav. 1996; 60: 403-409.


References


References


Owen AM, Evans AC, Petrides M. Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. Cerebral Cortex 1996a; 6: 31-38.


References


References


References

Robinson BW. Vocalisation evoked from the forebrain in Macaca mulatta. Physiol Behav. 1967; 2: 345-354.


References


Saver JL, Damasio AR. Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. Neuropsychologia 1991; 29: 1241-1249.


Shaffran, R. Children's Obsessive Compulsive Inventory. 2000. Ref Type: Personal Communication


References


Suzuki WA. Neuroanatomy of the monkey entorhinal, perirhinal and parahippocampal cortices: Organization of cortical inputs and interconnections with amygdala and striatum. Seminars in the Neurosciences 1996; 8: 3-12.


Teunisse, J. and Gelder, B. D. Impaired categorical perception of facial expressions in high-functioning adolescents with autism. 2001 Unpublished Work


Travis F. Cortical and cognitive development in 4th, 8th and 12th grade students The contribution of speed of processing and executive functioning to cognitive development. Biol.Psychol. 1998; 48: 37-56.


Tulving E. Memory and consciousness. Canadian Psychology 2001; 26: 1-12.


References


References


