J. Child Psychol. Psychiat. Vol. 32, No. 7, pp. 1081–1105, 1991
Printed in Great Britain.

0021-9630/91 $3.00 + 0.00
Pergamon Press plc


# Executive Function Deficits in High-Functioning Autistic Individuals: Relationship to Theory of Mind

Sally Ozonoff,* Bruce F. Pennington* and Sally J. Rogers†

*Abstract*—A group of high-functioning autistic individuals was compared to a clinical control group matched on VIQ, age, sex and SES. Significant group differences were found on executive function, theory of mind, emotion perception and verbal memory tests, but not on spatial or other control measures. Second-order theory of mind and executive function deficits were widespread among the autistic group, while first-order theory of mind deficits were found in only a subset of the sample. The relationship of executive function and theory of mind deficits to each other, and their primacy to autism, are discussed.

*Keywords:* Autism, executive function, theory of mind

## Introduction

Although autism was first described by Kanner (1943) nearly half a century ago, the nature of the core deficit(s) underlying this complex disorder is still unclear. The present study examined the relationship among three prominent, persistent deficits seen in autistic individuals, namely, emotion perception, theory of mind and executive function. Although previous studies have documented deficits in these domains, the present investigation is the first to examine all three in the same group of individuals. In addition, this study is the first to measure theory of mind abilities in non-retarded autistic subjects and the first to investigate executive functions in autistic children and adolescents.

Recent research has suggested that these three deficits are particularly promising in the search for the primary impairment(s) of autism. While other abilities, such as symbolic play (Wing, Gould, Yeates & Brierly, 1977) and pragmatic communication (Tager-Flusberg, 1981), also appear deficient in most autistic children, it has been argued that they are secondary to other, more fundamental, impairments (Baron-Cohen, 1988; Pennington & Ozonoff, in press; Tager-Flusberg, 1989).

*Accepted manuscript received* 9 *January* 1991

*University of Denver, Denver, Colorado, U.S.A.
†University of Colorado Health Sciences Center, Colorado, U.S.A.
*Requests for reprints to:* Dr Sally Ozonoff, Department of Psychology, 502 Social Behavioral Sciences, University of Utah, Salt Lake City, UT 84112, U.S.A.

In a disorder as complex and severe as autism, it may be misguided to search for one primary deficit capable of accounting for all manifestations of the syndrome. Instead, it may be more fruitful to consider autism a disorder of multiple primary deficits (Goodman, 1989). If a combination or configuration of deficits is primary to autism, we would expect that configuration to be specific to autism and universally found among autistic individuals. The present study examined whether emotion perception, theory of mind and executive function deficits might be components of this hypothesized primary configuration of deficits.

Deficits in emotion perception have been postulated to underlie the abnormal behavior of autism (see Hobson, 1989; Ozonoff, Pennington & Rogers, 1990, for a review of this topic). Several studies have suggested that autistic children are impaired at processing emotional cues (Hobson, 1986a, b; Langdell, 1981; Sigman, Ungerer, Mundy & Sherman, 1987; Weeks & Hobson, 1987). However, more recently, there have been several reports that emotion perception deficits may not be specific to autism, but may simply be a function of verbal mental age (Braverman, Fein, Lucci & Waterhouse, 1989; Ozonoff *et al.,* 1990; Prior, Dahlstrom & Squires, 1990).

While this may seem to be damaging evidence in the case for primacy of emotion perception deficits, it is not necessary that individual primary deficits be specific to autism. If autism is a disorder of multiple primary deficits, then what is important is that the configuration of deficits be specific to autism; it is not essential, however, that the individual deficits making up that configuration be specific. On the other hand, the lack of convergence across studies may indeed signal that emotion perception deficits are not primary to autism, but instead are correlates or secondary consequences of a more fundamental disability.

A good alternative candidate for this more fundamental impairment is the capacity to infer the mental states of others (e.g. their intentions, beliefs, desires); this ability has been referred to as a "theory of mind" (Premack & Woodruff, 1978). A deficit in this domain has been postulated to underlie autism and account for many symptoms of the disorder (Baron-Cohen, 1988; Frith, 1989; Pennington & Ozonoff; Tager-Flusberg, 1989).

Empirical support for theory of mind deficits has been provided by several recent investigations. In a series of elegantly designed studies employing a variety of paradigms, Baron-Cohen and colleagues have found that the ability of autistic children to attribute mental states to others is deficient. Baron-Cohen, Leslie and Frith (1985, 1986) found that 80% of autistic children were unable to correctly predict the beliefs of others, while most mentally retarded and normal controls of lower mental age were able to do so with facility. Baron-Cohen (1989a) reported that autistic children were also less able than controls to distinguish between mental and physical states, understand mental functions of the brain, and use their own mental states to judge aspects of the environment. Similarly, Harris and Muncer (1988) found that autistic children had difficulty predicting the desires of others when they ran counter to the subjects' own desires.

Finally, Baron-Cohen (1989b) investigated a group of autistic children who were able to make simple mental state attributions. When he administered a more difficult "second-order belief attribution" task, in which subjects had to predict what one person thought another person thought, all autistic subjects failed. In contrast, 90%

of normal controls and 60% of mentally retarded controls, both of lower mental age than the autistics, passed the test. This series of investigations suggests that theory of mind deficits are relatively specific to and universal among retarded autistic children; no studies of non-retarded or older autistic individuals have yet been performed.

Recently, several investigations within a different research tradition, that of cognitive neuropsychology, have identified a pattern of cognitive disabilities in autism that appear to reflect deficits in executive function. Executive function is defined as the ability to maintain an appropriate problem-solving set for attainment of a future goal; it includes behaviors such as planning, impulse control, inhibition of prepotent but irrelevant responses, set maintenance, organized search, and flexibility of thought and action.

Some features of autism are reminiscent of executive function deficits. The behavior of autistic people often appears rigid and inflexible: many autistic children become extremely distressed over trivial changes in the environment and insist on following routines in precise detail. They are often perseverative, focusing on one narrow interest or repetitively engaging in one stereotyped behavior. Their cognition often seems to lack executive functions; autistic individuals do not appear future-oriented, do not anticipate long-term consequences of behavior well, and have great difficulty self-reflecting and self-monitoring. They frequently appear impulsive, as if unable to delay or inhibit responses.

The first report of executive function deficits in autism came from a case study of an adult with residual state autism (Steel, Gorman & Flexman, 1984). On a neuropsychological test battery, the subject experienced great difficulty on measures of executive function; problem-solving strategies were reportedly extremely inflexible and perseverative. In contrast, much milder impairment was seen on tests of memory and language, while visuospatial and non-verbal analytic abilities were measured in the average to superior range.

The first empirical investigations of executive functions in a group of autistic subjects were undertaken by Rumsey and her collaborators (Rumsey, 1985; Rumsey & Hamburger, 1988, 1990). Rumsey (1985) compared the performance of a group of non-retarded, verbal men with documented childhood diagnoses of autism to sex-, age- and IQ-matched controls; she found that the autistic men were significantly more perseverative on the Wisconsin Card Sorting Test than the controls. Rumsey and Hamburger (1988) compared a similar group of high-functioning autistic men to matched controls and found that performance on the Wisconsin Card Sorting Test was again deficient, this time in terms of the number of categories achieved, as was the ability to complete a planned search task. In contrast to these executive function difficulties, several other abilities were only mildly affected, including language and memory, or entirely unaffected, such as visuospatial and sensory-perceptual capacities. Rumsey and Hamburger (1990) reported similar results when comparing autistic adults with severely dyslexic adults. Thus, several recent investigations have documented executive function deficits in autistic adults. The present study examined whether executive function impairment exists in younger autistic individuals, and if so, how it might be related to theory of mind deficits.

A group of high-functioning autistic people was compared to a clinical control group matched on verbal IQ, age, sex, race and socioeconomic status. A battery of emotion perception, theory of mind and executive function measures was administered, as well as discriminant tasks of memory and spatial ability. This design allowed us to address two questions: (1) do deficits in these domains exist in high-functioning autistic individuals, relative to matched controls? and (2) does one (or more) of the deficits appear primary to autism?

Given the strong empirical support for theory of mind deficits, we hypothesized that theory of mind impairment would be primary to autism. We felt that executive function deficits, if apparent, would be correlated symptoms arising from the same gross brain insult that caused theory of mind deficits, but involving a separate neurological system. Since correlated deficits are not found in all individuals with a given disorder, but vary according to the extent of brain damage, we predicted that executive function impairments would not be present in all of our autistic subjects. Since previous studies have only inconsistently found emotion perception impairments in autistic individuals, we hypothesized that they would be correlated deficits as well, present in only a subset of our sample. Finally, we predicted that the memory and spatial abilities of the autistic group would not be deficient compared to controls.

A subset of the autistic group met draft ICD-10 diagnostic criteria for Asperger's syndrome (World Health Organization, 1989), a disorder that is thought to represent a mild form of autism without associated mental retardation (Rutter & Schopler, 1987). One recent study found few differences between children with Asperger's syndrome and those with more classical high-functioning autism (Szatmari, Tuff, Finlayson & Bartolucci, 1990). Since no empirically based distinctions between Asperger's syndrome and high-functioning autism have yet been demonstrated (Schopler, 1985; Volkmar, Paul & Cohen, 1985), and since it has been argued that autism cannot be reliably distinguished from other pervasive developmental disorders (Fein, Waterhouse, Lucci & Snyder, 1985), there appeared to be justification for including both "subtypes" in the present study. Our sample size was large enough, however, that we were also able to divide the larger autistic group into two subgroups, one with more typical high-functioning autism, the other with Asperger's syndrome; the comparison of these subgroups is reported in the following companion paper (Ozonoff, Rogers & Pennington, 1991).

## Method

### *Subjects*

Two groups of subjects participated in the present study, an autistic group ($n = 23$) and a control group of non-autistic individuals ($n = 20$). The autistic subjects ranged in age from 8 to 20 years (mean CA = 12.05 years, SD = 3.19). To be included in the study, all autistic subjects had to obtain full-scale IQ (FSIQ) scores above 69 (prorated on five subtests of the WISC-R or WAIS-R). Three autistic individuals were excluded from the study because their FSIQs fell below this criterion. The mean FSIQ of the autistic group was 89.52 (SD = 15.17). Autistic subjects were recruited from a treatment program for autistic children at the University of Colorado Health Sciences Center, the Autism Society of America, the Denver Public School system and referrals by Denver-area clinicians. Table 1 contains descriptive characteristics of the sample.

Seventeen of the autistic subjects met DSM-III-R diagnostic criteria for autistic disorder (299.00). Six subjects met DSM-III-R criteria for pervasive developmental disorder not otherwise specified (PDDNOS; 299.80); each of the six manifested significant autistic-like symptoms, including avoidance

**Table 1. Descriptive characteristics of the sample**

| | Autistic group ($n = 23$)<br>$\bar{X}$ (SD)<br>range | Control group ($n = 20$)<br>$\bar{X}$ (SD)<br>range |
|---|---|---|
| CA | 12.05 (3.19)<br>8.08–20.92 | 12.39 (3.04)<br>8.58–19.50 |
| VIQ | 82.91 (18.08)<br>55–122 | 87.60 (17.79)<br>57–122 |
| PIQ | 98.35 (18.13)<br>68–140 | 97.00 (18.38)<br>63–140 |
| FSIQ | 89.52 (15.17)<br>70–128 | 91.30 (18.75)<br>55–134 |
| CARS | 34.65 (4.64)<br>25.0–42.5 | 18.40 (2.59)<br>15.0–23.0 |
| SES | 46.02 (11.71)<br>27.0–63.5 | 44.55 (14.69)<br>13.0–66.0 |
| SEX (M:F) | 21:2 | 18:2 |
| Hand laterality (R:L) | 17:6 | 19:1 |

of eye contact, difficulty with pragmatic communication, odd vocal inflection, abnormal use of non-verbal gestures, and peculiar interests and preoccupations. Since the aim of the present study was to test what deficit remained in high-functioning autistic individuals, we retained these subjects in the autistic group; if they introduced a bias into the test results, it would be a conservative bias that, if anything, would prejudice us against finding group differences.

All but two of the autistic subjects received a score in the autistic range (30 or above) on the Childhood Autism Rating Scale (CARS; Schopler, Reichler & Renner, 1986). The CARS scores of the two subjects who fell outside this range were higher than those of any controls. In addition, both of these individuals received scores indicating minimal to mild autism on the General Impressions Subscale of the CARS; in contrast, no control subject received a score in this range. The mean CARS score of the autistic group was 34.65 (SD = 4.64).

Twenty non-autistic individuals between 8 and 19 years of age comprised the clinical control group (mean CA = 12.39 years, SD = 3.04); diagnoses of controls included dyslexia, other learning disabilities, attention deficit hyperactivity disorder and mild mental retardation. The mean FSIQ of the control group was 91.30 (SD = 18.75). Subjects were recruited from the Special Education Department of the Denver Public School system and from the University of Denver Child Neuropsychology Laboratory. All controls obtained scores below 24 on the CARS, demonstrating that they did not fall in the range considered indicative of autism (mean CARS score = 18.40, SD = 2.59).

The 20 control subjects were matched with 20 of the 23 autistic subjects in a pairwise manner on chronological age (CA) and sex. The pairs were also matched on VIQ, which was prorated from three subtests of the WISC-R or WAIS-R (Information, Similarities and Vocabularly); these subtests were chosen because they load most highly on the verbal comprehension factor of the Wechsler tests (Sattler, 1988). The comprehension subtest of the WISC-R and WAIS-R is thought to measure social judgment and knowledge of social rules (Sattler, 1988); since we did not want to control for the very abilities the present study was attempting to examine, the Comprehension subtest was not used to prorate VIQ.

No matching control subjects were found for three of the autistic individuals. All three of these subjects obtained relatively low VIQ scores, but high PIQ scores; this pattern brought their FSIQs into the appropriate range for the study. However, due to the relative rarity of large verbal–performance IQ discrepancies in non-autistic individuals, no control subjects were found who matched the autistic subjects on VIQ and also met the minimum FSIQ criterion for inclusion in the study. Since the mean VIQ, PIQ and FSIQ of the autistic group ($n = 23$) and control group ($n = 20$) did not differ, the three additional autistic subjects were retained in the study, although they had no matching control subjects.

*Measures*

1. *Emotion perception domain*

*Emotion perception task.* A photograph of a target face displaying an emotional expression was presented to the subject; he was instructed to choose a second face that "felt the same way" from four choices. The stimuli were presented in a bound booklet that opened in an up-and-down manner. The target was mounted on the upper page and the four choices, one in each quadrant, were arranged on the lower page. The correct choice appeared an equal number of times in each quadrant to control for the effect of position on choice. Correct choices varied from the target in the identity of the model and the intensity of the expressed affect.

Half of the items ($n = 17$) contained distractors that shared similar perceptual features with correct choices; for example, a face expressing fear was used as a distractor for the target emotion of surprise, since both emotions share the perceptual feature of an open mouth. Nine emotions were depicted, four "simple emotions" (e.g. happiness, sadness, anger, fear) and five "complex emotions" (e.g. surprise, shame, disgust, interest, contempt). The photographs have been used extensively in studies of emotion perception (Izard, 1971, 1982).

The dependent variable was number correct and the maximum score possible was 34. Two training trials preceded the test items to teach task demands to subjects; training trials were not included in the number correct score. If subjects gave incorrect answers to either of the training items, the examiner corrected the subject. Regardless of the accuracy of responses to training items, verbal labels were provided for the emotions depicted in the training trials.

The internal consistency of the emotion perception task, as assessed by coefficient alpha, was .73. Task items varied in difficulty and thus were able to discriminate among a variety of ability levels.

2. *Theory of Mind Domain*

Details of the tasks and their scoring can be found in the primary sources cited below.

*Picture sequencing measure* (Baron-Cohen *et al.,* 1986). Subjects were given four pictures that were not in the correct order and instructed to arrange them so that the story made sense. Five different types of stories were used: the two "mechanical subscale" stories could be sequenced properly using only physical cause-and-effect knowledge, while the two "behavioral subscale" stories required knowledge of simple behavioral interactions to order correctly. Only the fifth "intentional subscale" required using mental state knowledge to correctly sequence the pictures. The measure was administered and scored as described by Baron-Cohen *et al.* (1986). The maximum score for each subscale was 6.

No information on reliability or validity of this measure was reported by Baron-Cohen *et al.* (1986). In the present study, the internal consistency of the five subscales was low, ranging from alpha = .33 for the behavioral 1 subscale to alpha = .61 for the behavioral 2 subscale; in general, alpha coefficients lower than .70 are not considered acceptable. On four of the five subscales, there was little variance in the difficulty of items, thus providing poor discrimination among subjects of low and high ability levels.

*Appearance–reality task* (Baron-Cohen, 1989a). This measure tested the ability of subjects to use their own mental states to judge aspects of the environment. Subjects were asked to distinguish between appearance and reality; they were given a deceptive object, such as a sponge painted to look like a rock, and asked, "What does this look like?" and "What is it really?" Baron-Cohen (1989a) argued that this test measures, among other things, the ability to attribute mental states to oneself: to pass this test, the subject must think, "This *looks like* a rock, but I *know* (mental state attribution) from handling it that it is really a sponge". The subject must then apply this mental state knowledge to override the object's superficial appearance.

This task was administered as described by Baron-Cohen (1989a). Subjects were shown four deceptive objects; thus, the range of scores on this test was 0 to 4. Errors were scored in three categories suggested by Baron-Cohen (1989a): (1) phenomenist errors, in which the subject was able to represent only the object's appearance and did not integrate his own knowledge into his answer (e.g. "it looks like a rock and it really is a rock"); (2) realist errors, in which the subject only attended to his knowledge of the object and did not simultaneously represent its appearance (e.g. "it looks like a sponge and it really is a sponge"); and (3) all other errors. If autistic individuals have difficulty representing mental state knowledge, we predicted that they would not only receive lower scores on the appearance–reality task than controls, but also produce more phenomenist errors, in which they failed to take into account their own mental states, than controls.

*Mental–physical distinction task* (Baron-Cohen, 1989a). This measure tested the autistic child's ability to make correct behavioral and sensory judgments about mental and physical events (e.g. physical states afford "behavioral-sensory" contact and can be seen and touched, while mental states do not). Baron-Cohen reasoned that if autism involves a difficulty appreciating mental states, autistic people should randomly attribute physical properties to mental and physical events.

Subjects were shown 4" x 6" color photographs of two children. They were then told stories in which one child had physical contact with an object, while the other had mental contact. For example, one story went as follows: "This is Kate. She likes cookies, but she is all alone, so she is thinking about a cookie. This is David. He likes cookies, so his mother gives him one". Four stories were told, each one dealing with a different mental state (thinking, remembering, dreaming, pretending). After each story, subjects were asked the behavioral contact question ("Who can eat the cookie?") and the sensory contact question ("Who can touch the cookie?"). They were also asked questions to control for memory of events and understanding of key words in the stories. The highest score possible on this test was 8, if the subject passed both the behavioral and sensory contact questions correctly for all four stories.

*Brain function task* (Baron-Cohen, 1989a). This task measured the ability to attribute mental, rather than behavioral, functions to the brain. Baron-Cohen reasoned that if autistic individuals understand mental phenomena poorly, they will be more likely to fail to understand any functions of the brain, or report purely behavioral functions, than mental age-matched controls. Subjects were asked to point to the location of the brain (using their own bodies as a reference) and asked "What does the brain do?" To control for general deficits in biological knowledge, they were also asked about the location and function of the heart.

On the brain function question, a score of 2 was given if the subject responded with a mental function (e.g. thinking, feeling, remembering, dreaming, learning), a score of 1 was given for a behavioral function (e.g. it moves your body, it makes you run) and a score of 0 was given for irrelevant or incorrect responses (e.g. it wiggles, it's squishy) or no response. The heart function question was scored in a pass/fail manner. A score of 1 was given if the subject indicated that the heart pushed blood through the body, kept you alive, etc.; incorrect responses were scored 0. Partially correct answers, such as "it beats", were further queried to see whether the subject understood the fundamental nature of the heart's role in pumping blood. If a correct response was given after the query, the subject was credited with a score of 1. The order of the brain and heart questions was counterbalanced across subjects.

*M&Ms false belief task* (Perner, Frith, Leslie & Leekam, 1989). Subjects were shown a box of M&Ms (sweets) and asked what it contained. All subjects replied "M&Ms". The box was then opened, to reveal that it actually contained a pencil. Subjects were then asked to predict what another child, who had never seen the box, would think it contained. A pass was scored if the subject responded "M&Ms" (or candy).

*Second-order belief attribution task.* This task was first developed by Perner and Wimmer (1985) and has been used with autistic children by Baron-Cohen (1989b). It is considered more difficult than the five theory of mind tasks described above; to pass the simpler tasks, subjects must attribute mental states to another person, making "first-order attributions". Second-order attributions, as in this task, require recursive thinking about mental states, in which subjects must predict one person's thoughts about another person's thoughts. While the previous first-order tasks begin to be mastered by age 4 in normally developing children (Flavell, Flavell & Green, 1983; Johnson & Wellman, 1982; Wellman & Estes, 1986), the ability to make second-order belief attributions develops later, closer to age 6 or 7 (Perner & Wimmer, 1985).

The test employed a toy town constructed from railroad miniatures. Subjects were told a story about two children, Mary and John, playing in a park (see Baron-Cohen, 1989b, for a depiction of the town and the story narrative). After hearing the story, subjects were asked to predict Mary's belief about John's whereabouts (the belief question) and then explain why Mary would hold this belief (the justification question). The belief question was scored in a pass/fail manner, after Baron-Cohen (1989b). The justification question was scored according to the order of mental state attributions made by the subject. If the subject made no mental state attributions, he received a score of 0. If mental states were attributed to only one character, a score of 1 was given. If the subject accounted for the mental states of both characters, a score of 2 was obtained.

The second-order belief attributions task was administered only if the subject passed three of the five simpler theory of mind tasks. If the measure was not administered, the subject received scores of zero on the two dependent measures, the belief question and the justification question.

3. *Executive function domain*

*Tower of Hanoi*. This task has been used to study the planning capacities of normal and retarded children and adults (Borys, Spitz & Dorans, 1982; Welsh, Pennington & Groisser, in press, a); efficient performance also requires the ability to inhibit prepotent, but irrelevant, responses. A deficit in this type of planning and inhibitory behavior has been demonstrated in frontal-damaged adults (Shallice, 1982).

Identical apparatuses (boards with three vertical pegs and three disks of different sizes and colours) were set up in front of the subject and the experimenter. The experimenter's disks were arranged on the experimenter's right-hand peg to form a tower, with the largest disk on the bottom and the smallest disk on the top. This disk arrangement represented the goal state that the subject was trying to achieve on all problems. The disks on the subject's board were placed on different pegs, in a prearranged sequence. The subject had to plan and carry out a sequence of moves that created a tower like that of the experimenter on the subject's left-hand peg. The primary rule governing the way in which the disks were moved was that larger disks could not be placed on top of smaller disks. All subjects exhibited their understanding of the task's rules by demonstrating legal and illegal moves before continuing on to the practice problem.

The Tower of Hanoi has been used successfully with preschoolers when a cover story explaining the task demands was used (Welsh, Pennington, Ozonoff, Rouse & McCabe, 1990). Specifically, the story described a ''jungle'', in which the pegs were called ''trees'', the disks were called ''monkeys'' and the sequence of moves was described as ''helping the monkeys jump from tree to tree''. In the present study, this cover story was used to reduce the complexity of task demands for subjects younger than age 10 or older subjects with IQs in the borderline or mildly retarded range.

On the practice problem, the subject was given a simple problem to complete; if he had difficulty, he was guided through the problem's solution and the rules of the task were reiterated. The practice item was not included in the final score.

Subjects were then administered several problems of increasing difficulty; difficulty was reflected in the minimum number of moves necessary to match the experimenter's disk configuration. Subjects were required to complete the problem correctly for two consecutive trials to receive credit; they were given six trials to achieve this criterion for each problem. Testing continued until the subject was unable to complete a problem within the six trials.

The scoring system for the Tower of Hanoi was developed by Borys *et al.* (1982). If a problem was solved in the first two trials, a score of 6 was given; point totals decreased with the number of trials required to solution, with only one point given for solution on trials five and six. The dependent variable was a ''planning efficiency'' score that was the sum of scores received on problems of differing move lengths.

*Wisconsin Card Sorting Test (WCST)*. This executive function task, originally developed by Grant and Berg (1948), measures set maintenance skills, the ability to flexibly modify incorrect strategies and the ability to inhibit prepotent, but incorrect, responses. It is able to discriminate adults with frontal lesions from those with non-frontal cerebral damage (Robinson, Heaton, Lehman & Stilson, 1980) and has been used successfully with children (Chelune & Thompson, 1987; Welsh *et al.*, in press).

Four cards, varying along the dimensions of color, shape and number, were placed in front of the subject. Subjects were given two decks of cards that varied along these same dimensions and asked to match the cards in the deck with one of the four "key" cards. The experimenter told the subject if he had placed a card correctly or incorrectly, but did not reveal the sorting strategy to the subject. Once the subject had categorized 10 consecutive cards correctly, the sorting principle was changed without the subject's knowledge. The previous sorting strategy then received negative feedback and the subject was expected to switch to the new categorization principle.

This measure provided several dependent variables. The number of categories the subject completed and the number of errors made were calculated as measures of conceptual problem-solving that highlighted the subject's understanding of the nature of the categorization task. The number of perseverative responses, in which the subject continued sorting by a previously correct category despite negative feedback (Heaton, 1981), was also calculated; this score is the best predictor of prefrontal dysfunction derived from the WCST (Heaton, 1981). Finally, the failure to maintain set variable was calculated; it was defined as the number of times the subject made five correct responses in a row but failed to sort the 10 cards necessary to complete a category. When this happened, the subject had shown insight into the correct sorting principle and had almost certainly had unambiguous correct responses reinforced, but yet had not been able to maintain the correct problem solving set consistently.

4. *Discriminant domain*

*Buschke Selective Reminding Test.* This verbal learning and memory test (Buschke, 1974) required subjects to learn a list of 10 animals. After the list was read aloud at a 2-second rate, subjects were prompted to recall as many items as possible. They were then reminded of words they failed to name and asked to recall all items of the list a second time. Testing was continued until the entire list was recalled on two consecutive trials or until trial 10 was reached. Two dependent variables, long-term storage and consistent long-term retrieval, were calculated, as described by Buschke (1974). Rumsey and Hamburger (1988) found that performance on this measure was preserved in high-functioning autistic adults.

*Children's Embedded Figures Test (CEFT).* This test measured visuospatial ability and attention to perceptual detail (Witkin, Oltman, Raskin & Karp, 1971). Previous studies have documented good reliability and validity (Hyde, Geiringer & Yen, 1975; Vernon, 1972). The CEFT was used as a discriminant task in the experimental battery of the present study, since it is well established that visuospatial abilities tend to be preserved in autism (Dawson, 1983; Fein *et al.,* 1985; Prior, 1979). One study has shown that performance on the CEFT is unimpaired in autistic children (Shah & Frith, 1983).

The CEFT required subjects to identify a simple geometric form embedded in a complex picture. Subjects were given cardboard models of the shapes they were to identify in the pictures. On all trials, the cardboard models remained present and subjects were allowed to lay the cut-outs on the hidden shapes, point to the embedded targets or trace around them with a finger. The dependent variable was number of correct responses (the maximum score was 25 and did not include the training trials).

*Procedure*

All subjects were tested at the University of Denver Child Neuropsychology Laboratory. Informed consent was obtained from both parents and subjects participating in the study. Occupational and educational status of parents was recorded to calculate SES (Hollingshead, 1975). The first author tested 60% of the subjects in the autistic group and 50% of the subjects in the control group. The remainder of the subjects were tested by one of two trained research assistants, who were blind to the diagnoses of the subjects and the hypotheses of the experiment.

All subjects in the autistic group were diagnosed as meeting DSM-III-R diagnostic criteria for either autistic disorder or PDDNOS by the first author. Subjects who were not directly tested by the first author in the present study ($n = 9$) had been seen by her formerly, either clinically or during participation in previous studies, and had been diagnosed at that time.

The WISC-R or WAIS-R was administered to all subjects in the autistic group. Control subjects had all been given the appropriate Wechsler test within 2 years of participation in the study and thus it was not re-administered. Measures of the experimental battery were then presented in one of two predetermined orders; the order was counterbalanced across subjects. The full testing battery took approximately 3 hours; the attention span and cooperation of all subjects was sufficient that testing was conducted in one session, with time for several breaks. Subjects were paid US $50 for participation in the study. Although all subjects were informed that they could discontinue testing at any time, all chose to complete the study.

### *Statistical analyses*

To reduce the number of statistical comparisons, composite scores were calculated to represent functioning in the different neuropsychological domains. The first-order theory of mind (TOM1) composite included tests that required only first-order belief attributions (e.g. the mental-physical distinction task, the appearance–reality task, the M&Ms false belief task, the brain function task and the intentional condition of the picture sequencing measure). The second-order theory of mind (TOM2) composite contained items requiring second-order belief attributions (e.g. the belief question and the justification question of the second-order belief attribution task).

The executive function (EF) composite was made up of the Tower of Hanoi efficiency score, the WCST number of perseverations and the WCST number of failures to maintain set (coded in the direction in which more failures to maintain set indicated better ability to shift set and thus, better executive functions). The emotion perception (EP) composite included only the number correct on the emotion perception measure. The verbal memory (V. MEM) composite contained the Buschke long-term storage and consistent long-term retrieval scores. The spatial (SPAT) composite included the Children's Embedded Figures Test and the WISC-R or WAIS-R Block Design and Object Assembly subtests.

Composite scores were constructed by converting raw scores on each test to $T$-scores (mean = 50, SD = 10), using the grand mean and standard deviation across both groups. The $T$-scores for the different measures within the domain were then averaged to obtain composite scores of average performance within the domain. These composite scores were the primary variables used to answer the questions of the present study.

$t$-Tests and chi-square tests were used to answer Question 1 concerning group differences. All $t$-tests reported in tables and text were independent (non-paired) sample $t$-tests ($n = 43$, df = 41).

To answer Question 2 regarding primacy, two types of analyses were undertaken. First, to examine the universality of deficits among the autistic subjects, the proportion of subjects scoring below the mean of the control group was calculated. Second, a discriminant function analysis was performed to determine which deficit was best able to discriminate between the groups and correctly classify subjects.

## Results

### *Preliminary analyses*

One-way analyses of variance were performed on all measures to examine possible effects of identity of tester and order of administration on performance. No significant main effects of tester or order were found. Consequently, data were collapsed across tester and order conditions for all subsequent analyses.

As described in the method section, six subjects in the present study met DSM-III-R criteria for PDDNOS, but not autistic disorder; all analyses were run both with and without these subjects. While the values of the test statistics and $p$-levels changed slightly, there were no differences in the statistical significance of the results when the subjects were included versus when they were excluded: all significant group differences remained significant, as did all non-significant differences. Thus, there did not appear to be justification for removing these subjects from the autistic group and they were retained for all statistical analyses.

### *Analyses to examine existence of deficits (Question 1)*

*Group differences on composites.* There were no significant differences between the groups on chronological age, sex, VIQ, PIQ, FSIQ, SES, race or handedness. As expected, the CARS score of the autistic group was significantly higher than that of the control group [$t(41) = 14.41$, $p < .0001$].

The autistic group performed significantly less well than the control group on the executive function composite [$t(41) = -5.58$, $p < .0001$], the first-order theory of mind composite [$t(41) = -2.51$, $p < .02$], the second-order theory of mind composite [$t(41) = -4.30$, $p < .0001$], the emotion perception composite [$t(41) = -3.12$, $p < .01$] and the verbal memory composite [$t(41) = -3.97$, $p < .0001$]. The groups did not differ on the spatial composite [$t(41) = -.38$, $p = .71$]. Table 2 contains means, standard deviations and group differences on the composite scores.

**Table 2. Performance data and group differences on composites**

| | Autistic group ($n = 23$) $\bar{X}$ (SD) | Control group ($n = 20$) $\bar{X}$ (SD) | $t$ (41) | $p$ |
|---|---|---|---|---|
| EF | 45.2 (7.0) | 55.4 (4.9) | − 5.58 | < .0001 |
| TOM1 | 47.5 (9.6) | 52.9 (3.5) | − 2.51 | .018 |
| TOM2 | 45.3 (6.1) | 55.4 (8.9) | − 4.30 | < .0001 |
| EP | 46.0 (10.5) | 54.5 (7.2) | − 3.12 | .003 |
| V. MEM | 45.5 (10.1) | 55.1 (5.3) | − 3.97 | < .0001 |
| Spatial | 49.5 (8.5) | 50.5 (8.4) | − 0.38 | .707 |

*Note:* All values are expressed in $T$-scores.

*Group differences on individual measures.* Independent sample $t$-tests ($n = 43$) and chi-square tests ($n = 43$) were performed to examine group differences on the individual measures making up the composite scores; performance data on the individual measures are presented in Table 3.

In the executive function domain, the groups differed significantly on the Tower of Hanoi efficiency score [$t(41) = -6.26$, $p < .0001$] and the WCST number of perseverative responses [$t(40) = 3.90$, $p < .001$], with the autistic group performing more poorly than the control group, as predicted. Although the mean number of failures to maintain set for both groups were within the average range according to published norms (Chelune & Baer, 1986), the control group had significantly *more* failures to maintain set than the autistic group [$t(41) = -2.11$, $p < .05$]. While this result is not surprising, given the strong perseverative tendencies of the autistic group, it is not an inevitable consequence of the WCST scoring procedures, since these two variables are logically independent. This result is further evidence that the executive function strategies of the two groups were different.

In contrast, there were no significant group differences on the WCST number of categories and number of errors; since these variables were thought to measure conceptual understanding of the task demands, rather than executive functions, group differences were not expected. Thus, the marked perseveration of the autistic subjects

**Table 3. Performance data and group differences, individual measures**

| | Autistics $\bar{X}$ (SD) | Controls $\bar{X}$ (SD) | Test statistic | $p$ |
|---|---|---|---|---|
| *Executive function* | | | | |
| WCST psv | 67.32 (46.5) | 24.75 (20.5) | $t = 3.90$ | .001 |
| WCST ftms | 0.52 (0.85) | 1.35 (1.57) | $t = -2.11$ | .044 |
| WCST categories | 3.25 (2.49) | 3.95 (1.96) | $t = -.98$ | .340 |
| WCST errors | 57.70 (31.5) | 46.30 (24.9) | $t = 1.21$ | .242 |
| Tower of Hanoi | 19.78 (9.85) | 37.45 (8.65) | $t = -6.26$ | $< .001$ |
| *Theory of mind 1* | | | | |
| Mental–physical | 5.87 (1.85) | 7.35 (1.18) | $t = -3.18$ | .003 |
| MP control | 6.74 (1.63) | 7.85 (0.49) | $t = -3.11$ | .004 |
| MP comprehension | 9.48 (1.41) | 10.00 (0.0) | $t = -1.77$ | .090 |
| Brain function | 1.57 (0.79) | 1.60 (0.50) | $\chi^2 = 10.1$ | .007 |
| Heart function | 0.91 (0.29) | 0.95 (0.22) | $\chi^2 = 0.23$ | .630 |
| M&Ms task | 0.65 (0.49) | 0.95 (0.22) | $\chi^2 = 6.46$ | .011 |
| Appearance–reality | 2.87 (1.60) | 3.90 (0.31) | $t = -3.02$ | .006 |
| Phenomenist errors | 0.26 (0.54) | 0.05 (0.22) | $\chi^2 = 2.73$ | .098 |
| Realist errors | 0.08 (0.92) | 0.05 (0.22) | $\chi^2 = 0.23$ | .630 |
| PS mechanical 1 | 5.61 (0.84) | 6.00 (0.0) | $t = -2.24$ | .036 |
| PS mechanical 2 | 5.26 (1.14) | 5.90 (0.45) | $t = -2.48$ | .019 |
| PS behavioral 1 | 4.13 (1.60) | 5.30 (0.92) | $t = -2.98$ | .005 |
| PS behavioral 2 | 4.43 (1.84) | 5.30 (1.03) | $t = -1.77$ | .085 |
| PS intentional | 3.39 (1.99) | 3.90 (1.97) | $t = -0.84$ | .406 |
| *Theory of Mind 2* | | | | |
| Belief question | 0.26 (0.45) | 0.65 (0.49) | $\chi^2 = 6.73$ | .009 |
| Justification question | 0.30 (0.47) | 1.25 (0.72) | $\chi^2 = 20.9$ | $< .001$ |
| *Emotion perception* | | | | |
| EP task | 20.17 (4.49) | 23.80 (3.09) | $t = -3.12$ | .003 |
| *Verbal memory* | | | | |
| Long-term storage | 7.47 (1.64) | 8.77 (0.87) | $t = -3.28$ | .002 |
| Consistent LTR | 4.80 (3.05) | 7.84 (1.62) | $t = -4.16$ | $< .001$ |
| *Spatial* | | | | |
| CEFT | 16.35 (5.77) | 18.75 (5.18) | $t = -1.44$ | .158 |
| Block design | 10.00 (4.00) | 9.40 (4.07) | $t = 0.49$ | .630 |
| Object assembly | 9.26 (3.92) | 9.30 (3.67) | $t = -0.03$ | .973 |

*Note:* All $t$-tests have 41 df, $\chi^2$ tests on $n = 43$.

cannot be attributed to a general performance deficit and instead appears to reflect specific difficulty shifting cognitive set.

In the first-order theory of mind domain, the autistic group performed significantly less well on the mental–physical distinction task [$t(41) = -3.18$, $p < .01$], the appearance–reality task [$t(41) = -3.02$, $p < .01$], the brain function task [$\chi^2$ (2, $N = 43) = 10.06$, $p < .01$], and the M&Ms false belief task [$\chi^2$ (1, $N = 43) = 6.46$, $p < .01$], as predicted. To demonstrate that these deficits were specific to theory of mind knowledge and did not just reflect general difficulties with the tasks, it was important to demonstrate that the groups performed equally well on control tasks.

As evidence of this, the groups did not differ in their ability to explain the function of the heart, which was a control question for the brain task; there were also no significant group differences on the comprehension questions of the mental–physical distinction task, which were used to control for understanding of the lexical terms used in the test. The control group did perform significantly better than the autistic group on the memory control questions of the mental–physical distinction task, however [$t(41) = -3.11$, $p < .01$]. Also against expectation, there were no group differences on the intentional subtest of the picture sequencing measure, while there were significant group differences on the mechanical 1 subtest [$t(41) = -2.24$, $p < .05$], the mechanical 2 subtest [$t(41) = -2.48$, $p < .05$] and the behavioral 1 subtest [$t(41) = -2.98$, $p < .01$], in each case with the autistic group performing less well than the control group.

On the appearance–reality task, the autistic group tended to make more phenomenist errors than the controls [$\chi^2$ (1, $N = 43) = 2.73$, $p < .10$]. However, since this difference did not reach statistical significance, these results only marginally supported our prediction that autistic individuals would be more likely than controls to fail to use their own mental states in answering appearance–reality questions. There were no group differences in the rate of realist errors, as predicted.

In the second-order theory of mind domain, the autistic group performed significantly less well than the control group on both the belief question [$\chi^2$ (1, $N = 43) = 6.73$, $p < .01$] and the justification question of the second-order belief attribution task [$\chi^2$ (2, $N = 43) = 20.90$, $p < .0001$].

There were also significant group differences on the emotion perception task [$t(41) = -3.12$, $p < .01$]. The percentage of errors that the autistic group made on items with obvious perceptual foils was significantly greater than the percentage made by controls [$t(41) = 2.07$, $p < .05$]; this suggests that the autistic subjects were using a different, more perceptually driven, matching strategy than the controls. The autistic group made significantly more errors in matching both simple [$t(41) = -2.41$, $p < .05$] and complex [$t(41) = -2.76$, $p < .01$] emotions than controls. A repeated measures analysis of variance revealed that there were no significant group by item type (simple or complex emotion) interaction effects.

Within the discriminant domain, contrary to our hypothesis, the control group performed significantly better than the autistic group on the Buschke long-term storage [$t(41) = -3.28$, $p < .01$] and consistent long-term retrieval scores [$t(41) = -4.16$, $p < .001$]. As expected, in the spatial domain, the two groups performed similarly on the CEFT and the Block Design and Object Assembly subtests of the WISC-R (or WAIS-R).

To summarize, these results demonstrated that the autistic group performed significantly less well than their matched controls in all domains predicted, as well as in the verbal memory domain, which was not expected. Since the autistic group did not perform more poorly than controls on all tasks, was matched with controls on IQ, had intact spatial abilities and performed well on most control tasks, this differential pattern of performance suggested that the autistic group had selective impairments in these domains.

*Analyses to examine primacy (Question 2)*

*Universality of deficits*. To examine whether deficient performance by the autistic group was determined by the very poor performance of a few individuals or was more widespread among autistic subjects, the proportion of the autistic group performing more poorly than the control group mean was calculated.

On the composites, the proportions of autistic subjects performing more poorly than the control group mean were: 96% on the executive function composite, 52% on the first-order theory of mind composite, 87% on the second-order theory of mind composite, 65% on the emotion perception composite and 78% on the verbal memory composite.

Exploratory chi-square goodness-of-fit tests revealed that significantly more autistic subjects performed below the mean of the control group on the executive function and second-order theory of mind composites than on the other composites. Thus, these two deficits were significantly more widespread among the autistic group than the other deficits.

While deficits on the executive function and second-order theory of mind composites were equally widespread among autistic subjects, second-order theory of mind deficits appeared less specific to the autistic group than executive function deficits. In the second-order theory of mind domain, while 56% of the autistic group failed all tasks, so did 15% of the control group, suggesting that deficits in this area were not specific to autism. In the executive function domain, on the other hand, nearly two standard deviations separated the lowest scoring autistic subject from the lowest scoring control.

*Discriminant analysis*. A discriminant function analysis was performed to evaluate how well the groups could be empirically distinguished from each other on the basis of test performance. When performance on the six composites was entered in the analysis, an overall classification accuracy rate of 87.5% was achieved; the group membership of 90% of the autistic subjects and 85% of the control subjects was correctly predicted by the discriminant function. The executive function and second-order theory of mind composites were most highly correlated with the function and, thus, best able to discriminate between the groups (see Table 4). The multivariate Wilks was highly significant (lambda = .47, $p = .0002$).

**Table 4. Correlations of composites with discriminant function**

| Composite | $r$ |
|---|---|
| Executive function | .75 |
| Second-order theory of mind | .60 |
| Verbal memory | .51 |
| First-order theory of mind | .45 |
| Emotion perception | .28 |
| Spatial | .02 |

*Relationships among variables*

To examine interrelationships among the variables, the composite scores were correlated with each other and with the independent variables using Pearson product-moment correlation coefficients. Table 5 contains the intercorrelation matrix of the autistic group, while Table 6 contains that of the control group.

**Table 5. Intercorrelation matrix, autistic group ($n = 23$)**

| | EF | TOM1 | TOM2 | EP | V.MEM | Spatial |
|---|---|---|---|---|---|---|
| TOM1 | .64**** | | | | | |
| TOM2 | .50** | .51** | | | | |
| EP | .34 | .44** | .14 | | | |
| V.MEM | .58*** | .61*** | .42** | .17 | | |
| Spatial | .43** | .04 | .07 | .21 | .17 | |
| CARS | −.71**** | −.73**** | −.57*** | −.25 | −.73**** | −.07 |
| VIQ | .75**** | .74**** | .68**** | .28 | .73**** | .20 |
| PIQ | .42** | .03 | .02 | .21 | −.02 | N/A |
| FSIQ | .76**** | .51** | .48** | .32 | .46** | N/A |
| CA | −.16 | −.10 | .06 | .04 | .22 | −.01 |
| SES | .21 | .20 | .40* | −.16 | .35 | −.03 |

*Note:* Correlation coefficients not computed between spatial composite and PIQ or FSIQ because they contained variables in common (Block Design and Object Assembly).

*$p < .10$; **$p < .05$; ***$p < .01$; ****$p < .001$.

**Table 6. Intercorrelation matrix, control group ($n = 20$)**

| | EF | TOM1 | TOM2 | EP | V.MEM | Spatial |
|---|---|---|---|---|---|---|
| TOM1 | .13 | | | | | |
| TOM2 | .39* | .44** | | | | |
| EP | .25 | .70**** | .46** | | | |
| V.MEM | −.13 | −.11 | .07 | .31 | | |
| Spatial | .29 | .40* | .33 | .43* | .06 | |
| CARS | −.16 | −.34 | −.46** | −.29 | −.27 | −.45* |
| VIQ | .35 | .43* | .50** | .48** | .20 | .78**** |
| PIQ | .12 | .41* | .38* | .42* | .07 | N/A |
| FSIQ | .25 | .45** | .47** | .47** | .15 | N/A |
| CA | .11 | −.21 | −.29 | .16 | .30 | −.33 |
| SES | .32 | −.03 | .12 | .09 | .01 | .57*** |

*Note:* Correlation coefficients not computed between spatial composite and PIQ or FSIQ because they contained variables in common (Block Design and Object Assembly).

*$p < .10$; **$p < .05$; ***$p < .01$; ****$p < .001$.

In general, the patterns of intercorrelation differed for the two groups. In the autistic group, the executive function composite was significantly correlated with most of the other composites; the first- and second-order theory of mind and verbal memory composites were also highly intercorrelated. The spatial composite was significantly related only to the executive function composite. Finally, the emotion perception composite was significantly correlated with the first-order theory of mind composite, but not with the second-order composite; thus, in the autistic group, these three abilities within the more general domain of social cognition were not as related as might have been expected.

In the control group, on the other hand, there were very few significant intercorrelations among the composites (with one important exception; see below). Thus, for controls, the composites appear to measure separate and distinct dimensions of performance and do not simply reflect a general intellectual factor.

The exception to the general lack of correlation in the control group was that the three social cognition composites (emotion perception, first- and second-order theory of mind) were significantly intercorrelated in the control group, in contrast to the autistic group. Thus, the control group appeared to be more homogeneous and integrated in its functioning in the social cognition domain than the autistic group.

When the composites were correlated with the independent variables, different patterns across the two groups again emerged. CARS scores were highly correlated with performance on the composites in the autistic group, but not in the control group. Performance on the executive function and verbal memory composites was independent of FSIQ in the control group, while this was not the case in the autistic group. There were no significant correlations with SES in the autistic group and only one (with the spatial composite) in the control group.

Since one goal of the present study was to examine the relationship of executive function and theory of mind abilities, intercorrelations among these constructs were examined more closely. First, the two theory of mind composites were combined to make one composite score. Given the high correlations of these composites with FSIQ, this variable was then partialled out of the correlation. Within the autistic group, the correlation between the executive function and theory of mind composites, covarying FSIQ, was significant [$r(20) = .46$, $p < .05$]. However, within the control group, the correlation between the executive function and theory of mind composites, controlling for FSIQ, was non-significant [$r(17) = .05$, $p = .84$].

In summary, the results of the intercorrelational analyses suggest that abilities that are unrelated in our non-autistic control group (e.g. executive function, social cognition and IQ) come to be associated in autistic individuals. It appears that something about the etiology of autism may cause normally independent dimensions of functioning to become highly related and interdependent.

## Discussion

### *Question 1, existence of deficits*

This study found that high-functioning autistic individuals have selective deficits in executive function, theory of mind, emotion perception and verbal memory. In contrast, they performed as well as controls on spatial tests, IQ measures and most control tasks (e.g. mental–physical comprehension questions, heart function task, WCST number of categories and number of errors). These results demonstrated that the autistic subjects were capable of understanding the demands of the tests and sustaining attention and motivation to complete them to the best of their ability.

This is particularly important, as an autism-specific deficit in motivation could account for some of the present findings, particularly those in the executive function domain. For example, one could argue that relatedness deficits caused autistic subjects to disregard the examiner's feedback on the WCST; thus, their poor score might reflect social or motivational factors more than executive function impairment. However, since the autistic group performed as well as controls on two of the four variables derived from the WCST (the two ''non-executive function'' variables), as well as on most control tasks within the other domains, the motivational argument is not a compelling one.

The present results replicate and extend the findings of previous studies. Earlier investigations focused on theory of mind abilities in autistic individuals who were also mentally retarded; it was not known if autistic subjects with relatively good cognitive capacities would exhibit the same deficits. We found that many of the present subjects evidenced great difficulty inferring the mental states of others, despite the fact that their FSIQs fell in the non-retarded range of functioning.

The notion that autistic individuals have some degree of executive function impairment is relatively new. Until recently, only three empirical investigations had included measures of executive function (Rumsey, 1985; Rumsey & Hamburger, 1988, 1990). More recently, Szatmari and colleagues (Szatmari, Bartolucci, Bremner, Bond & Rich, 1989) reported that performance on the WCST was among the most sensitive predictors of outcome in a group of high-functioning autistic men. The present study extended these findings by examining executive functions in children and adolescents, rather than adults, and in finding group differences across three measures of executive function (Tower of Hanoi planning efficiency, WCST perseverations, WCST failures to maintain set), rather than just one measure. These deficits, occurring as they did in individuals of normal intelligence, in comparison with controls stringently matched on age, VIQ, sex, race and SES, were all the more striking.

While, in general, the present results are consistent with previous work, this study did fail to replicate some earlier investigations. For example, the finding of significantly poorer emotion perception abilities in this sample of autistic individuals was somewhat surprising, as several previous studies have failed to document emotion perception deficits when autistic children were matched with controls on verbal mental age (Braverman *et al.*, 1989; Ozonoff *et al.*, 1990; Prior *et al.*, 1990). Thus, the relation of emotion perception deficits to both autism and verbal mental age must be further studied.

The present investigation also failed to replicate Baron-Cohen *et al.*'s (1986) findings that autistic children performed more poorly than controls on the intentional condition of the picture sequencing measure. Rather, we found that our autistic subjects performed more poorly on the control conditions that did not require inference of mental states (e.g. the mechanical 1, mechanical 2 and behavioral 1 conditions), but not on the critical intentional condition requiring mental state attributions. Recently, Oswald and Ollendick (1989) also failed to find group differences on this measure. These two non-replications, as well as the poor psychometric properties of the task, call into question its utility as a measure of mental state knowledge. Further research is necessary to more fully understand the potential assets and problems of this task.

Finally, the results of performance on the Buschke Selective Reminding Test failed to replicate the findings of Rumsey and Hamburger (1988) that high-functioning autistic men were not impaired compared to controls on long-term storage or consistent long-term retrieval of verbal material. Several explanations are possible for this non-replication. First, the verbal component of the Buschke may complicate the task for autistic subjects, who, despite being matched with controls on VIQ, appeared to have greater difficulty initiating verbal responses and completing tasks that were verbally demanding. Thus, poor performance on the Buschke may have measured specific verbal or social deficits, such as verbal initiative and communicative intent, more than it measured memory problems. This explanation, however, does not account for the success of the subjects in the Rumsey and Hamburger (1988) study.

A second possibility is that deficits on the Buschke may be secondary to executive function deficits. Previous studies have found impairment in consistent long-term retrieval in patients with frontal lobe lesions (Luria, 1981) and it has been suggested that selective reminding procedures require executive function mediation (Delis, Kramer, Kaplan & Ober, 1987; Loring & Papanicolau, 1987). However, this explanation does not account for why the subjects of Rumsey and Hamburger's (1988) study were proficient on this task, despite exhibiting executive function impairments.

A final explanation, which could coexist with either of the two above, is that autistic children "grow out of" their verbal memory deficits with development; improvement with age is seen in many of the behaviors and deficits associated with autism. Support for this possibility was provided by our finding that verbal memory abilities were highly positively correlated with mental age, suggesting that deficits undergo a developmental progression and improve with age. This explanation would also help integrate our findings with those of Boucher and Lewis (1989), who found verbal memory impairments in autistic children with a mean MA of 7 years, and Rumsey and Hamburger (1988), who found no impairments on the Buschke in autistic adults with a mean MA of 27 years.

*Question 2, pattern of deficits and primacy*

Three criteria can be used to evaluate the primacy of a deficit. They are: (1) universality of the deficit among autistic individuals; (2) specificity of the deficit to autism; and (3) causal precedence of the deficit in development. The present study was not able to evaluate the final criterion, but provided some information regarding the first two.

The specificity criterion is the least important of the three, particularly if we expect to find multiple primary deficits accounting for autism, rather than just one deficit. While we know from previous research that emotion perception (Novic, Luchins & Perline, 1984) and executive function deficits (Chelune, Ferguson, Koon & Dickey, 1986) are not specific to autism, either might still be primary to the disorder, in combination with other deficits. Results of the present study also suggested that theory of mind impairment may not be specific to autism, as deficits in this domain were found among the control group as well.

The question of universality was examined by calculating the proportion of autistic subjects performing below the mean of the control group; this analysis explored whether deficient performance was widespread among the autistic subjects or present in only a few individuals. Our original hypothesis was that theory of mind impairment was primary to autism, while executive function impairment was a correlated deficit on the basis of neuroanatomic proximity; thus, we predicted that deficits on both first- and second-order theory of mind tasks would be relatively universal among the autistic group, while executive function deficits would be present in only a subset of the group, those with more severe brain dysfunction. The present results did not support this overall view, however. While executive function and second-order theory of mind deficits were significantly more widespread among the autistic subjects than the other deficits, impairment on first-order theory of mind tasks was present in only a subset of the group.

In addition, as discussed in the following companion paper (Ozonoff *et al.*, 1991), executive function deficits were the only deficits found in both subjects with classic high-functioning autism and those with Asperger's syndrome. If we believe that both subtypes lie on the autistic continuum, as most researchers do (Schopler, 1985; Szatmari *et al.*, 1990; Volkmar *et al.*, 1985; Wing, 1981), then we would expect that they would both share the same primary deficit(s). The universality of executive function deficits in the present sample suggests that it might be a primary deficit of autism.

How might executive function deficits be related to the deficits in theory of mind found in a subset of the present sample and in previous studies of lower-functioning samples (Baron-Cohen *et al.*, 1985, 1986; Baron-Cohen, 1989a, b)? Several possibilities exist, among which are: (1) one deficit is primary and causes the other, which is secondary; (2) one deficit is primary, but does not cause the other, which is a correlated deficit caused by brain damage to a neuroanatomically proximal system; (3) a third deficit is primary and causes both the executive function and theory of mind impairments; and (4) both executive function and theory of mind deficits are independent primary deficits of autism.

Regarding the first possibility, poor performance on the theory of mind tasks could be accounted for by the perseverative, inflexible problem-solving strategies characteristic of executive function impairment. For example, on the M&Ms false belief task, if autistic individuals perseverate or inappropriately focus on what *they* know is in the box, they would respond incorrectly when asked what another person would think was in the box. Like prefrontal patients with executive function deficits, who can verbalize the correct sorting strategy on the WCST, but not act on it behaviorally, autistic individuals might be aware of what another person would think was in the M&Ms box, but be unable to overcome their perseverative tendencies to indicate this knowledge to the examiner. A similar argument can be made for the appearance–reality measure, in which cognitive inflexibility and difficulty shifting set may account for what appears to be deficient mental state knowledge.

One difficulty with this argument, however, is that the Asperger's syndrome subgroup had executive function deficits, yet did not exhibit impairment in the theory of mind domain (see Ozonoff *et al.*, 1991). If one deficit causes the other, we would expect them to co-occur. It is also unlikely that theory of mind deficits cause executive function impairment, since only a subset of subjects demonstrated theory of mind deficits, but all were impaired on executive function tasks.

Perhaps theory of mind deficits are correlated with executive function deficits, but neither one causes the other. They are both caused by brain damage, but to different neural systems. The two deficits would thus be related by neuroanatomic proximity, but one would not cause the other.

Another alternative is that impairment at a deeper level of analysis causes both executive function and theory of mind deficits. A possible candidate is widespread prefrontal impairment. Damage to dorsolateral regions of the prefrontal cortex is associated with typical executive function deficits, such as perseveration, planning difficulties and impulsivity (Damasio, 1985; Stuss & Benson, 1984). Lesions of orbitofrontal cortex result in social isolation, decreased affiliative behavior, impaired social communication, shallow affect and lack of appreciation of social rules (Damasio & Van Hoesen, 1983; Deutsch, Kling & Steklis, 1979; Franzen & Myers, 1973). Since prefrontal cortex appears to be intimately involved in regulating both executive functions and emotional behavior, it may be that prefrontal impairment is the underlying deficit in autism, capable of explaining both cognitive and social symptoms of the disorder.

Goldman-Rakic (1987) has hypothesized that prefrontal cortex is called into use whenever stored information is used to govern behavior. Different areas of prefrontal cortex access different types of representations generated elsewhere in the brain to guide voluntary behavior. This theory provides a means of integrating deficits in executive function with theory of mind impairment in autism. To perform well on a theory of mind task, subjects must appreciate and act upon the mental states of others, using internal representations of these mental states to guide their responses; to perform well on cognitive executive function tasks, they must use internal representations of hypothetical configurations to guide their planning and problem-solving.

Indirect support for the association of theory of mind abilities with prefrontal function comes from a recent study of two individuals who sustained widespread prefrontal damage early in development (Price, Daffner, Stowe & Mesulam, 1990). While neither was autistic, both exhibited severe deficits in interpersonal role-taking, leading the authors to speculate that the frontal lobes might be particularly important in the development of this skill. If this view is correct, it suggests that theory of mind abilities may be mediated by the prefrontal cortex.

Others have drawn an even more explicit connection between prefrontal impairment and the symptoms of autism. Damasio and Maurer (1978) theorized that dysfunction of the ring of mesolimbic cortex located in the mesial frontal lobes causes autism. They felt that language disturbances (specifically in communicative intent), stereotypies, resistance to change and social relatedness deficits in autistic people are similar to those seen in patients with frontal lobe damage. In summary, Damasio and Maurer state, "near-normal autistic individuals strikingly resemble patients with some types of frontal lobe defects. They share a similar lack of initiative, a similar concreteness in thought and language, an inability to focus attention, shallowness of affect, and lack of empathy".

While no direct evidence of frontal lobe abnormalities in autism has been obtained to support this analogy, few neuroanatomic studies of autism have examined this region of the brain. The neurological abnormalities that have been most consistently documented in autistic individuals, those in the limbic system and cerebellum, are at least consistent with prefrontal damage. All three brain regions are intimately connected (Damasio & Van Hoesen, 1983; Heath, Dempsey & Fontana, 1980), so that early damage to one could change both the structure and function of the others. Moreover, developmental studies indicate that two of the structures undergo late neurogenesis, making them similarly vulnerable to genetic or environmental insults *in utero* occurring at that time (Jacobson, 1978).

In summary, moving to a new level of analysis and considering the possibility that an underlying deficit in prefrontal function may be primary to autism allows us to account for many symptoms of the disorder, both cognitive and affective. This perspective may help the field of autism research move away from the sometimes misleading cognitive vs affective dichotomy that has been a focus of recent theoretical debate (Baron-Cohen, 1988).

A prefrontal hypothesis of autism is also helpful in explaining the different patterns of intercorrelation seen between the autistic and control groups. In the control group, most composites were not significantly correlated and seemed to represent differentiable, independent domains of functioning. However, in the autistic group, most composites were highly intercorrelated. Widespread prefrontal impairment is potentially capable of causing dysfunction in a wide variety of neuropsychological domains. Thus, abilities that are normally independent would become related through the etiology of autism and the connections between the brain systems involved.

If future investigations bear out the importance of prefrontal (or executive function) abnormalities to autism, the implications for educational remediation will be great. Building the autistic individual's repertoire of planning and organizational strategies and teaching him to use feedback from mistakes and shift away from ineffective problem-solving strategies could become a focus of treatment.

However, it would be simplistic to assume that all of autism can be explained by prefrontal impairment. If that were the case, then why do children with early frontal lesions not appear autistic (Price *et al.*, 1990)? Perhaps additional subcortical abnormalities must be present to cause all the deficits typical of the disorder. Clearly, further refinement of a prefrontal account of autism is needed. If, as speculated in the Introduction, multiple primary deficits are necessary to cause autism, prefrontal impairment may be a necessary, but not a sufficient, condition for development of the disorder. The present study may have measured only one of several core underlying deficits that, together, cause autism. Thus, while these results provide important new information about possible underlying causes of autism and suggest several areas for further investigation, there is still much work to be done before we fully understand the complexities of this disorder.

*Acknowledgements*—This research was supported in part by an award to the first author from the Developmental Psychobiology Research Group Endowment Fund provided by the Grant Foundation. The second author's work was supported by the following grants: an NIMH RSDA (MH00419) and project grants from NIMH (MH38820), NICHD (HD19423), and the March of Dimes (12-135). The authors would like to thank the individuals and parents who participated in the study, the administrators of the Denver and Cherry Creek Public School Systems for recruiting subjects, Larry Epstein and Chris Saiz for testing subjects, Simon Baron-Cohen for sharing his measures, and the anonymous reviewers for their helpful and thought-provoking comments.

## References

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

Baron-Cohen, S. (1989a). Are autistic children behaviorists? An examination of their mental–physical and appearance–reality distinctions. *Journal of Autism and Developmental Disorders,* **19**, 579–600.

Baron-Cohen, S. (1989b). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry,* **30**, 285–297.

Baron-Cohen, S., Leslie, A. M. & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition,* **21**, 37–46.

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

Borys, S. V., Spitz, H. H. & Dorans, B. A. (1982). Tower of Hanoi performance of retarded young adults and nonretarded children as a function of solution length and goal state. *Journal of Experimental Child Psychology,* **33**, 87–110.

Boucher, J. & Lewis, V. (1989). Memory impairments and communication in relatively able autistic children. *Journal of Child Psychology and Psychiatry,* **30**, 99–122.

Braverman, M., Fein, D., Lucci, D. & Waterhouse, L. (1989). Affect comprehension in children with pervasive developmental disorders. *Journal of Autism and Developmental Disorders,* **19**, 301–316.

Buschke, H. (1974). Components of verbal learning in children: analysis by selective reminding. *Journal of Experimental Child Psychology,* **18**, 488–496.

Chelune, G. J. & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology,* **8**, 219–228.

Chelune, G. J., Ferguson, W., Koon, R. & Dickey, T. O. (1986). Frontal lobe disinhibition in attention deficit disorder. *Child Psychiatry and Human Development,* **16**, 221–234.

Chelune, G. J. & Thompson, L. L. (1987). Evaluation of the general sensitivity of the Wisconsin Card Sorting Test among younger and older children. *Developmental Neuropsychology,* **3**, 81–89.

Damasio, A. R. (1985). The frontal lobes. In K. M. Heilman & E. Valenstein (Eds), *Clinical neuropsychology* (pp. 339–375). New York: Oxford University Press.

Damasio, A. R. & Maurer, R. G. (1978). A neurological model for childhood autism. *Archives of Neurology,* **35**, 777–786.

Damasio, A. R. & Van Hoesen, G. W. (1983). Emotional disturbances associated with focal lesions of the limbic frontal lobe. In K. M. Heilman & P. Satz (Eds), *The neuropsychology of human emotion* (pp. 85–110). New York: Guilford.

Dawson, G. (1983). Lateralized brain dysfunction in autism: evidence from the Halstead–Reitan neuropsychological battery. *Journal of Autism and Developmental Disorders,* **13**, 269–286.

Delis, D. C., Kramer, J. H., Kaplan, E. & Ober, B. A. (1987). *California verbal learning test manual* (research edn). New York: Psychological Corporation.

Deutsch, R. D., Kling, A. & Steklis, H. D. (1979). Influence of frontal lobe lesions on behavioral interactions in man. *Research Communications in Psychology, Psychiatry and Behavior,* **4**, 415–431.

Fein, D., Waterhouse, L., Lucci, D. & Snyder, D. (1985). Cognitive subtypes in developmentally disabled children: a pilot study. *Journal of Autism and Developmental Disorders,* **15**, 77–95.

Flavell, J. H., Flavell, E. & Green, F. L. (1983). Development of the appearance–reality distinction. *Cognitive Psychology,* **15**, 95–120.

Franzen, E. A. & Myers, R. E. (1973). Neural control of social behavior: prefrontal and anterior temporal control. *Neuropsychologia,* **11**, 141–157.

Frith, U. (1989). A new look at language and communication in autism. *British Journal of Disorders of Communication,* **24**, 123–150.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In V. B. Mountcastle, F. Plum & S. R. Geiger (Eds), *Handbook of physiology: the nervous system* (pp. 373–417). Bethesda, MD: American Physiological Society.

Goodman, R. (1989). Infantile autism: a syndrome of multiple primary deficits? *Journal of Autism and Developmental Disorders,* **19**, 409–424.

Grant, D. A. & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigle-type card sorting problem. *Journal of Experimental Psychology,* **32**, 404–411.

Harris, P. L. & Muncer, A. (1988). Autistic children's understanding of beliefs and desires. Paper presented at the British Psychological Society Section Conference, Coleg Harlech.

Heath, R. G., Dempsey, C. W. & Fontana, C. J. (1980). Feedback loop between cerebellum and septal-hippocampal sites: its role in emotion and epilepsy. *Biological Psychiatry,* **15**, 541–556.

Heaton, R. K. (1981). *Wisconsin card sorting test manual.* Odessa, FL: Psychological Assessment Resources.

Hobson, R. P. (1986a). The autistic child's appraisal of expressions of emotion. *Journal of Child Psychology and Psychiatry,* **27**, 321–342.

Hobson, R. P. (1986b). The autistic child's appraisal of expressions of emotion: a further study. *Journal of Child Psychology and Psychiatry,* **27**, 671–680.

Hobson, R. P. (1989). Beyond cognition: a theory of autism. In G. Dawson (Ed.), *Autism: new perspectives on diagnosis, nature and treatment* (pp. 22–48). New York: Guilford.

Hollingshead, A. B. (1975). *Four factor index of social status.* New Haven, CT: Yale University Press.

Hyde, J. S., Geiringer, E. R. & Yen, W. M. (1975). On the empirical relation between spatial ability and sex differences in other aspects of cognitive performance. *Multivariate Behavioral Research,* **10**, 289–309.

Izard, C. E. (1971). The development of emotion recognition and emotion labeling in normal children. In C. E. Izard (Ed.) *The face of emotion* (pp. 317–349). New York: Appleton-Century-Crofts.

Izard, C. E. (1982). *Measuring emotion in infants and children.* New York: Cambridge University Press.

Jacobson, M. (1978). Histogenesis and morphogenesis of the central nervous system. In *Developmental neurobiology* (pp. 57–114). New York: Plenum.

Johnson, C. N. & Wellman, H. M. (1982). Children's developing conceptions of the mind and brain. *Child Development,* **53**, 222–234.

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

Langdell, T. (1981). Face perception: an approach to the study of autism. Unpublished doctoral dissertation, London University College, London.

Loring, D. W. & Papanicolau, A. C. (1987). Memory assessment in neuropsychology: theoretical considerations and practical utility. *Journal of Clinical and Experimental Neuropsychology,* **9**, 340–358.

Luria, A. R. (1981). *Higher cortical functions in man.* New York: Basic Books.

Novic, J., Luchins, D. J. & Perline, R. (1984). Facial affect recognition in schizophrenia: is there a differential deficit? *British Journal of Psychiatry,* **144**, 533–537.

Oswald, D. P. & Ollendick, T. H. (1989). Role taking and social competence in autism and mental retardation. *Journal of Autism and Developmental Disorders,* **19**, 119–127.

Ozonoff, S., Pennington, B. F. & Rogers, S. J. (1990). Are there specific emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry,* **31**, 343–361.

Ozonoff, S., Rogers, S. J. & Pennington, B. F. (1991). Asperger's syndrome: evidence of an empirical distinction from high-functioning autism. *Journal of Child Psychology and Psychiatry,* **32**, 1107–1122.

Pennington, B. F. & Ozonoff, S. (1991). A neuroscientific perspective on continuity and discontinuity in developmental psychopathology. In D. Cicchetti & S. Toth (Eds), *Rochester Symposium on Developmental Psychopathology* (Vol. III). Hillsdale, NJ: Erlbaum (in press).

Perner, J., Frith, U., Leslie, A. M. & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: knowledge, belief and communication. *Child Development,* **60**, 689–700.

Perner, J. & Wimmer, H. (1985). "John thinks that Mary thinks that. . ." Attribution of second-order beliefs by 5–10 year old children. *Journal of Experimental Child Psychology,* **39**, 437–471.

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences,* **1**, 515–526.

Price, B. H., Daffner, K. R., Stowe, R. M. & Mesulam, M. M. (1990). The comportmental learning disabilities of early frontal lobe damage. *Brain,* **113**, 1383–1393.

Prior, M. R. (1979). Cognitive abilities and disabilities in infantile autism: a review. *Journal of Abnormal Child Psychology,* **7,** 357–380.

Prior, M. R., Dahlstrom, B. & Squires, T. (1990). Autistic children's knowledge of thinking and feeling states in other people. *Journal of Child Psychology and Psychiatry,* **31,** 587–601.

Robinson, A. L., Heaton, R. K., Lehman, R. A. W. & Stilson, D. W. (1980). The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *Journal of Consulting and Clinical Psychology,* **48,** 605–614.

Rumsey, J. M. (1985). Conceptual problem-solving in highly verbal, nonretarded autistic men. *Journal of Autism and Developmental Disorders,* **15,** 23–36.

Rumsey, J. M. & Hamburger, S. D. (1988). Neuropsychological findings in high-functioning men with infantile autism, residual state. *Journal of Clinical and Experimental Neuropsychology,* **10,** 201–221.

Rumsey, J. M. & Hamburger, S. D. (1990). Neuropsychological divergence of high-level autism and severe dyslexia. *Journal of Autism and Developmental Disorders,* **20,** 155–168.

Rutter, M. & Schopler, E. (1987). Autism and pervasive developmental disorders: concepts and diagnostic issues. *Journal of Autism and Developmental Disorders,* **17,** 159–186.

Sattler, J. M. (1988). *Assessment of children's intelligence and special abilities* (3rd edn). Boston: Allyn & Bacon.

Schopler, E. (1985). Convergence of learning disability, higher-level autism and Asperger's syndrome. *Journal of Autism and Developmental Disorders,* **15,** 359.

Schopler, E., Reichler, R. J. & Renner, B. R. (1986). *The childhood autism rating scale (CARS) for diagnostic screening and classification of autism.* New York: Irvington.

Shah, A. & Frith, U. (1983). An islet of ability in autistic children: a research note. *Journal of Child Psychology and Psychiatry,* **24,** 613–620.

Shallice, T. (1982). Specific impairments in planning. In D. E. Broadbent & L. Weiskrantz (Eds), *The neuropsychology of cognitive function* (pp. 199–209). London: The Royal Society.

Sigman, M., Ungerer, J. A., Mundy, P. & Sherman, T. (1987). Cognition in autistic children. In D. J. Cohen, A. M. Donnellan & R. Paul (Eds), *Handbook of autism and pervasive developmental disorders* (pp. 103–120). New York: Wiley.

Steel, J. G., Gorman, R. & Flexman, J. E. (1984). Neuropsychiatric testing in an autistic mathematical idiot-savant: evidence for nonverbal abstract capacity. *Journal of the American Academy of Child Psychiatry,* **23,** 704–707.

Stuss, D. T. & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin,* **95,** 3–28.

Szatmari, P., Bartolucci, G., Bremner, R., Bond, S. & Rich, S. (1989). A follow-up study of high-functioning autistic children. *Journal of Autism and Developmental Disorders,* **19,** 213–225.

Szatmari, P., Tuff, L., Finlayson, A. J. & Bartolucci, G. (1990). Asperger's syndrome and autism: neurocognitive aspects. *Journal of the American Academy of Child and Adolescent Psychiatry,* **29,** 130–136.

Tager-Flusberg, H. (1981). On the nature of linguistic functioning in early infantile autism. *Journal of Autism and Developmental Disorders,* **11,** 45–56.

Tager-Flusberg, H. (1989). An analysis of discourse ability and internal state lexicons in a longitudinal study of autistic children. Paper presented at the Society of Research in Child Development, Kansas City, MO, 28 April, 1989.

Vernon, P. E. (1972). The distinctiveness of field independence. *Journal of Personality,* **40,** 366–391.

Volkmar, F. R., Paul, R. & Cohen, D. J. (1985). The use of "Asperger's syndrome". *Journal of Autism and Developmental Disorders,* **15,** 437–439.

Weeks, S. J. & Hobson, R. P. (1987). The salience of facial expression for autistic children. *Journal of Child Psychology and Psychiatry,* **28,** 137–152.

Wellman, H. M. & Estes, D. (1986). Early understanding of mental entities: a re-examination of childhood realism. *Child Development,* **57,** 910–923.

Welsh, M. C., Pennington, B. F. & Groisser, D. B. (1991). A normative-developmental study of executive function: a window on prefrontal function in children. *Developmental Neuropsychology* (in press).

Welsh, M. C., Pennington, B. F., Ozonoff, S., Rouse, B. & McCabe, E. R. B. (1990). Neuropsychology of early-treated phenylketonuria: specific executive function deficits. *Child Development,* **61,** 1697–1713.

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine,* **11,** 115–129.

Wing, L., Gould, J., Yeates, S. R. & Brierly, L. M. (1977). Symbolic play in severely mentally retarded and autistic children. *Journal of Child Psychology and Psychiatry,* **18,** 167–178.

Witkin, H. A., Oltman, P. K., Raskin, E. & Karp, S. (1971). *A manual for the embedded figures test.* Los Angeles: Consulting Psychologists Press.

World Health Organization (1989). *International classification of diseases* (10th edn, draft version). Geneva: World Health Organization.