

Performance on Cambridge Neuropsychological Test Automated Battery Subtests Sensitive to Frontal Lobe Function in People with Autistic Disorder: Evidence from the Collaborative Programs of Excellence in Autism Network

Sally Ozonoff,^{1,9,10} Ian Cook,⁹ Hilary Coon,¹ Geraldine Dawson,² Robert M. Joseph,³ Ami Klin,⁴ William M. McMahon,¹ Nancy Minshew,⁵ Jeffrey A. Munson,² Bruce F. Pennington,⁶ Sally J. Rogers,^{7,9} M. Anne Spence,⁸ Helen Tager-Flusberg,³ Fred R. Volkmar,⁴ and Debora Wrathall¹

Recent structural and functional imaging work, as well as neuropathology and neuropsychology studies, provide strong empirical support for the involvement of frontal cortex in autism. The Cambridge Neuropsychological Test Automated Battery (CANTAB) is a computer-administered set of neuropsychological tests developed to examine specific components of cognition. Previous studies document the role of frontal cortex in performance of two CANTAB subtests: the Stockings of Cambridge, a planning task, and the Intradimensional/Extradimensional Shift task, a measure of cognitive set shifting. To examine the integrity of frontal functions, these subtests were administered to 79 participants with autism and 70 typical controls recruited from seven universities who are part of the Collaborative Programs of Excellence in Autism network. The two groups were matched on age, sex, and full-scale IQ. Significant group differences were found in performance on both subtests, with the autism group showing deficits in planning efficiency and extradimensional shifting relative to controls. Deficits were found in both lower- and higher-IQ individuals with autism across the age range of 6 to 47 years. Impairment on the CANTAB executive function subtests did not predict autism severity or specific autism symptoms (as measured by the ADI-R and ADOS), but it was correlated with adaptive behavior. If these CANTAB subtests do indeed measure prefrontal function, as suggested by previous research with animals and lesion patients, this adds to the accumulating evidence of frontal involvement in autism and indicates that this brain region should remain an active area of investigation.

KEY WORDS: Autism; executive function; planning; set shifting; frontal lobes; CANTAB.

¹ Department of Psychiatry, University of Utah School of Medicine, Salt Lake City, Utah.

² Center on Human Development and Disability, University of Washington, Seattle, Washington.

³ Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston, Massachusetts.

⁴ Child Study Center, Yale University School of Medicine, New Haven, Connecticut.

⁵ Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

⁶ Department of Psychology, University of Denver, Denver, Colorado.

⁷ Department of Psychiatry, University of Colorado Health Sciences Center, Denver, Colorado.

⁸ Department of Pediatrics, University of California–Irvine, Irvine, California.

⁹ M.I.N.D. Institute, University of California–Davis School of Medicine, Sacramento, California.

¹⁰ Correspondence should be addressed to Sally Ozonoff, M.I.N.D. Institute, University of California–Davis 2825 50th Street, Sacramento, CA 95817; e-mail: sally.ozonoff@ucdmc.ucdavis.edu

INTRODUCTION

One of the most consistently replicated cognitive deficits in individuals diagnosed with autism is executive dysfunction (see Pennington & Ozonoff, 1996, for a review). Executive function is a broadly defined cognitive construct originally used to describe the deficits associated with focal frontal lobe lesions. The executive function domain includes the many skills required to prepare for and execute complex behavior, including planning, inhibition, organization, self-monitoring, mental representation of tasks and goals, cognitive flexibility, and set-shifting. Executive functions are thought to be driven by prefrontal cortex (Duncan, 1986). There are many empirical reports of executive impairment in individuals with autism spectrum disorders, across wide age ranges and functioning levels (Benvenuto, Pennington, & Rogers, 1996; Ozonoff & Jensen, 1999; Prior & Hoffman, 1990; Rumsey & Hamburger, 1988; Russell, 1997). Clinical similarities between people with autism and adults with frontal lesions have been noted (Damasio & Maurer, 1978). There is also a growing body of evidence of frontal involvement from functional imaging (Baron-Cohen *et al.*, 1999; Horwitz, Rumsey, Grady, & Rapoport, 1988; Luna *et al.*, 2002; Minschew, Luna, & Sweeney, 1999) and neuropathology investigations (Casanova, Buxhoeveden, Switala, & Roy, 2002).

The Cambridge Neuropsychological Test Automated Battery (CANTAB) is a computer-administered, nonverbal (visually presented) set of tasks developed to examine specific components of cognition, particularly those associated with frontal and medial temporal regions of the brain (Robbins *et al.*, 1994). Several subtests have been used in functional imaging, animal, and human lesion studies (Baker *et al.*, 1996; Dias, Robbins, & Roberts, 1996; Owen, Doyen, Petrides, & Evans, 1996; Roberts, Robbins, & Everitt, 1988), permitting both cross-species comparisons and inferences about the underlying neural circuitry involved in task performance. Subtests are graded in difficulty, minimizing floor and ceiling effects and allowing use with a wide variety of ages and diagnoses (Fray, Robbins, & Sahakian, 1996; Luciana & Nelson, 1998; Robbins *et al.*, 1994, 1998).

Two CANTAB subtests, the Stockings of Cambridge (SOC) and the Intradimensional/Extradimensional Shift tasks, were developed to preferentially tap frontal functions, with imaging data providing validation of the role of prefrontal cortex in their performance (see following). A study using CANTAB to assess the executive functions of neurosurgical patients who had undergone frontal lobe excision, temporal lobe excision, or amygdalo-

hippocampectomy revealed deficits in the frontal lobe patients but not in the other groups (Owen *et al.*, 1991), further supporting the role of frontal cortex in performance of these tasks. Thus, CANTAB offers the opportunity to better understand both the cognitive impairments associated with autism and their neural underpinnings.

CANTAB's SOC subtest was designed to be similar to other Tower tasks (e.g., Tower of London and Hanoi) and is thought to measure planning efficiency. Three colored balls are arranged at the top of the computer screen in a specific configuration. Participants see three identical balls, in a different configuration, at the bottom of the computer screen, which they need to match with the goal set. They are told the minimum number of moves necessary to match the goal configuration and are instructed to use as few moves as possible. This demands planning and executing an optimal set of moves that transforms the initial ball configuration to the goal configuration.

"Tower" planning tasks have been widely used in studies of autism, and very large group differences are routinely reported (Pennington & Ozonoff, 1996). The differences between the Tower of Hanoi and Tower of London (and CANTAB's SOC) tasks are subtle, but important. In the Tower of Hanoi, a set of doughnut-like discs graded in size to form a pyramid-like structure must be moved from one of three equally sized pegs to another, according to the following constraints: only one disc can be moved at a time; if there is more than one disc on a peg, only the top disc can be moved; and no disc can be placed above a smaller disc. The number of moves required for a "perfect" (minimum move) solution in the Tower of Hanoi is a function of the number of discs involved in the problem. The key measure for the Tower of Hanoi task is the number of attempts the individual requires to achieve a perfect solution. The task therefore taps not only planning ability but also rule-following and procedural learning.

The simplified variant of the Tower of Hanoi, the Tower of London task, however, provides a purer test of planning ability. In this task the discs are replaced by three differently colored balls (removing much of the need for rule-constraints), and the pegs vary in length, allowing them to hold three balls, two balls, or just one ball, considerably reducing the size of the problem space. Moreover, rather than requiring a full tower-to-tower transfer, a graded set of problems requiring a minimum of 2, 3, 4, or 5 moves is presented; this set of novel subproblems allows the subject to gain familiarity with the task but still be presented with novel problems, thereby minimizing practice effects. Finally, the computerized version of the Tower of London in the

CANTAB includes a yoked set of “follow” tasks in which the participant is presented with his or her own solution, move by move, at the top of the screen and is simply required to follow these moves on the lower half of the screen. By subtracting response times in the “follow” tasks from those on the corresponding full tasks, planning and movement time can be estimated separately. A study using positron emission tomography (PET) demonstrated activation in dorsolateral and rostralateral prefrontal cortex associated with the planning demands of this task (Baker *et al.*, 1996).

A second subtest thought to require prefrontal function is CANTAB’s Intradimensional/Extradimensional Shift (ID/ED) task, which taps the executive component of set-shifting or cognitive flexibility. A series of compound stimuli composed of colored shapes and lines is presented. Participants learn, through trial and error with computer-generated feedback, to respond to the shape; the line is effectively an irrelevant dimension. Once training to the shape is complete, the necessity of performing two kinds of shifts takes place. In the first kind, the intradimensional shift, new shapes and lines are introduced, but shape remains the relevant response dimension. In the later, extradimensional shift, the contingencies change, with the line becoming the salient stimulus and the previously trained shape now becoming irrelevant. Only the extradimensional shift requires conceptual flexibility (i.e., shifting from one concept or cognitive set to another); the intradimensional shift only requires perceptual flexibility, or shifting from one exemplar to another exemplar within the same cognitive set (e.g., shape). This task is functionally similar to the category shifts required by the Wisconsin Card Sorting Test, but conceptually simpler and with multiple manipulations built in to control for other sources of impairment, such as inhibitory dysfunction or discrimination learning deficits. Experiments on marmoset monkeys with prefrontal lesions indicate that both orbital and lateral regions of the prefrontal cortex are involved in the extradimensional shift (Dias *et al.*, 1996).

Three recent investigations have used these CANTAB subtests with individuals with autism spectrum disorders (Hughes, Russell, & Robbins, 1994; Ozonoff, South, & Miller, 2000; Turner, 1997). Hughes *et al.* (1994) employed the SOC planning task with participants with autism who were also mentally retarded. The group with autism took significantly more moves to solve the problems than did the mentally retarded and normally developing controls. Only 13% of the group with autism, but 49% of the mentally retarded and 65% of the normal control group, solved the problems in the most efficient manner (i.e., in the minimum number of moves). A later study failed to detect any planning

impairments on the CANTAB SOC task, however (Ozonoff *et al.*, 2000). The primary difference between the two studies appears to be the intellectual level of the participants. Whereas the ability of the Hughes *et al.* (1994) sample fell in the mentally retarded range, the participants in the Ozonoff *et al.* (2000) study were quite high-functioning, with a group mean full-scale IQ of 111.

A similar pattern has been found with CANTAB’s ID/ED subtest. Hughes *et al.* (1994) documented intact performance during the discrimination learning and intradimensional shifting phases of the task, but impairment at the extradimensional shift, in participants with both mental retardation and autism, relative to appropriately matched mentally retarded controls without autism. Two later investigations had more mixed results, however. Turner (1997) replicated the extradimensional shifting deficit in mentally retarded individuals with autism, but found no deficits in non-retarded participants with autism, relative to outpatient psychiatric controls. Similarly, Ozonoff *et al.* (2000) found no extradimensional shifting difficulties in individuals with high-functioning autism relative to typically developing controls matched on IQ.

Thus, previous studies using CANTAB indicate that executive function impairments may be more prominent in individuals who have both autism and mental retardation than in those with autism who function in the average range of intelligence. However, this pattern was not found in previous executive function research, using tasks such as the Tower of Hanoi and Wisconsin Card Sorting Test. Previous studies have largely relied on average IQ groups and have consistently documented executive deficits (Benvenuto *et al.*, 1996; Ozonoff, 1995; Ozonoff & McEvoy, 1994; Ozonoff, Pennington, & Rogers, 1991; Prior & Hoffman, 1990; Rumsey, 1985; Rumsey & Hamburger, 1988, 1990; Szatmari, Tuff, Finlayson, & Bartolucci, 1990). Thus, one rationale for the present study was to compare performance on these CANTAB subtests with a much larger, multisite sample with a wide IQ range. The second goal of the study was to examine the integrity of frontal cortex in autism, using tasks whose neural underpinnings may be better understood than other neuropsychological measures.

METHOD

Participants

Recruitment and testing of participants with Autistic Disorder took place at seven sites participating in the NICHD/NIDCD-funded Collaborative Programs

of Excellence in Autism (CPEA) network: Boston University, University of California at Irvine, University of Colorado Health Sciences Center, University of Pittsburgh, University of Utah, University of Washington, and Yale University. All participants in the autism group met stringent research-based criteria for Autistic Disorder, according to both history (as collected by the Autism Diagnostic Interview-Revised; Lord, Rutter, & Le Couteur, 1994) and observation of current function (using the Autism Diagnostic Observation Schedule-Generic; Lord *et al.*, 2000). No participants met criteria for Asperger Disorder.

Participants in the control group were pronounced free of neurodevelopmental conditions, psychopathology, or learning disability based on extensive interviewing and assessment at each site. Three sites (University of Colorado Health Sciences Center, University of Pittsburgh, and University of Utah) recruited and tested all controls.

An initial sample of 151 participants with autism and 101 typical controls meeting all inclusion criteria was collected from across the seven CPEA sites. However, the autism and control groups differed significantly on many important variables, including age, verbal, performance, and full-scale IQ, and sex. The control group was significantly older and more intellectually capable than the autism group and contained significantly more females. Therefore, a matching procedure was undertaken, blind to results of the CANTAB testing, to construct a matched sample. Specifically, the oldest and highest-IQ controls were dropped, as were the youngest and lowest-IQ participants with autism. This resulted in a well-matched sample of 79 individuals with autism and 70 individuals with typical development who did not differ significantly in age, IQ, or sex. Participants ranged in age from 6 to 47 years and in full-scale IQ from 71 to 142 points. There was a preponderance of males in both groups (91% in the autism group, 83% in the control group), reflecting the skewed gender ratio typical of autism. Characteristics of the sample are detailed in Table I.

Measures

Autism Diagnostic Interview-Revised

The Autism Diagnostic Interview-Revised (ADI-R; Lord *et al.*, 1994), a parent report measure, collects information about behaviors relevant to the diagnosis of Pervasive Developmental Disorders. It contains three scales that correspond with the three Diagnostic and Statistical Manual (DSM-IV) (4th edition)

Table I. Characteristics of the Sample

	Autism (n = 79)		Control (n = 70)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age (years)	15.7	8.7	16.0	7.6
Verbal IQ	104.9	17.9	106.1	11.6
Performance IQ	106.0	16.0	105.0	12.0
Full Scale IQ	106.3	16.3	106.0	11.5
ADOS Social + Communication	14.4	2.9	—	—
ADOS Play	1.0	0.7	—	—
ADOS Stereotyped Behavior	1.6	1.6	—	—
ADI-R Social	22.7	5.2	—	—
ADI-R Communication	16.9	4.4	—	—
ADI-R Stereotyped Behavior	7.0	2.7	—	—
VABC	58.9	18.2	—	—

Note: VABC, Vineland Adaptive Behavior Composite (n = 57); ADOS, Autism Diagnostic Observation Schedule; ADI-R, Autism Diagnostic Interview-Revised.

(APA, 1994) symptom categories of Social Interaction, Communication, and Repetitive/Stereotyped Behavior impairments. Each scale contains detailed questions about both current functioning and early development. Responses are coded on a 4-point scale according to the quality and severity of symptoms (0 = normal for developmental level, 3 = severely autistic). Scores in each of the three domains are summed according to a research-derived algorithm that distinguishes between individuals with and without autism. Separate summary scores for current behavior and behavior during the 4-to-5-year age period are obtained. The ADI-R has excellent reliability and validity when used by trained examiners (Lord *et al.*, 1994).

Autism Diagnostic Observation Schedule

The Autism Diagnostic Observation Schedule (ADOS-G; Lord *et al.*, 2000) is a standardized interview and observational assessment that measures social and communicative behaviors diagnostic of Pervasive Developmental Disorders. Algorithm scores corresponding to DSM-IV criteria are obtained. The ADOS-G demonstrates good reliability and validity when used by trained examiners and differentiates well between individuals with autism and those with other developmental disabilities (Lord *et al.*, 2000).

Wechsler Intelligence Scales

Depending on the age of the participant, either the Wechsler Intelligence Scale for Children—Third

Edition or the Wechsler Adult Intelligence Scale—Third Edition was administered. Both yield three intelligence quotients: a Verbal IQ (VIQ), a Performance (or nonverbal) IQ (PIQ), and a Full-Scale IQ (FSIQ).

Vineland Adaptive Behavior Scales

The Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984), a parent-report measure of adaptive behavior, was administered to a parent of participants under age 18. It assesses social, communication, motor, and daily living skills. It is normed for use with infants to adults and provides standard scores and age equivalents. It was included in the study to examine potential relationships between adaptive behavior and executive function.

Cambridge Neuropsychological Test Automated Battery

Two subtests from the CANTAB (Robbins *et al.*, 1994) computerized battery were administered. Directions are presented via computer, as is accuracy feedback. Participants are given multiple training trials to learn the requirements of each task, and responses are recorded directly with a touch-sensitive screen. All participants were able to sustain attention and comply with task demands.

Stockings of Cambridge Subtest. In this computerized version of the Tower of London task, three colored balls are arranged at the top of the computer screen in a specific configuration. In the “plan and move” condition, participants see three identical balls, in a different configuration, displayed in the bottom half of the computer screen, which they need to match with the goal set. They are told the minimum number of moves necessary to match the goal configuration (between two and five moves) and are instructed to use as few moves as possible. Participants are also instructed to wait to begin moving balls until they have planned their moves (e.g., “don’t start until you think you know which moves to make”). Several problems requiring between two and five moves are then administered in a block. Following this phase, a second condition that controls for motor performance, the “follow” condition, is administered. Participants are presented with their own solutions to problems in the “plan and move” condition, seen move by move, at the top of the screen and are simply required to follow these moves on the lower half of the screen. By subtracting response times in the “follow” condition from those in the “plan and move” condition, it is possible to separately measure planning and movement times.

Several performance variables are obtained. The basic measure of planning efficiency is the “Minimum Moves” variable, which is the total number of test problems completed in the fewest possible number of moves. The “Mean Moves” variable describes the mean number of moves required by the subject to solve a test problem. The “Initial Thinking Time” variable is the difference in time taken to select the first ball for the same problem under the “plan and move” and “follow” conditions. The “Subsequent Thinking Time” variable is obtained by taking the difference in time between selecting the first ball and completing the problem under the “plan and move” and “follow” conditions and dividing it by the number of moves made. This measure reflects the subject’s speed of movement after the initial move has been made.

Intradimensional/Extradimensional Shift Subtest. This task measures the ability to attend to specific attributes of compound stimuli, shifting attention from one attribute to another when required. Participants are presented with a series of multidimensional stimuli consisting of shapes and lines. In stages 1 through 5 of the task, the discrimination and learning stages, participants learn through trial and error to respond selectively to one specific shape, ignoring the other shape and the lines. In stage 6, the intradimensional shift, new shapes and lines are introduced, but shape continues to be the salient response dimension. In stage 7, the intradimensional reversal, the previously nonreinforced shape now becomes the correct response. The shifts at stages 6 and 7 are not thought to be primary measures of flexibility, as participants continue to respond to the same rule or set as in previous trials. At stage 8, during the critical extradimensional shift, however, the correct rule now changes to the other dimension (e.g., the line) that has been irrelevant for the preceding dozens of trials. Finally, in stage 9, the extradimensional reversal, participants must respond to the previously nonreinforced line. Research on monkeys indicates that only stages 8 and 9 require prefrontal function, with the extradimensional shift using dorsolateral prefrontal cortex and the extradimensional reversal tapping orbitofrontal cortex (Dias *et al.*, 1996). The primary variables of interest on the ID/ED task were the number of errors committed and the number of trials taken to achieve criterion on stages 6 through 9. When participants failed to achieve criterion (six consecutive correct responses) at a given stage, the test was failed and the maximum number of errors (25) was recorded for all subsequent stages not administered. If autism involves selective deficits in prefrontal function, a dissociation between performance at stages 6/7 and stages 8/9, relative to controls, was predicted.

RESULTS

Stockings of Cambridge Task

An independent samples *t*-test was conducted on the “minimum moves” variable. As displayed in Table II, this analysis revealed significant group differences, $t(136) = -5.1, p < .001$, with the control group solving significantly more problems in the most efficient manner possible than the group with autism. The autism group took more moves to solve each of the three-, four-, and five-move problems than the control group, as measured by the “Mean Moves” variables. These differences were significant for the three- and five-move problems ($p < .001$) but not for the four-move problem. The two groups did not differ significantly in initial thinking time for the three-move [$t(136) = -1.5, p > .12$], four-move [$t(136) = 1.4, p > .15$], or five-move [$t(136) = 1.8, p > .08$], problems. There were significant group differences in subsequent thinking time, however. The autism group took longer (after the initial move) to solve the three-move [$t(136) = -2.4, p < .05$], four-move [$t(136) = -2.0, p < .05$], and five-move [$t(136) = -2.4, p < .05$], problems. See Table II.

Intradimensional/Extradimensional Shift Task

The primary variables of interest were the number of trials taken to reach criterion at stages 6 (intradimensional shift), 7 (reversal of the intradimensional shift), 8 (extradimensional shift), and 9 (reversal of the extradimensional shift) of the task. A repeated measures analysis of variance was conducted, with stage as the within-subjects factor and group as the

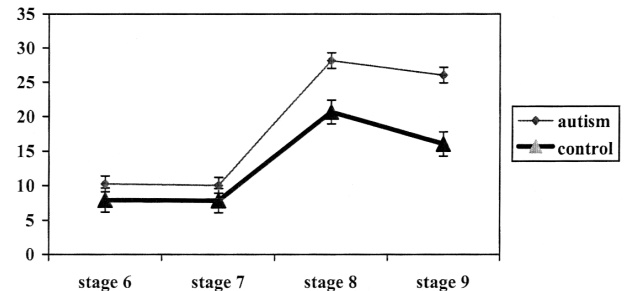


Fig. 1. Intradimensional/Extradimensional Shift trials to criterion as a function of group and stage.

between-subjects factor. This analysis revealed a significant group effect [$F(1,134) = 12.0, p < .001$], a significant stage effect [$F(3,134) = 71.4, p < .001$], and a significant group by stage interaction effect [$F(3,134) = 4.7, p < .05$]. Follow-up contrasts to explore the source of the interaction effect revealed a lack of significant group differences in performance at both stages 6 and 7 ($p > .05$), but significant group differences at stage 8 [$F(1,134) = 7.0, p < .01$], and stage 9 [$F(1,134) = 9.9, p < .01$]. At both stages 8 and 9, the group with autism needed significantly more trials to reach criterion than the control group. See Table III and Figure 1.

A similar analysis was undertaken to explore group differences in the number of errors made at stages 6 through 9. The repeated measures ANOVA again revealed a significant effect of both group [$F(1,134) = 13.3, p < .001$] and stage [$F(3,134) = 72.4, p < .001$], as well as a significant group by stage interaction [$F(3,134) = 5.5, p < .01$]. The interaction was again caused by a lack of significant group differences at

Table II. Performance on SOC Variables as a Function of Diagnostic Group

	Autism group (n = 79)		Control group (n = 70)	
	<i>M</i>	SD	<i>M</i>	SD
Problems solved in minimum moves	94.4	18.2	110.5	19.0***
Mean moves (three-move problems)	66.1	53.0	92.7	20.0***
Mean moves (four-move problems)	91.6	15.3	92.3	17.0
Mean moves (five-move problems)	85.5	20.2	101.9	17.9***
Mean initial thinking time (three-move)	101.9	20.7	106.6	8.2
Mean initial thinking time (four-move)	107.1	12.6	104.3	12.1
Mean initial thinking time (five-move)	104.3	16.0	99.8	14.4
Mean subsequent thinking time (three-move)	94.1	47.6	108.5	2.8*
Mean subsequent thinking time (four-move)	103.8	14.6	108.4	9.6*
Mean subsequent thinking time (five-move)	100.8	17.6	107.4	10.2*

Note: All SOC variables in standard scores ($M = 100, SD = 15$), based on age norms provided by CANTAB.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table III. Performance on ID/ED Shift Variables as a Function of Diagnostic Group

	Autism group (n = 79)		Control group (n = 70)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ID/ED stage 6 trials to criterion ^a	10.3	10.7	7.9	4.6
ID/ED stage 7 trials to criterion ^a	10.0	8.9	7.8	2.4
ID/ED stage 8 trials to criterion ^a	28.2	18.3	20.7	14.3**
ID/ED stage 9 trials to criterion ^a	26.0	20.7	16.0	16.1**
ID/ED stage 6 errors to criterion ^b	2.3	6.0	.95	2.0
ID/ED stage 7 errors to criterion ^b	2.5	4.9	1.4	1.2
ID/ED stage 8 errors to criterion ^b	13.6	11.5	8.4	8.8**
ID/ED stage 9 errors to criterion ^b	11.7	11.7	5.8	8.9***

Note: All ID/ED variables expressed in raw scores.

^a Minimum to achieve criterion = 6; maximum trials before failing stage = 50.

^b Maximum errors before failing stage = 25.

** *p* < .01.

*** *p* < .001.

stages 6 and 7 (*p* > .05), but significantly divergent performance at stage 8 [*F*(1,134) = 8.7, *p* < .01] and stage 9 [*F*(1,134) = 19.7, *p* < .001]. At both stages 8 and 9, the group with autism committed significantly more errors than the control group. See Table III.

Effect Sizes

Statistical significance is jointly determined by sample size and effect size (e.g., magnitude of the group differences). Because the size of the current sample was much larger than most previous studies of autism, it is important to calculate effect sizes (*d*). For all variables in which statistically significant group differences were found, all values of *d* fell in the medium to large range according to Cohen (1988; for SOC minimum moves, *d* = .87; for ID/ED stage 8 trials to criterion, *d* = .46; and for ID/ED stage 9 trials to criterion, *d* = .54). The mean value of *d* across these variables was .62. In contrast, mean *d* for performance at ID/ED stages 6 and 7 was .35, which falls in the small range. Therefore, the statistically significant group differences found in this study were not solely a function of the large sample size, but were also the result of robust effect sizes.

Correlations

The SOC minimum moves variable and the ID/ED trials to criterion at stage 8 variable were correlated with a variety of demographic and other variables within the autism group. Two-tailed Pearson product-moment correlation coefficients are reported in Table IV. The patterns of intercorrelations for the two

CANTAB subtest variables were similar but not entirely overlapping. ID/ED trials to criterion at stage 8 was significantly correlated with PIQ and FSIQ (*p* < .001), whereas the SOC minimum moves variable was correlated with PIQ but not FSIQ. Neither subtest was significantly correlated with VIQ, chronological age, or autistic symptoms (as measured by ADOS-G and ADI-R algorithm scores). SOC, but not ID/ED, performance was related to adaptive behavior, as measured by the Vineland (*p* < .01). The two CANTAB subtests were not significantly correlated with each other [*r*(59) = -.17, *p* > .18].

Table IV. Intercorrelations among CANTAB and Demographic Variables within the Autism Group (n = 79 unless otherwise indicated)

	SOC Minimum moves	ID/ED Trials to criterion stage 8
Age	-.08	-.10
Verbal IQ	.14	-.20
Performance IQ	.29*	-.55***
Full-Scale IQ	.23	-.42***
VABC	.42**	.14
ADOS-G Social + Communication	-.07	.05
ADOS-G Play	.04	.001
ADOS-G Stereotyped Behavior	-.04	.20
ADI-R Social	-.09	.14
ADI-R Communication	-.04	.20
ADI-R Stereotyped Behavior	-.09	.13

Note: VABC, Vineland Adaptive Behavior Composite (n = 57); ADOS, Autism Diagnostic Observation Schedule; ADI-R, Autism Diagnostic Interview-Revised.

* *p* < .05.

** *p* < .01.

*** *p* < .001.

The control group was not administered the ADI-R, ADOS-G, or Vineland, but the correlation coefficients among SOC and ID/ED performance and age and IQ were not remarkably different from the autism group. Performance was not significantly correlated with age on either CANTAB subtest. IQ was not significantly correlated with performance on the SOC subtest, but was moderately related to ID/ED performance at stage 8, VIQ: $r(63) = -.29, p < .05$; PIQ: $r(63) = -.31, p < .05$; FSIQ: $r(63) = -.30, p < .05$. Finally, the two CANTAB subtests were not significantly correlated with each other [$r(62) = -.10, p > .40$], as in the autism group.

Effect of IQ

In previous studies, group differences have been found between lower-IQ participants with and without autism (Hughes *et al.*, 1994; Turner, 1997), but not in higher-IQ groups (Ozonoff *et al.*, 2000; Turner, 1997). To examine this potential pattern in our study, we split the autism and control groups into two subgroups on the basis of FSIQ. SOC and ID/ED performance of these subgroups are displayed in Table V. Because our sample is largely high-functioning, the higher-IQ subgroup was larger than the lower-IQ subgroup. The higher-IQ subgroup was made up of 36 participants from the autism group (mean FSIQ = 111.5, mean age = 16.1 years) and 36 participants from the control group (mean FSIQ = 112.5, mean age = 16.2 years). These two subgroups together formed a subsample comparable in intellectual function to the Ozonoff *et al.* (2000) sample, which had a mean IQ of 111. The lower-IQ subgroup comprised 13 participants with autism

(mean IQ = 86.7, mean age = 19.3 years) and 12 participants from the control group (mean IQ = 86.5, mean age = 21.9 years). Whereas the lower-IQ subgroup had lower intelligence, with IQs ranging from 71 to 95, none of its members had mental retardation, and direct comparison with previous mentally retarded samples (Hughes *et al.*, 1994; Turner, 1997) is not appropriate. For this reason and because of the small sample size, caution should be used in drawing conclusions from the lower-IQ subgroup analyses reported below.

The higher-IQ autism and control subgroups did not differ significantly in IQ [$t(70) = -.55, p > .5$] or age [$t(70) = -.07, p > .9$]. Using repeated measures analysis of variance, a significant group by stage interaction effect was found on the ID/ED task [$F(3,64) = 5.3, p < .01$]. Follow-up contrasts to explore the source of this interaction revealed a lack of significant group differences at stages 6 and 7 ($ps > .25$) but significant group differences at stage 8 [$F(1,64) = 7.3, p < .01$] and stage 9 [$F(1,64) = 9.9, p < .01$]. An independent-samples *t*-test revealed that the higher-IQ autism subgroup differed from the higher-IQ control subgroup on the SOC minimum moves variable as well [$t(68) = -2.53, p < .05$]. Thus, the results of the higher-IQ subgroup analyses were consistent with the full group analysis. These results do not support earlier findings of attenuated group differences in samples of average intelligence (Ozonoff *et al.*, 2000; Turner, 1997).

The lower-IQ autism and control subgroups did not differ significantly in age [$t(23) = -.60, p > .50$] or FSIQ [$t(23) = .07, p > .90$]. Statistically significant group differences were found on the SOC minimum moves variable [$F(1,18) = 5.74, p < .05$], replicating

Table V. Performance as a Function of IQ-Level Subgroup

	Lower-IQ Autism (n = 13)		Lower-IQ Control (n = 12)		Higher-IQ Autism (n = 36)		Higher-IQ Control (n = 36)	
	M	SD	M	SD	M	SD	M	SD
SOC problems solved in minimum moves ^a	79.8	13.5	102.3	25.3*	100.5	19.2	110.8	15.5*
ID/ED Stage 6 Trials to criterion ^b	12.8	12.2	7.3	2.1	9.9	10.7	8.3	6.0
ID/ED stage 7 trials to criterion ^b	9.9	4.2	8.2	2.5	10.3	10.5	8.1	3.0
ID/ED stage 8 trials to criterion ^b	38.8	14.5	29.2	18.2	26.4	17.6	16.5	11.5**
ID/ED stage 9 trials to criterion ^b	37.9	19.6	26.6	19.9	23.9	20.0	11.3	11.3**

Note: SOC, Stockings of Cambridge; ID/ED, Intradimensional/Extradimensional Shift.

^a In standard scores ($M = 100, SD = 15$).

^b Minimum to achieve criterion = 6; maximum trials before failing stage = 50.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

the pattern found in both the whole sample and the higher-IQ subsample analyses. However, a significant group by stage interaction effect was not found in the lower-IQ subsample on the ID/ED task [$F(3,19) = .63$, $p > .50$], in contrast to the larger-sample and higher-IQ analyses. Because of the very small size of the lower-IQ subgroups, the effect sizes of group differences at each stage were calculated, rather than indices of statistical significance. Effect size calculations revealed that the mean differences between the two groups fell in the moderate to large range across all four stages (effect size, d , of .77, .51, .59, and .57 across stages 6 through 9, respectively; Cohen, 1988). These effect sizes indicate that the lower-IQ subgroup with autism experienced broad difficulties with both intra- and extradimensional shifting, rather than the specific impairments in the latter process seen in the higher-IQ subgroup with autism.

To summarize the analyses of the effect of IQ on performance, we found that the overall group differences on the SOC and ID/ED tasks were not predominantly the result of the very poor performance of the lower-IQ individuals with autism. Poorer performance by the autism group relative to the control group was found at both IQ levels on both CANTAB subtests.

Effect of Age

If a deficit is secondary, one prediction is that it worsens with age. To examine the prediction that executive function deficits are secondary to other primary impairments, the autism and control groups were split into three age subgroups (under 12 years, 12–19 years, and 20 years and over), and analyses of variance specifying diagnostic group and age group as between-subject factors were conducted on the SOC minimum moves and ID/ED trials to criterion at stage 8 variables. On SOC, the age group main effect was not significant [$F(2,129) = 1.2$, $p > .2$], but the age by diagnosis interaction effect was [$F(2,129) = 4.0$, $p < .05$]. This interaction effect was caused by a lack of significant autism–control group differences in problems solved in minimum moves in the under-12 subgroup [$F(1,42) = .6$, $p > .4$], but by large and significant group differences in the teenage [$F(1,54) = 21.2$, $p < .001$] and adult [$F(1,33) = 7.2$, $p < .05$] subgroups. Planned contrasts revealed that SOC performance improved with age in the controls, but did not develop significantly with age in the autism group.

A diagnostic group by age group analysis of variance on the extradimensional shift variable, trials to criterion at stage 8, revealed significant main effects of

both diagnostic group [$F(1,126) = 4.6$, $p < .05$] and age group [$F(2,126) = 3.6$, $p < .05$], but a nonsignificant interaction effect. Significant differences existed between the autism and control groups at each age level ($p < .05$), with the deficit in the autism group relatively constant across the three age groups. Thus, whereas the SOC age subgroup analysis supports the prediction that executive function deficits worsen with age, the ID/ED analysis does not.

DISCUSSION

Recent research strongly indicates frontal lobe involvement in autism (Baron-Cohen *et al.*, 1999; Casanova *et al.*, 2002; Luna *et al.*, 2002; Minshew *et al.*, 1999). There are also numerous empirical reports of executive dysfunction in autism (see Pennington & Ozonoff, 1996; Russell, 1997, for reviews), but evidence that the executive function tests used in these studies selectively tap prefrontal cortex is mixed (Anderson, Bigler, & Blatter, 1995; Mountain & Snow, 1993). In contrast, the two CANTAB subtests administered in this study were developed, using lesioned animals and human neuroimaging studies, to preferentially tap the prefrontal cortex (Baker *et al.*, 1996; Dias *et al.*, 1996; Owen *et al.*, 1991, 1996). Thus, performance on the measures used in this study may get closer to measuring prefrontal function than previous investigations of executive abilities, although any neuropsychological test is an indirect measure of brain function, and results need to be replicated with neuroimaging studies. A second justification for this study was the examination of whether executive function abilities predicted specific autism symptoms, adaptive behavior, or intellectual level.

This study used CANTAB to investigate the integrity of prefrontal cortical function in individuals with autism, using a large, multisite sample well-matched on age and intelligence. Significant deficits in planning efficiency and set shifting were found, replicating a large body of earlier work demonstrating executive dysfunction in autism. Subgroup analyses were conducted to compare group differences in executive function at the higher and lower ends of the IQ range of our sample. Deficits were found in both lower- and higher-IQ individuals with autism, contrary to some previous studies (Ozonoff *et al.*, 2000; Turner, 1997).

A second subgroup analysis, dividing the sample into three age ranges, was conducted to investigate the developmental course of the executive function deficit in autism. Two previous studies have failed to find

executive deficits in young children with autism that differentiate them from mental age-matched controls (Dawson *et al.*, 2002; Griffith, Pennington, Wehner, & Rogers, 1999), leading to the suggestion that differential executive impairment is acquired later and may be secondary to other, more primary, cognitive difficulties. One prediction of this hypothesis is that the magnitude of executive deficits and their dissociation from typical development should increase with age. Although the size of the group difference in set-shifting ability remained constant across the three age groups, the size of the planning efficiency deficit appeared to increase with age. The difference between the autism and control group was not significant in the youngest age group (under 12 years of age), but it became substantial and significant during the teen years (age 12–19 years) and leveled off in adulthood (age 20 years and older). However, inspection of group means indicated that the pattern was more the result of age-related improvements in planning efficiency in the typical controls than of worsening executive deficits with age in those with autism. This is consistent with the suggestion that frontal lobe functions typically mature after age 12 years, attenuating group differences in the younger-age group analyses. Correlations did not reveal a significant relationship between age and either the SOC measure of planning efficiency or the ID/ED measure of set-shifting ability. Thus, these data do not support the suggestion that executive deficits worsen with age and may be secondary to other, more primary impairments.

The relationship of executive dysfunction to other abilities is also of interest. Prefrontal cortex appears to be involved in the regulation of social behavior, emotional reactions, and social discourse (Dennis, 1991; Grattan, Bloomer, Archambault, & Eslinger, 1990; Price, Daffner, Stowe, & Mesulam, 1990; Stuss, 1992). Previous studies have found significant correlations between executive function abilities and social-communication skills in children with autism (McEvoy, Rogers, & Pennington, 1993; Ozonoff & McEvoy, 1994). A significant relationship was found between planning efficiency and a composite measure of adaptive behavior (encompassing social, communication, and daily living skills). However, the predicted relationship with set shifting (ID/ED performance) was not found, and correlations between performance on both CANTAB subtests and ADOS-G and ADI-R social and communication algorithm scores were not significant either. One explanation may be lack of variability in the social and communication variables. Because all participants had to meet ADI-R and ADOS-G criteria for Autistic Disorder, the range of scores on these

measures was necessarily truncated. Indeed, other studies have combined autism, developmentally delayed, and typical control groups in their correlational analyses to increase the range of scores and enhance power (McEvoy *et al.*, 1993). This was not possible in the present study, however, as the control group was not administered the ADI-R, ADOS-G, or Vineland Adaptive Behavior Scales.

Future investigations would benefit from a large sample size, as achieved in this study, but preferably with an even wider range of intellectual functioning that will permit examination of potential IQ level effects on group differences. Turner (1995, as cited in Turner, 1997) found a different pattern of group differences in participants with autism with and without mental retardation on a set-shifting task. Those with mental retardation performed significantly less well than controls in the perseveration condition, but not in the learned irrelevance condition. In contrast, there were no significant differences between the higher-functioning group with autism and a comparison group in either condition (as cited in Turner, 1997). Studies that include participants from across the spectrum of symptom presentation are also needed and may prove helpful in answering basic questions about the universality of executive function impairments and the external validity of subtypes (e.g., Asperger syndrome).

In conclusion, these results replicate previous findings of deficits in planning and flexibility in people with autism, using a relatively new computerized measure of executive function. The strengths of this study lie in its large sample size, its carefully matched comparison group, and the numerous controls built into the design of its executive function tasks. Previous neuroimaging and animal work, as well as examination of focal lesions in humans, provide strong support for the role of prefrontal cortex in performance of the tasks used in this study. Previous work implicates prefrontal cortex in performance at only two stages of CANTAB's ID/ED subtest. Dias *et al.* (1996) found that extradimensional shifting (stage 8) uses dorsolateral prefrontal cortex in marmoset monkeys, whereas extradimensional reversals (stage 9) require orbitofrontal cortex. Our sample of individuals with autism experienced significant difficulties relative to controls at both stages 8 and 9, but not at earlier stages requiring discrimination learning and intradimensional shifting. Thus, this study indicates that not all types of attention shifting are impaired in autism—only those that require prefrontal cortical function. At the cognitive level, shifting within a category or rule does not appear

impaired, whereas shifting between categories, sets, or rules is deficient.

These results contribute to the accumulating evidence of frontal lobe involvement in autism. The neural circuitry that causes the symptoms of autism is likely widely distributed throughout the brain. This study indicates that prefrontal cortex is involved in these circuits at some level and that this brain region should remain an area of active investigation in the future.

REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*. (4th ed.). Washington, DC: Author.
- Anderson, C. V., Bigler, E. D., & Blatter, D. D. (1995). Frontal lobe lesions, diffuse damage, and neuropsychological functioning in traumatic brain-injured patients. *Journal of Clinical and Experimental Neuropsychology*, *17*, 900–908.
- Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., Dolan, R. J., Frackowiak, R. S. J., & Robbins, T. W. (1996). Neural systems engaged by planning: A PET study of the Tower of London task. *Neuropsychologia*, *34*, 515–526.
- Baron-Cohen, S., Ring, H., Wheelwright, S., Bullmore, E., Brammer, M., Simmons, A., & Williams, S. (1999). Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience*, *11*, 1891–1898.
- Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Development*, *67*, 1816–1835.
- Casanova, M. F., Buxhoeveden, D. P., Switala, A. E., & Roy, E. (2002). Minicolumnar pathology in autism. *Neurology*, *58*, 428–432.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Damasio, A. R., & Maurer, R. G. (1978). A neurological model for childhood autism. *Archives of Neurology*, *35*, 777–786.
- Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Development*, *73*, 700–717.
- Dennis, M. (1991). Frontal lobe function in childhood and adolescence: A heuristic for assessing attention regulation, executive control, and the intentional states important for social discourse. Special Issue: Developmental consequences of early frontal lobe damage. *Developmental Neuropsychology*, *7*, 327–358.
- Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of attentional and affective shifts. *Nature*, *380*, 69–72.
- Duncan, J. (1986). Disorganization of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, *3*, 271–290.
- Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuropsychiatric applications of CANTAB. *International Journal of Geriatric Psychiatry*, *11*, 329–336.
- Grattan, L. M., Bloomer, R., Archambault, F. X., & Eslinger, P. J. (1990). Cognitive and neural underpinnings of empathy. *The Clinical Neuropsychologist*, *4*, 279.
- Griffith, E. M., Pennington, B. F., Wehner, E. A., & Rogers, S. J. (1999). Executive functions in young children with autism. *Child Development*, *70*, 817–832.
- Horwitz, B., Rumsey, J. M., Grady, C. L., & Rapoport, S. I. (1988). The cerebral metabolic landscape in autism: Intercorrelations of regional glucose utilization. *Archives of Neurology*, *45*, 749–755.
- Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*, 477–492.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205–223.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview—Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 659–685.
- Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*, 273–293.
- Luna, B., Minshew, N. J., Garver, K. E., Lazar, N. A., Thulborn, K. R., Eddy, W. F., & Sweeney, J. A. (2002). Neocortical system abnormalities in autism: An fMRI study of spatial working memory. *Neurology*, *59*, 834–840.
- McEvoy, R. E., Rogers, S. J., & Pennington, B. F. (1993). Executive function and social communication deficits in young autistic children. *Journal of Child Psychology and Psychiatry*, *34*, 563–578.
- Minshew, M. J., Luna, B., & Sweeney, J. A. (1999). Oculomotor evidence for neocortical systems but not cerebellar dysfunction in autism. *Neurology*, *52*, 917–922.
- Mountain, M. A., & Snow, W. G. (1993). Wisconsin card sorting test as a measure of frontal pathology: A review. *The Clinical Neuropsychologist*, *7*, 108–118.
- Owen, A. M., Doyon, J., Petrides, M., & Evans, A. C. (1996). Planning and spatial working memory: A positron emission tomography study in humans. *European Journal of Neuroscience*, *8*, 353–364.
- Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., & Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, *29*, 993–1006.
- Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology*, *9*, 491–500.
- Ozonoff, S., & Jensen, J. (1999). Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders*, *29*, 171–177.
- Ozonoff, S., & McEvoy, R. E. (1994). A longitudinal study of executive function and theory of mind development in autism. *Development and Psychopathology*, *6*, 415–431.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *32*, 1081–1105.
- Ozonoff, S., South, M., & Miller, J. N. (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral, and early history differentiation from high-functioning autism. *Autism*, *4*, 29–46.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathologies. *Journal of Child Psychology and Psychiatry*, *37*, 51–87.
- Price, B. H., Daffner, K. R., Stowe, R. M., & Mesulam, M. M. (1990). The compartmental learning disabilities of early frontal lobe damage. *Brain*, *113*, 1383–1393.
- Prior, M., & Hoffmann, W. (1990). Brief report: Neuropsychological testing of autistic children through an exploration with frontal lobe tests. *Journal of Autism and Developmental Disorders*, *20*, 581–590.

- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. M. A. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society, 4*, 474–490.
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): A factor analytic study of a large sample of normal elderly volunteers. *Dementia, 5*, 266–281.
- Roberts, A., Robbins, T. W., & Everitt, B. J. (1988). Extra- and intradimensional shifts in man and marmoset. *Quarterly Journal of Experimental Psychology, 40B*, 321–342.
- 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.
- Russell, J. (1997). *Autism as an executive disorder*. New York: Oxford University Press.
- Stuss, D. T. (1992). Biological and psychological development of executive functions. Special Issue: The role of frontal lobe maturation in cognitive and social development. *Brain and Cognition, 20*, 8–23.
- 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.
- Turner, M. (1997). Towards an executive dysfunction account of repetitive behavior in autism. In J. Russell (Ed.), *Autism as an executive disorder* (pp. 57–100). New York: Oxford University Press.