1   Investigating Frank Autism: Clinician Initial Impressions and Autism Characteristics
2   Rebecca R. Canale[1], M.S.; Caroline Larson, Ph.D.[2,3]; Rebecca P. Thomas, Ph.D.[4]; Marianne
3   Barton, Ph.D.[1]; Deborah Fein, Ph.D.[1]; Inge-Marie Eigsti, Ph.D. [1]
4
5   [1]Department of Psychological Sciences, University of Connecticut, 406 Babbidge Road, Storrs,
6   CT 06269, USA
7   [2] Department of Speech, Language, and Hearing Sciences, University of Missouri, Columbia,
8   MO, USA
9   [3] Interdisciplinary Neuroscience Program, University of Missouri, Columbia, MO, USA
10   [4]Developmental Medicine Center, Boston Children's Hospital, 300 Longwood Avenue, Boston,
11   MA 02115, USA
12
13   **Corresponding authors**: Rebecca Canale, M.S., University of Connecticut; Department of
14   Psychological Sciences; 406 Babbidge Road, Storrs, CT 06269, USA. Email:
15   rebecca.canale@uconn.edu; Telephone: 860-486-6021
16   Inge-Marie Eigsti, Ph.D., University of Connecticut; Department of Psychological Sciences; 406
17   Babbidge Road, Storrs, CT 06269; USA. Email: inge-marie.eigsti@uconn.edu; Telephone: 860-
18   486-6021
19   **Running head**: Clinician initial impressions and autism characteristics

Abstract

Background: "Frank autism," recognizable through the first minutes of an interaction, describes a behavioral presentation of a subset of autistic individuals that is closely tied to social communication challenges, and may be linked to so-called "prototypical autism." To date, there is no research on frank autism presentations of autistic adolescents and young adults, nor individuals diagnosed with autism spectrum disorder (ASD) in childhood who do not meet diagnostic criteria during or after adolescence (loss of autism diagnosis, LAD). In addition, there are currently no data on the factors that drive frank autism impressions in these adolescent groups.

Methods: This study quantifies initial impressions of autistic characteristics in 24 autistic, 24 LAD and 26 neurotypical (NT) individuals ages 12 to 39 years. Graduate student and expert clinicians completed five-minute impressions, rated confidence in their own impressions, and scored the atypicality of behaviors associated with impressions; impressions were compared with current gold-standard diagnostic outcomes.

Results: Overall, clinicians' impressions within the first five minutes generally matched current gold-standard diagnostic status (clinical best estimate), were highly correlated with ADOS-2 CSS, and were driven primarily by prosodic and facial cues. However, this brief observation did not detect autism in all cases. While clinicians noted some subclinical atypicalities in the LAD group, impressions of the LAD and NT groups were similar.

Limitations: The brief observations in this study were conducted during clinical research, including some semi-structured assessments. While results suggest overall concordance between initial impressions and diagnoses following more thorough evaluation, findings may not generalize to less structured, informal contexts. In addition, our sample was demographically homogeneous and comprised only speaking autistic participants. They were also unmatched for sex, with more females in the non-autistic group. Future studies should recruit samples that are diverse in demographic variables and ability level to replicate these findings and explore their implications.

Conclusions: Results provide insights into the behavioral characteristics that contribute to the diagnosis of adolescents and young adults and may help inform diagnostic decision making in the wake of an increase in the demand for autism evaluations later than childhood. They also substantiate claims of an absence of apparent autistic characteristics in individuals who have lost the diagnosis.

Keywords: five-minute impressions, autism diagnosis, autism in adulthood, loss of autism diagnosis, optimal outcomes, prototypical autism

Background

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by persistent deficits in social-emotional reciprocity as well as the presence of restricted and repetitive behaviors and interests (RRBI) (1). In the *Diagnostic and Statistical Manual* (*DSM-V*), ASD is defined as a spectrum of behaviors marked by heterogenous core and co-occurring features (1), diagnosable according to behaviorally defined criteria (2) by trained clinicians. A full diagnostic assessment involves several hours of expert clinician time and standardized assessments; however, studies indicate that nearly all clinicians (97%) report forming an immediate, strong impression of diagnostic status in many cases (3). Such impressions are generally consistent with gold-standard diagnoses(4). This impression represents the distinct behavioral presentation of a subset of autistic individuals, dubbed "frank autism," purportedly recognizable within minutes. The current study evaluates the consistency of frank autism impressions in adolescents and young adults, including impressions of a group of individuals diagnosed early in life who no longer display autism symptoms, and compares them to gold-standard diagnostic classifications. We also characterize the behavioral factors that contribute to clinician impressions of autism, to better understand how frank autism impressions relate to enduring, core symptoms, and discuss how the presence of "frank autism" relates to the construct of "prototypical autism" (5). Note that person-first and identity-first language will both be used in this manuscript to acknowledge the diverse preferences within the autism community (6,7); in addition, we limit the use of language consistent with the medical model of autism (e.g., "symptoms" and "deficits") in our discussion of diagnostic criteria.

Autism diagnoses and the role of frank autism impressions

DSM-V ASD diagnostic criteria can differentiate autistic individuals from non-autistic individuals and capture the wide variability within the autism spectrum (8). Indeed, autism is one of the most reliable diagnoses in the DSM-V (9) with Kappa values from 0.60 to $\geq$ 0.80 (considered moderate to substantial agreement). A study assessing expert clinician team-based diagnoses reported that some clinicians relied heavily on "feeling autism in the encounter," along with quality of collateral report from parents, to inform diagnostic decisions (10). However, a study utilizing expert clinician consensus to evaluate the reliability of evaluations performed by community clinicians without ASD-specific expertise found suboptimal agreement on diagnostic status (11). Of the 87 children and young adults, ages 2-25 years, with community ASD diagnoses, 23% were classified by expert clinician consensus as not autistic, illustrating a discrepancy in diagnostic judgments based on available resources (e.g., time, access to diagnostic tools, expert consultation, etc.) and clinician training.

In one study of frank autism, licensed psychologists watched 10-minute video clips of another clinician administering the Autism Diagnostic Observation Schedule (12) to 42 toddlers who had been flagged with possible ASD during screening (13). After watching a 10-minute clip, clinicians indicated whether they would refer the child for a comprehensive ASD-specific assessment; two referral decisions were made per child by different clinicians (84 total videos). In this sample, 17 (61%) were referred by one or both clinicians for further ASD assessment and ultimately diagnosed with ASD; seven (25%) were referred by one or both clinicians, but ultimately diagnosed with language delays but not autism; and three (11%) were referred by one or both clinicians, but ultimately found to be typically developing, indicating sensitivity of 0.61 and specificity of 0.82. Of the 57 videos for which neither clinician recommended an ASD-specific evaluation, 11 (39%) were ultimately diagnosed with ASD (i.e., these cases were missed by the observing clinicians). These findings suggest that trained clinicians can identify and

102   distinguish autistic symptoms from characteristics of other developmental delays in toddlers with
103   *some* accuracy, based on 10 minutes of behavioral observation. However, the high number of
104   false negatives suggests that this information alone is insufficient, at least in the case of toddlers.
105   Furthermore, the determination needed in a clinical evaluation requires not just ruling autism in
106   or out, but also differentiating between autism and other conditions – a significantly more
107   challenging endeavor.
108   A related study explored the initial impressions of trained clinicians for a sample of 294
109   children ages 1-4 years who were referred for a diagnostic evaluation after being flagged as at-
110   risk for autism on a brief parent-report screener (4). After five minutes of interaction during the
111   diagnostic evaluation, clinicians paused and indicated their initial diagnostic impression (ASD or
112   non-ASD) and rated their confidence in this initial impression. Results showed that 238 (81%)
113   initial clinical impressions were concordant with the final diagnosis; the autism cases were
114   judged more accurately than the non-ASD cases, with 86 (92%) of the ASD impressions
115   ultimately receiving an autism diagnosis, consistent with a frank autism phenotype. There was a
116   high false negative or "missed cases" rate: 49 (24%) cases initially viewed as *not* autism
117   ultimately received an ASD diagnosis; false positive rates were far lower (7%). Clinicians were
118   confident in their initial impressions, particularly for non-autistic cases, with an average
119   confidence rating of 3.74 out of 5. These results highlight the ability of trained clinicians to
120   detect ASD from brief behavioral observation, but underscore that some young autistic children
121   (e.g., 18% in this sample) would be missed by an initial diagnostic impression
122   A recent study by the same group (14) further explored what behavioral characteristics
123   informed diagnostic impressions within the first five minutes of interaction with 55 toddlers
124   (mean age = 22.9 months) referred for a developmental evaluation due to parent or pediatrician
125   concerns for autism-related behaviors. Junior (e.g., graduate student) and senior (e.g., PhD level)
126   clinicians were asked to rate their diagnostic impression (autistic or non-autistic), their
127   confidence in this impression, and what behaviors contributed to their impression. Consistent
128   with prior findings, clinicians rated 63% of cases that ultimately received an autism diagnosis as
129   autistic and 100% of cases that did *not* receive an autism diagnosis as non-autistic. Both junior
130   and senior clinicians relied on social reciprocity, nonverbal communication, and eye contact to
131   form accurate initial impressions. Additionally, senior clinicians relied on the child's focus of
132   attention in forming accurate impressions of both autistic and non-autistic children, whereas
133   junior clinicians only relied on this behavior in forming accurate impressions of non-autistic
134   children. These results are the first to explore the behaviors that contribute to diagnostic
135   impressions during brief clinical interactions with young children.
136   ***Autism evaluations in adulthood***
137   The prevalence of first-time diagnostic evaluations of adolescents and adults has
138   significantly increased in the past decade, in part because of changes in awareness, diagnostic
139   criteria, and professional practice (15,16). The assessment of older individuals provides a unique
140   set of challenges that are not present when assessing young children. Typical diagnostic practice
141   relies heavily on parent or caregiver report of the early developmental history of the individual,
142   which may be difficult to obtain or inaccurately recalled years later (15,17,18); it can also
143   display "telescoping" effects, such that caregivers of individuals who currently display stronger
144   adaptive skills are more likely to recall more strengths and fewer delays in early development
145   (19). This lack of clear developmental history may force clinicians to rely more heavily on
146   current behavioral observation alone. This, in combination with the evidence that some clinicians
147   rely on less operationalized behavioral observations, by "feeling autism in the encounter" (10),

148 may lead some autistic adults to receive an official diagnosis of autism more readily than others.
149 To date, no studies have explored the behavioral factors that impact clinician impressions of
150 autistic adults.

151 *Autistic characteristics and their impact on impressions*
152   Despite clinical and empirical evidence regarding the good reliability of brief initial
153 impression, the specific factors that contribute to this impression in adolescents and adults are
154 unknown. The initial study proposing frank autism (3) surveyed 151 clinicians with autism-
155 specific expertise about their representation and usage of this construct. Results showed that
156 nearly all (97%) believed that something like frank autism exists, and that they could determine
157 whether an individual fits the phenotype of frank autism in roughly the first ten minutes of
158 interaction or observation. The clinicians who were familiar with the construct estimated that
159 roughly 40% of the ASD population exhibits the frank autism phenotype. Clinicians also
160 reported that the most common specific behaviors associated with this phenotype included
161 impairments or atypicalities in reciprocity, vocal prosody, eye contact, motor mannerisms (such
162 as stereotypies), and gait or posture. These findings highlight factors that may impact initial
163 impressions during ASD diagnostic decision making. To date, no studies have empirically tested
164 the endorsements of these behaviors associated with correct or incorrect frank autism
165 impressions in adults.
166   Gestures, facial expressions, eye contact, vocal prosody, and social reciprocity have each
167 been implicated as atypical in autistic individuals, and relevant for difficulties with social
168 functioning. Compared to neurotypical peers, autistic individuals produce semantically,
169 pragmatically, and motorically atypical gestures (20–25), as well as atypical facial expressions
170 (26,27), eye contact (28), and vocal prosody (29,30). Together, these characteristics may
171 negatively impact autistic individuals' social interactions and elicit impressions of social
172 awkwardness from naïve observers (26,27,29,31). Initial impressions for expert clinicians and
173 naïve laypeople may reflect similar processes, despite differences in rater goals (e.g., motivation
174 to engage in future social interaction, versus clinical motivation to arrive at an accurate
175 diagnosis) and the nature of ratings (Likert scales measuring the likeliness that an individual has
176 friends versus binary diagnostic ratings).
177   In summary, the construct of frank autism is widely assumed in clinical practice and is
178 relevant for initial impressions of behavioral atypicalities in non-clinical settings. As such, it is
179 important to establish which behavioral factors contribute to this impression, as they likely have
180 implications for diagnostic decision making (e.g., who is ultimately diagnosed with autism), as
181 well as clinical management.

182 *Loss of autism diagnosis*
183   Although developmental disorders are typically seen as life-long conditions, a series of
184 studies has identified and characterized a group of individuals who were diagnosed with autism
185 in childhood but who no longer meet DSM-V criteria in adolescence, based on ADOS-2
186 observations, parent and child symptom report, and clinical best estimate. Estimates suggest that
187 3-25% of children diagnosed with ASD in early childhood fall into this category (32) by
188 adolescence, although a recent study reported that 37% of toddlers diagnosed with ASD lost the
189 diagnosis by early school age (33). Our research team has extensively studied these types of
190 individuals (34). Findings indicate that in early development, the "loss of autism diagnosis"
191 (LAD) group had milder symptoms in the social domain, compared to an age-matched currently
192 autistic group, but equally significant difficulties with communication and repetitive behaviors,
193 including the presence of early language delays. Tests of current functioning indicated that,

194 compared to age- and IQ-matched children with a current autism diagnosis and with neurotypical
195 (NT) children with no history of autism, the LAD group had typical or above-average scores on
196 standardized and experimental assessments of language (35–39), social skills (40,41), and
197 restricted and repetitive behaviors (42). To date, no studies have explored frank autism in LAD,
198 and whether these individuals present with subtle or overt frank autism behaviors during initial
199 interactions; findings would help to establish the degree to which these individuals continue to
200 display subtle behavioral characteristics of autism. More broadly, understanding frank autism in
201 LAD may be useful in addressing controversies about the nature of the autism diagnosis (43,44).
202 For example, Mottron and colleagues have suggested that developing more constrained
203 diagnostic criteria for autism, informed by strong developmental history data, would facilitate
204 clinical ascertainment and homogeneity of research samples (5,45).

205 ***The current study***
206   The current study had three pre-registered aims (see
207 https://osf.io/5tkrn/?view_only=1f0b6bf70d7d4bab9ebf22da7603e647). Our first aim was to
208 evaluate group (autism, LAD, NT) differences in frank autism impressions made by seven
209 graduate-level (clinical psychology PhD student) and two expert PhD-level clinicians as a
210 predictor of current gold-standard diagnosis in an adolescent and young adult sample. Based on
211 prior studies of LAD and autism, we predicted significantly reduced ASD-like impressions in the
212 LAD and NT groups relative to the autism group, and significant positive correlations between
213 initial impressions of frank autism and ADOS-2 Calibrated Severity Scores (CSS).
214   Second, drawing on the prior frank autism studies of young children, we asked which
215 behaviors were the most salient contributors to frank autism impressions, by assessing rates of
216 atypicality in gesture, eye contact, motor mannerisms, prosody, facial expressions, attentional
217 focus, and shifting attention (including perseverative thinking and distractibility), social
218 reciprocity, and social initiations. We predicted significantly higher (more atypical) ratings for
219 gesture, eye contact, motor mannerisms, prosody, facial expressions, and social reciprocity. We
220 also predicted that attentional focus and social initiations would be similar across groups, as prior
221 literature typically implicates these more infrequent behaviors (that may be difficult to perceive
222 during a brief encounter) as less consistently associated with a frank autistic presentation.
223   Third, we hypothesized high overall confidence (e.g., 3 or above on a scale of 1-5) in
224 initial impressions, with higher ratings for NT individuals that had never received an autism
225 diagnosis (based on Wieckowski et al., 2021). We also predicted that the confidence ratings for
226 the LAD group would be significantly lower than both the ASD and NT groups due to possible
227 subclinical social impairments. We hypothesized that higher confidence would be significantly
228 associated with eye contact, motor mannerisms, prosody, and social reciprocity, but not gesture,
229 facial expressions, focus/shifting of attention, or social interactions.
230   **Methods**
231 ***Participants***
232   This study included participants from a larger study of long-term outcomes in autism.
233 Participants who had completed the Autism Diagnostic Observation Schedule-2 (46) were
234 included in the present study. The sample included currently autistic participants ($n = 24$; 7
235 females), participants with a history of ASD who no longer met diagnostic criteria (LAD; $n = 24$;
236 5 females), and participants with a neurotypical developmental history (NT; $n = 26$; 15 females).
237 Participant details are summarized in Table 1. The groups did not differ on age, race/ethnicity,
238 mean household income, verbal skills as measured by Penn Verbal Analogies, or nonverbal skills
239 as measured by Penn Matrix Reasoning.

240   As expected, the ASD group had higher ADOS-2 scores. The LAD group had marginally
241 higher ADOS-2 CSS scores than the NT group, though the means of both groups fell well below
242 the autism threshold. The ASD and LAD groups had more males than the NT group; no
243 participants identified with a gender other than male or female.
244 **Table l.**
245 *Participant characteristics*

	ASD (*n* = 24)	**LAD** (*n* =24)	**NT** (*n* =26)	**F/χ^2**	**Post-hoc comparison**
Age (yrs)	21.22(4.50)	22.72(3.71)	22.74(6.41)	1.02	
M:F *	17:7	19:5	11:15	8.14	ASD=LAD >NT
Race	Native Amer = 0 Asian/Pacific Islander = 1 African Amer = 0 White = 22 Multiracial = 1 Not reported = 1	Native Amer = 0 Asian/Pacific Islander = 0 African Amer = 0 White = 21 Multiracial = 1 Not reported = 1	Native Amer = 0 Asian/Pacific Islander = 1 African Amer = 0 White = 21 Multiracial = 0 Not reported = 4	2.01	
Ethnicity	Latinx = 1 Not Latinx = 18 Not reported = 6	Latinx = 0 Not Latinx = 19 Not reported = 4	Latinx = 2 Not Latinx = 18 Not reported = 6	2.05	
Household income ($)	90,833(19,497)	98,611(5,892)	87,333(26,795)	10.65	
Penn Matrix Reasoning	18.85(4.34)	20.03(4.14)	20.32(2.64)	1.04	
Penn Verbal Analogies	7.07(1.94)	7.71(1.44)	7.59(1.84)	1.01	
ADOS-2 CSS***	7.54(1.67) 6-10	1.75(0.85) 1-3	1.27(0.60) 1-3	235.70	ASD> LAD>NT

246 *Note*. Data are presented as M(SD), range, or as count variables. Amer = American. Penn Matrix
247 Reasoning and Penn Verbal Analogy scores represent "efficiency," a composite of accuracy and
248 RT. ADOS-2 CSS= Autism Diagnostic Observation Schedule-2 Calibrated Severity Score. ‡ *p <*
249 .10, **p* < .05, ***p* < .01, ****p* < .001.
250
251   Inclusion criteria were based on the aims of a larger study and thus reflect goals of that
252 project (not discussed here). Criteria thus required: no history of intellectual disability, per parent
253 report; current cognitive abilities in the normal range, and scores > 77 on the Vineland
254 Adaptative Behavior Scales-3 (47), a parent-report measure of adaptive functioning; no
255 uncorrected visual or hearing impairments; and no severe psychiatric disorders (e.g., bipolar
256 disorder or schizophrenia). Participants with other, less severe, co-morbid psychiatric disorders
257 such as anxiety, depression, and/or attention deficit/hyperactivity disorder were included in the
258 study. Diagnostic evaluation of such conditions was completed via structured clinical interview

259   and self-report as part of the larger diagnostic battery; further discussion of these data is outside
260   the scope of this paper. Inclusion in the autism group required an ASD diagnosis prior to age five
261   years, documented in a written report by a clinician specializing in autism, as well as the
262   presence of early language delay (first words after age 18 months or first phrases after 24
263   months). In addition, participants in the autism group had to meet criteria for current ASD based
264   on ADOS-2 scores and best estimate clinical judgement. Inclusion in the LAD group required
265   similar early diagnostic criteria as for the autism group; in addition, participants could exhibit no
266   or minimal current symptoms of ASD, as measured by the ADOS-2 and expert clinical
267   judgement, and had to participate in mainstream educational or occupational environment with
268   no ASD-related accommodations. Inclusion in the NT group required no history of
269   developmental disorder per parent report, no first-degree relatives with an ASD diagnosis, and
270   no or minimal current symptoms of ASD based on the ADOS-2 and expert clinical judgement.
271   Participants were recruited via their participation in prior studies of ASD, through
272   clinician referrals, posts on social media, flyers distributed at schools and organizations that offer
273   services for autistic individuals and their families, at local schools, libraries, and community
274   centers, a university registry of diverse community members interested in research participation,
275   and by snowball recruitment (e.g., asking participants to nominate other potential participants).
276   ***Procedures***
277   Participants completed a comprehensive testing battery to confirm diagnostic status,
278   including the ADOS-2 (46) and a parent interview. ADOS-2 administrations were conducted in
279   person or via a validated online protocol (48), and were recorded for later review. In-person
280   participants completed the standard ADOS-2 Module 4 administration, while online participants
281   completed a modified version that excluded the puzzle task and the break. Autism diagnosis
282   required an ADOS-2 raw score of 8 or greater and expert clinical judgement of autism based on
283   behavioral observation. All ADOS-2 recordings were reviewed by a licensed clinical
284   psychologist with autism expertise to confirm diagnostic status. Participants completed
285   additional measures (including a detailed psychiatric interview) not relevant to the current study.
286   To measure frank autism impressions, seven graduate students (Clinical Psychology
287   Ph.D. students) and two expert Ph.D.-level clinicians reviewed the recording of the first five
288   minutes of the diagnostic session, comprising discussion of the visit agenda, set up, small talk,
289   and, in some cases, a minute of the first structured ADOS-2 activity (the Tuesday story). The
290   graduate clinicians all established ADOS-2 reliability with a research-reliable licensed
291   psychologist; this group also included a post-doctoral speech-language pathologist fellow with
292   autism experience. The graduate clinicians also conducted the ADOS-2 assessments, though it is
293   important to note that they did not complete frank autism impression for any participant for
294   which they conducted the diagnostic study visit. Expert Ph.D.-level clinicians were faculty
295   members with decades of experience in autism assessment and diagnosis. Each ADOS-2
296   recording was reviewed by two graduate clinicians and one expert clinician, for a total of three
297   raters per recording; clinicians did not watch their own administrations. Graduate clinicians each
298   reviewed 21-23 recordings, and expert clinicians each viewed 37-38 recordings. All clinicians
299   were blind to group status prior to viewing the recordings. After reviewing the five-minute video,
300   clinicians completed a Five-Minute Impressions Form (49); see Appendix A. The form captured
301   participant details (e.g., date of evaluation, date of review, identity of examining clinician and
302   rater), as well as eight behaviors: gesture, eye contact, motor mannerisms, prosody and
303   vocalizations, facial expressions, attention focus and shifting (including perseverative thinking
304   and distractibility), social reciprocity, and social initiations. Each item was assigned an item-

305   level Score on a 0 - 2 Likert scale, with "0" representing typical or expected behavior in the
306   category, "1" representing mildly atypical behavior, and "2" representing definitely atypical
307   behavior; this scoring structure is analogous to that used in the ADOS-2. Raters were instructed
308   to respond to all items. If an item did not inform their impression, they were instructed to score
309   that item as 0 and make a note; this score was subsequently converted to a 9, indicating that the
310   item did not contribute to the overall impression, and was not included in the total score
311   calculation. Items were scored based on the rater's observations and their clinical knowledge of
312   typical age-appropriate behavior in that context. Items were summed to form an initial
313   impression total score ranging from 0 (no atypical behaviors detected) to 16 (definitely atypical
314   behavior in eight items). Raters provided initial impressions (autistic or non-autistic) and rated
315   their confidence in this diagnosis from 1 (not very confident) to 5 (extremely confident). Initial
316   impressions from the three raters were averaged, with 0.0 indicating non-autism and 1.0
317   indicating autism; intermediate scores thus indicated disagreement among the three raters.
318   Confidence scores were also averaged across the three raters.
319   The Five-Minute Impressions Form was based on Wieckowski et al.,(2021) and Thomas
320   et al., (2024) and modified for use in adolescent and young adult populations via discussions
321   with a large study team including expert clinicians and clinicians in training. A pilot form was
322   employed in evaluations of 10 participants, and further refined via discussion with the study
323   team. Inter-rater reliability scoring for initial impressions of three participants was exceptionally
324   high (Cronbach's alpha =1.0). Training was performed to ensure adequate agreement on the
325   definition of each item-level behavior. Inter-rater reliability for item-level characteristics was not
326   performed during this initial validation, as we expected variability across individual raters as to
327   which behaviors contributed to their initial impressions.

328   ***Measures***
329   Participants completed a battery of measures as part of the larger study, a subset of which
330   were included in the present study analyses. Participants and parents or caregivers completed an
331   online Qualtrics survey probing sociodemographic information of the participant including race
332   and ethnicity, sex assigned at birth, and yearly gross family income.
333   *ADOS-2*. The Autism Diagnostic Observation Schedule, Second Edition (46) Module 4,
334   served to confirm diagnosis and to provide a measure of autism-related behavioral
335   characteristics. The ADOS-2 consists of a series of semi-structured tasks designed to elicit
336   social, communicative, and repetitive and stereotyped behaviors relevant to the ASD diagnosis.
337   Module 4, chosen based on developmental level, includes 32 scorable items, scored from 0
338   (typical) to 3 (definitely atypical). Item scores are used to calculate two Domain Scores (Social
339   Affect and Restricted and Repetitive Behaviors), and to calculate the ADOS-2 Overall Total and
340   Calibrated Severity Score (CSS). A CSS of eight or more suggests an ASD diagnosis. The
341   ADOS-2 CSS is a reliable index of the severity of autism symptoms for each module and yields
342   greater sensitivity and specificity than the ADOS-2 Module 4 raw scores (sensitivity = 89.6 (raw
343   scores) and 90.5 (CSS); specificity = 72.2 (raw scores) and 82.2 (CSS); (50). ADOS-2
344   classifications have good concurrent validity with clinical best estimate of ASD diagnoses (50).
345   This study used the CSS from administration of the Module 4 revised algorithm, along with
346   information from participant evaluations performed early in development to inform clinical
347   judgment.
348   *Cognitive ability*. Participants completed two tasks from the Penn Computerized
349   Neurocognitive Battery (CNB); (51). The Penn CNB, modeled on standardized
350   neuropsychological tests, provides a reliable online estimate of cognitive functioning in a five-to-

351 eight-minute test. Cronbach's alphas for subtests range from moderate to high (0.78-0.97) with
352 high internal consistency for speed (alpha = 0.78-0.98) and moderate internal consistent for
353 accuracy (alpha = 0.55-0.95); (52). Subtests included in this study were the Abbreviated Verbal
354 Reasoning Test, in which participants answer multiple choice questions about verbal analogies,
355 and the Matrix Reasoning Test, in which participants complete visual puzzles. The Abbreviated
356 Verbal Reasoning test has high concordance with the full Penn verbal reasoning battery,
357 R^2=0.90-0.92 (53). The Matrix Reasoning Test forms part of the Nonverbal Reasoning domain; it
358 was found to load appropriately in both exploratory and confirmatory bifactor analyses (loading=
359 0.32-0.49); (54). Following standard procedures (52), we transformed accuracy and RT into an
360 efficiency score, calculated as percent accuracy divided by log RT, to yield individually
361 interpretable scores.

362 ***Planned analyses***
363   We used frequentist statistics in R-Studio (55) to test each research question. Significance
364 values were Bonferroni corrected for multiple comparisons. Across analyses, we evaluated 3-
365 group comparisons (group), as well as comparing the autism group to the groups with a non-
366 autism outcome (LAD, NT). First, *t*-tests were used to assess significant differences in accuracy
367 (e.g., concordance of initial impression with final diagnostic classification of ASD or non-ASD)
368 for graduate versus expert PhD level clinicians. Sensitivity, specificity, positive predictive value,
369 and negative predictive value were calculated and *t*-tests were used to evaluate differences in
370 accuracy and total score by diagnostic status (e.g. ASD, non-ASD). Analysis of variance models
371 (ANOVA) evaluated group differences (e.g., ASD, LAD, NT) in total score and accuracy, with
372 post-hoc two-way comparisons for any significant results. The relationship between total score
373 and ADOS-2 CSS was assessed using a generalized linear model collapsed across groups.
374   To evaluate behavioral factors that contribute to initial impression, analysis of variance
375 models were used to compare item-level factors (e.g., gesture, motor mannerisms, eye contact,
376 prosody and vocalizations, facial expressions, focus/shifting of attention, social reciprocity, and
377 social initiations) by group, with post-hoc two-way comparisons for any significant results.
378 Within groups, Pearson correlations assessed the relationship between item-level and initial
379 impression. To assess which item-level best predicted initial impression scores, a generalized
380 linear model was used with each item-level score added as a predictor in the model.
381   Analysis of variance models were used to compare initial impression confidence across
382 groups, with post-hoc group two-way comparisons for any significant results. There was no
383 missing data across all variables utilized in the proposed analyses.

384   **Results**
385 ***Preliminary analyses***
386   We compared the impression accuracy of graduate and expert clinicians (79% and 78%,
387 respectively), which did not differ, $t = 0.107$, $p = 0.92$, $d = 0.02$. Total scores also did not differ
388 by expertise, $t = 0.447$, $p = 0.66$, *d = 0.07*. As such, all ratings were collapsed, and mean ratings
389 were used for all subsequent analyses. Similarly, we evaluated accuracy and total score as a
390 function of modality (videoconference, *n*=68, versus in-person, *n*=6). There was no difference
391 by modality for accuracy, $t = 0.389$, $p = 0.7$, $d = 0.22$, or total score, $t = -1.081$, $p = 0.32$, $d = 0.58$, and session modality was collapsed for all subsequent analyses.
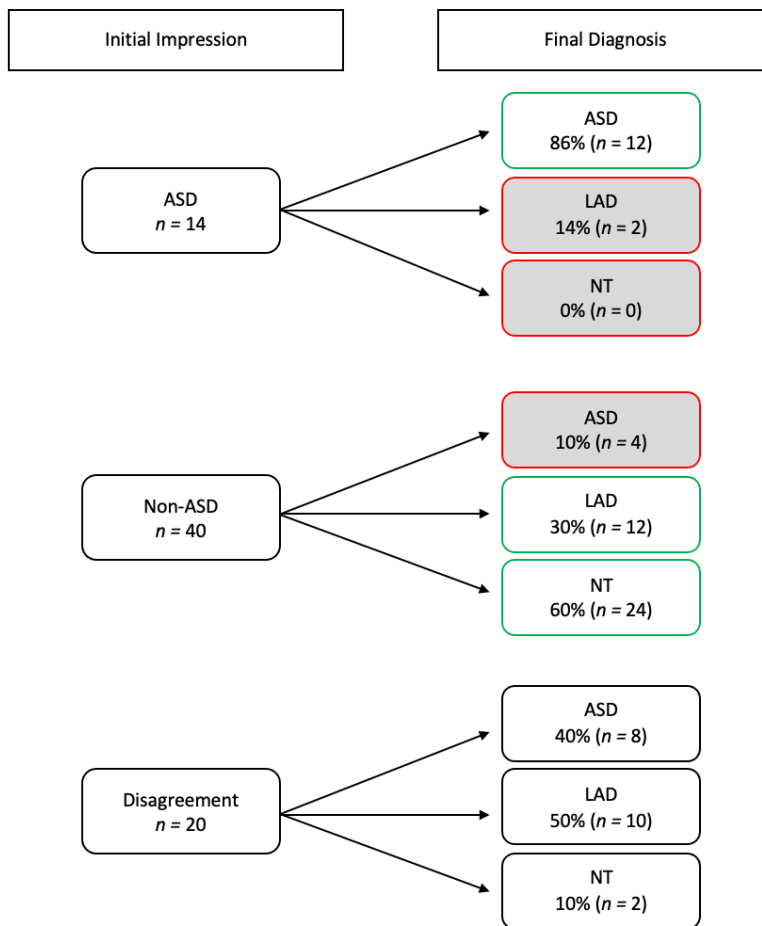392 0.58, and session modality was collapsed for all subsequent analyses.
393 ***Accuracy of initial impressions***
394   Initial impressions across the three raters were generally concordant with final diagnosis
395 (79% accurate). Of those with an initial impression of ASD (*n* = 14), 86% were in the final ASD
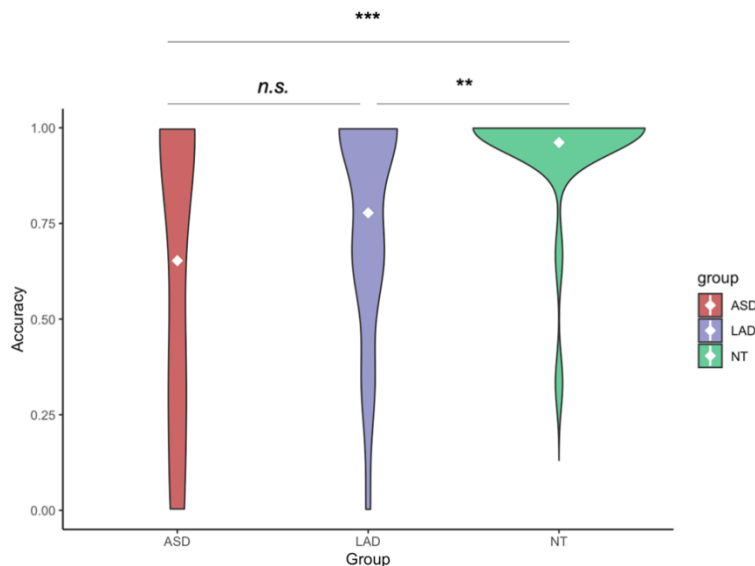
396   group, 14% were in the LAD group, and 0% were in the NT group, resulting in an overall
397   accuracy of 86%; Figure 1. Of those with a non-ASD initial impression ($n = 40$), 10% were in
398   the ASD group, 30% were in the LAD group, and 60% were in the NT group, resulting in an
399   overall accuracy of 90% (LAD and NT groups combined). Of those who received initial
400   impressions of both ASD and non-ASD from the three raters (e.g., those for whom clinicians
401   disagreed in initial impressions; $n = 20$), 40% were in the ASD group, 50% were in the LAD
402   group, and 10% were in the NT group. Accuracy was greater for non-autistic participants, $t = -7.319$, $p = 0.03$, $d = 0.63$. The three groups differed in accuracy, $F(2, 71) = 6.9$, $p = 0.002$.
403   7.319, $p = 0.03$, $d = 0.63$. The three groups differed in accuracy, $F(2, 71) = 6.9$, $p = 0.002$.
404   Specifically, initial impressions were significantly more accurate for the NT group compared to
405   the ASD group, $t = -3.582$, $p = 0.001$, $d = 1.05$, and the LAD group, $t = -3.122$, $p = 0.004$, $d = 0.23$; see Figure 2. Accuracy for ASD and LAD groups did not differ, $t = -0.793$, $p = 0.43$, $d = 0.91$.
406   0.23; see Figure 2. Accuracy for ASD and LAD groups did not differ, $t = -0.793$, $p = 0.43$, $d = 0.91$.
407   0.91.
408   **Figure 1.**
409   *Accuracy of initial impressions*

410
411   *Note.* Matches between clinician's initial impressions and group are highlighted in green with an
412   unshaded box; mismatches are shown in red with shaded boxes. Disagreement = differences in
413   initial diagnostic impression (ASD versus non-ASD) among the three raters.

414   **Figure 2.**
415   *Initial impression accuracy by group*

416
417   *Note.* Violin plot of initial impression accuracy by group. White diamonds indicate the mean.
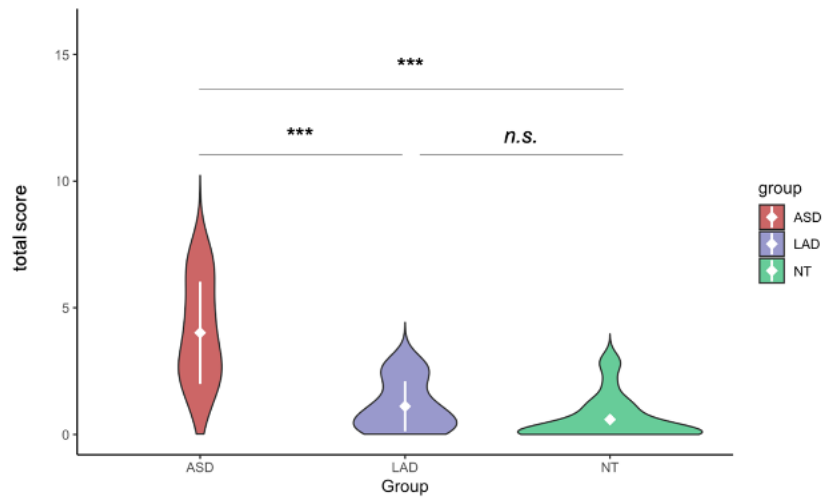418   ** $p < 0.01$, *** $p < 0.001$.
419
420   Each individual rater's initial impressions were used to calculate sensitivity (the
421   proportion of individuals with a final diagnosis of autism who were judged autistic on initial
422   impression; that is, initial impression true positives divided by final diagnoses of autism),
423   specificity (the proportion of individuals with a final diagnosis of non-autism who were judged
424   non-autistic on initial impression), positive predictive value (PPV; the likelihood that an
425   individual with initial impressions of autism would receive a final diagnosis of autism), negative
426   predictive value (NPV; the likelihood that an individual who received initial impression of non-
427   autism received a final diagnosis of non-autism), false negatives (e.g., the proportion of
428   individuals who received an initial impression of non-autism but ultimately received an autism
429   diagnosis), and false positives (e.g., the proportion of individuals who received an initial
430   impression of autism but ultimately received a diagnosis of non-autism). These calculations were
431   performed to compare the outcome of brief impressions to the outcome of a longer, more
432   thorough clinical diagnostic evaluation.
433   Results indicated that sensitivity was 66.7%, considered moderately low. Specificity was
434   calculated at 88.0%, considered high. The false positive rate was 34.7%, and the false negative
435   rate was 14.7%. Similarly, PPV was moderate (72.7%) and NPV was high (84.6%). These
436   findings were broadly consistent with results from Wieckowski et al. (2021), which reported
437   sensitivity = 64% and specificity = 96% (PPV and NPV were not reported).
438   ***Total scores***
439   Total scores of atypical behaviors differed by group, $F(2,71) = 44$, $p < 0.001$. Total
440   scores were significantly higher for the ASD group, M(SD) = 4.01(2.02), compared to both the
441   LAD group, M(SD) = 1.11(0.99), $t = 6.319$, $p < 0.001$, $d = 1.82$, and NT group, M(SD) =
442   0.56(0.85), $t = 7.698$, $p < 0.001$, $d = 2.24$; the LAD and NT groups did not differ, $t = 1.989$, $p =$
443   0.053, $d = 0.57$; see Figure 3. Total scores and ADOS-2 CSS scores were significantly and
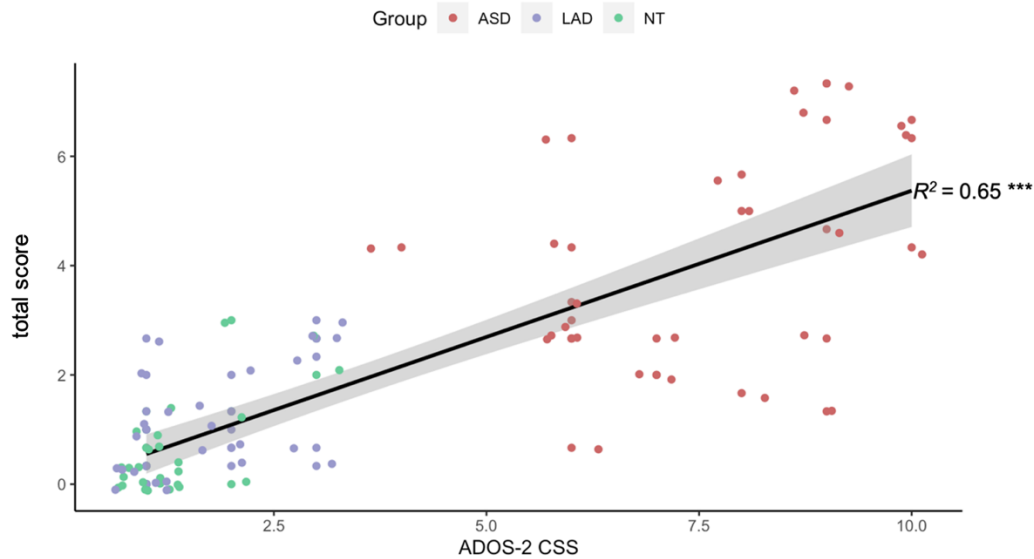444   strongly correlated, $R^2 = 0.65$, $p < 0.001$; see Figure 4.
445   **Figure 3.**

446   *Total score by group*

447
448   *Note*. White diamonds and lines indicate *M(SD)*. ***p* < 0.001.

449   **Figure 4.**

450   *Relationship between initial impression scores and ADOS-2 CSS scores*

451

452   *Note*. Black line represents the line of best fit; grey shading indicates SE. ***$p < 0.001$.

453

454   ***Factors contributing to initial impressions***

455   Item-level scores on gesture, eye contact, prosody, facial expressions, social reciprocity,

456   and social initiations differed significantly by group. In all cases, item-level Scores were

457   significantly higher for the ASD group compared to both the LAD and NT, which did not differ;

458   see Table 2. There were no group differences for motor mannerisms or attention.

459   **Table 2.**

460   ***Item-level scores by group***

	ASD M(SD)	LAD M(SD)	NT M(SD)	F	Post-hoc comparison	d	Correlation (r) with initial impression Scores		
							ASD	LAD	NT
Prosody	1.06 (0.63)	0.36 (0.47)	0.06 (0.21)	**29.35***	ASD>NT, LAD	2.13 1.24 0.83	**0.82***	**0.80***	**0.79 ***
Facial Expressions	0.67 (0.38)	0.21 (0.27)	0.14 (0.21)	**22.85***	ASD>NT, LAD	1.72 1.38 0.27	**0.73***	**0.55***	**0.68 ***
Gesture	0.60 (0.50)	0.19 (0.33)	0.10 (0.18)	**13.26***	ASD>NT, LAD	1.33 0.96 0.32	0.18	0.15	**0.69 ***
Eye contact	0.55 (0.42)	0.07 (0.16)	0.06 (0.13)	**25.75***	ASD>NT, LAD	1.57 1.49 0.04	0.30	0.14	**0.56 ***
Social Reciprocity	0.45 (0.36)	0.13 (0.29)	0.06 (0.16)	**13.34***	ASD>NT, LAD	1.40 0.99 0.26	**0.58***	0.10	-0.11

Social Initiations	0.51 (0.44)	0.14 (0.21)	0.08 (0.18)	**15.06***	ASD>NT, LAD	1.30 1.09 0.28	0.51	0.11	0.13
Attention	0.20 (0.31)	0.05 (0.13)	0.01 (0.07)	**6.24***	No differences	0.85 0.63 0.34	0.31	-0.20	-0.05
Motor Mannerisms	0.17 (0.27)	0.09 (0.24)	0.07 (0.18)	1.19	No differences	0.42 0.30 0.10	0.29	-0.04	0.06

461 *Note.* Data are represented as *M(SD)*. * Indicates significance after correcting for multiple
462 comparisons of ANOVA (0.05/8; $p<0.006$), t-tests (0.05/27; $p<0.002$), and Pearson's
463 correlations (0.05/9; $p<0.006$). Significant associations are **bolded**. Cohen's d values are listed
464 in the following order: ASD vs. NT, ASD vs. LAD, and LAD vs. NT.

466   The relationship between item-level scores and initial impression varied by diagnostic
467 status (autism, non-autism) and group (ASD, LAD, NT). For those in the ASD group, initial
468 impressions were significantly and strongly correlated with prosody, facial expressions, and
469 social reciprocity. In the LAD group, initial impressions were significantly and strongly
470 correlated with prosody and facial expressions. In the NT group, initial impressions were
471 significantly and strongly correlated with gesture, eye contact, prosody, and facial expressions;
472 see Table 2. Collapsed across group, initial impression scores were most strongly predicted by
473 item-level scores on prosody and facial expressions; Table 3.

474 **Table 3.**
475 *Generalized linear model of item-level scores and initial impression*

	Initial Impression			
Predictors	**Estimate**	**Std. Error**	***t***	***p***
Prosody	0.448	0.044	10.129	<0.001***
Facial Expressions	0.300	0.087	3.439	0.001**
Social Initiations	0.120	0.084	1.419	0.161
Social Reciprocity	0.032	0.082	0.389	0.698
Eye Contact	0.003	0.079	0.035	0.973
Intercept	-0.003	0.026	-0.138	0.891
Motor Mannerisms	-0.007	0.093	-0.071	0.943
Attention	-0.058	0.102	-0.569	0.571
Gesture	-0.111	0.062	-1.780	0.080

476 *Note.* All predictors were added to the model simultaneously. AIC = -54.819.

478 ***Confidence in initial impressions***
479   Confidence was high overall (mean of 3.06 out of 5) and ranged from 2.91-3.43, with the
480 highest confidence ratings for the NT group, though rating differences by group were not
481 significant. As noted above, expert and graduate clinicians did not differ in confidence. Overall,
482 clinicians were significantly more confident with initial impressions that were correct
483 (confidence = 3.43) compared to those that were incorrect (confidence = 2.83), $t = 3.249$,
484 $p=0.006$, $d = 0.77$, or in disagreement (confidence = 2.40), $t = 4.969$, $p < 0.001$, $d = 1.30$.

Discussion

485   The present study evaluated frank autism impressions in adolescents and adults,
486 compared them to current gold-standard diagnostic group classifications, and characterized the
487 behavioral factors that contributed to clinician impressions and confidence in initial impression.
488 Overall, the specificity and NPV of initial impressions of a combined group of graduate and
489 expert clinicians with specialized training in autism were high (88.0% and 84.6%, respectively),
490 indicating a low false positive rate; that is, clinicians were highly likely to identify an individual
491 who was non-autistic as such on initial impression. In contrast, the sensitivity and PPV of initial
492 impressions were lower (66.7% and 72.7%, respectively), indicating a **high false negative rate**;
493 about a third of the autistic individuals were misidentified as non-autistic. These results were
494 highly consistent with findings from three previous studies assessing initial impressions of
495 autistic toddlers (4,13,14), adding evidence that brief clinical observations provide valuable
496 insight about diagnostic status in *some* but not *all* cases of autism, and are more useful in ruling
497 *out* the presence of autism symptoms than in ruling them *in*. That is, *all* individuals who gave a
498 consistent initial impression of autism also met diagnostic criteria for autism upon full
499 evaluation. In addition, these results are novel in suggesting that clinicians are able to detect
500 autism symptoms after a brief observation of adolescents and adults, not just young children.
501 This may indicate that frank autism presentations may persist through adulthood, though the
502 specific behaviors that contribute to this impression vary over development.
503   These results represent a novel assessment of whether individuals who no longer present
504 with symptoms of autism (e.g., have lost the ASD diagnosis) present with subclinical, subtle
505 autistic characteristics. The LAD and NT groups did not differ in initial impression, with both
506 groups showing significantly lower scores than the ASD group, suggesting that the LAD group
507 overall presents as neurotypical. However, although misclassifications were relatively infrequent,
508 the accuracy of initial impressions (e.g., alignment with final diagnosis) did differ by group.
509 Raters were significantly more accurate for the NT group compared to both the ASD and LAD
510 groups; clinicians were more likely to have an initial impression of ASD for LAD relative to NT
511 participants. This result suggests that some individuals with LAD have subtle persistent autistic
512 behaviors.
513   In line with the sensitivity and specificity values, frankness of initial symptom
514 presentation (initial impression score) was consistent with symptom severity as observed through
515 a lengthy diagnostic observation (ADOS-2 CSS). Importantly, in this study, impression ratings
516 were provided by clinicians who did not perform the diagnostic evaluation, suggesting that the
517 identification of frank autism is not a result of confirmation bias (i.e., a tendency for clinicians to
518 align their final diagnosis with their initial impressions).

519 *Autistic characteristics and initial impressions*
520   To expand our understanding of how specific behaviors contribute to initial impressions
521 of frank autism in adolescents and adults, we assessed eight behaviors that have been described
522 as most central to impressions in anecdotal (3) and empirical (4,13,14) research. Results
523 indicated that ratings within the first five minutes on prosody and facial expressions were the
524 best predictors of initial impression across final diagnosis (autistic, non-autistic) and group
525 (ASD, LAD, NT). That is, atypical prosody and facial expressions appear to be the most salient
526 and reliably available indicators of ASD diagnoses in brief observations of adults; if these
527 behaviors are seen as typical, the individual is more likely to receive an initial impression and a
528 final determination as non-autistic. Because of the low false positive rate in this relatively small
529 sample, further analyses were not performed to determine which behaviors contributed to group

531 differences in accuracy; this should be examined in future research. Prosody and facial
532 expressions were significantly correlated with initial impression scores when evaluated within
533 each of three groups. In addition, for the ASD group, social reciprocity was a strong predictor of
534 impressions, and for the NT group, gestures and eye contact were strong predictors. This
535 indicates that, while prosody and facial expressions are the most prominent behavioral features
536 of frank autism impressions in adults, clinicians rely on other behavioral factors as well.
537   This finding was also reflected in clinician confidence in initial impressions, in that
538 higher scores on gesture, eye contact, prosody, facial expressions, and social reciprocity were all
539 significantly negatively correlated with confidence for the non-ASD diagnostic status (LAD and
540 NT). That is, if a clinician formed an initial impression of non-autism, but observed mild
541 atypicality in one or more of behavioral domains, they might still settle on a non-ASD
542 impression, but with reduced confidence. Even though overall confidence did not vary by group,
543 the absence of significant relationships between behavioral factors and confidence ratings for the
544 autism group may indicate that confidence in initial impressions is more susceptible to change
545 based on contradictory evidence (i.e., presence of atypical behaviors for a generally non-autism
546 impression) in LAD and NT groups as compared to contradictory evidence (i.e., absence of
547 atypical behaviors for a generally autistic impression) in the ASD group; this is consistent with
548 the broad finding that the presence of evidence is more salient than the absence of evidence (56).
549 ***Clinical implications***
550   One clinical implication of the current study is that *brief observations alone are not*
551 *sufficient to detect all cases of autism accurately.* Expert and graduate clinicians with specialized
552 autism-specific training did not have an initial impression of autism for roughly 33% of autistic
553 individuals. *Longer structured assessments by trained clinicians provide invaluable information*
554 *about the nature and severity of autism symptoms.* In practice, clinician judgment should be
555 considered an integral, but not the sole, factor in diagnosis.
556   The current study also provides important information about what behaviors inform
557 diagnosis of adolescents and adults. Typical diagnostic practices rely heavily on parent or
558 caregiver report of the developmental history of the individual, which may be difficult to obtain
559 and less accurately recalled years later (15,17,18). The absence of a clear developmental history
560 may force clinicians to rely more heavily on behavioral observations alone. This, in combination
561 with the evidence that some clinicians rely on more abstract behavioral observations, may lead
562 some autistic adults to receive an official diagnosis of autism more readily than others. For
563 example, if prosody and facial expressions are the most salient diagnostic cues, an individual
564 who presents with less frankly autistic behaviors in those domains may receive an initial
565 impression of being non-autistic. When combined with a limited developmental history, such a
566 presentation may lead to an incorrect non-ASD Impression. This factor may also contribute to
567 under-diagnosis of autistic females (57).
568   ***Implications for our understanding of autism as a diagnosis***
569   There has been a resurgence of the debate about the loosening of diagnostic criteria for
570 autism in the DSM-V, and the resulting increased prevalence and heterogeneity of the diagnosis
571 (43,45). While these DSM changes present challenges for how best to identify and diagnose
572 participants for autism research, we note that the *clinical* question of how to diagnose autism
573 requires, by definition and practice, a reliance on current DSM or ICD criteria, informed by a
574 detailed semi-structured clinical interview in combination with a thorough developmental
575 history. In research, Mottron and colleagues have proposed that cohorts of autistic individuals
576 should be selected based on the "prototypicality" of their autism symptoms, as judged by expert

577   clinicians (45). This manuscript asks how clinicians' initial diagnostic impressions ("five-minute
578   impressions" in the current study) correspond to a full diagnostic impression; the current study
579   also tests whether brief five-minute impressions provide additional information about the clinical
580   presentation of autism in individuals who have lost the autism diagnosis. As reported by
581   deMARHCNEA CITE, many expert clinicians report a strong initial impression of frank autism;
582   it is *clinically* and *diagnostically* relevant to contrast these initial impressions with the findings of
583   a full diagnostic evaluation, informed by clinical history and by a semi-structured interview, as
584   one's initial impressions cannot help but inform one's subsequent evaluation. The current results
585   indicate that initial impressions of frank autism, which likely overlaps with the construct of
586   "prototypical" autism, are highly sensitive; clinicians rarely "felt" autism that turned out *not* to
587   be autism on full evaluation. However, results also suggested that some cases that did not yield
588   an initial frank or prototypical impression of autism during a brief five-minute interaction, were
589   judged to be autism on full evaluation. This is not surprising, particularly in a population with
590   strong verbal and cognitive abilities. These results suggest that, even for adult individuals who
591   present as frankly autistic, the behavioral factors that contribute to this impression vary,
592   encompassing prosody, facial expressions, and other behaviors. Comparing behaviors used for
593   initial impressions in older adolescents and adults versus those used for toddlers (14) also
594   suggests that quite different behaviors are used for different ages or functioning levels, in part
595   because young children may not have enough language to judge prosody or other more mature
596   behaviors.
597   The current study established a strong association between initial impression and ADOS-
598   2 CSS, indicating that individuals who present as frankly autistic may have more apparent or
599   severe autism symptoms. However, those who present as less frankly autistic during initial
600   impressions may still meet clinical criteria when observed through a full diagnostic assessment.
601   This evidence suggests that the ADOS-2 or other structured observation plays a critical role in
602   eliciting behaviors that are core to the autism phenotype that may or may not be present within
603   the first few minutes of observation or interaction. Rather than focusing only on frank autism
604   characteristics, clinical researchers must carefully evaluate individual differences.
605   There has been ongoing debate about the stability of autism as a lifelong condition.
606   Several research groups have documented individuals who met criteria for autism in childhood
607   but no longer displayed clinical levels of ASD later in development (33,34,58–60). This subset
608   of the autism spectrum may range from 3 to 37% of children diagnosed with ASD in early
609   childhood (32,33). The effect of the transition to independence in late adolescence and adulthood
610   have yet to be characterized in this group of individuals. The current study indicates that, overall,
611   LAD individuals are indistinguishable from their NT peers on initial impression, but that some of
612   them present with subtle subclinical autistic features that are recognizable to trained clinicians
613   and that contribute to reduced accuracy of initial impressions. Importantly, these features do not
614   give rise to clinical level impairments in social functioning, as evidenced by non-autistic range
615   ADOS-2 CSS scores but may have subtle implications for daily functioning that have yet to be
616   fully explored.

617   ***Implications for social functioning and quality of life of autistic adolescents and adults***
618   Previous studies have established that naïve raters are sensitive to atypical behaviors,
619   including prosody and facial expressions, that negatively impact initial impressions of autistic
620   individuals and that may lead to a decrease in the rater's willingness to engage with the autistic
621   individual in hypothetical circumstances, such as sharing a meal (26,27,31). The current study
622   provided evidence that clinicians utilize similar cues in clinical diagnosis, indicating that prosody

and facial expressions are highly salient for both clinical and social initial impressions. These results have important implications when assessing the social functioning of autistic adolescents and adults. If naïve individuals are less willing to interact with their autistic peers due to an increased perception of social awkwardness, this may exacerbate the social difficulties of autistic individuals, which may in turn impact their quality of life and vulnerability to anxiety or depression. Future studies should explore relationships among initial impressions and other functional domains to better understand how we can best support autistic individuals, as well as provide further evidence to support societal acceptance of neurodiversity.

Limitations

The brief observations in this study were conducted in a clinical research setting, including some portion of structured testing. While results suggest concordance with naïve impressions during typical social interactions, findings may not generalize to less structured, informal contexts. In addition, the content of interactions varied, with some interactions comprising mainly informal "chit chat" during equipment setup, and others comprising a mix of structured and unstructured activities. The ADOS-2 activity at the beginning of the recording was not consistent across participants; in-person participants began with a puzzle task (Construction activity), whereas online participants began with a story activity (Tuesday Story). Although results revealed no modality-specific differences, this variability could have impacted interactions on a case-by-case basis. That said, the consistency of results across modalities is a testament to the stability and potential clinical utility of initial impressions, given that they seem to generalize to a range of settings. Our sample was relatively small, demographically homogeneous, and comprised of only speaking autistic participants; furthermore, there were significantly more females in the NT group compared to ASD or LAD groups. We were underpowered to compare which behaviors led to disagreement among raters, and to test gender difference, precluding the discovery of which factors played a critical role in the formation of accurate initial impressions. Future studies should recruit larger sample sizes that are diverse across demographic variables and ability level to thoroughly explore these nuances.

Conclusions

The results of the current study indicate that, while clinicians' initial impressions made within the first five minutes of observation of a diagnostic evaluation generally matched current gold-standard diagnostic status and were highly correlated with ADOS-2 CSS, this brief observation was not sufficient to detect autism in all cases. Clinicians relied heavily on atypical prosody and facial expressions when forming an initial impression of autism, indicating that these cues are extremely salient even within a brief observation. Lastly, the results of the current study further established that LAD individuals no longer exhibit clinically significant autism symptoms, but that some individuals in this group may continue to display subtle autism characteristics that lead to more variable initial impressions. Future research is needed to explore the impact of sex and gender on initial impression, as well as the impact of frank autism presentations on other domains of functioning, such as social functioning in everyday life (e.g., making and maintaining friendships and romantic relationships), quality of life, life satisfaction, and comorbid psychopathology.

<div align="center">**Declarations**</div>

669   ORCID: Rebecca R. Canale: 0000-0001-5892-2364; Caroline Larson: 0000-0001-7940-2528;
670   Rebecca P. Thomas: 0000-0002-7159-7532; Marianne Barton: 0000-0001-5940-7591; Deborah
671   Fein: 0000-0002-8478-5938; Inge-Marie Eigsti: 0000-0001-7898-1898

686
687   **References**

688   1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th
689   ed. 2013.

690   2. Volkmar FR, McPartland JC. From Kanner to DSM-5: Autism as an Evolving Diagnostic
691   Concept. Annu Rev Clin Psychol. 2014 Mar 28;10(1):193–212.

692   3. De Marchena A, Miller J. "Frank" presentations as a novel research construct and element of
693   diagnostic decision-making in autism spectrum disorder. Autism Res. 2017 Apr;10(4):653–62.

694   4. Wieckowski AT, De Marchena A, Algur Y, Nichols L, Fernandes S, Thomas RP, et al. The
695   first five minutes: Initial impressions during autism spectrum disorder diagnostic evaluations
696   in young children. Autism Res. 2021 Sep;14(9):1923–34.

697   5. Mottron L, Gagnon D. Prototypical autism: New diagnostic criteria and asymmetrical
698   bifurcation model. Acta Psychol (Amst). 2023 Jul;237:103938.

699   6. Anderson-Chavarria M. The autism predicament: models of autism and their impact on
700   autistic identity. Disabil Soc. 2021;37:1321–41.

701   7. Kenny L, Hattersley C, Molins B, Buckley C, Povey C, Pellicano E. Which terms should be
702   used to describe autism? Perspectives from the UK autism community. Autism Int J Res Pract.
703   2016 May;20(4):442–62.

704   8. Frazier TW, Youngstrom EA, Speer L, Embacher R, Law P, Constantino J, et al. Validation of
705   Proposed DSM-5 Criteria for Autism Spectrum Disorder. J Am Acad Child Adolesc
706   Psychiatry. 2012 Jan;51(1):28-40.e3.

707   9. Rice CE, Carpenter LA, Morrier MJ, Lord C, DiRienzo M, Boan A, et al. Defining in Detail
708   and Evaluating Reliability of DSM-5 Criteria for Autism Spectrum Disorder (ASD) Among
709   Children. J Autism Dev Disord. 2022 Dec;52(12):5308–20.

710   10. Hayes J, McCabe R, Ford T, Russell G. Drawing a line in the sand: affect and testimony in
711   autism assessment teams in the UK. Sociol Health Illn. 2020 May;42(4):825–43.

712   11. Hausman-Kedem M, Kosofsky BE, Ross G, Yohay K, Forrest E, Dennin MH, et al.
713   Accuracy of Reported Community Diagnosis of Autism Spectrum Disorder. J Psychopathol
714   Behav Assess. 2018 Sep;40(3):367–75.

715   12. Lord C, Rutter M, DiLavore PC, Risi S. Autism Diagnostic Observation Schedule (ADOS).
716   Los Angeles: Western Psychological Services; 2002.

717   13. Gabrielsen TP, Farley M, Speer L, Villalobos M, Baker CN, Miller J. Identifying Autism in a
718   Brief Observation. PEDIATRICS. 2015 Feb 1;135(2):e330–8.

719   14.Thomas RP, De Marchena A, Wieckowski AT, Stahmer A, Milan S, Burke JD, et al.
720   Accuracy of initial diagnostic impressions of autism in toddlers and behaviors that inform
721   these impressions. Autism Res. 2024 Mar;17(3):568–83.

722   15. Huang Y, Arnold SR, Foley KR, Trollor JN. Diagnosis of autism in adulthood: A scoping
723   review. Autism. 2020 Aug;24(6):1311–27.

724   16. Jensen CM, Steinhausen HC, Lauritsen MB. Time Trends Over 16 Years in Incidence-Rates
725   of Autism Spectrum Disorders Across the Lifespan Based on Nationwide Danish Register
726   Data. J Autism Dev Disord. 2014 Aug;44(8):1808–18.

727   17. Lai MC, Baron-Cohen S. Identifying the lost generation of adults with autism spectrum
728   conditions. Lancet Psychiatry. 2015 Nov;2(11):1013–27.

729   18. Rutherford M, McKenzie K, Forsyth K, McCartney D, O'Hare A, McClure I, et al. Why are
730   they waiting? Exploring professional perspectives and developing solutions to delayed
731   diagnosis of autism spectrum disorder in adults and children. Res Autism Spectr Disord. 2016
732   Nov;31:53–65.

733   19. Hus V, Taylor A, Lord C. Telescoping of caregiver report on the Autism Diagnostic
734   Interview - Revised: Telescoping on the ADI-R. J Child Psychol Psychiatry. 2011
735   Jul;52(7):753–60.

736   20. De Marchena A, Kim ES, Bagdasarov A, Parish-Morris J, Maddox BB, Brodkin ES, et al.
737   Atypicalities of Gesture Form and Function in Autistic Adults. J Autism Dev Disord. 2019
738   Apr;49(4):1438–54.

739   21. De Marchena A, Eigsti I. Conversational gestures in autism spectrum disorders: Asynchrony
740   but not decreased frequency. Autism Res. 2010 Dec;3(6):311–22.

741   22. Morett LM, O'Hearn K, Luna B, Ghuman AS. Altered Gesture and Speech Production in
742   ASD Detract from In-Person Communicative Quality. J Autism Dev Disord. 2016
743   Mar;46(3):998–1012.

744   23. Silverman LB, Bennetto L, Campana E, Tanenhaus MK. Speech-and-gesture integration in
745   high functioning autism. Cognition. 2010 Jun;115(3):380–93.

746   24. Taverna EC, Huedo-Medina TB, Fein DA, Eigsti IM. The interaction of fine motor, gesture,
747   and structural language skills: The case of autism spectrum disorder. Res Autism Spectr
748   Disord. 2021 Aug;86:101824.

749   25. Trujillo JP, Özyürek A, Kan CC, Sheftel-Simanova I, Bekkering H. Differences in the
750   production and perception of communicative kinematics in autism. Autism Res. 2021
751   Dec;14(12):2640–53.

752   26. Grossman RB. Judgments of social awkwardness from brief exposure to children with and
753   without high-functioning autism. Autism. 2015 Jul;19(5):580–7.

754   27. Sasson NJ, Faso DJ, Nugent J, Lovell S, Kennedy DP, Grossman RB. Neurotypical Peers are
755   Less Willing to Interact with Those with Autism based on Thin Slice Judgments. Sci Rep.
756   2017 Feb 1;7(1):40700.

757   28. Neumann D, Spezio ML, Piven J, Adolphs R. Looking you in the mouth: abnormal gaze in
758   autism resulting from impaired top-down modulation of visual attention. Soc Cogn Affect
759   Neurosci. 2006 Dec 1;1(3):194–202.

760   29. Weed E, Fusaroli R, Simmons E, Eigsti IM. Different in Different Ways: A Network-
761   Analysis Approach to Voice and Prosody in Autism Spectrum Disorder. Lang Learn Dev.
762   2023 Apr 25;1–18.

763   30. Grossman RB, Bemis RH, Plesa Skwerer D, Tager-Flusberg H. Lexical and Affective
764   Prosody in Children With High-Functioning Autism. J Speech Lang Hear Res. 2010
765   Jun;53(3):778–93.

766   31. Sasson NJ, Morrison KE. First impressions of adults with autism improve with diagnostic
767   disclosure and increased autism knowledge of peers. Autism. 2019 Jan;23(1):50–9.

768   32. Helt M, Kelley E, Kinsbourne M, Pandey J, Boorstein H, Herbert M, et al. Can Children with
769   Autism Recover? If So, How? Neuropsychol Rev. 2008 Dec;18(4):339–66.

770   33. Harstad E, Hanson E, Brewster SJ, DePillis R, Milliken AL, Aberbach G, et al. Persistence of
771   Autism Spectrum Disorder From Early Childhood Through School Age. JAMA Pediatr. 2023
772   Nov 1;177(11):1197.

773   34. Fein D, Barton M, Eigsti I, Kelley E, Naigles L, Schultz RT, et al. Optimal outcome in
774   individuals with a history of autism. J Child Psychol Psychiatry. 2013 Feb;54(2):195–205.

775 35. Tyson K, Kelley E, Fein D, Orinstein A, Troyb E, Barton M, et al. Language and Verbal
776   Memory in Individuals with a History of Autism Spectrum Disorders Who Have Achieved
777   Optimal Outcomes. J Autism Dev Disord. 2014 Mar;44(3):648–63.

778 36. Canfield AR, Eigsti IM, De Marchena A, Fein D. Story Goodness in Adolescents With
779   Autism Spectrum Disorder (ASD) and in Optimal Outcomes From ASD. J Speech Lang Hear
780   Res. 2016 Jun;59(3):533–45.

781 37. Suh J, Eigsti IM, Naigles L, Barton M, Kelley E, Fein D. Narrative Performance of Optimal
782   Outcome Children and Adolescents with a History of an Autism Spectrum Disorder (ASD). J
783   Autism Dev Disord. 2014 Jul;44(7):1681–94.

784 38. Irvine CA, Eigsti IM, Fein DA. Uh, Um, and Autism: Filler Disfluencies as Pragmatic
785   Markers in Adolescents with Optimal Outcomes from Autism Spectrum Disorder. J Autism
786   Dev Disord. 2016 Mar;46(3):1061–70.

787 39. Fitch A, Fein DA, Eigsti IM. Detail and Gestalt Focus in Individuals with Optimal Outcomes
788   from Autism Spectrum Disorders. J Autism Dev Disord. 2015 Jun;45(6):1887–96.

789 40. Orinstein AJ, Suh J, Porter K, De Yoe KA, Tyson KE, Troyb E, et al. Social Function and
790   Communication in Optimal Outcome Children and Adolescents with an Autism History on
791   Structured Test Measures. J Autism Dev Disord. 2015 Aug;45(8):2443–63.

792 41. Suh J, Orinstein A, Barton M, Chen CM, Eigsti IM, Ramirez-Esparza N, et al. Ratings of
793   Broader Autism Phenotype and Personality Traits in Optimal Outcomes from Autism
794   Spectrum Disorder. J Autism Dev Disord. 2016 Nov;46(11):3505–18.

795 42. Troyb E, Orinstein A, Tyson K, Eigsti IM, Naigles L, Fein D. Restricted and Repetitive
796   Behaviors in Individuals with a History of ASDs Who Have Achieved Optimal Outcomes. J
797   Autism Dev Disord. 2014 Dec;44(12):3168–84.

798 43. Fein D, Eigsti I, Barton M. Response to "A radical change in our autism research strategy is
799   needed: Back to prototypes" by Mottron et al. (2021). Autism Res. 2021 Oct;14(10):2237–8.

800 44. Eigsti I, Fein D, Larson C. Editorial Perspective: Another look at 'optimal outcome' in
801   autism spectrum disorder. J Child Psychol Psychiatry. 2023 Feb;64(2):332–4.

802 45. Mottron L. A radical change in our autism research strategy is needed: Back to prototypes.
803   Autism Res. 2021 Oct;14(10):2213–20.

804 46. Lord C, Luyster J, Gotham K, Guthrie W. Autism Diagnostic Observation Schedule, Second
805   Edition (ADOS-2). Torrance, CA: Western Psychological Services; 2012.

806 47. Sparrow SS, Cicchetti DV. The Vineland Adaptive Behavior Scales. C.S. Newmark (Ed.);
807   1989.

808 48. Eigsti I, Thomas RP, Stabile M, Mohan A, Dieckhaus MFS, Crutcher J, et al. Online
809   administration of the ADOS for research with adolescents and adults in response to the
810   pandemic. Autism Res. 2022 Oct;15(10):1909–16.

811 49. Thomas R, Canale R, Larson C. Five-Minute Initial Impressions Form. 2021.

812 50. Hus V, Lord C. The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm
813   and Standardized Severity Scores. J Autism Dev Disord. 2014 Aug;44(8):1996–2012.

814 51. Gur R. Computerized Neurocognitive Scanning: I. Methodology and Validation in Healthy
815   People. Neuropsychopharmacology. 2001 Nov;25(5):766–76.

816 52. Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB, et al. A cognitive
817   neuroscience-based computerized battery for efficient measurement of individual differences:
818   Standardization and initial construct validation. J Neurosci Methods. 2010 Mar;187(2):254–
819   62.

820 53. Bilker WB, Hansen JA, Brensinger CM, Richard J, Gur RE, Gur RC. Development of
821   Abbreviated Nine-Item Forms of the Raven's Standard Progressive Matrices Test.
822   Assessment. 2012 Sep;19(3):354–69.

823 54. Moore TM, Reise SP, Gur RE, Hakonarson H, Gur RC. Psychometric properties of the Penn
824   Computerized Neurocognitive Battery. Neuropsychology. 2015 Mar;29(2):235–46.

825 55. RStudio Team. RStudio: Integrated Development for R. Boston, MA: PBC; 2020.

826 56. Altman DG, Bland JM. Statistics notes: Absence of evidence is not evidence of absence.
827   BMJ. 1995 Aug 19;311(7003):485–485.

828 57. Cola ML, Plate S, Yankowitz L, Petrulla V, Bateman L, Zampella CJ, et al. Sex differences
829   in the first impressions made by girls and boys with autism. Mol Autism. 2020 Dec;11(1):49.

830 58. Anderson DK, Liang JW, Lord C. Predicting young adult outcome among more and less
831   cognitively able individuals with autism spectrum disorders. J Child Psychol Psychiatry. 2014
832   May;55(5):485–94.

833 59. Moulton E, Barton M, Robins DL, Abrams DN, Fein D. Early Characteristics of Children
834   with ASD Who Demonstrate Optimal Progress Between Age Two and Four. J Autism Dev
835   Disord. 2016 Jun;46(6):2160–73.

836 60. Fountain C, Winter AS, Bearman PS. Six Developmental Trajectories Characterize Children
837   With Autism. Pediatrics. 2012 May 1;129(5):e1112–20.

838

839   **Appendix A.**

840   *Five-Minute initial impressions Form*

Coding manual: ALTOS 5-minute diagnostic impression form (Thomas, R., Canale, R., Larson, C., 2021)

Pt. ID number: Clinician name:	☐ Telehealth ☐ In-person	PT wearing mask? ☐ Yes ☐ No	Setting used for impression (check all that apply)? ☐ Video call ☐ Waiting room ☐ Walk to eval room ☐ Structured activity ☐ Other _____
5 min Dx impression ☐ ASD ☐ Non-ASD	Did you have access to information before completing this form? ☐ Yes ☐ No	Type of information accessed before completing this form ☐ Reviewed records ☐ Spoke with parent ☐ PT self-disclosed ☐ Other	Confidence in 5-min impression ☐ 1: Not very confident ☐ 2 ☐ 3: Confident ☐ 4 ☐ 5: Extremely confident

Complete form after first 5 minutes with Pt: 1) Indicate what information contributed to your initial Dx impression (left column) and 2) your confidence (+ -- confident) in the scoring of that component of your impression (right column). Please note behavioral observations that contributed to your impression.

Gesture: +   --
   ☐ 2- No use of age/context appropriate gestures
   ☐ 1- Some use of age/context-appropriate gestures
   ☐ 0- Normal use of age/context-appropriate of gestures

Eye contact: +   --
   ☐ 2- Abnormal quality or amount based on context
   ☐ 1- Somewhat abnormal quality or amount based on context
   ☐ 0- Normal quality and amount based on context

Motor mannerisms: +   --
   ☐ 2- Frequent use of atypical mannerisms
   ☐ 1- Some use of atypical mannerisms
   ☐ 0- No use of atypical mannerisms

Prosody and vocalizations: +   --
   ☐ 2- Frequent use of unusual vocal prosody and/or repetitive language
   ☐ 1- Some use of unusual vocal prosody and/or repetitive language
   ☐ 0- No use of unusual vocal prosody and/or repetitive language

Facial expressions: +   --
   ☐ 2- Mostly abnormal for context
   ☐ 1- Sometimes abnormal for context
   ☐ 0- Normal for context

Focus/shifting of attention (including perseverative thinking and distractibility):   +   --
   ☐ 2- Difficulty focusing or shifting attention
   ☐ 1- Some difficulty focusing or shifting attention
   ☐ 0- No difficulty focusing and shifting attention

Social reciprocity (responses):   +   --
   ☐ 2- Impaired
   ☐ 1- Somewhat impaired
   ☐ 0- Normal

Interactions with others (initiations):   +   --
   ☐ 2- Impaired
   ☐ 1- Somewhat impaired
   ☐ 0- Normal

Other   +   --
   ☐ _____

COMPLETE AFTER FINISHING APPT
Confidence in final diagnosis
   ☐1 (Not very confident) ☐2   ☐3 (Confident)   ☐4   ☐5 (Extremely confident)

841