Child Neuropsychology, 2013
http://dx.doi.org/10.1080/09297049.2013.799644

Routledge
Taylor & Francis Group

Executive functioning in individuals with a history of ASDs who have achieved optimal outcomes

Eva Troyb[1], Michael Rosenthal[2], Inge-Marie Eigsti[1], Elizabeth Kelley[3], Katherine Tyson[1], Alyssa Orinstein[1], Marianne Barton[1], and Deborah Fein[1,4]

[1]Department of Psychology, University of Connecticut, Storrs, Connecticut, USA
[2]Child Mind Institute, New York, USA
[3]Department of Psychology, Queens University, Kingston, Ontario, Canada
[4]Department of Pediatrics, University of Connecticut, Farmington, Connecticut, USA

Executive functioning (EF) is examined among children and adolescents once diagnosed with an autism spectrum disorder (ASD), but who no longer meet diagnostic criteria. These individuals have average social and language skills, receive minimal school support and are considered to have achieved "optimal outcomes" (OOs). Since residual impairments in these individuals might be expected in deficits central to autism, and in developmentally advanced skills, EF was examined in 34 individuals who achieved OOs, 43 individuals with high-functioning autism (HFA), and 34 typically developing (TD) peers. Groups were matched on age ($M = 13.49$), gender, and nonverbal IQ (NVIQ) but differed on verbal IQ (VIQ; HFA < TD, OO). On direct assessment, all three groups demonstrated average EF; however, the OO and HFA groups exhibited more impulsivity and less efficient planning and problem-solving than the TD group, and more HFA participants exhibited below average inhibition than did OO and TD participants. Parent-report measures revealed average EF among the OO and TD groups; however, the OO group exhibited more difficulty than the TD group on set-shifting and working memory. HFA participants demonstrated more difficulty on all parent-reported EF domains, with a clinical impairment in attention-shifting. Results suggest that EF in OO appears to be within the average range, even for functions that were impaired among individuals with HFA. Despite their average performance, however, the OO and TD groups differed on measures of impulsivity, set-shifting, problem-solving, working memory, and planning, suggesting that the OO group does not have the above-average EF scores of the TD group despite their high-average IQs.

Keywords: Autism; High-functioning autism; Executive functioning; Outcome; Optimal outcome.

Address correspondence to Eva Troyb, Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269, USA. E-mail: eva.troyb@uconn.edu

The authors have no financial conflict of interest to declare.

We would like to thank Drs. Letitia Naigles, Michael Stevens, and Robert Schultz for their assistance with this project, Molly Helt for her assistance in collecting the data for this study, our undergraduate research assistants who were immensely helpful with data entry, the parents and children who participated in this study and the funding agency, and the National Institutes of Mental Health (R01MH076189).

INTRODUCTION

Autism Spectrum Disorders (ASDs) are generally considered lifelong disorders; however, several studies have indicated that a small percentage of individuals diagnosed with ASDs in childhood respond particularly well to early intervention and experience a marked reduction in autism symptoms as they mature, such that they no longer meet diagnostic criteria for ASD (for a review of these studies, see Helt et al., 2008).

Lovaas first introduced the idea of "recovery" from ASDs in 1987, when he reported that 9 of 19 children diagnosed with ASD in his treatment sample achieved normal intellectual and educational functioning following early intensive, behavioral intervention (Lovaas, 1987). In subsequent years, the finding that some children with ASD can lose the diagnosis was replicated, although often rates of "recovery" were much lower than reported by Lovaas (Cohen, Amerine-Dickens, & Smith, 2007; Harris & Handleman, 2000; Sallows & Graupner, 2005; Weiss, 1999; Zachor, Ben-Itzchak, Rabinovich, & Lahat, 2007). Definitions of "best outcome" or "recovery" varied, but usually included scoring in the average range on tests of IQ, language, and adaptive functioning, and placement in mainstream education settings.

Mundy (1993) raised an important concern regarding previous studies examining "recovery" by pointing out that high IQ and regular educational placement does not in itself indicate "recovery" because high-functioning children with ASDs can achieve both of these outcomes. To address this concern, the current study and several others (e.g., Kelley, Naigles, & Fein, 2010; Kelley, Paul, Fein, & Naigles, 2006; Sutera et al., 2007) have required that the children no longer meet diagnostic criteria for ASD. Several longitudinal studies exploring outcomes in ASD have indicated that between 3% and 25% of their samples did not meet diagnostic criteria for ASD by the conclusion of the study (Howlin, Goode, Hutton, & Rutter, 2004; Rutter, 1970; Seltzer, Shattuck, Abbeduto, & Greenberg, 2004; Sigman & Ruskin, 1999; Szatmari, Bartolucci, Bremner, Bond, & Rich, 1989; Venter, Lord, & Schopler, 1992), although most reported that these individuals, even while no longer meeting full criteria, still showed marked impairments in language or social functioning (e.g., Piven, Harper, Palmer, & Arndt 1996; Turner & Stone, 2007).

Whereas few studies have reported the presence of children who experience such a marked reduction in the ASD symptoms within their cohorts, even fewer have gone on to characterize their functioning in detail. Kelley et al. (2006, 2010) examined language functioning of a group of children diagnosed with ASD in early childhood, who no longer met diagnostic criteria for ASD in middle childhood, had average IQs and were mainstreamed in regular classrooms without extra assistance. These children exhibited grammatical abilities (measured by the Clinical Evaluation of Language Fundamentals) that were largely indistinguishable from their peers, but experienced some residual difficulties in pragmatic and semantic language (i.e., understanding second-order theory of mind, using mental state verbs, inductive reasoning, and narrative production). Fein, Dixon, Paul, and Levin (2005) described several cases of children who were diagnosed with ASD in childhood, but who evolved into clear-cut cases of ADHD from early to middle childhood as they lost their ASD diagnoses. These children continued to show residual impairments in attention and impulsivity, but no longer met diagnostic criteria for ASD. The authors speculated that inattention features may be central to ASDs and may be harder to remediate with early behavioral intervention than the core language and social symptoms.

The results reported by Fein et al. (2005) suggest that individuals who achieve positive outcomes following an ASD diagnosis continue to exhibit difficulties in attention and perhaps in other frontally-mediated cognitive abilities often referred to as executive

functions. Executive functions refer to a loosely defined set of processes that involve set-shifting, inhibition, working memory, fluency, planning, and organization. Although executive functioning (EF) has not been explored in children who lose their ASD diagnoses as they mature, the presence of such deficits would not be surprising, given multiple reports indicating that the EF of children and adults with ASDs is impaired (e.g., Damasio & Maurer, 1978; Koshino et al., 2005; Luna, Doll, Hegedus, Minshew, & Sweeney, 2007; O'Hearn, Asato, Ordaz, & Luna, 2008; Pennington & Ozonoff, 1996; Schuh & Eigsti, in press). Specifically, studies exploring EF in high-functioning adolescents and adults with ASDs revealed impaired performance on tasks that involve set-shifting, working memory, planning, and fluency (see Eigsti, 2011, for a review). Some theories of ASDs have even suggested that deficits in EF may account for many behavioral features of ASD (e.g., Pennington & Ozonoff, 1996). On the other hand, if EF deficits account for the behavioral features of ASD, one might predict that EF deficits may not be evident among a sample of children who no longer present with behavioral features of ASD. In other words, it is possible that a resolution of the behavioral symptoms of ASD experienced by children who lose their ASD diagnosis would not be possible without a concomitant improvement in EF.

The present study seeks to describe the EF of children and adolescents with a history of ASD who no longer meet diagnostic criteria for any ASD, have good social functioning, are functioning in the normal range on standardized measures of cognitive and adaptive skills, and receive limited or no special education services. The purpose of this study is to further our understanding of possible residual EF difficulties and to provide direction for continued intervention or support. Performance of these individuals on measures of EF is compared to that of age- and IQ-matched children and adolescents with high-functioning autism (HFA) and to typically developing (TD) peers. These participants were recruited as part of a larger study examining OOs that was described recently by Fein et al. (2013). The results of this larger study supported the existence of a group of individuals who at a young age were diagnosed with an ASD, but who no longer met diagnostic criteria for ASD and whose current social and communication skills were similar to their TD peers. Fein and colleagues also reported that in early childhood the OO and HFA groups exhibited a similar degree of ASD symptomatology within the communication and repetitive behaviors domain according to parent interviews (i.e., Autism Diagnostic Interview, Revised and Social Communication Questionnaire, Lifetime). However, within the socialization domain, the OO group appeared to demonstrate milder symptoms of ASD than the HFA group during early childhood.

We hypothesize that, despite having achieved favorable outcomes, individuals who lost their ASD diagnosis (those with OO) will continue to exhibit some residual deficits on measures of EF that will not be observed in typically developing peers, as these abilities have been shown to be deficient in children with ASDs. However, we expect that the performance of individuals who achieved OO will be superior to same-aged individuals with HFA.

METHODS

Participants

Thirty-four individuals with a history of ASD who have achieved OO, 43 high-functioning individuals with a current ASD diagnosis (HFA), and 34 typically developing peers (TD) were tested. Participants ranged from 8 years, 1 month to 21 years, 8 months. The groups were matched on age, gender, and nonverbal IQ (NVIQ), but were significantly

Table 1 Participant Characteristics.

	HFA	OO	TD	F or $\chi 2$	p	η_p^2	Post Hoc
n	43	34	34				
Sex	39 M; 4 F	27 M; 7 F	31 M; 3 F	2.83	.24		
Age	13.85 (2.68) (8.6–20.0)	12.77 (3.45) (8.1–21.2)	13.87 (2.58) (9.9–21.7)	1.32	.27	.03	
WASI: VIQ	105.43 (14.38) (81–142)	112.65 (13.72) (80–137)	112.00 (11.17) (93–138)	4.15	.02	.08	OO>HFA
WASI: NVIQ	110.19 (12.76) (78–147)	110.29 (15.07) (81–142)	112.79 (11.32) (89–139)	0.35	.71	.01	
Vineland: Communication	82.70 (13.86) (42–108)	98.30 (12.66) (79–122)	93.32 (9.35) (77–119)	14.82	<.001	.22	G-H: OO, TD>HFA
Vineland: Socialization	75.51 (16.02) (46–109)	102.03 (8.44) (80–118)	101.74 (8.56) (86–120)	57.41	<.001	.53	G-H: OO, TD>HFA
Vineland: Daily Living	75.40 (14.26) (46–110)	92.30 (15.88) (65–120)	88.76 (9.26) (74–115)	15.34	<.001	.23	G-H: OO, TD>HFA
ADOS: Communication	3.50 (1.42) (2–7)	0.47 (0.62) (0–2)	0.41 (0.56) (0–2)	122.70	<.001	.70	G-H: HFA> OO, TD
ADOS: Socialization	6.77 (2.21) (4–13)	1.09 (1.31) (0–4)	0.50 (0.75) (0–2)	174.86	<.001	.77	G-H: HFA> OO, TD
ADOS: Stereotyped Behaviors and Restricted Interests	1.04 (1.17) (0–4)	0.18 (0.46) (0–2)	0.03 (0.17) (0–1)	19.76	<.001	.26	HFA> OO, TD

Notes. Table reports means, followed by standard deviations, and ranges. WASI (Wechsler Abbreviated Scale of Intelligence); and Vineland (Vineland Adaptive Behavior Scales); subtest Mean = 100, Standard deviation = 10. Unless otherwise indicated, Tukey's post hoc test was used; "G-H" indicates the use of the Games-Howell post hoc test to account for violations in homogeneity of variance.

different on verbal IQ (VIQ), with the OO and TD groups having a VIQ about seven points higher than the HFA group (See Table 1). Six HFA participants and 3 OO participants were evaluated at Queens University in Kingston, Ontario, Canada. Their performance did not significantly differ from the other participants on any measure. The participants tested at the University of Connecticut were primarily from the northeast United States. Participants were predominantly Caucasian, with three OO individuals, two HFA individuals, and three TD individuals reporting other races or ethnicities. The parents of a subset of participants in each group completed a questionnaire that assesses EF. No significant group differences were observed in the demographic characteristics of participants whose parents completed the parent-report measure of EF and those who did not. The study was approved by the Institutional Review Boards of the University of Connecticut, the Institute of Living Hartford Hospital, Children's Hospital of Philadelphia, and Queens University. See Figure 1 for a flow chart of inclusion and exclusion criteria. Recruitment was done through media outlets (newspaper stories, radio interviews), private practices, and clinic referrals. In some cases, therapists contacted parents of children known to have achieved

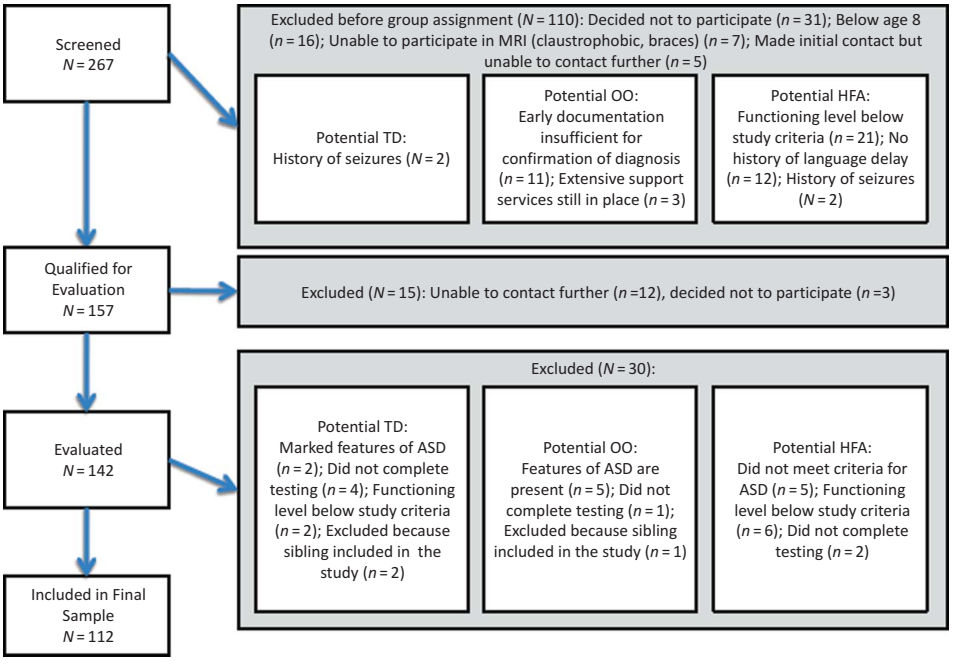

Figure 1 Flow chart of participant inclusion.

OOs, and in some cases, parents saw media reports and contacted the investigators. Data were collected as part of a larger study of OOs, reported in detail in Fein et al. (2013).

Inclusion Criteria. All participants had verbal, nonverbal, and full-scale IQ standard scores greater than 77 (within 1.5 standard deviations of the average of 100). Each group had additional specific inclusion criteria.

For the OO group:

1. Participants had a documented ASD diagnosis made by a physician or psychologist specializing in autism before the age of 5, verified in a written diagnostic report provided by parents. Early language delay (no words by 18 months or no phrases by 24 months) documented in the report was required. As a second step in confirming diagnosis, the report was edited to remove information about diagnosis, summary, and recommendations but leaving descriptions of behavior. One of the co-investigators (MB), an expert in diagnosis of ASD and Director of the University of Connecticut Psychological Services Clinic, reviewed these reports, blind to early diagnosis and current group membership. In addition to potential OO participants, she reviewed 24 "foil" reports for children with non-ASD diagnoses, such as global delay or language disorder. Four potential OO participants were rejected for insufficient early documentation and were dropped from the study. All 24 foils were correctly rejected.
2. Participants could not currently meet criteria for any ASD according to the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) administered by a research-reliable interviewer. In addition, a clinician with more than 15 years of autism diagnostic experience (IME, MB, or DF) reviewed the ADOS of all potential OO cases

to confirm that ADOS scores were below ASD thresholds. In addition, these clinicians reviewed the ADOS, the participant's developmental history and parent responses to the Autism Diagnostic Interview, Revised for all potential OO participants and made an expert clinical judgment that an ASD was not present.
3. Participants' scores on the Communication and Socialization domains of the Vineland (see below) had to be greater than 77 (within 1.5 standard deviations of the mean of 100) (see Table 1).
4. Participants had to be fully included in regular education classrooms with no one-on-one assistance and no special education services to address autism deficits (e.g., no social skills training). However, participants could be receiving limited special education services or psychological support to address impairments not specific to ASDs, such as attention or academic difficulties.

For the HFA group:

1. Following Collaborative Programs of Excellence in Autism diagnostic guidelines (Luyster et al., 2005), participants had to meet criteria for ASD on the ADOS (both Social and Communication domains and total score) and according to best estimate clinical judgment.

For the TD group:

1. Participants could not meet criteria for any ASD at any point in their development by parent report.
2. Participants could not have a first-degree relative with an ASD diagnosis.
3. Participants could not meet current diagnostic criteria for an ASD on the ADOS, or by clinical judgment (see Table 1). There was no attempt to exclude TD children for other learning or psychiatric disorders (but see general exclusion criteria).
4. Scores on the Communication and Socialization domains of the Vineland had to be greater than 77 (see Table 1).

Exclusion Criteria. Potential participants for any group were excluded from the study if (a) at the time of the telephone screening they exhibited symptoms of major psychopathology (e.g., active psychotic disorder) that would impede full participation, (b) they had severe visual or hearing impairments, or (c) they had a history of seizure disorder, Fragile X syndrome, or significant head trauma with loss of consciousness. Two participants in the TD group and 2 in the HFA group were excluded because of possible seizure disorder; none were excluded for other reasons (See Figure 1).

Procedure

Phone screenings based on study criteria were conducted with parents of each potential participant. Those who passed screening were scheduled for an assessment. Informed consent and assent were obtained, as appropriate, prior to testing. The evaluation was administered in a quiet room over the course of two or three testing sessions at the University of Connecticut, the Institute of Living of Hartford Hospital, Queens University, or in the home. Testing lasted approximately 6 hours. In most cases, parent interviews were conducted concurrently by a second examiner and lasted approximately 3 hours for the OO

and HFA groups and 1.5 hours for the TD group. Interviews with the parent gathered the participant's developmental history, history of ASD symptomatology, and adaptive functioning. Parents of the participants in the HFA and OO groups received a more extensive interview assessing ASD symptomatology than did parents of participants in the TD group, which accounted for the difference in the length of the parent interviews across groups. Participants received a monetary incentive for participation, even if the testing could not be completed.

Measures

Enrollment Criteria Measures: Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) (module 3 or 4, depending on age) was used to determine current diagnostic status for the OO and HFA groups, to rule out autistic features in the TD group, and to compare social interaction in the OO and TD groups. The individuals who administered the ADOS were not blind to group status of the participant. However, a rater blind to group membership coded videotapes of the ADOS administrations for 5 participants in each of the three groups and high interrater reliability was found for both the algorithm (86.7%) and total items (85.7%). Cognitive abilities were measured using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). The Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984) is a parent-report measure used to evaluate adaptive functioning in Communication, Daily Living Skills, and Socialization.

EF was assessed using three subtests of the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001). The D-KEFS measures set-shifting, inhibition, working memory, fluency, planning, and the ability to follow rules in individuals aged 8–89 years. The *Color-Word Interference* subtest is similar to the well-known Stroop test and involves four conditions: *Color Naming* requires participants to name patches of colors as rapidly as possible; *Word Reading* requires participants to read names of colors printed in black ink; *Inhibition* requires participants to name ink colors of color words printed in an incongruent color, inhibiting the more automatic task of reading the word; *Inhibition/Switching* requires participants to either name the ink color or to read color words printed in incongruent colors depending on whether the word is in a box or not, requiring inhibition as well as rapid set-shifting. Completion times on each condition, contrasts comparing completion times of the four conditions, and the number of errors made during each condition are evaluated.

The *Verbal Fluency* subtest asks the participant to list as many words as possible that begin with a particular letter of the alphabet (*Letter Fluency*) or that fit in a semantic category (*Category Fluency*). The child is also asked to name objects rapidly while switching between different two semantic categories (*Category Switching*). This subtest evaluates the number of words generated during each condition, comparisons in performance across the conditions, the number of words generated during each of the 15-second intervals across the conditions, and the number of set-loss and repetition errors made.

The *Tower* subtest asks the participant to move a series of wooden discs on three pegs in order to replicate a presented figure using the fewest number of moves possible following specific rules. This subtest assesses planning, working memory, and inhibition. A participant receives a *Total Achievement* score based on both the number of items completed and the number of extraneous moves made. The subtest also evaluates latency to first move (*Mean First-Move Time*), average move time (*Time-Per-Move-Ratio*), number of additional moves made as compared to the minimum necessary (*Move Accuracy Ratio*),

and the number of rule violations per item (*Rule-Violation-Per-Item Ratio*). These subtests produce scaled scores with a mean of 10 and standard deviation of 3.

In addition to the D-KEFS, parent report of the participants' EF was assessed using the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). The BRIEF is a questionnaire completed by parents that assesses the following executive functions: *Inhibit* (inhibit impulses), *Shift* (shift attention between activities), *Emotional Control* (regulate emotional reactions), *Initiate* (begin activities), *Working Memory* (hold information in mind to complete a task), *Plan/Organize* (set goals for anticipated future events and plan steps to carry them out), *Organization of Materials* (orderliness of work and play area), and *Monitor* (review one's behavior and work in search of mistakes). These scales comprise three Index scores: *Behavioral Regulation Index* (BRI; reflects performance on *Inhibit*, *Shift*, and *Emotional Control*), *Metacognitive Index* (MI; reflects performance on *Working Memory*, *Plan/Organize, Organization of Materials*, and *Monitor*), and the *Global Executive Composite* (GEC; overall summary score). The scales of the BRIEF have a mean of 50 and standard deviation of 10. Higher scores indicate weaker EF and scores at or above 65 suggest clinically significant impairment.

RESULTS

Scores on most measures met the assumptions of normality of data or homogeneity of variance required for parametric statistical tests. Although most parametric techniques are robust enough to accommodate non-normal distributions (Stevens, 1996), nonparametric test equivalents (Kurskal-Wallis and Mann-Whitney tests) were conducted to confirm the results of parametric tests when the normality assumption was violated. In all cases, nonparametric equivalents confirmed the findings of parametric techniques. When homogeneity of variance was violated, the Games-Howell post hoc test was used; in all other cases, the Tukey post hoc test was used. One-way univariate analyses of variance (ANOVAs) were conducted for each of the dependent variables of interest (selected subtest scores from the D-KEFS and BRIEF) with the three cohorts designated as independent variables. These analyses were repeated with VIQ included as a covariate. Chi-square tests were used to examine proportions of individuals within each group who scored more than 1.5 standard deviations below the mean on the measures. Finally, the Sequential Bonferroni Correction was employed to safeguard against Type II error (Holm, 1979). Four variables were initially significant, but did not survive the correction for multiple comparisons (D-KEFS Color-Word Interference subtest: Inhibition/Switching vs. Combined Naming Reading Contrast, Color Naming Total Errors Percentile; D-KEFS Verbal Fluency: Category Switching vs. Category Fluency comparison; D-KEFS Tower: Total Achievement score). The *p* values reported below were uncorrected for multiple comparisons.

Behavioral Assessment of Executive Functioning (D-KEFS)

Color-Word Interference Subtest. This subtest is similar to the Stroop test and measures inhibition and set-shifting. All three groups performed within the average range on all of the conditions and contrast scores on this subtest (see Table 2). The only group mean score that fell in the low end of the average range was obtained by the OO group on the contrast comparing scores on the *Inhibition/Switching* and *Word Reading* conditions.

The only significant group difference in the completion times of the conditions on this subtest was found on the *Word Reading* condition (see Table 2 for F, p, and η_p^2 values).

Table 2 Performance on D-KEFS Color-Word Interference Subtest.

	HFA	OO	TD	*F*	*p*	η_p^2	Post hoc
n	43	34	34				
Completion Times:							
Color Naming	9.67 (2.99) (3–15)	10.24 (3.06) (5–15)	10.29 (3.04) (3–15)	0.51	.61	.01	
Word Reading	(3.22) (2–15)	(2.09) (7–16)	(2.45) (3–15)	4.46	.01	.08	G-H: OO>HFA
Inhibition	9.65 (3.61) (1–15)	10.06 (3.51) (1–15)	10.62 (2.4) (5–16)	0.84	.44	.02	
Inhibition/Switching	8.79 (3.19) (1–13)	9.12 (3.29) (1–15)	10.32 (1.92) (6–14)	2.83	.06	.05	TD>HFA
Combined Naming + Reading	9.95 (2.95) (4–15)	11.18 (2.33) (7–15)	10.65 (2.58) (4–15)	2.05	.13	.04	
Contrast Measures:							
Inhibition vs. Color Naming	9.98 (2.76) (3–14)	9.82 (3.01) (3–17)	10.32 (1.98) (7–14)	0.32	.73	.01	
Inhibition/Switching vs. Color Naming	9.12 (2.73) (5–16)	8.97 (3.50) (1–13)	10.03 (2.39) (5–16)	1.37	.26	.03	
Inhibition/Switching vs. Word Reading	9.07 (2.83) (4–17)	7.56 (2.90) (1–13)	9.85 (1.93) (6–13)	6.81	.002	.11	TD, HFA>OO
Inhibition/Switching vs. Inhibition	9.95 (3.24) (2–18)	8.76 (3.17) (1–14)	9.79 (1.87) (4–14)	1.26	.29	.02	
Inhibition/Switchi ng vs. Combined Naming + Reading	8.84 (2.64) (4–16)	8.00 (3.02) (1–13)	9.68 (1.87) (6–14)	3.65	.03*	.06	TD>OO
Error Analysis:							
Color Naming: Total Errors Percentile	62.83 (36.88) (2–100)	47.79 (38.86) (1–100)	71.38 (36.11) (2–100)	3.51	.03*	.06	TD>OO
Word Reading: Total Errors Percentile	69.95 (41.99) (1–100)	72.56 (38.13) (2–100)	77.38 (38.69) (2–100)	0.33	.72	.01	
Inhibition: Total Errors Scaled Score	9.93 (2.99) (1–14)	9.41 (3.13) (1–13)	10.35 (2.19) (5–14)	0.95	.39	.02	
Inhibition/Switching: Total Errors Scaled Score	9.47 (2.68) (1–13)	9.44 (3.20) (1–13)	10.15 (2.06) (6–13)	0.78	.46	.01	

Notes. Table reports means, followed by standard deviations, and ranges. D-KEFS subtest Mean = 10, Standard deviation = 3. Unless otherwise indicated, Tukey's post hoc test was used; "G-H" indicates the use of the Games-Howell post hoc test to account for violations in homogeneity of variance.

*A statistically significant *F*-test that did not survive the Sequential Bonferroni Correction for multiple comparisons.

This group difference remained significant after controlling for VIQ and was attributable to the HFA group taking significantly more time than the OO group to complete the task (Games-Howell: $p = .01$). The completion time of the TD group on this condition did not differ significantly from the other two groups.

This subtest includes contrast scores that compare the completion times of the four conditions of this subtest. The contrast comparing scores on the *Inhibition/Switching* and *Word Reading* conditions differed significantly across groups (see Table 2), even when VIQ was included as a covariate. This group difference reflected a significantly lower score of the OO group as compared to both the HFA ($p = .001$) and TD ($p = .04$) groups, while the performance of the HFA and TD groups did not differ significantly from each other. No significant group differences were found in the number of errors made during any condition. All of the groups scored at or above average on their error scores.

Group means can obscure different patterns of results for individuals within each group. Chi-square tests showed that the HFA group contained significantly more participants scoring lower than 1.5 standard deviations below the mean on the *Word Reading* condition than the OO and TD groups ($\chi^2(2, N = 111) = 8.86$, $p = .01$; TD = 3%, OO = 0%, HFA = 16%). The number of participants scoring below the average range (more than 1.5 standard deviations below the mean) did not differ across groups on the other conditions of this subtest.

Verbal Fluency Subtest. This subtest measures initiation of verbal responses, lexical fluency, and set-shifting. All three groups scored solidly in the average range on all conditions and contrast measures of this subtest (see Table 3). In addition, no group differences were evident in the number of words produced during any of the conditions on this subtest.

All three groups produced an average number of words during each of the four 15-seconds intervals across the conditions. However, the groups differed significantly on the number of words generated during the first interval across the conditions of this subtest (see Table 3 for F, p, and η_p^2 values), even when VIQ was included as a covariate. During this interval, the OO group scored in the high end of the average range and produced more words than the HFA group (Tukey: $p = .01$), while the score of the TD group did not differ significantly from the other two groups.

The number of errors made by each group on this subtest fell in the average range and scores did not differ significantly across groups (see Table 3). In addition, the groups did not differ on the number of participants scoring lower than 1.5 standard deviations below the mean on the conditions on this subtest.

Tower Subtest. This subtest measures aspects of planning, working memory, inhibition, and problem-solving strategy. The three groups performed in the average range on all scores on this subtest (see Table 4). It is important to note that both the OO and HFA groups scored in the low average range on the ratio examining the additional number of moves compared to the minimum necessary to complete the items (*Move Accuracy Ratio*).

No group difference was evident on the *Total Achievement* score or in the number of items administered to participants (TD: 9.00, OO: 8.94, HFA: 8.83; see Table 4 for F, p, and η_p^2 values). Latency to first move also differed significantly across groups, and again the TD group received a higher score (i.e., took less time) than both of the

Table 3 Group Performance on the D-KEFS Verbal Fluency Subtest.

	HFA	OO	TD	*F*	*p*	η_p^2	Post hoc
n	43	34	34				
Primary Measures:							
Letter Fluency	10.23 (3.14) (5–17)	11.68 (3.53) (6–19)	11.00 (3.32) (6–19)	1.83	.17	.03	
Category Fluency	10.60 (3.79) (3–19)	12.44 (3.65) (5–19)	11.06 (2.95) (6–19)	2.72	.07	.05	OO > HFA ($p = .06$)
Category Switching – Total Correct Resp.	11.12 (3.16) (3–17)	10.94 (2.86) (6–17)	11.06 (3.35) (1–19)	0.03	.97	.01	
Category Switching – Accuracy	10.57 (3.12) (4–17)	10.74 (2.61) (5–16)	11.24 (3.07) (5–17)	0.50	.61	.01	
Letter Fluency vs. Category Fluency	10.05 (3.39) (3–19)	9.24 (2.80) (3–14)	9.91 (2.86) (1–16)	0.73	.48	.01	
Category Switching vs. Category Fluency	10.47 (3.37) (4–19)	8.50 (3.04) (3–18)	9.82 (3.17) (2–16)	3.62	.03*	.06	HFA > OO
Optional Measures:							
First Interval, Total Correct	10.69 (3.63) (4–19)	13.09 (3.67) (6–19)	11.82 (2.66) (7–19)	4.75	.01	.08	OO > HFA
Second Interval, Total Correct	10.26 (3.99) (3–18)	11.21 (3.50) (5–19)	10.82 (3.52) (5–19)	0.63	.54	.01	
Third Interval, Total Correct	10.38 (3.27) (4–17)	11.44 (3.73) (6–19)	10.97 (3.02) (6–19)	0.96	.39	.02	
Fourth Interval, Total Correct	10.64 (2.86) (5–16)	11.06 (3.30) (5–17)	10.44 (3.60) (4–18)	0.32	.72	.01	
Set-Loss Errors	9.74 (3.34) (1–13)	9.68 (3.46) (1–13)	11.26 (2.11) (3–13)	3.05	.05	.05	G-H: TD>HFA; TD>OO ($p = .07$)
Repetition Errors	8.40 (2.55) (1–12)	8.88 (2.53) (1–12)	8.47 (2.31) (1–11)	0.39	.68	.01	

Notes. Table reports means, followed by standard deviations, and ranges. D-KEFS subtest Mean = 10, Standard deviation = 3. Unless otherwise indicated, Tukey's post hoc test was used; "G-H" indicates the use of the Games-Howell post hoc test to account for violations in homogeneity of variance.

*A statistically significant *F*-test that did not survive the Sequential Bonferroni Correction for multiple comparisons.

other groups (TD: 2.12 seconds, OO: 3.56 seconds, HFA: 3.25 seconds; Tukey: TD > OO: $p = .01$, TD > HFA: $p = .02$). A significant group difference was also noted on the rule violations measure, even after controlling for VIQ, and again the TD group scored higher (that is, violated fewer rules) than the other two groups (Tukey: TD > OO: $p = .01$, TD > HFA: $p = .03$). While all scores were within the average range, the

Table 4 Group Performance on the D-KEFS Tower Subtest.

	HFA	OO	TD	*F*	*p*	η_p^2	Post Hoc
n	43	33	33				
Total Achievement	9.67 (2.58) (3–17)	9.42 (2.19) (5–14)	11.00 (2.29) (6–17)	4.27	.02*	.07	TD > OO, HFA
Mean First-Move Time	10.76 (1.71) (8–13)	10.55 (1.66) (7–13)	11.79 (1.24) (9–14)	6.06	.01	.10	TD > OO, HFA
Time-Per-Move-Ratio	10.05 (1.99) (3–13)	9.33 (2.23) (1–12)	10.33 (2.12) (5–14)	1.99	.14	.04	
Move Accuracy Ratio	7.85 (3.07) (1–13)	7.67 (2.81) (2–14)	8.73 (2.00) (3–12)	1.49	.23	.02	
Rule-Violation-Per-Item Ratio	9.57 (1.64) (2–11)	8.73 (2.53) (1–11)	10.27 (0.67) (9–11)	6.30	.003	.11	G-H: TD > OO, HFA

Notes. Table reports means, followed by standard deviations, and ranges. D-KEFS subtest Mean = 10, Standard deviation = 3. Unless otherwise indicated, Tukey's post hoc test was used; "G-H" indicates the use of the Games-Howell post hoc test to account for violations in homogeneity of variance.

*A statistically significant *F*-test that did not survive the Sequential Bonferroni Correction for multiple comparisons.

high average TD scores were more commensurate with the group's high average IQ while the HFA and OO mean scores were slightly below average despite above average mean IQs.

Chi-square tests revealed no group differences in the frequency of participants scoring lower than 1.5 standard deviation below the mean on this subtest.

Parent Report of Executive Functions (BRIEF)

Both the TD and OO groups scored in the average range on all scale and index scores of the BRIEF (see Table 5). The HFA group received clinically significant scores (scores $\geq$ 65) on the *Shift* scale, the *Behavioral Regulation Index*, and the *Global Executive Composite*. Their scores were above 60 (sometimes considered "at risk") on six of the other scales.

Significant group differences were detected on all of the scales and index scores of the BRIEF (see Table 5 for *F* and η_p^2 values; all *p*s $< .001$). Group differences remained significant when VIQ was included as a covariate. Post hoc tests revealed that the HFA group scored significantly higher (indicative of weaker EF) on all scale and index scores of the BRIEF than both the OO and TD groups (Tukey: all *p*s $< .01$). The OO group did not differ from their TD peers on any measures except the *Shift* and *Working Memory* subscales, and the *Behavioral Regulation Index* (Games-Howell: $p = .01$, $p = .03$, and $p = .01$, respectively).

The HFA group contained significantly more participants scoring in the clinically significant range (scores $\geq$ 65) than the TD and OO groups on all of the scale and index scores of the BRIEF (see Table 6). The difference in the number of participants scoring in

Table 5 Group Performance on the BRIEF.

	HFA	OO	TD	*F*	*p*	η_p^2	Post Hoc
n	38	25	32				
Inhibit	62.13 (14.72) (42–94)	51.00 (10.10) (40–72)	45.63 (7.10) (37–72)	19.02	<.001	.32	G-H: HFA > TD, OO
Shift	69.24 (13.56) (41–95)	49.60 (9.45) (38–71)	42.94 (5.97) (36–61)	59.89	<.001	.59	G-H: HFA > OO > TD
Emotional Control	61.13 (11.53) (41–89)	48.56 (9.69) (37–76)	42.88 (8.31) (36–73)	30.23	<.001	.41	G-H: HFA > TD, OO
Initiate	60.68 (11.97) (39–86)	49.04 (9.74) (35–70)	45.59 (8.16) (35–63)	20.85	<.001	.31	HFA > TD, OO
Working Memory	62.50 (11.90) (40–90)	52.72 (12.30) (36–79)	45.19 (7.74) (36–63)	22.60	<.001	.36	G-H: HFA > OO > TD
Plan/Organize	60.78 (10.59) (41–80)	48.76 (11.22) (33–77)	45.97 (7.89) (33–63)	21.54	<.001	.33	HFA > TD, OO
Org. of Materials	57.03 (9.91) (36–72)	50.44 (8.53) (37–72)	47.78 (7.22) (37–63)	10.43	<.001	.19	HFA > TD, OO
Monitor	63.95 (8.83) (47–78)	49.32 (9.50) (27–66)	46.19 (9.68) (28–68)	36.23	<.001	.45	HFA > TD, OO
Behavioral Regulation Index	66.05 (12.72) (46–96)	49.68 (9.09) (36–68)	42.91 (7.36) (35–65)	46.89	<.001	.52	G-H: HFA > OO > TD
Metacognition Index	62.57 (10.09) (41–81)	50.36 (10.82) (33–72)	45.47 (7.77) (31–60)	29.20	<.001	.40	HFA > TD, OO
Global Executive Composite	65.11 (10.93) (45–89)	50.04 (9.36) (34–68)	44.06 (7.56) (32–62)	45.19	<.001	.51	HFA > TD, OO

Notes. Table reports means, followed by standard deviations, and ranges. BRIEF Mean = 50, Standard deviation = 10. Higher scores indicate weaker EF; scores ≥ 65 fall in clinically significant range. Unless otherwise indicated, Tukey's post hoc test was used; "G-H" indicates the use of the Games-Howell post hoc test to account for violations in homogeneity of variance.

the clinically significant range between the OO and TD groups reached significance on the *Initiate*, $\chi^2(2, N = 57) = 4.053$, $p = .04$, and *Working Memory* scales, $\chi^2(2, N = 57) = 7.02$, $p = .01$, as well as *Metacognition Index*, $\chi^2(2, N = 57) = 5.51$, $p = .02$, although the percent of the OO group scoring low was small for each scale (12–20%).

In summary, the results indicated that the all three groups performed in the average range on standardized assessments of EF. Comparisons that reached statistical significance indicated that on the Stroop-like task, the OO group read color words faster than the HFA group, which contained more participants who scored in the below-average range on this task than both the OO and TD groups. The OO group scored lower than the TD and HFA groups on the contrast comparing the *Inhibition/Switching* condition

Table 6 Percentage of Participants Scoring in the Clinically Significant Range on the BRIEF.

	HFA	OO	TD	χ^2	*p*
n	38	25	32		
Inhibit	37	8	3	15.70	<.001
Shift	66	8	0	43.91	<.001
Emotional Control	34	4	3	16.17	<.001
Initiate*	37	12	0	16.85	<.001
Working Memory*	47	20	0	21.57	<.001
Plan/Organize	42	8	0	23.46	<.001
Org. of Materials	29	4	0	15.48	<.001
Monitor	53	4	3	30.92	<.001
Behavioral Regulation Index	55	12	3	29.00	<.001
Metacognition Index*	47	16	0	23.70	<.001
Global Executive Composite	50	4	0	33.08	<.001

Notes. BRIEF Mean = 50, Standard deviation = 10, scores ≥ 65 fall in clinically significant range.
*A statistically significant difference between the OO and TD groups at $p < .05$.

with the *Word Reading* condition; however, the effect size of this group differences was small and likely reflects the OO group's high score on the *Word Reading* condition. On *Verbal Fluency*, the OO group produced more words than the HFA group during the first 15-second interval of the conditions. On the *Tower*, the TD group took less time to make their first move and made fewer rule violations than both the OO and HFA groups.

Parent report of EF on the BRIEF revealed average scores for the OO and TD groups on all domains. Significant differences between the OO and TD group were noted in set-shifting and working memory with the TD group having less difficulty. The HFA group demonstrated more difficulty on all EF domains and received clinically significant scores on attention-shifting and an EF composite score. In addition, the HFA group contained a greater proportion of participants who scored in the clinically significant range on all domains of the BRIEF than did the other groups. Finally, the OO group had a greater proportion of participants scoring in the clinically significant range on initiation, working memory, and the *Metacognition Index* than did the TD group, but proportions were still quite low (12–20%).

DISCUSSION

This study investigated the EF of individuals diagnosed with ASD in early childhood, but who currently do not meet diagnostic criteria for any ASD. Participants were selected to have average IQs and adaptive skills and were included in regular-education classrooms with no support and minimal special education services. The results of this study indicate that, contrary to predictions, the OO group's performance on all measures of EF fell solidly in the average range.

All three groups performed in the average range on direct assessment of EF. The results on the *Color Word Interference* subtest indicated that all three groups possess average ability to inhibit automatic verbal responses and average cognitive flexibility. The measure of verbal fluency indicated that the three groups exhibited average initiation of verbal responses, lexical fluency, cognitive flexibility (i.e., shifting between two

semantic categories), vocabulary knowledge, and simultaneous processing (i.e., participants have to follow task rules while generating words). Finally, performance on the *Tower* test indicated that all three groups exhibited average spatial planning, learning of effective problem-solving strategies, inhibition of impulsive responding, as well as establishment and maintenance of cognitive set.

Parent report of EF revealed average EF in the OO and TD groups and confirmed a lack of residual EF deficits in the OO group. The OO group's average scores on measures of EF that are sensitive to attentional difficulties (e.g., working memory) are particularly interesting because they suggest that this OO group may not be exhibiting attentional difficulties as frequently as was evident in the OO sample described by Fein et al. (2005). It is possible that because this sample was older than that described by Fein et al., this OO group may have learned to compensate for attentional difficulties or grew out of these difficulties as they aged. However, despite the OO group's average EF scores, significantly more participants in the OO group exhibited clinically significant weaknesses in working memory and initiation than did participants in the TD group. This result suggests that even within this older sample, attentional difficulties may be more common in this OO sample as compared to the TD group. To determine the rate of clinically significant attentional deficits in the OO group more directly, a future paper will examine current and past psychiatric symptoms within this OO sample.

On two domains of EF, the OO group scored more poorly (but still within the average range) than the TD group, but significantly better than the HFA group, according to parent report: attention shifting and working memory. The solidly average performance of the OO group on this measure suggests that these differences are not the result of EF deficits within the OO group, but are due to the above-average performance of the TD group. The TD group's above-average EF was commensurate with the group's above-average intellectual ability, indicating a more consistent cognitive profile of developed intellectual and executive functioning. In comparison, the OO group had higher than average intellectual ability but only average EF. Because previous studies have suggested a significant association between intellectual ability and EF among children with high-average IQs (Ardila, Pineda, & Rosselli, 2000; Arffa, 2007; Arffa, Lovell, Podell, & Goldberg, 1998; Baron, 2004; Mahone et al., 2002), the observed discrepancy between EF and intellectual ability within the OO group may indicate that the EF abilities of the OO participants are not as developed as might be expected given the group's above-average intellectual abilities, and might reflect a weakness in working memory and attention-shifting, which, as noted in the introduction, are frequently found in ASD.

Across all of the measures of EF, the OO scored significantly more poorly than both the TD and HFA groups on only one contrast score of the *Color Word Interference* subtest of the D-KEFS. Specifically, the OO group received a significantly lower score than the TD and HFA groups on a contrast comparing the response time on a reading ability task and a task requiring inhibition and set-shifting. It is likely that this low comparison score of the OO group reflects the groups' strong word-reading ability, which when compared to the groups' average inhibition and set-shifting ability, resulted in below average scores. This finding is consistent with our prediction that the OO group will generally outperform the HFA group on measures of EF.

The OO group's performance on the D-KEFS Tower task demonstrated the only instances in which the OO group scored similarly to the HFA group and performed significantly below the TD group. Despite their average performance, the OO and HFA

groups took significantly longer than the TD groups to make their first move (though still well within the average range), suggesting that the OO and HFA groups required more time to plan and initiate their problem-solving. The OO and HFA groups also violated the rules of the task more frequently than the TD group, suggesting some relative difficulty in the inhibition of impulsive responding.

On two other variables assessed, the OO group's score differed only from the HFA group, while the TD group did not differ significantly from either of the groups. These included word-reading ability, with the HFA group scoring lower than the OO group and containing more participants who scored below the average range, as compared to the OO and TD groups. This finding suggests that the HFA group contains considerable variability in reading speed of high-frequency words that is not evident in the other groups and that is masked by the group's overall average score on the word-reading task. The OO group also generated more words during the first 15 seconds of the verbal fluency conditions than the HFA group, scoring in the high-average range compared to population norms. These results indicate that the OO group's high-average ability to initiate verbal responses may have supported the achievement of the positive outcome seen in this group.

Parent report indicated that the HFA group had more difficulty than the OO and TD groups in *all* domains of EF measured. The HFA group exhibited clinically significant deficits on the overall EF composite and on the indices measuring emotion and behavior regulation and set-shifting. The HFA group also scored in the clinically significant range on the ability to shift activities. Furthermore, the HFA group contained significantly more participants who scored in the clinically significant range within each domain of EF measured than did the OO or TD groups. It possible that, unlike the HFA group, individuals who achieved OO did not experience the same degree of EF deficits originally or EF deficits in the OO group may have been ameliorated through intensive early intervention or maturation. It is important to note that parent report of EF revealed considerably more differences in the performance of the HFA group as compared to the other two groups, than did direct testing of EF. This discrepancy may indicate that individuals with HFA are able to demonstrate age-appropriate EF tasks under optimal testing conditions, but show difficulty with these activities in everyday situations. This discrepancy may also reflect parental bias, in that parents of individuals with ASDs may over- or underreport current symptoms relative to their prior functioning. This study would have benefitted from the inclusion of a teacher's rating on the BRIEF in order to limit parental bias and to assess EF in school settings.

There are several additional limitations to the current study. First, the participants were homogeneous in terms of functioning level, race, and socioeconomic status. As a result, the generalizability of the current findings to the broader autism community may be limited, as was the study's power to detect small effects and subtle group differences.

The generalizability of the current findings may also be limited by the decision to require that the OO group have a minimum IQ of 77, in so limiting the extent to which these results would apply to a broader community of individuals who may have lost their ASD diagnosis. The decision to only include a high-functioning sample of individuals with ASD reflected this study's goal to characterize a group of individuals with a history of ASD who did not meet diagnostic criteria for ASD as they aged and who had average cognitive functioning and adaptive social and communication skills. We recognize this inclusion criteria may have excluded some individuals who had a history of ASD and who no longer meet diagnostic criteria for ASD but continue to exhibit cognitive, social, and behavioral deficits. While this group warrants further research, it is not the central goal of

this study, which aimed to describe an optimal outcome that included cognitive as well as symptomatic functioning.

In addition, it is possible that by restricting the cognitive functioning and adaptive socialization and communication skills of the OO and HFA groups to a minimum of 77, we limited the extent to which these groups may have differed from controls on EF. However, the functioning level specified in this study (IQ > 77) is consistent with criteria used by other studies of EF in high-functioning individuals with ASD that have found EF deficits (e.g., Kenworthy et al., 2005; Landa & Goldberg, 2005; Szatmari, Tuff, Finlayson, & Bartolucci, 1990). Further, no participants had to be excluded from the OO group because their cognitive or adaptive functioning level fell below the study's criteria (see Figure 1), suggesting that the study's enrollment criteria do not account for the OO group's high-functioning level or their average EF. It remains a possibility that the HFA group scored in the average range on all direct measures of EF because they were selected to have strong cognitive abilities. Consequently, the average EF performance of this HFA group may not generalize to other individuals with HFA.

In addition, the study is not prospective but rather uses parent interview and early reports to confirm a history of ASD in the OO group. A prospective study would allow for a standardized assessment of the initial ASD diagnosis and would allow for a more thorough analysis of factors that could predict such a positive outcome. However, because a prediction cannot be made about which children will go on to achieve an OO, a prospective study would have to longitudinally follow a very large sample of children with ASD in order to collect a large enough sample of children who will achieve OO. Consequently, such a prospective study would be prohibitively large and time consuming. To address the lack of prospective design, multiple steps were taken, including blind review of early reports and a requirement for the initial diagnosticians' expertise in ASDs, in order to ensure that all participants in the OO group were accurately diagnosed with an ASD in early childhood.

Evaluators who administered the ADOS were not blind to group membership of the participant, which may have biased the results. However, a rater blind to group membership coded videotapes of the ADOS administration for 5 participants in each of the three groups and high interrater reliability was found for both the algorithm and total items, indicating that ADOS scores were not influenced by the administrator's knowledge of the subject's group status.

The study was also limited because direct assessment of EF was conducted under optimal conditions in this study. The participants were tested in settings that limited distractions and an adult was present to monitor and redirect the participant back to tasks. While this means that performance on the standardized measures is probably reliable, it is possible that deficits in the OO and HFA group may have been masked by this testing approach, and that more deficits would be apparent if the testing environment more closely mirrored the participants' everyday life. That said, the BRIEF was employed to reflect EF in "everyday life."

The current study is one of the first to characterize executive functioning in a group of children with a history of ASD who have achieved favorable outcomes and no longer meet diagnostic criteria for an ASD. The results of this study indicate that children and adolescents who were once diagnosed with an ASD but who have achieved OO exhibit average EF abilities, even within the domains that are weak among similarly aged, high-functioning individuals who have retained their ASD diagnosis. However, a few differences in EF were found between individuals who achieved OO and age- and IQ-matched TD

peers and these differences included impulsivity, working memory, efficiency in problem-solving, and shifting of attention. Future studies should explore the predictors of such positive outcomes and would benefit from prospective studies that could include analyses of early developmental history, including treatment approaches and intensity.

Original manuscript received November 9, 2012
Revised manuscript accepted April 23, 2013
First published online June 4, 2013

REFERENCES

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, *15*(1), 31–36. doi:10.1093/arclin/15.1.31

Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology*, *22*, 969–978. doi:10.1016/j.acn.2007.08.001

Arffa, S., Lovell, M., Podell, K., & Goldberg, E. (1998). Wisconsin Card Sorting Test performance in above average and superior school children: Relationship to intelligence and age. *Archives of Clinical Neuropsychology*, *13*, 713–720. doi:10.1016/S0887-6177(98)00007-9

Baron, I. (2004). *Neuropsychological evaluation of the child*. New York, NY: Oxford University Press.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2007). Early intensive behavioral treatment: Replication of the UCLA model in a community setting. *Developmental and Behavioral Pediatrics*, *27*(2), S145–S155. doi:10.1097/00004703-200604002-00013

Damasio, A. R., & Maurer, R. G. (1978). A neurological model for childhood autism. *Archives of Neurology*, *35*, 777–786. doi:10.1001/archneur.1978.00500360001001

Delis, D. C., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan Executive Function System*. San Antonio, TX: Psychological Corporation.

Eigsti, I. M. (2011). Executive functions. In D. A. Fein (Ed.), *Neuropsychology of autism* (pp. 185–204). New York, NY: Oxford University Press.

Fein, D., Barton, M., Eigsti, I. M., Kelley, E., Naigles, L, Schultz, R., . . . Tyson, K. (2013). Optimal outcome in individuals with a history of autism. *Journal of Child Psychology and Psychiatry*, *54*(2), 195–205. doi: 10.1111/jcpp.12037.

Fein, D., Dixon, P., Paul, J., & Levin, H. (2005). Brief report: Pervasive Developmental Disorder can evolve into ADHD: Case illustrations. *Journal of Autism and Developmental Disorders*, *35*, 525–534. doi:10.1007/s10803-005-5066-3

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function*. Odessa, FL: Psychological Assessment Resources.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four- to six-year follow-up. *Journal of Autism and Developmental Disorders*, *30*(2), 137–142. doi:10.1023/A:1005459606120

Helt, M., Kelley, E., Kinsbourne, M., Pandey, J., Boorstein, H., Herbert, M., & Fein, D. (2008). Can children with autism recover? If so, how? *Neuropsychology Review*, *18*(4), 339–366. doi:10.1007/s11065-008-9075-9

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, *45*, 212–229. doi:10.1111/j.1469–7610.2004.00215.x

Kelley, E., Naigles, L. R., & Fein, D. (2010). An in-depth examination of optimal outcome children with a history of autism spectrum disorders. *Research in Autism Spectrum Disorders*, *4*, 526–538. doi:10.1016/j.rasd.2009.12.001

Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual language deficits in OO children with a history of autism. *Journal of Autism and Developmental Disorders*, *36*, 807–828. doi:10.1007/s10803-006-0111-4

Kenworthy, L. E., Black, D. O., Wallace, G. L., Ahluvalia, T., Wagner, A. E., & Sirian, L. M. (2005). *Developmental Neuropsychology*, *28*, 809–827. doi: 10.1207/s15326942dn2803_4

Koshino, H., Carpenter, P. A., Minshew, N. J., Cherkassky, V. L., Keller, T. A., & Just, M. A. (2005). Functional connectivity in an fMRI working memory task in high-functioning autism. *NeuroImage*, *24*(3), 810–821. doi:10.1016/j.neuroimage.2004.09.028

Landa, R. J., & Goldberg, M. C. (2005). Language, social, and executive functions in high functioning autism: A continuum of performance. *Journal of Autism and Developmental Disorders*, *35*, 557–573. doi: 10.1007/s10803-005-0001-1

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., . . . 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. doi:10.1023/A:1005592401947

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, *55*, 3–9. doi:10.1037/0022-006X.55.1.3

Luna, B., Doll, S. K., Hegedus, S. J., Minshew, N. J., & Sweeney, J. A. (2007). Maturation of executive function in autism. *Biological Psychiatry*, *61*(4), 474–481. doi:10.1016/j.biopsych.2006.02.030

Luyster, R., Richler, J., Risi, S., Hsu, W., Dawson, G., Bernier, R., . . . Lord, C. (2005). Early regression in social communication in Autism Spectrum Disorders: A CPEA Study. *Developmental Neuropsychology*, *27*, 311–336. doi:10.1207/s15326942dn2703_2

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R. . . . Denckla, H. S. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, *17*, 643–662.

Mundy, P. (1993). Normal versus high-functioning status in children with autism. *American Journal on Mental Retardation*, *97*(4), 381–384.

O'Hearn, K., Asato, M., Ordaz, S., & Luna, B. (2008). Neurodevelopment and executive function in autism. *Development and Psychopathology*, *20*(4), 1103–1132. doi:10.1017/S0954579408000527

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*, 51–87. doi:10.1111/j.1469–7610.1996.tb01380.x

Piven, J., Harper, J., Palmer, P., & Arndt, S. (1996). Course of behavioral change in autism: A retrospective study of high-IQ adolescents and adults. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*, 523–529.

Rutter, M. (1970). Autistic children: Infancy to adulthood. *Seminars in Psychiatry*, *2*, 435–450.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal of Mental Retardation*, *110*(6), 417–438. doi:10.1352/0895-8017(2005)110[417:IBTFCW]2.0.CO;2

Schuh, J. M., & Eigsti, I. M. (in press). Working memory, language skills, and autism symptomatology. *Behavioral Sciences*.

Seltzer, M. M., Shattuck, P., Abbeduto, L., & Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *10*, 234–247. doi:10.1002/mrdd.20038

Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development*, Serial No. 256, *64*. Chicago, IL: University of Chicago Press.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales* (Interview ed.). Circle Pines, MN: American Guidance Service.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Sutera, S., Pandey, J., Esser, E., Rosenthal, M., Wilson, L., Barton, M., & Fein, D. (2007). Predictors of optimal outcome in toddlers diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(1), 98–107. doi:10.1007/s10803-006-0340-6

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*(2), 213–225. doi:10.1007/BF02211842

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

Turner, L. M., & Stone, W. L. (2007). Variability in outcome for children with an ASD diagnosis at age 2. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *48*(8), 793–802.

Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry*, *33*, 489–597. doi:10.1111/j.1469-7610.1992.tb00887.x

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. New York, NY: The Psychological Corporation.

Weiss, M. J. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behavioral Interventions*, *14*, 3–22. doi:10.1002/(SICI)1099-078X(199901/03)14:1<3::AID-BIN25>3.0.CO;2-F

Zachor, D. A., Ben-Itzchak, E., Rabinovich, A. L., & Lahat, E. (2007). Change in autism core symptoms with intervention. *Research in Autism Spectrum Disorders*, *1*, 304–317. doi:10.1016/j.rasd.2006.12.001