

Social Function and Communication in Optimal Outcome Children and Adolescents with an Autism History on Structured Test Measures

Alyssa J. Orinstein · Joyce Suh · Kaitlyn Porter · Kaitlin A. De Yoe · Katherine E. Tyson · Eva Troyb · Marianne L. Barton · Inge-Marie Eigsti · Michael C. Stevens · Deborah A. Fein

© Springer Science+Business Media New York 2015

Abstract Youth who lose their ASD diagnosis may have subtle social and communication difficulties. We examined social and communication functioning in 44 high-functioning autism (HFA), 34 optimal outcome (OO) and 34 typically developing (TD) youth. Results indicated that OO participants had no autism communication symptoms, no pragmatic language deficits, and were judged as likable as TD peers. Some group differences were found: OO youth had less insight into social relationships and poorer friendship descriptions than TD youth. OO participants had attention, self-control, and immaturity difficulties that may impact social abilities. However, OO participants were most engaged, friendliest, warmest, and most approachable. Overall, OO participants had no social and communicative impairments, although some exhibited mild social difficulties that often accompany attentional problems.

Keywords Autism spectrum disorder · Outcome · Optimal outcomes · Social · Communication

Introduction

Autism spectrum disorder (ASD) is generally considered a lifelong disorder; however, studies of autism in adolescence and adulthood have demonstrated a reduction of symptoms (Gilchrist et al. 2001), with improvements most frequently occurring during the pre-adolescent through early adolescent period (Kobayashi et al. 1992). More significant improvement also occurs most commonly in individuals with higher IQs (McGovern and Sigman 2005). Social and communicative aspects of language often improve (Ballaban-Gil et al. 1996; Piven et al. 1996). Some communication deficits are more likely to fully remit with age in individuals with ASD, including use of idiosyncratic language (Seltzer et al. 2003). Other skills may improve but are likely to remain impaired, such as nonverbal communication, pragmatic language, atypical prosody, stereotyped or repetitive language, and asking inappropriate questions (Gilchrist et al. 2001; Orsmond et al. 2004; Seltzer et al. 2003; Whitehouse et al. 2009). In regards to social skills, a substantial portion of adolescents and adults with ASD do not display many of the inappropriate social behaviors typical of younger individuals with ASD, including using others' bodies as tools, making atypical social overtures, and being unable to comfort others. Skills that are likely to remain impaired include engaging in reciprocal social interactions, forming and maintaining relationships, sharing enjoyment with others, making appropriate eye contact, and showing a range of or appropriate facial expressions (Howlin et al. 2000; Seltzer et al. 2003).

Individuals with high-functioning autism (HFA) are generally impaired in their spontaneous speech and conversational ability (Eales 1993; Freitag et al. (2006). Compared to typically developing peers, adolescents with

A. J. Orinstein (✉) · J. Suh · K. Porter · K. A. De Yoe · K. E. Tyson · E. Troyb · M. L. Barton · I.-M. Eigsti · D. A. Fein
Unit 1020, Department of Psychology, University of Connecticut, Storrs, CT 06269, USA
e-mail: alyssa.orinstein@uconn.edu

M. C. Stevens
Olin Neuropsychiatry Research Center, Institute of Living, Hartford Hospital, Hartford, CT, USA

M. C. Stevens
Yale University School of Medicine, New Haven, CT, USA

HFA have difficulty with topic management and reciprocity, have unusual speech characteristics, and use less appropriate gaze, facial expressions, and gestures (Paul et al. 2009). Younger children with HFA fail to take turns appropriately in conversation, persevere on topics, and fail to clarify ambiguities (Lam and Young 2012) and these pragmatic deficits in reciprocal social interactions and relationships tend to persist into adolescence and adulthood (Howlin et al. 2000; Seltzer et al. 2003). Friendship quality has also consistently remained impaired in adolescence and adulthood in individuals with HFA (Howlin 2003; Howlin et al. 2000; Shattuck et al. 2007; Whitehouse et al. 2009). Some studies have reported very low percentages of those who report having friends, ranging between 0 and 15.8 % (Howlin 2003; Howlin et al. 2000; Orsmond et al. 2004; Shattuck et al. 2007; Whitehouse et al. 2009). Eaves and Ho (2008) had more promising findings, with 33 % of young adults with ASD reporting at least one close friendship. Seltzer et al. (2003) found that quantity and quality of friendships was unlikely to change over time, as only 4.4 % who did not have true friendships between the ages of ten and fifteen did so at the time of the study (mean age of 22). Shattuck et al. (2007), similarly, found that the increase of individuals with ASD who had friendships was only 7.5 % over 4.5 years.

Thus, there is considerable evidence that, despite well-documented gains, the social and communication symptoms and delays of ASD are likely to persist into adolescence and adulthood.

However, some studies have described individuals who actually lose their ASD diagnosis, suggesting an even greater amelioration of symptoms. The first published study noting “recovery” in autism was conducted by Lovaas (1987). He reported that after receiving intensive behavioral intervention, nine of 19 children in the study “recovered,” as indicated by completion of first grade in a regular classroom and by achieving an average or above IQ score. However, this study did not indicate whether autism symptomatology had been completely resolved. Since then, studies have found somewhat lower rates of “recovery,” which will be referred to as optimal outcome (OOs), generally between 3 and 25 %, using varied criteria (see Helt et al. 2008 for a review). It is important to note that these later studies were not treatment studies, so it may not be appropriate to directly compare the rates of “recovery” or “optimal outcome.”

A few recent studies have examined the current behavioral presentation of OO children. Fein et al. (2005) reported on a number of ASD children in whose early childhood clinical presentations evolved into behaviors more characteristic of Attention-Deficit/Hyperactivity Disorder (ADHD) by age eight. Some of the children continued to display mild social awkwardness or delays,

but their social difficulties were more consistent with those typically found in children with ADHD. Specifically, the children tended to be impulsive, aggressive, or immature, rather than withdrawn or odd (Fein et al. 2005). Kelley et al. (2006) focused on language functioning in a group of 14 OO children, ages 5–9; these children, although within the normal range on all standardized language measures, continued to show subtle difficulties in semantic and pragmatic areas of language. A later study of 13 of these children at age 8–13, found language, adaptive, and problem behavior scores within the average range (Kelley et al. 2010). A recent study by Fein et al. (2013) examined a group of OO children and adolescents between the ages of 8 and 21 and compared them to children and adolescents with HFA and typical development (TD). The authors found that, based on parent report of early history, the OO group had somewhat milder social symptoms than the HFA group, but did not differ in communication or repetitive behavior symptoms. In addition, results showed that, at the time of the study, the OO participants did not differ from the TD participants on summary measures of socialization, communication, face recognition, or most language subscales. Anderson et al. (2014) found eight adolescents out of 85 in their study who had obtained ‘very positive outcomes’ and lost their ASD diagnosis. These adolescents had adaptive functioning scores well within the average range and had clear social strengths. They also had less repetitive behavior, hyperactivity, irritability, and depression compared with high-functioning adolescents who retained their ASD diagnosis (Anderson et al. 2014). Based on these findings, more research is needed on what areas of social relationships may be challenging to individuals with a history of a diagnosis of ASD.

The current study presents a more detailed analysis of social and communicative abilities in the participants described in Fein et al. (2013). The aim is to determine whether OO children and adolescents exhibit subtle residual symptoms not apparent on summary scores by examining specific communication and social behaviors in more detail. Since such social communication impairments are generally considered the core of autism symptomatology, and tend to persist in individuals with HFA as they enter adolescence and young adulthood, it might be expected that OO individuals, even while falling within the normal range on standardized measures of social and communication functioning, would show subtle impairments in these areas with more sensitive measures. Such impairments might be appropriate targets for continued intervention. In particular, it was anticipated that, compared to TD individuals, the OO individuals would have a few mild, residual ASD symptoms in the communication and social domains, display more pragmatic language problems, provide poorer friendship descriptions, be

judged as less likable, and have more psychiatric symptoms, such as inattention, that may relate to struggles in communication and social areas. Importantly, in all areas, both OO and TD individuals were anticipated to perform significantly better than the HFA individuals. In addition, it was hypothesized that for all groups, more autism communication symptoms and poorer adaptive communication and social functioning would be correlated with worse performance on the communication and social measures in the current study.

Methods

Participants

The current study used the participants and testing procedures as described in Fein et al. (2013). Thirty-four individuals with a history of ASD and OO, 44 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. Groups were matched on age, gender, and nonverbal IQ, but were significantly different on verbal IQ (see Table 1). Six HFA participants and three OO participants were evaluated at Queens University in Ontario, Canada. Their performance did not differ from the remaining participants on any measure. The participants tested at the University of Connecticut were primarily from the northeast US. Participants were mostly White, with three OO individuals,

two HFA individuals, and three TD individuals reporting other races or ethnicities.

The study was approved by Institutional Review Boards at the University of Connecticut, the Institute of Living of Hartford Hospital and Queens University. See Fein et al. (2013) for a flow chart of participant inclusion and exclusion.

Inclusion Criteria

All participants had verbal, nonverbal, and full-scale IQ standard scores greater than 77 (within 1.5 SD of the average of 100). Additional OO criteria were:

1. ASD diagnosis before the age of 5 by a physician or psychologist specializing in autism, in a written report. Documented early language delay (no words by 18 months or no phrases by 24 months) was required. 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.

Table 1 Participant characteristics

Mean (SD) Range	HFA N = 44	OO N = 34	TD N = 34	<i>F</i> / χ^2	<i>p</i>	Post-hoc
Gender (male:female)	40:4	27:7	31:3	2.92	.23	
Age	13.9 (2.7) 8.63–20.04	12.8 (3.5) 8.12–21.24	13.9 (2.6) 9.93–21.7	1.66	.20	
WASI VIQ	105.4 (14.4) 81–142	112.7 (13.7) 80–137	112.0 (11.2) 93–138	3.62	.03	HFA < OO
WASI NVIQ	110.2 (12.8) 78–147	110.3 (15.1) 81–142	112.8 (11.3) 89–139	0.45	.64	
Vineland—communication	82.7 (13.9) 42–108	98.3 (12.7) 79–122	93.3 (9.3) 74–119	15.8	<.001	HFA < OO, TD
Vineland—social	75.5 (16.0) 46–109	102.0 (8.4) 80–118	101.7 (8.6) 86–120	62.0	<.001	HFA < OO, TD
ADOS—communication	3.50 (1.42) 2–7	0.47 (0.62) 0–2	0.41 (0.56) 0–2	124.2	<.001	OO, TD < HFA
ADOS—social	6.77 (2.21) 4–13	1.09 (1.31) 0–4	0.50 (0.75) 0–2	183.7	<.001	OO, TD < HFA

2. On the phone screening, parents had to report that the participant had typically developing friends. During evaluation, participants could not meet criteria for any ASD on the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2000). In addition, the ADOS videotapes of all potential OO cases were reviewed by a clinician with more than 15 years of autism diagnostic experience (IME, MB, or DF) who confirmed that ADOS scores were below ASD thresholds and that, in their expert clinical judgment, an ASD was not present. Five potential OO participants were judged to have social impairments with an autistic quality and were excluded. These five children were borderline cases, as they had an autistic quality but would not have met criteria for the HFA group.
3. Participants' scores on the Communication and Socialization domains of the Vineland Adaptive Behavior Scales (Vineland; Sparrow et al. 1985) had to be greater than 77 (within 1.5 SDs of the mean of 100).
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 ASD, such as attention or academic difficulties.

To be included in the HFA group:

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

To be included in 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. 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.

Exclusion Criteria

Potential participants for any group were excluded if (1) at the time of the telephone screening they exhibited

symptoms of major psychopathology that would impede full participation, (2) they had severe visual or hearing impairments, or (3) they had a seizure disorder, Fragile X syndrome, or significant head trauma with loss of consciousness. Two in the TD group and two in the HFA group were excluded because of possible seizure disorder based on parent report; none were excluded for other reasons.

Procedure

Potential participants who passed the telephone screening were scheduled for an assessment. For participants under 18, parent consent and child assent were obtained prior to testing. For participants 18 and over, informed consent was obtained. Participants received a monetary incentive for participation, even if the testing could not be completed.

Measures

The *Wechsler Abbreviated Scale of Intelligence* (WASI; Wechsler 1999) was used to assess verbal and nonverbal cognitive abilities. The *Vineland Adaptive Behavior Scales* (Vineland; Sparrow et al. 1985) assessed Communication and Socialization skills via parent interview. Modules 3 or 4 of the *Autism Diagnostic Observation Schedule* (ADOS; Lord et al. 2000), a structured play and interview session, were used to assess autistic features in the Communication and Reciprocal Social Interaction domains. ADOS administrations were videotaped and five administrations per group were coded by a rater blind to group status, with high inter-rater reliability for both algorithm (86.7 %) and total items (85.7 %).

Test of Language Competence—Expanded Edition (TLC-E; Wiig and Secord 1989) *Making Inferences* subtest assesses the ability to understand verbal descriptions of situations and to generate multiple plausible inferences. The *Figurative Language* subtest assesses the ability to comprehend and interpret metaphors.

Clinical Evaluation of Language Fundamentals (CELF-4; Semel et al. 2003) *Pragmatics Profile* (PP) was available for 25 OO, 27 HFA, and 21 TD participants. The PP contains a checklist of 52 items in three domains. The *Rituals and Conversational Skills* (*Conversation*) domain includes turn-taking, introduction and maintenance of topics, and appropriate strategies for getting attention. The *Asking For, Giving, and Responding to Information* (*Information*) domain includes giving and responding to advice or suggestions appropriately, agreeing and disagreeing using appropriate language, and initiating and responding to verbal and nonverbal negotiations. The *Nonverbal Communication Skills* (*Nonverbal*) domain includes facial cues, body language, and tone of voice.

Based on videotaped evaluations of the ADOS, individual behaviors were scored by a trained undergraduate student rater, blind to group membership, on a 1–4 Likert Scale (1 = “never observed”; 4 = “always observed”). Twelve items were excluded from analysis because the items/behaviors were not applicable to situations presented in the ADOS. An additional two items were excluded because they could not be coded from the video. An item could also be scored as “Not Observed,” which means that although the behavior was potentially applicable to situations presented in the ADOS, the participant did not display the behavior. The distribution and number of “Not Observed” items varied for each participant. Therefore, a mean score was calculated based on coded items for each participant in each domain. Ten percent of the tapes were double-scored to establish inter-rater reliability. Intraclass correlations (ICCs) for mean domain scores were 0.88 for Conversation, 0.82 for Information, and 0.62 for Nonverbal, which is considered high moderate to good reliability (Doi and Williams 2013).

The *Friendship Description Rating Scale* was developed for the current study based on the five categories of friendship quality assessed on the Friendship Qualities Scale (Bukowski et al. 1994): companionship, security-intimacy and trust, closeness, help, and conflict. Each question was scored using a Likert scale format, from 1 (not at all) to 5 (very much). Friendship descriptions from the ADOS were transcribed and used to complete the Friendship Description Rating Scale. Thirty undergraduate students from the University of Connecticut, blind to group membership and the purpose of the study, rated each participant’s description of friendship from the ADOS, with each student rating nine descriptions. These data were available for 15 participants from each group, closely matched on age and of a smaller age range (9–15), to reduce differences based on developmental level. The scores for each participant were averaged across raters to create an average score for each item. A total score for the Friendship Description Rating Scale was computed by re-coding the reverse scored items and summing the average scores for each item to create a total friendship description quality score for each participant.

Reysen Likability Scale (Reysen 2005) is used to assess a participant’s likability, including factors such as knowledge and attractiveness, as well as requiring raters to imagine the participant as part of their lives in roles such as friend, roommate, or coworker. The 11 item scale was modified for the present study to include three additional items, which asked the rater to judge the likelihood that the participant has a group of friends, has a best friend, and is popular. One five to seven minute segment of the ADOS videos, in which the participant talks about emotions and tells the story of a cartoon, was used to complete the modified Reysen Likability Scale. Five undergraduate research assistants, blind to group membership,

watched the ADOS segments and coded the modified Reysen Likability Scale. Because likability is based on individual subjective preferences, inter-rater reliability was not calculated for this measure. Each question was scored using a Likert scale format, from 1 (very strongly disagree) to 7 (very strongly agree). The raters were told to rate the items as if they were the same age as the participant. All items are positively scored, with higher scores representing greater likability or social status of the participant. The scores for each participant were averaged across observers to create an average score for each item. A total score for modified Reysen Likability Scale was computed by summing the average scores for each item to create a total likability score for each participant.

Semistructured Clinical Interview for Children & Adolescents (SCICA) Observation Form (McConaughy and Achenbach 2001) allows for the rating of observations of participants’ behavior, affect, and interaction style. The SCICA Observation form was used in a non-standard manner, as the coding was based on the video-taped ADOS, rather than the SCICA interview. Some items were excluded for the current study because they were unlikely to be elicited/observed in the context of the ADOS interview. The items on the observation form of the SCICA combine to produce syndrome scales, including: Anxious (13 items), Withdrawn/Depressed, (17 items) Language/Motor Problems (9 items), Attention Problems (10 items), and Self-Control Problems (9 items). 58 items were scored in total for the current study. The domains were proposed by the authors of the test, based on factor analyses. The Language/Motor items seem more reflective of the child’s Immaturity; therefore, the domain will be referred to as such in this paper. The first 30 min of the ADOS videos was again utilized to complete the SCICA Observation Form. Each behavior on the SCICA Observation Form was coded on a four point scale, with 0 being no occurrence, 1 being slight or ambiguous occurrence, 2 being definite occurrence with mild to moderate intensity, and 3 being definite occurrence with severe intensity. A total score for each of the five scales was calculated by summing the individual item scores within that scale. A trained undergraduate observer, blind to group status, watched the ADOS interview, coding all SCICA Observation items one time, at the end of the 30 min. Two trained undergraduate observers coded eighteen percent of the videos for reliability purposes. The use of the SCICA as a video-coded measure is experimental, so a larger percentage of videos were coded for reliability than for the ADOS itself. ICCs for the domain total scores ranged from 0.63 to 0.81, which is considered high moderate to good reliability (Doi and Williams 2013).

Results

The scores on many measures did not meet the statistical assumptions of normality or homogeneity. When Levene’s

test for homogeneity of variance was violated, the Games-Howell post hoc test was used; in all other cases, the Tukey HSD post hoc test was used. There is no non-parametric test equivalent for a MANOVA, but, to increase confidence in the results, non-parametric test equivalents were also conducted for all ANOVAs and *t* tests, with results displayed in the tables. Significant and nonsignificant findings from parametric and non-parametric tests were identical and significance levels were very similar.

In this study, the goal was to elucidate all possible social and communication deficits in the OO individuals, rather than prematurely concluding that they lost all such symptoms; therefore, it was considered most conservative in this case not to correct for multiple comparisons, so that even small to moderate effect sizes would be preserved.

Communication

As previously reported in Fein et al. (2013), there were no differences between the OO and TD groups on either the Communication domain of the Vineland or the Communication domain of the ADOS, suggesting globally intact communication functioning in the OO group (see Table 1). As expected, the HFA group had significantly worse Communication scores on both the Vineland and ADOS (see Table 1).

ADOS Items

The current analyses looked at the nine individual communication behaviors common to ADOS Modules 3 and 4. A MANOVA indicated a significant group difference on the combined dependent variables, $F(18, 198) = 12.6$, $p < .001$, Wilks' $\lambda = 0.22$; $d = 0.50$. Follow-up ANOVAs on each of the dependent variables revealed two variables (overall language level and echolalia) that did not differ by group. *t* tests on the remaining seven items indicated that there were no communication items on which the OO group scored higher (worse) than the TD group, suggesting that there were no residual deficits in autism communication symptoms; see Table 2. In fact, the individuals in OO group asked the examiner for information more frequently and showed a trend to offer more information than the TD group (also confirmed via Mann-Whitney U non-parametric tests); see Table 2. The HFA group scored significantly higher (worse) than the OO group on these items.

TLC-E Subtests

On Making Inferences, the groups differed significantly $F(2, 105) = 8.29$, $p < .001$, $d = 0.56$. Post-hocs indicated that OO and TD groups did not differ. HFA scores ($M = 8.29$, $SD = 2.94$) were significantly lower than the

TD ($M = 11.0$, $SD = 2.77$) but not the OO ($M = 9.82$, $SD = 2.86$) scores; see Table 3.

On Figurative Language, the groups differed significantly $F(2, 105) = 20.58$, $p < .001$, $d = 0.89$. Post-hocs indicated that the HFA scores ($M = 7.46$, $SD = 2.56$) were lower than both the OO ($M = 9.91$, $SD = 2.80$) and TD ($M = 11.2$, $SD = 2.30$) scores, which did not differ; see Table 3.

CELF-4 Pragmatics Profile

There was a main effect of group on each of the three PP domains (see Table 3). Post-hocs indicated that the HFA group scored worse than the TD and OO groups on each of the three domains; the TD and OO groups did not differ. These results were confirmed by non-parametric Kruskal-Wallis tests.

Socialization

As discussed in Fein et al. (2013), the OO and TD groups did not differ on the Social domain of the Vineland or the ADOS, suggesting globally intact social functioning (see Table 1). As expected, the HFA group had significantly worse scores.

ADOS Items

A MANOVA examined group differences on ten ADOS social items common to Modules 3 and 4. There was a main effect of group on the combined variables, $F(20, 200) = 16.3$, $p < .001$, Wilks' $\lambda = 0.14$; $d = 0.57$.

When evaluated individually, all of the social items showed significant group differences. The OO group scored significantly higher (worse) than the TD group on insight into the nature of social relationships, with 12/35 (35 %) of the OO participants displaying mild abnormality; see Table 2. For example, one 12-year-old OO participant stated that people get married because "the human race cannot survive without being married," and an 8-year-old OO participant stated that a friend is "when people stick to my games." The OO and TD groups did not differ on any remaining social items (confirmed via Mann-Whitney U non-parametric tests); see Table 2. The HFA group scored significantly higher than the OO group on all ADOS social items (Table 2).

Friendship Description Rating Scale

An ANOVA on total scores was statistically significant (see Table 4). The TD group scored the highest (best), the HFA group scored the lowest (worst), and the OO group scored in the middle, different from both other groups.

Table 2 ADOS items

	HFA		OO		TD		t tests: OO versus TD				Mann-Whitney U tests: OO versus TD				t tests: OO versus HFA				Mann-Whitney U tests: OO versus HFA			
	Mean (SD)		Mean (SD)		Mean (SD)		t	p	Cohen's d	U	Z	p	R	t	p	Cohen's d	U	Z	p	R		
<i>Communication items</i>																						
Asks for information	1.58 (0.73)	0.73 (0.63)	1.09 (0.75)	-2.13	.04	-0.52	413	-2.03	.04	.25	-5.37	<.001	1.25	294	-4.66	<.001	-.53					
Offering information	0.65 (0.57)	0.03 (0.17)	0.15 (0.36)	-1.72	.09	-0.43	510	-1.70	.09	.21	-6.75	<.001	1.47	310	-5.20	<.001	-.59					
Reporting of events	0.48 (0.55)	0.03 (0.17)	0.09 (0.29)	-1.02	.31	-0.25	544	-1.02	.31	.12	-5.10	<.001	1.11	430	-4.17	<.001	-.47					
Use of emphatic gestures (Mod 4)	1.40 (1.27)	0.14 (0.36)	0.20 (0.41)	-0.42	.68	-0.16	132	-0.42	.67	.07	-4.18	<.001	1.35	55	-3.28	.001	-.56					
Stereotyped or idiosyncratic language	0.89 (0.81)	0.15 (0.36)	0.12 (0.33)	0.35	.73	0.09	561	-0.36	.72	.04	-5.39	<.001	1.18	369	-4.34	<.001	-.49					
Use of gestures	0.70 (0.59)	0.12 (0.33)	0.09 (0.29)	0.39	.70	0.10	561	-0.40	.69	.05	-5.56	<.001	1.21	354	-4.60	<.001	-.52					
Conversation	1.07 (0.63)	0.15 (0.36)	0.09 (0.29)	0.75	.46	0.18	544	-0.75	.46	.09	-7.66	<.001	1.79	204	-6.01	<.001	-.68					
Presence of speech abnormalities	1.32 (0.60)	0.24 (0.50)	0.15 (0.36)	0.84	.40	0.21	542	-0.68	.50	0.08	-8.71	<.001	1.20	166	-6.27	<.001	-.71					
Language level	0.02 (0.15)	0.00 (0.00)	0.00 (0.00)	-	-	-	578	0.00	1.00	0.00	-.88	.38	0.19	731	-0.88	.38	-.10					
Echolalia	0.07 (0.26)	0.00 (0.00)	0.00 (0.00)	-	-	-	578	0.00	1.00	0.00	-1.77	.08	0.38	697	-1.54	.12	-.17					
<i>Social items</i>																						
Language production and linked nonverbal communication	0.30 (0.46)	0.00 (0.00)	0.06 (0.24)	-1.44	.16	-0.35	544	-1.36	.17	0.16	-4.25	<.001	0.92	527	-3.45	<.001	-.39					
Shared enjoyment	0.45 (0.59)	0.03 (0.17)	0.12 (0.33)	-1.40	.17	-0.34	527	-1.38	.17	0.17	-4.55	<.001	0.97	463	-3.85	<.001	-.44					
Amount of reciprocal social communication	0.89 (0.69)	0.09 (0.29)	0.15 (0.36)	-0.75	.46	-0.18	544	-0.75	.46	0.09	-6.94	<.001	1.51	275	-5.39	<.001	-.61					
Quality of social response	0.82 (0.45)	0.03 (0.17)	0.03 (0.17)	0.00	.99	0.00	578	0.00	1.00	0.00	-10.75	<.001	2.32	145	-6.66	<.001	-.75					
Quality of social overtures	0.89 (0.39)	0.06 (0.24)	0.03 (0.17)	0.58	.56	0.14	561	-0.59	.56	0.07	-11.61	<.001	2.56	145	-6.97	<.001	-.79					
Empathy or comment on others' emotions	1.32 (0.71)	0.29 (0.46)	0.21 (0.41)	0.83	.41	0.18	527	-0.83	.40	0.10	-7.31	<.001	1.72	215	-5.74	<.001	-.65					
Quality of rapport with examiner	0.80 (0.55)	0.21 (0.41)	0.12 (0.33)	0.98	.33	0.24	527	-0.98	.33	0.12	-5.20	<.001	1.22	348	-4.58	<.001	-.52					
Communication of own affect (Mod 4)	1.00 (0.73)	0.29 (0.47)	0.15 (0.37)	0.95	.35	0.33	121	-0.95	.34	0.16	-3.23	.003	1.16	62	-3.04	.002	-.52					
Facial expressions directed to others	0.84 (0.48)	0.21 (0.41)	0.09 (0.29)	1.67	.18	0.34	510	-1.43	.15	0.17	-6.17	<.001	1.41	300	-5.16	<.001	-.58					
Responsibility for own actions (Mod 4)	0.70 (0.57)	0.07 (0.27)	0.00 (0.00)	1.00	.34	0.37	130	-1.20	.23	0.21	-4.30	<.001	1.41	59	-3.13	.001	-.54					
Appropriateness of eye contact	1.27 (0.97)	0.18 (0.58)	0.00 (0.00)	1.79	.083	0.44	527	-1.76	.07	0.21	-6.12	<.001	1.36	338	-4.87	<.001	-.55					

Table 2 continued

	HFA		OO		TD		t tests: OO versus TD			Mann-Whitney U tests: OO versus TD			t tests: OO versus HFA			Mann-Whitney U tests: OO versus HFA			
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	t	p	Cohen's d	U	Z	p	t	p	Cohen's d	U	Z	p	
Insight into social relationships	1.18 (0.69)	0.41 (0.61)	0.15 (0.36)	0.15 (0.36)	2.18	.033	2.18		0.52	454	-2.02	.04	0.24	-5.14	1.18	330	-4.52	<.001	-0.51

For most items, there were 34 participants in the OO group and 34 participants in the TD group. Use of emphatic gestures, communication of own affect, and responsibility for own actions are only on ADOS Module 4 and therefore contained a smaller sample of 14 OO participants and 20 TD participants
 Bold values are statistically significant $p < .10$

A MANOVA examined group differences on twelve individual items. There was a main effect of group on the combined variables, $F(24, 62) = 2.91, p < .001$, Wilks' $\lambda = 0.22; d = 0.43$. When items were considered separately, 10 of the 12 items differed by group. The two items that did not differ were about jealousy and annoyance. The remaining items were probed with ANOVAs. The HFA group was rated the poorest on overall description of a friend and on unusualness of the description, while the TD group was rated the best on both items. The OO group showed a trend for a worse overall description than the TD group but showed no difference from the TD group in unusualness of the description (see Table 4). For several items (time spent with friends, closeness, trust, reliability, and bond), the HFA participants' descriptions were rated the lowest, the TD participants' descriptions were rated the highest, with OO participants' descriptions in the middle, differing from both other groups. For two other items (affection and helpfulness), the HFA group was rated lower than the other two groups; OO and TD did not differ from each other. Finally, the HFA group was rated as having more conflict in their friendship descriptions than the TD group; the OO group did not differ from either of the other groups.

Modified Reysen Likability Scale

A repeated measures ANOVA was conducted to compare the average Likability score for each rater (raters 1–5) for each group (HFA, OO, TD). There was a significant mean effect of rater, Wilks' Lambda = 0.476, $F(4, 36) = 9.91, p < .001$. More importantly, there was no rater x group interaction, Wilks' Lambda = 0.719, $F(8, 72) = 1.62, p = .136$. Therefore, raters' scores were averaged for the group analyses.

All of the items on the Modified Reysen Likability Scale were highly intercorrelated (r s between .40 and .91), suggesting that all items are tapping into the same construct. For the 14 individual items (Table 5), MANOVA indicated a significant group difference on the combined dependent variables, $F(28, 118) = 3.04, p < .001$, Wilks' $\lambda = 0.34; d = 0.32$. An ANOVA of the Total score indicated that the OO (Mean = 64.04, SD = 7.19) and TD (Mean = 62.10, SD = 7.60) groups did not differ from each other and scored significantly higher than the HFA (Mean = 51.29, SD = 7.19) group, $F(2, 72) = 23.60, p < .001$.

All 14 individual variables differed by group. There were no items on which the OO group scored lower (worse) than the TD group, suggesting that OO participants were perceived as being at least as likable as the TD participants (see Table 5). In fact, the OO group was rated as significantly friendlier, warmer, and more approachable than the TD group. The HFA group scored significantly lower

Table 3 TLC-E and CELF-4 pragmatics profile scores

Mean (SD)	HFA (N = 41)	OO (N = 33)	TD (N = 34)	F	p	Post-hoc	Effect size
TLC-E listening comprehension	8.29 (2.94)	9.82 (2.86)	10.97 (2.77)	8.29	<.001	HFA < TD	OO and HFA = 0.53 OO and TD = 0.41 HFA and TD = 0.94
TLC-E figurative language	7.46 (2.56)	9.91 (2.80)	11.18 (2.30)	20.58	<.001	HFA < OO, TD	OO and HFA = 0.91 OO and TD = 0.50 HFA and TD = 1.53
Mean (SD)	HFA (N = 28)	OO (N = 24)	TD (N = 21)	F	p	Post-hoc	Effect size
CELF-4 PP conversation	3.46 (0.42)	3.79 (0.26)	3.85 (0.17)	11.77	<.001	HFA < OO, TD	OO and HFA = 0.94 OO and TD = 0.27 HFA and TD = 1.22
CELF-4 PP information	3.66 (0.31)	3.95 (0.09)	3.98 (0.07)	18.76	<.001	HFA < OO, TD	OO and HFA = 1.27 OO and TD = 0.37 HFA and TD = 1.42
CELF-4 PP nonverbal	3.45 (0.46)	3.87 (0.24)	3.88 (0.15)	14.46	<.001	HFA < OO, TD	OO and HFA = 1.14 OO and TD = 0.05 HFA and TD = 1.26

(worse) than the OO group on all 14 likability items and significantly worse than the TD group on 13/14 likability items (friendly was the only exception) (see Table 5).

SCICA Observation Form

Anxious Scale A MANOVA was conducted to determine whether there were group differences on the 13 items that make up the Anxious scale of the SCICA and was not significant for the combined dependent variables, [$F(26, 100) = 1.02, p = .449, \text{Wilks' } \lambda = 0.62; \text{partial eta squared} = 0.21$]; therefore, no further analyses were conducted.

Withdrawn/Depressed Scale Another MANOVA was conducted to determine whether there were group differences on the 17 items that make up the Withdrawn/Depressed scale of the SCICA. The MANOVA indicated significant group differences on the combined dependent variables, [$F(34, 106) = 1.93, p = .006, \text{Wilks' } \lambda = 0.38; \text{partial eta squared} = 0.38$]. There was also a group effect on the Withdrawn/Depressed total score, which was created by summing the individual items (see Table 6), with the HFA group scoring higher than the OO and TD groups. In addition, because the objective of this study was to discover more subtle social difficulties in the OO group, non-conservative, exploratory independent sample *t* tests were conducted on OO versus TD Withdrawn/Depressed total score; they did not differ ($t = 1.51, p = .14, \text{Cohen's } d = 0.44$).

When the individual Withdrawn/Depressed items in the MANOVA were considered separately, 7 items differed by

group (see Table 7). Post-hocs showed that the HFA group scored significantly higher (worse) than the OO group on all items and significantly higher than the TD group on several items (avoids eye contact, reluctant to discuss feelings, and stares blankly); the OO group did not score worse than the TD group on any item but did score significantly better than the TD group on one item (limited fantasy or imagination). Exploratory planned comparison *t* tests for these seven items, which were conducted for the reason described above, showed that the OO group scored higher (worse) than the TD group on one item (stares blankly), with a medium effect size, and lower (better) than the TD group on one other item (limited fantasy or imagination), also with a medium effect size (see Table 8).

Immaturity Scale A MANOVA indicated significant group differences on the 9 combined dependent variables, [$F(18, 128) = 3.19, p < .001, \text{Wilks' } \lambda = 0.47; \text{partial eta squared} = 0.31$]. There was a significant group effect for total score (see Table 6), with the HFA group scoring higher than the OO and TD groups. In addition, an exploratory planned comparison *t* test showed that the OO group's mean score was significantly higher than the TD group, with a medium effect size, $t = 2.12, p = .035, \text{Cohen's } d = 0.62$.

When the results for the individual Immaturity items were considered separately, 4 items differed by group (see Table 7). Post-hocs showed that the HFA group scored significantly higher (worse) than the OO group on only one item (acts too young for age) and significantly higher than the TD group on several items (acts too young for age, lapses

Table 4 Friendship description rating scale mean scores

Mean (SD)	HFA (N = 15)	OO (N = 15)	TD (N = 15)	F	p	Post-hoc	Effect size
Total	33.67 (8.09)	42.52 (4.31)	47.49 (4.08)	21.94	<.001	HFA < OO < TD	OO and HFA = 1.37 OO and TD = 1.18 HFA and TD = 2.16
<i>Overall description</i>							
Description of a friend	2.31 (1.13)	3.25 (0.68)	3.78 (0.51)	12.39	<.001	HFA < OO, TD; Trend for OO < TD	OO and HFA = 1.01 OO and TD = 0.88 HFA and TD = 1.68
Unusual description	3.16 (0.98)	2.27 (0.49)	1.91 (0.46)	13.28	<.001	OO, TD < HFA	OO and HFA = 1.15 OO and TD = 0.75 HFA and TD = 1.63
<i>Companionship</i>							
Time spent	2.31 (0.93)	3.03 (0.53)	3.70 (0.64)	13.96	<.001	HFA < OO < TD	OO and HFA = 0.95 OO and TD = 1.14 HFA and TD = 1.74
Closeness	2.17 (0.78)	3.13 (0.45)	3.66 (0.50)	24.07	<.001	HFA < OO < TD	OO and HFA = 1.51 OO and TD = 1.11 HFA and TD = 2.27
<i>Security-intimacy and trust</i>							
Trust	2.50 (0.88)	3.40 (0.43)	3.81 (0.47)	17.17	<.001	HFA < OO < TD	OO and HFA = 1.30 OO and TD = 0.91 HFA and TD = 1.86
Reliable	2.27 (0.93)	3.20 (0.48)	3.76 (0.49)	19.17	<.001	HFA < OO < TD	OO and HFA = 1.26 OO and TD = 1.15 HFA and TD = 2.00
<i>Closeness</i>							
Bond	2.29 (0.82)	3.22 (0.47)	3.69 (0.52)	19.62	<.001	HFA < OO < TD	OO and HFA = 1.39 OO and TD = 0.95 HFA and TD = 2.04
Affection	2.25 (0.87)	3.16 (0.46)	3.55 (0.51)	16.20	<.001	HFA < OO, TD	OO and HFA = 1.31 OO and TD = 0.80 HFA and TD = 1.82
<i>Help</i>							
Helpfulness	2.40 (0.89)	3.43 (0.49)	3.76 (0.56)	16.99	<.001	HFA < OO, TD	OO and HFA = 1.43 OO and TD = 0.63 HFA and TD = 1.83
Jealousy (reverse coded)	1.64 (0.60)	1.52 (0.45)	1.34 (0.24)	1.64	.207		
<i>Conflict</i>							
Annoyance	1.76 (0.59)	1.65 (0.67)	1.36 (0.33)	2.16	.129		
Conflict	2.27 (0.70)	1.87 (0.63)	1.62 (0.42)	4.61	.015	TD < HFA	OO and HFA = 0.60 OO and TD = 0.47 HFA and TD = 1.13

in attention, and needs repetition of instructions or questions). The OO group scored significantly worse than the TD group on one item (giggles too much). *t* tests for these four items showed that the OO group scored higher (worse) than the TD group on giggles too much (large effect size) and acts too young for age (medium effect size) (see Table 8).

Attention Problems Scale A MANOVA indicated significant group differences on the 10 combined dependent variables, [$F(18, 128) = 2.19, p = .006, \text{Wilks' } \lambda = 0.59; \text{partial eta squared} = 0.24$]. Attention Problems total score also showed a group effect (see Table 6), with the HFA group scoring higher than the OO and TD groups. In

Table 5 Modified Reysen likability scale—all participants

Mean (SD)	HFA (N = 28)	OO (N = 23)	TD (N = 24)	<i>F</i>	<i>p</i>	Post-hoc	Effect size
Total	51.29 (7.19)	64.04 (7.19)	62.10 (7.60)	23.60	<.001	HFA < TD, OO	OO and HFA = 1.77 OO and TD = 0.26 HFA and TD = 1.46
Friendly	4.56 (0.60)	5.22 (0.37)	4.81 (0.60)	9.00	<.001	HFA, TD < OO	OO and HFA = 1.32 OO and TD = 0.82 HFA and TD = 0.46
Likeable	4.20 (0.63)	5.08 (0.55)	4.84 (0.64)	14.51	<.001	HFA < TD, OO	OO and HFA = 1.49 OO and TD = 0.40 HFA and TD = 1.01
Warm	3.87 (0.68)	4.68 (0.40)	4.16 (0.73)	10.57	<.001	HFA < TD < OO	OO and HFA = 1.45 OO and TD = 0.88 HFA and TD = 0.41
Approachable	4.08 (0.69)	5.00 (0.38)	4.57 (0.65)	15.16	<.001	HFA < TD < OO	OO and HFA = 1.65 OO and TD = 0.81 HFA and TD = 0.73
Would ask for advice	3.27 (0.59)	4.44 (0.70)	4.24 (0.67)	24.26	<.001	HFA < TD, OO	OO and HFA = 1.81 OO and TD = 0.29 HFA and TD = 1.54
Would work on a school project	3.27 (0.65)	4.49 (0.75)	4.47 (0.63)	28.35	<.001	HFA < TD, OO	OO and HFA = 1.74 OO and TD = 0.03 HFA and TD = 1.87
Would like as a roommate	2.94 (0.59)	4.02 (0.80)	3.88 (0.65)	19.51	<.001	HFA < TD, OO	OO and HFA = 1.54 OO and TD = 0.19 HFA and TD = 1.51
Would like to be friends	3.73 (0.51)	4.62 (0.57)	4.45 (0.59)	19.17	<.001	HFA < TD, OO	OO and HFA = 1.65 OO and TD = 0.29 HFA and TD = 1.31
Physically attractive	3.44 (0.63)	4.10 (0.87)	4.21 (0.68)	8.66	<.001	HFA < TD, OO	OO and HFA = 0.87 OO and TD = 0.14 HFA and TD = 1.17
Similar to me	3.15 (0.52)	4.10 (0.61)	4.13 (0.57)	25.73	<.001	HFA < TD, OO	OO and HFA = 1.68 OO and TD = 0.05 HFA and TD = 1.80
Knowledgeable	4.03 (0.68)	4.67 (0.61)	4.64 (0.58)	8.65	<.001	HFA < TD, OO	OO and HFA = 0.99 OO and TD = 0.05 HFA and TD = 0.97
Likely has a group of friends	4.02 (0.69)	4.91 (0.58)	4.90 (0.58)	17.92	<.001	HFA < TD, OO	OO and HFA = 1.40 OO and TD = 0.02 HFA and TD = 1.38
Likely has a best friend	3.99 (0.69)	4.79 (0.46)	4.83 (0.62)	16.02	<.001	HFA < TD, OO	OO and HFA = 1.36 OO and TD = 0.07 HFA and TD = 1.28
Likely to be popular	2.73 (0.73)	3.93 (0.80)	3.96 (0.80)	21.71	<.001	HFA < TD, OO	OO and HFA = 1.57 OO and TD = 0.04 HFA and TD = 1.61

Table 6 SCICA observation form

Mean (SD)	HFA (N = 28)	OO (N = 23)	TD (N = 25)	F	p	Post-hoc	Effect size
Withdrawn/depressed	12.14 (7.16)	4.04 (4.59)	6.76 (7.45)	10.09	<.001	HFA > OO, TD	OO and HFA = 1.35 OO and TD = 0.44 HFA and TD = 0.74
Immaturity	4.57 (3.14)	2.50 (2.50)	1.26 (1.32)	12.20	<.001	HFA > OO, TD	OO and HFA = 0.73 OO and TD = 0.62 HFA and TD = 1.37
Attention problems	6.52 (5.01)	3.26 (4.15)	1.40 (1.88)	11.42	<.001	HFA > OO, TD	OO and HFA = 0.71 OO and TD = 0.58 HFA and TD = 1.35
Self-control problems	4.16 (3.51)	3.15 (1.75)	1.62 (2.67)	5.49	.006	HFA > TD	OO and HFA = 0.36 OO and TD = 0.68 HFA and TD = 0.81

Table 7 SCICA items

Mean (SD)	HFA (N = 28)	OO (N = 23)	TD (N = 25)	F	p	Post-hoc
<i>Withdrawn/depressed</i>						
Avoids eye contact	1.66 (1.07)	0.65 (0.78)	0.78 (0.95)	8.82	<.001	HFA > OO, TD
Limited conversation	1.39 (1.14)	0.48 (0.79)	0.92 (1.09)	5.02	.009	HFA > OO
Limited fantasy or imagination	0.91 (1.05)	0.09 (0.42)	0.60 (0.87)	6.08	.004	HFA, TD > OO
Reluctant to discuss feelings or personal issues	1.29 (1.19)	0.28 (0.75)	0.28 (0.58)	11.10	<.001	HFA > OO, TD
Says “don’t know” a lot	1.02 (1.00)	0.33 (0.56)	0.48 (0.78)	5.22	.008	HFA > OO
Slow to respond verbally	0.91 (1.16)	0.13 (0.46)	0.36 (0.67)	5.93	.004	HFA > OO
Stares blankly	1.39 (1.16)	0.67 (0.91)	0.22 (0.41)	11.62	<.001	HFA > OO, TD
<i>Immaturity</i>						
Giggles too much	0.34 (0.69)	0.70 (0.93)	0.14 (0.34)	3.98	.023	OO > TD
Acts too young for age	1.13 (1.12)	0.30 (0.56)	0.02 (0.10)	15.67	<.001	HFA > OO, TD
Lapses in attention	0.61 (0.82)	0.35 (0.78)	0.06 (0.30)	4.30	.017	HFA > TD
Needs repetition of instructions or questions	1.34 (0.82)	0.78 (0.95)	0.38 (0.62)	9.56	<.001	HFA > TD
<i>Attention problems</i>						
Complains of tasks being too hard/upset by tasks	0.39 (0.88)	0.00 (0.00)	0.04 (0.20)	4.16	.019	Trend for HFA > OO
Does not concentrate or pay attention for long	0.66 (0.75)	0.39 (0.84)	0.04 (0.20)	5.91	.004	HFA > TD
Does not sit still, restless, or hyperactive	1.13 (1.18)	0.74 (1.05)	0.34 (0.64)	4.12	.020	HFA > TD
Easily distracted by external stimuli	0.75 (0.98)	0.44 (0.79)	0.06 (0.22)	5.66	.005	HFA > TD
Frequently off-task	0.61 (0.92)	0.30 (0.64)	0.02 (0.10)	5.23	.008	HFA > TD
Out of seat	0.63 (1.02)	0.39 (0.94)	0.06 (0.30)	3.09	.051	N/A
Wants to quit or does quit tasks	0.91 (1.09)	0.00 (0.00)	0.04 (0.20)	15.52	<.001	HFA > OO, TD
<i>Self-control problems</i>						
Defiant, talks back, or sarcastic	0.56 (0.80)	0.04 (0.21)	0.04 (0.20)	8.84	<.001	HFA > OO, TD
Impulsive or acts without thinking	0.45 (0.88)	0.17 (0.49)	0.02 (0.10)	3.46	.037	HFA > TD
Laughs inappropriately	0.50 (0.87)	0.35 (0.65)	0.02 (0.10)	3.82	.027	HFA > TD
Strange behavior	0.29 (0.66)	0.00 (0.00)	0.00 (0.00)	4.50	.014	HFA > OO, TD

addition, a *t* test between the OO and TD group showed that OO group’s total score on the Attention Problems scale was marginally higher than the TD group, with a medium effect size, ($t = 1.97$, $p = .058$, Cohen’s $d = 0.58$).

Seven individual items differed by group (see Table 7). Post-hocs showed that the HFA group scored significantly higher (worse) than the TD group on five items and marginally higher than the OO group on only one item

Table 8 SCICA scales: OO versus TD *t* tests (significant items)

	OO (N = 23)	TD (N = 25)	<i>p</i>	Effect size
Limited fantasy or imagination	0.09 (0.42)	0.60 (0.87)	.012	−0.75
Stares blankly	0.67 (0.91)	0.22 (0.41)	.048	0.64
Giggles too much	0.70 (0.93)	0.14 (0.34)	.011	0.80
Acts too young for age	0.30 (0.56)	0.02 (0.10)	.024	0.70
Easily distracted by external stimuli	0.44 (0.79)	0.06 (0.22)	.037	0.66
Frequently off-task	0.30 (0.64)	0.02 (0.10)	.045	0.61
Laughs inappropriately	0.35 (0.65)	0.02 (0.10)	.025	0.71

Table 9 SCICA clinical T-scores

N (%)	HFA (N = 28)	OO (N = 23)	TD (N = 25)	Chi square	<i>p</i> value	Cramer’s V	Post-hoc
Anxious	3 (11 %)	0 (0 %)	0 (0 %)	5.35	.069	0.265	
Withdrawn/depressed	19 (68 %)	3 (24 %)	7 (28 %)	17.71	<.001	0.483	HFA > OO, TD
Attention problems	12 (43 %)	4 (17 %)	0 (0 %)	14.86	.001	0.442	HFA, OO > TD

(complains of tasks being too hard). The OO group did not differ from the TD group on any item. *t* tests for these items showed that the OO group scored higher (worse) than the TD group on being easily distracted by external stimuli (medium effect size) and frequently off-task (medium effect size) (see Table 8).

Self-Control Problems Scale Finally, a MANOVA indicated significant group differences on the 9 combined dependent variables, [$F(18, 124) = 2.54, p = .001, Wilks' \lambda = 0.53; \text{partial eta squared} = 0.27$] and there was a statistically significant difference between the groups on the Self-Control Problems total score (see Table 6), with the HFA group scoring higher than the TD group. *t* test indicated that the OO group total score was significantly higher than the TD score, with a medium effect size, $t = 2.33, p = .024, \text{Cohen's } d = 0.68$.

When the results for the individual Self-Control Problems dependent variables in the MANOVA were considered separately, 4 items differed by group. Post-hocs showed that the HFA group scored significantly higher (worse) than the TD group on all four items and significantly higher than the OO group on two items (defiant and strange behavior). The OO group did not differ from the TD group on any items. Exploratory *t* tests for these four items showed that the OO group had more inappropriate laughter than the TD group, with a medium effect size (see Table 8).

Potentially Severe Problems Because the normative data for the SCICA is from samples of clinically referred children ages 6–11 and 12–18, clinical T scores > 55 (>69th percentile) are considered to indicate potentially severe problems. Chi square tests were conducted to determine whether the frequency of potentially severe problems differed between groups. This was only conducted for the

Anxious, Withdrawn/Depressed, and Attention Problems scales because not all items were coded in the other domains. The groups were not significantly different on the frequency of potentially severe problems on the Anxious domain, although there was a trend with a small to medium effect size, as 11 % of HFA participants were above the cutoff compared with 0 % in the other two groups (see Table 9). There was a significant difference between the groups on the frequency of participants with potentially severe problems in the Withdrawn/Depressed domain, with a medium to large effect size. Post-hoc tests revealed that significantly more participants in the HFA group (68 %) had problems related to being withdrawn/depressed, compared with 24 % in the OO group and 28 % in the TD group (see Table 9). There was also a significant difference between the groups on the frequency of participants with potentially severe problems in the Attention Problems domain, with a medium to large effect size. Post-hoc tests revealed that significantly more participants in the HFA group (43 %) and OO group (17 %) had attention problems, compared with 0 % in the TD group (see Table 9).

Correlations Between Measures

When examining the OO group, the Friendship Description Rating Scale and Modified Reysen Likability Scale Total score were both negatively correlated with ADOS communication and/or socialization scores, which suggests that, unsurprisingly, youth who exhibited *more* symptoms of autism in the communication and social domains scored more poorly on the other measures (see Table 10). The ADOS was not significantly correlated with any other measure.

When examining the HFA group, the TLC-E Figurative Language subtest and the Friendship Description Rating

Table 10 Correlation matrices between measures—OO group

	ADOS soc.	Vineland comm.	Vineland soc.	TLC-E listening comp.	TLC-E Fig. lang.	CEL4 PP convo	CEL4 PP info	CEL4 PP NV	Friendship desc. rating	Likability total	SCICA anxious	SCICA withdrawn/ depression	SCICA immaturity	SCICA attn problems	SCICA self-control problems
ADOS communication	0.511**	0.002	-0.201	-0.253	-0.301	0.025	-0.126	0.053	-0.557*	-0.548**	-0.111	0.143	-0.069	0.238	-0.196
ADOS socialization		-0.058	-0.236	-0.110	-0.234	-0.398	-0.355	-0.261	-0.409	-0.453*	-0.077	0.292	0.275	0.378	0.130
Vineland communication			0.349*	0.266	0.145	0.001	-0.033	0.180	0.418	0.047	-0.032	-0.385	-0.139	-0.078	0.189
Vineland socialization				0.284	0.399*	-0.159	0.111	-0.100	0.429	0.113	-0.209	-0.398	-0.065	-0.407	0.427*
TLC-E listening comp.					0.623*	-0.048	-0.193	0.085	0.213	-0.143	0.033	-0.193	-0.007	-0.203	0.105
TLC-E figurative lang.						-0.054	0.003	0.006	0.099	0.006	-0.114	-0.319	-0.208	-0.422*	0.118
CEL4 PP conversation							0.680**	0.874**	0.000	0.622**	0.109	-0.292	-0.687**	-0.546*	-0.423
CEL4 PP information								0.617**	0.334	0.311	-0.138	-0.358	-0.367	-0.692**	-0.226
CEL4 PP nonverbal									-0.011	0.479*	0.194	-0.223	-0.771**	-0.721**	-0.444*
Friendship description rating										0.142	-0.131	-0.484	-0.221**	-0.581*	0.619*
Likability total											0.004	-0.344	-0.507*	-0.364	0.062
SCICA anxious												0.403	0.229	0.071	-0.024
SCICA withdrawn/ depression													0.559*	0.420*	-0.280
SCICA immaturity														0.662**	0.228
SCICA attn problems															0.175

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

Scale score were negatively correlated with the ADOS communication score, while the SCICA Immaturity scale was positively correlated with the ADOS communication score. The Likability Scale Total score was negatively correlated and the SCICA Immaturity scale score was positively correlated with the ADOS socialization score. This suggests that, even more so than for the OO group, HFA youth who exhibited *more* symptoms of autism in the communication and/or social domains scored more poorly on other measures (see Table 11).

When examining the TD group, ADOS scores were uncorrelated with any of the other communication and social measures, indicating no clear relationship between autism symptomatology and more subtle social and communicative behaviors. However, it is possible that the limited variability in ADOS scores within the TD group at least partially explains this finding (Table 12).

Despite predictions, Vineland scores were generally not correlated in any meaningful way for any of the measures in the OO, HFA, or TD groups. For the HFA group, all three subscales of the CELF-4 PP were negatively correlated with Vineland communication scores. This would suggest that better adaptive communication functioning was associated with poorer pragmatic language ratings. This finding is certainly unexpected and no clear explanation can be provided.

Discussion

Summary of Results

1. On the ADOS, there were no communication items on which the OO group scored worse than the TD groups. In fact, the individuals in the OO group asked the examiner for information more frequently and showed a trend to offer more information than the TD group. The only significant social item on the ADOS between the OO and TD groups was ability to describe nature of typical social relationships. The HFA group had more evidence of symptoms on all ADOS items, as expected.
2. The OO group was generally rated as having poorer descriptions of friendship than the TD group, in the categories of time spent with friends, closeness, trust, reliability, and bond. The HFA group was rated as having poorer descriptions of friendship than both the OO and TD groups in most categories.
3. Participants in the OO group were judged to be as likable as participants in the TD group. Participants in the HFA group were judged to be less likable than participants in the OO and TD groups.
4. The HFA group had higher (worse) Withdrawn/Depressed, Immaturity, Attention Problems, and Self-Control Problems scores than the TD group, as measured on the SCICA. The OO group did not differ from the TD group for Withdrawn/Depressed scores, but had at least marginally higher scores on the other scales. The OO group scored worse than TD group for several items: giggles too much, lapses in attention, easily distracted by external stimuli, frequently off-task, and laughs inappropriately.
5. On a structured assessment of pragmatic language (TLC-E), the HFA group scored the lowest, and there were no TD-OO differences; all three groups scored in the average range.
6. Spontaneous pragmatic abilities (CELF-4 PP) showed no deficits in the OO group relative to the TD group, but the HFA group was impaired.
7. For OO and HFA but not TD participants, *more* communication and social symptoms of autism were correlated with *lower* performance or ratings on other measures of social and pragmatic functioning.

Strikingly, individuals in the OO group did not exhibit *any* subtle measurable deficits, relative to TD peers, on any ADOS communication item. In fact, the OO participants were significantly more likely than TD participants to spontaneously offer information about thoughts, feelings, or experiences and ask the examiner about his/her thoughts, feelings, or experiences. Extensive experience with therapists in the OO group may have contributed to this, as asking for and offering information may have been targeted behaviors. In addition, their intervention histories may have offered them relatively greater familiarity and comfort with adults. The high rate of asking for and offering information in the OO group, scored as more “normal” than that of the TD group, may have masked an inappropriate quality that was not captured in the coding.

On two measures of pragmatic language (the TLC-E and the CELF-4 PP), the OO group was indistinguishable from the TD group, while the HFA group performed worse than both other groups. Despite scoring in the average range, the HFA group was at the lowest end of average, which is not commensurate with their IQ, suggesting continued difficulty with pragmatic language. Both standardized tasks and behavioral observation thus indicate no pragmatic language deficits in individuals who have achieved OO, unlike the persistent pragmatic language impairments found in HFA.

In the social domain of the ADOS, the OO group had generally fully intact social skills. An exception was their poorer insight into the nature of typical social relationships, with 12 participants showing mild abnormality in this area. This finding was consistent with the results that the OO group’s friendship descriptions were rated as poorer than those of the TD group, suggesting that the OO group either had a poorer understanding of typical social relationships,

Table 11 Correlation matrices between measures—HFA group

	ADOS soc.	Vineland comm.	Vineland soc.	TLC-E listening comp.	TLC-E fig. lang.	CEL-F 4 PP convo	CEL-F 4 PP info	CEL-F 4 PP NV	Friendship desc. rating	Likability total	SCICA anxious	SCICA withdrawn/depression	SCICA immaturity	SCICA attn problems	SCICA self-control problems
ADOS communication	0.459**	-0.119	0.204	-0.259	-0.328*	-0.069	-0.163	-0.242	-0.530*	-0.223	-0.012	0.058	0.425*	0.232	0.255
ADOS socialization	0.144	0.263	0.263	-0.161	-0.273	-0.161	-0.341	-0.215	-0.076	-0.374*	0.089	0.334	0.495**	0.289	0.105
Vineland communication		0.560**	0.560**	0.098	0.253	-0.414*	-0.414*	-0.377*	0.102	-0.268	0.028	-0.030	0.220	0.073	0.228
Vineland socialization				-0.089	-0.202	-0.211	-0.156	-0.087	0.193	-0.079	-0.077	-0.057	0.004	-0.045	0.245
TLC-E listening comp.					0.434**	-0.027	0.073	-0.100	0.426	0.153	0.066	-0.126	-0.342	-0.086	-0.098
TLC-E figurative lang.						-0.316	-0.054	-0.192	-0.040	0.162	-0.032	-0.264	-0.082	-0.210	-0.094
CEL-F 4 PP conversation							0.628**	0.819**	0.365	0.227	-0.042	0.165	-0.323	-0.248	-0.384
CEL-F 4 PP information								0.632**	0.495	0.251	0.198	0.137	-0.436*	-0.454*	-0.474*
CEL-F 4 PP nonverbal									0.484	0.164	0.027	0.133	-0.330	-0.458	-0.309
Friendship description rating										0.407	0.294	-0.239	-0.267	-0.282	-0.106
Likability total											0.055	-0.474*	-0.332	-0.237	-0.079
SCICA anxious												-0.092	0.212	-0.052	-0.183
SCICA withdrawn/depression													0.177	0.024	-0.187
SCICA immaturity														0.632**	0.474*
SCICA attn problems															0.428*

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

Table 12 Correlation matrices between measures—TD group

	ADOS soc.	Vineland comm.	Vineland soc.	TLC-E listening comp.	TLC-E fig. lang.	CEL-4 PP convo	CEL-4 PP info	CEL-4 PP NV	Friendship desc. rating	Likability total	SCICA anxious	SCICA withdrawn/depression	SCICA immaturity	SCICA attn problems	SCICA self-control problems
ADOS communication	0.509**	0.011	-0.142	0.067	-0.106	-0.208	0.195	-0.168	-0.061	-0.335	0.067	0.488	0.275	0.011	0.105
ADOS socialization		-0.033	-0.196	0.153	-0.193	0.238	0.153	0.249	-0.098	-0.396	-0.076	0.371	0.301	-0.003	0.035
Vineland communication			0.398*	0.106	-0.011	-0.157	-0.068	-0.075	-0.264	-0.178	0.400*	0.110	0.221	0.423*	0.087
Vineland socialization				0.211	-0.079	0.192	0.359	0.172	0.132	0.254	0.054	-0.086	-0.086	0.192	-0.039
TLC-E listening comp.					0.291	0.038	-0.197	0.106	-0.005	0.093	0.094	-0.116	0.046	-0.099	0.099
TLC-E figurative lang.						0.097	0.107	0.230	-0.006	0.361	0.208	-0.214	-0.310	-0.193	-0.146
CEL-4 PP conversation							0.693**	0.935**	0.173	0.210	0.208	0.016	-0.174	0.392	-0.652**
CEL-4 PP information								0.610**	-	0.016	0.186	0.121	-0.410	0.199	-0.874*
CEL-4 PP nonverbal									0.411	0.330	0.294	-0.050	-0.088	0.487*	-0.524
Friendship description rating										0.582*	0.019	-0.336	0.028	0.218	0.292
Likability total											0.047	-0.582**	-0.445*	0.052	-0.091
SCICA anxious												-0.250	0.157	0.366	0.012
SCICA withdrawn/depression													0.212	-0.209	-0.339
SCICA immaturity														0.447*	0.549**
SCICA attn problems															0.427*

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

and/or that they had trouble expressing their understanding. The present study could not directly observe whether this lack of understanding of the subtleties of social relationships actually translated into poorer social relationships with peers or adults.

The likability data suggest that the OO participants were at least as likable as the TD participants when judged by a rater based on a brief video clip. It is noteworthy that not only were there no differences in likability between the OO and TD participants, but the scores of the OO participants were quite similar to those of the TD participants. In fact, OO participants were rated as friendlier, warmer, and more approachable than the TD participants, with large effects, suggesting that this difference is meaningful and real. These results are consistent with the findings on the communication domain of the ADOS in that the OO participants were more engaged with the examiner. Ratings of interactions with peers would be needed to see if these likability ratings would generalize. The HFA group, in general, was rated as less likable than the OO and TD participants, suggesting that their ASD symptoms interfere with their perceived social functioning and competence.

Despite previous research showing high rates of anxiety disorders in children and adolescents with ASD (Gjevik et al. 2011; Joshi et al. 2010; Kim et al. 2000; Mattila et al. 2010; Muris et al. 1998; Simonoff et al. 2008), few symptoms of anxiety were found in any of the groups based on behavioral ratings in the current sample. However, given that anxiety disorders are internalizing disorders, it is possible that the ASD groups may have had higher rates of subjective anxiety. Nonetheless, the goal of the current study was to determine if observable symptoms of anxiety impacted the social functioning of the OO participants. Since no differences were found in regards to behaviorally observable symptoms of anxiety (e.g., appearing nervous, preoccupation with certain thoughts, making self-deprecating remarks, easily embarrassed, etc.), it is unlikely that an anxious presentation could be significantly affecting the social interactions for the participants in the study.

Difficulties with depression and withdrawal were very common in the HFA participants, at a level higher than other studies that have reported depression in ASD (Kim et al. 2000). However, this was based on behavioral presentation, rather than reports of internal feelings of depressed mood/sadness, which is vital for a depression diagnosis. In fact, the behavioral symptoms of depression/withdrawal that were more prevalent in the HFA group were generally those that overlap with the symptoms of ASD, such as avoiding eye contact, stares blankly, limited conversation skills, limited fantasy or imagination, and reluctance to discuss feelings. Thus, the SCICA Withdrawn/Depressed scale cannot differentiate ASD from true depression. The OO group did not display an elevation in

these symptoms, regardless of whether they stemmed from ASD itself or depression/withdrawal. As a result, withdrawal does not appear to be negatively impacting the social relatedness of the OO participants.

Ratings on the Immaturity scale of the SCICA suggest that the participants in both the OO and HFA groups, on average, presented as more socially immature than participants in the TD group. Specifically, some OO and HFA participants acted too young for their age and giggled too much. Similarly, on the Self-Control Problems scale, the OO and HFA participants were more likely to laugh inappropriately. On the Attention Problems scale, the OO and HFA participants were more likely to be easily distracted by external stimuli and frequently off-task than the TD participants, indicating a greater difficulty sustaining attention appropriately. In addition, the HFA participants were more likely than the OO participants to complain of tasks being too hard and wanting to give up easily. The lack of these difficulties in the OO group suggests that they may not avoid, dislike, or be reluctant to do tasks that require mental effort. The negative correlations between the Immaturity, Self-Control Problems, and Attention Problems scales and ADOS social scores indicate that more symptoms in these areas are related to poorer social functioning.

The results of the current study, taken together, suggest that the some individuals in the OO group have mild social difficulties that are suggestive of attentional difficulties. Consistent with the current findings in the OO group, previous studies have demonstrated that children and adolescents with ADHD are more frequently off-task, and socially immature or inappropriate (Barkley 2006; Sibley et al. 2010; Wehmeier et al. 2010). These results provide evidence that, at least in some cases, ASD early in life resolves into a constellation of clinically subtle deficits that include attention-based deficits. Attention symptoms are typically part of the ASD behavioral presentation, though not included in the diagnostic criteria, and may potentially be more difficult to extinguish through intervention. Although this is an intriguing observation, caution must be used before assuming the similarities of presentation represent common etiological factors between attention symptoms in children and adolescents with a history of ASD and children and adolescents with a diagnosis of ADHD. It is impossible to conclude from the current results whether the attention symptoms in OO children and adolescents are a lingering feature of ASD or a due to a comorbid disorder, such as ADHD, that persists. Future research should examine the similarities and differences of children and adolescents with OO and those with ADHD. Tracking children with ASD longitudinally to discern how their attention symptoms change over time might shed light on the trajectory of attentional symptoms in children with ASD who achieve OOs.

Limitations and Future Directions

A significant limitation of the study is that initial ASD diagnosis was assessed retrospectively. Care was taken to obtain documentation of ASD symptoms in early diagnostic reports. There was no difference between the two groups in early communication or repetitive behaviors (Fein et al. 2013). While the combination of early diagnostic reports and parental recall enhance confidence in early presentation of ASD, the cross-sectional nature of the study does not address *how* communication and social skills change over time. Because the children were not followed longitudinally, we also cannot report the age at which the OO participants no longer qualified for a diagnosis of ASD. The current study could not fully assess how intervention played a role in improvement. Intervention history was collected and reported in depth in a separate paper (Orinstein et al. 2014). In sum, results suggested that children in the OO group had earlier and more intensive intervention than those in the HFA group. Substantially more children with OO than HFA received applied behavior analysis (ABA) therapy, although for children who received ABA, the intensity did not differ between the groups (Orinstein et al. 2014). A careful specification of the relationship of intervention to outcome will require large scale prospective, longitudinal study.

The participants in the present study were predominantly white, with less than 10 % belonging to other racial or ethnic groups. All three groups were high functioning, with mean nonverbal IQs in the high average range. Thus, these findings may not generalize well to other racial or ethnic groups, or to a broader spectrum of intellectual functioning. Future studies should include a more diverse sample.

A wide age range of participants was included in the current study. This approach was necessary in order to obtain a large enough sample of OO participants. However, the age range of 8–21 years spans developmental levels, with different communication and social demands. Thus, including such a large age range may have prevented finding differences at specific developmental levels. However, analyses showed no differences between younger and older subsamples or correlations with age. In addition, there were different numbers of participants in each group and for each measure, which potentially complicates analysis and interpretation.

There were limits to the measures used in the current study. The ADOS was designed to help clinicians and researchers detect the symptoms and diagnose ASD and was not intended to be utilized as a measure of subtle symptomatology. Furthermore, a concern is that the inclusion criteria limited the possibility to detect differences on the ADOS, as there were inclusion cutoffs on the ADOS in

order to clearly define the groups. However, the OO and TD groups were not deliberately matched on these criteria, which still allowed room for differences between the groups to appear. The ADOS was conducted in a one-on-one setting with minimal distractions by adult examiners experienced in working with children and adolescents with autism. The participants may have performed differently in a more natural environment or with naive adults or peers. Furthermore, the examiner was not naive to group membership when administering the ADOS. However, a coder blind to group status watched and scored five ADOS videos per group, with high inter-rater reliability. Nonetheless, the ADOS administration was the basis for much of the results, which is problematic because the ADOS limitations impact multiple measures and shared method variance may be a concern. The concern may be particularly problematic if the OO individuals have had more experience with the ADOS than the TD individuals. Future research should use additional, broader samples of behavior in order to strengthen confidence in the findings of the current study. An additional limitation related to the measures is that several tools were designed for the current study or used in an experimental manner. The Friendship Description Rating Scale was created and the Reysen Likability Scale was modified for the current study so there is no research reliability on either measure. The SCICA was also used in a non-traditional manner, as the ratings were based on a video-taped measure (the ADOS), rather than the SCICA interview that is generally used, so the results may need to be interpreted with caution.

An additional limitation was lack of direct measures of peer interaction. Given the considerable difficulty individuals with HFA have with friendship (Howlin et al. 2000; Orsmond et al. 2004), close examination of friendship quantity and quality in the OO group is warranted. Based on the results of the current study, it is not possible to conclude whether the quantity and quality of friendships in the OO group actually different from TD peers in every day life. Future research should examine this more closely as it is important to know whether an early diagnosis of ASD negatively affects later social behavior or just the verbal report of issues related to socialization and friendships.

Conclusions

The OO participants in this study clearly lost their ASD diagnosis and are functioning well in the communication and social domains. They have very few symptoms of ASD, show no deficits in pragmatic language, and are judged to be as likable as their typically developing peers. The OO participants in this study had some symptoms related to attention, self-control, and immaturity that may

have impacted their social abilities. Although OO participants exhibited somewhat poorer ability to express insight into the nature of typical social relationships and to describe friendships compared to TD peers, they displayed no other measurable communicative and social functioning difficulties. In fact, OO participants were the most engaged in the interactions with the examiner, and were therefore rated as friendlier, warmer, and more approachable and were more likely to ask and offer information to the examiner. In sum, the OO group is doing quite well in the social and communication domains, although some OO youth exhibit mild social difficulties that seem to be the result of attentional difficulties.

Acknowledgments This research was supported by the National Institutes of Mental Health (R01MH076189 to Deborah Fein). We are deeply grateful to participants and families and our tireless undergraduate research assistants.

Conflict of interest There are no conflicts of interest to declare.

Ethical statement All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

- Anderson, D. K., Liang, J. W., & Lord, C. (2014). Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *55*(5), 485–494.
- Ballaban-Gil, K., Rapin, I., Tuchman, R., & Shinnar, S. (1996). Longitudinal examination of the behavioral, language, and social changes in a population of adolescents and young adults with autistic disorder. *Pediatric Neurology*, *15*(3), 217–223.
- Barkley, R. A. (2006). *Attention-deficit/hyperactivity disorder: A handbook for diagnosis and treatment* (3rd ed.). New York, NY: Guilford Press.
- Bukowski, W. M., Hoza, B., & Boivin, M. (1994). Measuring friendship quality during pre- and early adolescence: The development and psychometric properties of the Friendship Qualities Scale. *Journal of Social and Personal Relationships*, *11*(3), 471–484.
- Doi, S. A. R., & Williams, G. M. (2013). *Methods of clinical epidemiology*. Berlin: Springer.
- Eales, M. J. (1993). Pragmatic impairments in adults with childhood diagnoses of autism or developmental receptive language disorder. *Journal of Autism and Developmental Disorders*, *23*(4), 593–617.
- Eaves, L. C., & Ho, H. H. (2008). Young adult outcome of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(4), 739–747.
- Fein, D., Barton, M., Eigsti, I. M., Kelley, E., Naigles, L., Schultz, R. T., et al. (2013). Optimal outcome in individuals with a history of autism. *Journal of Child Psychology and Psychiatry*, *54*(2), 195–205.
- 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*(4), 525–534.
- Freitag, C. M., Kleser, C., & von Gontard, A. (2006). Imitation and language abilities in adolescents with Autism Spectrum Disorder without language delay. *European Child and Adolescent Psychiatry*, *15*(5), 282–291.
- Gilchrist, A., Green, J., Cox, A., Burton, D., Rutter, M., & Le Couteur, A. (2001). Development and current functioning in adolescents with Asperger syndrome: A comparative study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 227–240.
- Gjevik, E., Eldevik, S., Fjæran-Granum, T., & Sponheim, E. (2011). Kiddie-SADS reveals high rates of DSM-IV disorders in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*(6), 761–769. doi:10.1007/s10803-010-1095-7.
- 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.
- Howlin, P. (2003). Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *33*(1), 3–13.
- Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and developmental receptive language disorder—A follow-up comparison in early adult life. II: Social, behavioural, and psychiatric outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(5), 561–578.
- Joshi, G., Petty, C., Wozniak, J., Henin, A., Fried, R., Galdo, M., et al. (2010). The heavy burden of psychiatric comorbidity in youth with autism spectrum disorders: A large comparative study of a psychiatrically referred population. *Journal of Autism and Developmental Disorders*, *40*(11), 1361–1370. doi:10.1007/s10803-010-0996-9.
- 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.
- Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders*, *36*(6), 807–828.
- Kim, J. A., Szatmari, P., Bryson, S. E., Streiner, D. L., & Wilson, F. J. (2000). The prevalence of anxiety and mood problems among children with autism and Asperger syndrome. *Autism: The International Journal of Research and Practice*, *4*(2), 117–132.
- Kobayashi, R., Murata, T., & Yoshinaga, K. (1992). A follow-up study of 201 children with autism in Kyushu and Yamaguchi areas, Japan. *Journal of Autism and Developmental Disorders*, *22*(3), 395–411.
- Lam, Y. G., & Young, S. S. S. (2012). Towards a convergent account of pragmatic language deficits in children with high-functioning autism: Depicting the phenotype using the Pragmatic Rating Scale. *Research in Autism Spectrum Disorders*, *6*, 792–797.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr, Leventhal, B. L., DiLavore, P. C., et al. (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*(3), 205–223.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, *55*(1), 3–9.
- Luyster, R., Richler, J., Risi, S., Hsu, W. L., Dawson, G., Bernier, R., et al. (2005). Early regression in social communication in autism spectrum disorders: A CPEA Study. *Developmental neuropsychology*, *27*(3), 311–336. doi:10.1207/s15326942dn2703_2.

- Mattila, M. L., Hurtig, T., Haapsamo, H., Jussila, K., Kuusikko-Gauffin, S., Kielinen, M., et al. (2010). Comorbid psychiatric disorders associated with Asperger syndrome/high-functioning autism: A community- and clinic-based study. *Journal of Autism and Developmental Disorders*, *40*(9), 1080–1093. doi:10.1007/s10803-010-0958-2.
- McConaughy, S. H., & Achenbach, T. M. (2001). *Manual for the semistructured clinical interview for children and adolescents* (2nd ed.). Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- McGovern, C. W., & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*(4), 401–408.
- Muris, P., Steerneman, P., Merckelbach, H., Holdrinet, I., & Meesters, C. (1998). Comorbid anxiety symptoms in children with pervasive developmental disorders. *Journal of Anxiety Disorders*, *12*(4), 387–393.
- Orinstein, A. J., Helt, M., Troyb, E., Tyson, K. E., Barton, M. L., Eigsti, I. M., et al. (2014). Intervention for optimal outcome in children and adolescents with a history of autism. *Journal of Developmental and Behavioral Pediatrics (JDBP)*, *35*(4), 247–256. doi:10.1097/DBP.0000000000000037.
- Orsmond, G. I., Krauss, M. W., & Seltzer, M. M. (2004). Peer relationships and social and recreational activities among adolescents and adults with autism. *Journal of Autism and Developmental Disorders*, *34*(3), 245–256.
- Paul, R., Orlovski, S. M., Marcinko, H. C., & Volkmar, F. (2009). Conversational behaviors in youth with high-functioning ASD and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *39*(1), 115–125.
- 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 and Adolescent Psychiatry*, *35*(4), 523–529.
- Reysen, S. (2005). Construction of a new scale: The Reysen likability scale. *Social behavior and personality*, *33*(2), 201–208.
- Seltzer, M. M., Krauss, M. W., Shattuck, P. T., Orsmond, G., Swe, A., & Lord, C. (2003). The symptoms of autism spectrum disorders in adolescence and adulthood. *Journal of Autism and Developmental Disorders*, *33*(6), 565–581.
- Semel, E., Wiig, E. H., & Secord, W. (2003). *CELF-4: Clinical evaluation of language fundamentals examiner's manual* (4th ed.). Toronto: PsychCorp.
- Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., et al. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(9), 1735–1747.
- Sibley, M. H., Evans, S. W., & Serpell, Z. N. (2010). Social cognition and interpersonal impairment in young adolescents with ADHD. *Journal of Psychopathology and Behavioral Assessment*, *32*, 193–202.
- Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: Prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*(8), 921–929. doi:10.1097/CHI.0b013e318179964f.
- Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1985). *Vineland Adaptive Behavior Scales* (Interview ed.). Circle Pines, MN: American Guidance Service.
- Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence (WASI)*. New York, NY: The Psychological Corporation.
- Wehmeier, P. M., Schacht, A., & Barkley, R. A. (2010). Social and emotional impairment in children and adolescents with ADHD and the impact on quality of life. *The Journal of Adolescent Health*, *46*(3), 209–217. doi:10.1016/j.jadohealth.2009.09.009.
- Whitehouse, A. J., Watt, H. J., Line, E. A., & Bishop, D. V. (2009). Adult psychosocial outcomes of children with specific language impairment, pragmatic language impairment and autism. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists*, *44*(4), 511–528.
- Wiig, E. H., & Secord, W. A. (1989). *Test of language competence-expanded edition*. San Antonio, TX: The Psychological Corporation.