Overlap and Differences of Autism and ADHD: Digital Phenotyping of Movement and Communication During Development

Aimar Silvan[1], Adriana Di Martino[2], Michael P. Milham[2], Lucas C. Parra[1], Jens Madsen[1]

[1] Department of Biomedical Engineering, City College of New York, New York, NY, 10031, USA
[2] Child Mind Institute, New York, NY, 10022, USA

Corresponding author: parra@ccny.cuny.edu

Abstract

Attention-Deficit/Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD) often co-occur, which complicates diagnosis across development. Here, we test whether automatic analysis of naturalistic video can disentangle shared from distinct behavioral signatures of the two disorders. We analyzed videos of 2,341 youths (ages 5–22) from a community sample while they described an emotive short film. Multivariate models revealed that language deficits previously attributed to ADHD were largely explained by age. Difficulties in understanding and recalling a social narrative, sometimes attributed to ADHD, were uniquely predicted by ASD. Increased motor activity was a specific marker of the hyperactive-impulsive domain of ADHD. Conversely, ASD showed structurally intact language but significant impairments in narrative ability and perspective-taking, coupled with a unique vocal profile of higher pitch, intensity, and altered voice quality. These findings suggest that despite substantial comorbidity, ADHD and ASD exhibit separable behavioral profiles that can be measured objectively at scale.

Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD) affect a significant portion of youth and often create long-term challenges in education, social relationships, and overall quality of life [1–4]. Diagnostic criteria define ADHD as persistent, significantly impairing inattention and/or hyperactivity across multiple settings, while ASD is characterized by impairments in social reciprocity skills accompanied by restricted, repetitive behaviors and interests [5]. However, this distinction is often blurred by substantial clinical overlap. The disorders frequently co-occur, and individuals with either condition exhibit heterogeneity in their symptoms [6–12], a complexity echoed by shared genetic risk factors [13–16] and converging atypicalities in large-scale brain networks [17–22].

Current diagnostic assessments for ASD and ADHD predominantly rely on clinical observation and informant reports, methods that are inherently subjective and resource-intensive but also struggle to reliably differentiate the two disorders or account for their comorbidity [23–25]. This reliance on manual, expert-driven analysis has traditionally limited research to smaller, often narrowly defined samples that do not capture the vast heterogeneity inherent in ADHD and ASD in more natural settings. This highlights the need for objective, quantifiable, and scalable markers derived from naturalistic behavior. Such markers may more clearly reveal the complex interplay of symptoms in ASD and ADHD as they manifest in real-world interactions.

Effective communication depends on both structural and pragmatic language skills. Structural language encompasses fundamental aspects of language production, such as vocabulary and syntax, and pragmatic language involves using these skills appropriately in social contexts to convey meaning, maintain topic coherence, and respond relevantly [26]. While deficits in social communication are a defining characteristic of ASD, they are not considered a core diagnostic feature of ADHD [5]. Despite this, studies using informant-based assessments suggest a significant degree of overlap, with both groups demonstrating difficulties in both pragmatic and structural language [27–29].

Another aspect that is central to successful communication in a social setting is perspective-taking, often described as Theory of Mind (ToM), i.e., the ability to infer the mental and emotional states of others [30], which are also essential for discourse and narrative comprehension and construction [31]. Difficulties in perspective-taking have long been associated with ASD [30], but the investigation of such abilities has progressed independently for ADHD and ASD. Natural language processing (NLP) as an objective tool has identified distinct linguistic patterns in smaller-scale ASD studies, such as repetitive speech [31], and atypical narratives [31,32], while separate studies have linked ADHD to narrative disorganization [33]. To date, studies using automated methods have typically used small sample sizes with fewer than fifty participants [34]. This limits the power to identify meaningful and replicable linguistic and socio-cognitive profiles that scale to larger cohorts in naturalistic settings. Additionally, conducting analyses of youth with ASD or ADHD separately has prevented the delineation of shared versus distinct domains. This study addresses these gaps by exploring both language (structural and pragmatic) and higher-order social-cognitive (memory, narrative comprehension and construction, perspective-taking) abilities to develop a detailed model of impairment in both ADHD and ASD.

Aside from difficulties in social communication, atypical prosody is a commonly reported clinical feature of ASD [32]. While early research using subjective ratings struggled to conceptually define the clinically observed "unusual" prosody in ASD [33], recent systematic reviews consistently identify higher mean pitch and greater pitch variability as features of ASD [34]. However, for other acoustic features like vocal intensity, speech rate, and voice quality, findings have remained inconsistent, likely due to the small sample sizes, diversity of methods, and clinical heterogeneity across studies [34,35]. Importantly, the potential influence of comorbid ADHD, which may also be associated with altered speech prosody [36,37], remains poorly understood. Although recent automated methods have shown promise for ASD severity identification or ASD diagnosis based on vocal patterns [38–40], their application to small, constrained samples leaves a gap in our ability to isolate a true, disorder-specific vocal signature for either condition.

Motor behavior offers another important domain for differentiating these disorders. Research in ASD has characterized a wide range of atypical motor behaviors [41,42]. In contrast, objective motor analysis in ADHD has remained comparatively sparse, focusing primarily on global activity levels via actigraphy [43]. Additionally, to date, no study has used the same objective movement measures to directly compare ADHD and ASD while simultaneously modeling their high rate of comorbidity.

Further complicating matters, language, speech, and motor coordination develop rapidly with age, and symptom presentations can shift [44]. For instance, hyperactivity in ADHD often decreases while inattention persists into adolescence [45]. Similarly, the co-occurrence of ASD and ADHD symptoms emerges in middle childhood and appears to peak during adolescence, a crucial and understudied period of social and cognitive change [44]. Without rigorously modeling the effect of age, along with other key sources of variability in ASD and ADHD, such as different prevalence in boys and girls [46,47], or wide ranges of cognitive abilities as measured by IQ [48,49], most behavioral markers may simply reflect demographic and cognitive differences or typical developmental trajectories rather than underlying psychopathology.

We address these gaps by leveraging artificial intelligence (AI) to extract objective markers of language, speech, and movement from audio and video recordings of conversations from the Healthy Brain Network cohort (HBN) [50]. This large community sample includes youths with a variety of mental health and neurodevelopmental conditions, including ADHD and ASD, along with neurotypical controls. This large-scale approach links computationally-derived behaviors to gold-standard, consensus-based clinical diagnoses, their differential presentation, and co-occurrence, overcoming the limitations of smaller studies that rely on subjective instruments. We test two questions central to the field: (1) Do ASD and ADHD show distinct or overlapping behavioral signatures? and (2) How are these signatures shaped by development, sex, and IQ? This work represents the first large-scale validation of digital phenotyping to deconstruct the behavioral heterogeneity of these two overlapping neurodevelopmental disorders.

Results

A large-scale developmental cohort for digital phenotyping

To measure naturalistic behaviors, 2,341 participants (aged 5-22) of the HBN cohort were video-recorded during a semi-structured interview after watching "The Present", an emotive animated short film. During the interview participants are asked identical questions (Extended Data Table S1 & Extended Data Fig. S1) and the adult interviewer follows up with encouraging prompts when necessary to elicit responses. This conversation serves as a social-emotional probe, eliciting responses across broad domains, including narrative production and emotional expression, while also probing social-cognitive functions relevant to ASD, such as narrative comprehension and perspective-taking (Theory of Mind) (see Extended Data Table S2 for an overview of related measures). We used state-of-the-art computational tools to derive objective, quantitative metrics from the video-recorded interviews, covering expressive language, vocal prosody, as well as facial and body movements (Fig. 1A).

The participant cohort was diagnostically rich and clinically complex (Fig. 1B, C & Table 1). It included 1,455 youths with clinician-confirmed ADHD DSM-5 diagnoses. Among these, 653 had

an inattentive presentation, 93 a hyperactive-impulsive presentation, and 625 met criteria for a combined presentation. A subset (N = 84) did not meet full diagnostic criteria (categorized in the DSM-5 as ADHD other/unspecified) and were excluded from analysis. The sample also included 349 participants with mild to moderate ASD (see Extended Data Fig. S2 for ASD-specific clinical instrument scores). Due to the nature of the conversational task, only verbal participants took part in the interview, thus limiting severe cases of ASD. Consistent with the clinical literature in child neurodevelopmental health, the cohort is characterized by significant diagnostic comorbidities, with 81% of individuals with ASD also having a comorbid ADHD diagnosis (Fig. 1B). The cohort also included 168 typically developing (TD) youths and 653 participants with other clinical diagnoses, most commonly anxiety and learning disorders (Fig. 1C) without ASD and ADHD. The sample exhibited an overall male-to-female ratio of approximately 2:1 (Fig. 1D), with lower than chance prevalence in females (17% of the ASD group and 23-27% of the ADHD group; Fig. 1E). Note that in Fig. 3E, and in all our analyses, ADHD-Combined presentation is coded as both Inattentive and Hyperactive. To distinguish this coding with two binary variables (Inattentive, Hyperactive) from the three DSM-5 ADHD presentations (Inattentive, Hyperactive, Combined), we will refer to these as diagnostic *status* and *presentation*, respectively.

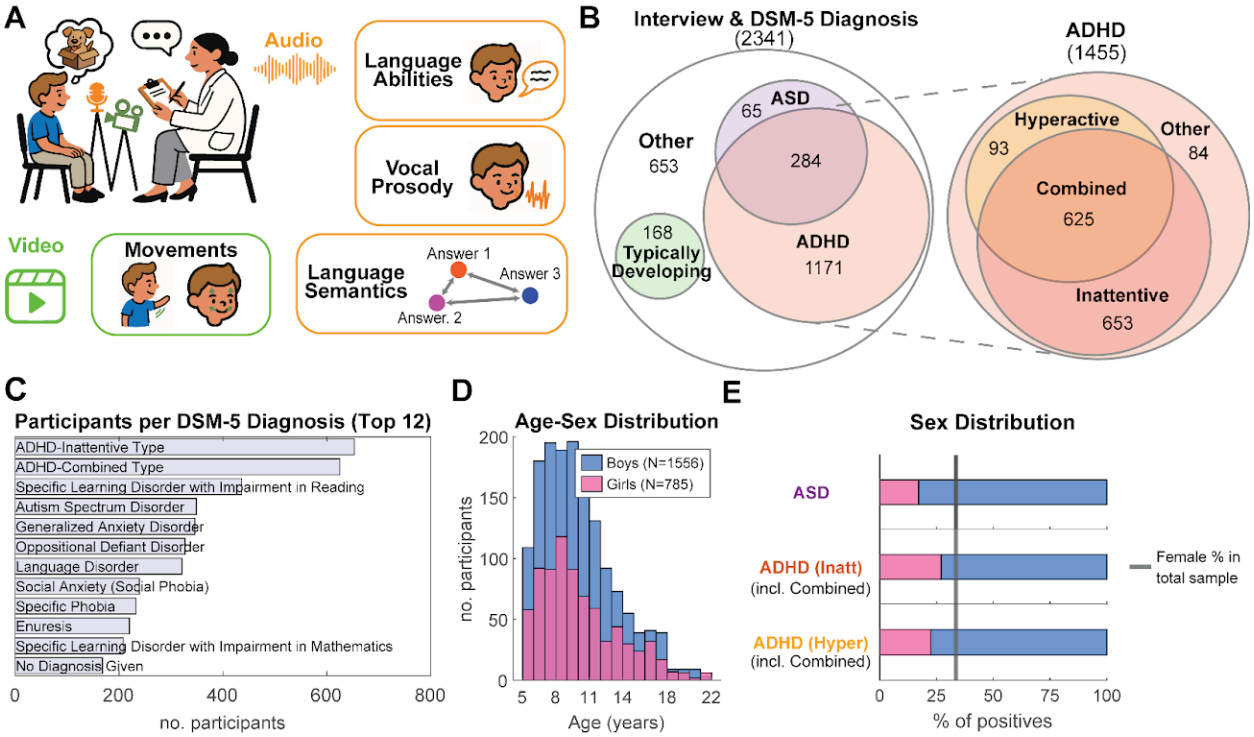


Figure 1. Overview of the study cohort. **A)** Using AI for objective digital phenotyping. Video recordings were processed to quantify body and facial movements, while audio was transcribed and analyzed to extract metrics of language structure, semantic content, and vocal prosody. **B)** Venn diagram showing the distribution and overlap of the primary diagnostic groups. **C)** The 12 most common clinician-assigned DSM-5 diagnoses in the sample. **D)** Age and sex distribution of the full cohort (N=2,341). **E)** Sex distribution within the children with ASD and/or ADHD diagnostic, ADHD-Combined are here considered positive in both hyperactive and inattentive diagnostic status. The black line indicates chance occurrence based on the sex ratio in the cohort.

Table 1. Demographic, clinical, and interview characteristics

Characteristic	ADHD-Inattentive (incl. Combined) (N=1278)	ADHD-Hyperactive (incl. Combined) (N=718)	ASD (N=349)	Typically Developing (N=168)	Full Sample (N=2341)
Age (in years), mean (SD)	10.1 (3.22)	9.1 (2.91)	10.3 (3.69)	9.4 (3.23)	10.1 (3.39)
Sex, N (%)					
Female	346 (27)	162 (23)	60 (17)	74 (44)	785 (34)
Male	932 (73)	556 (77)	289 (83)	94 (56)	1556 (66)
Co-morbidities, N (%)					
ADHD-Inattentive	N/A	625 (87)	264 (76)	0 (0)	1278 (55)
ADHD-Hyperactive	625 (49)	N/A	166 (48)	0 (0)	718 (31)
Autism Spectrum Disorder	264 (21)	166 (23)	N/A	0 (0)	349 (15)
Clinical Instruments, mean (SD)					
Inattention (SWAN)	1.2 (0.9)	1.1 (0.9)	1.1 (1.0)	-0.3 (1.0)	0.7 (1.1)
Hyperactivity (SWAN)	0.7 (1.0)	1.2 (0.8)	0.8 (0.9)	-0.3 (1.0)	0.4 (1.1)
Autism (ASSQ)	8.5 (9.2)	10.2 (10.1)	17.5 (10.9)	2.3 (3.1)	7.1 (8.6)
Depression (MFQ)	10.0 (8.8)	9.8 (8.0)	10.3 (8.8)	4.7 (4.9)	9.0 (8.6)
Anxiety (SCARED)	14.7 (12.1)	14.6 (12.0)	16.5 (12.9)	9.4 (8.0)	14.2 (11.8)
Externalizing (CBCL)	12.6 (9.8)	15.5 (9.8)	12.4 (9.5)	5.6 (5.7)	10.9 (9.5)
Internalizing (CBCL)	10.5 (8.3)	10.5 (8.0)	12.2 (8.3)	5.3 (5.0)	9.8 (8.2)
IQ, mean (SD)*					
Full-Scale	98.5 (16.3)	100.3 (16.2)	96.7 (18.9)	106.5 (14.3)	99.4 (16.6)
Verbal	103.0 (16.3)	104.1 (16.2)	100.8 (19.4)	108.8 (14.3)	103.2 (16.5)
Interview length (minutes), mean (SD)	4.5 (1.51)	4.5 (1.47)	4.7 (1.62)	4.5 (1.95)	4.5 (1.54)

* IQ was measured using WISC-V [51], with exceptions for early participants using WASI [52], and children under 6 or with IQ below 70 using KBIT [53] . See 'IQ Score' section in Methods for details.

Modeling commorbidity and clinical manifestations of ADHD and ASD.

This community-cohort recapitulates the well-known clinical overlap between ADHD and ASD [12]. A diagnostic status of either ADHD-Inattentive or ADHD-Hyperactive significantly increased the odds of a co-occurring ASD (Adjusted Odds Ratio (OR) = 2.3 and 1.6, respectively) (Fig. 2A). This overlap extends to clinical questionnaires, where trait measures of autism (indexed by the parent ratings of ASSQ [54]), inattention and hyperactivity (indexed by the parent rating of the SWAN-inattention and hyperactive scales [55], respectively) are not only correlated to each other, but also correlate with other clinical traits, such as anxiety (SCARED [56]), depression (MFQ [57]), as well as age, and sex (Fig. 2B). For instance, the hyperactivity severity decreases with age ($\rho(2194) = -0.28$, $p<0.001$, Fig. 2C), consistent with typical development [45]. The same is also reflected in the effect of age on the prevalence of ADHD-Hyperactive status (Fig. 1A, OR = 0.83). Given the well-established effects of development, this "common cause" (age) could induce spurious correlations with ADHD-Hyperactive status for a number of behavioral measures. Similarly, some established associations with ASD may be due to the co-occurring ADHD, or vice-versa.

To isolate the unique contribution of each diagnostic and demographic factor, we employed a multivariate regression model (Fig. 2D), where the clinician-confirmed diagnostic status is coded as binary variables (ASD, ADHD-Inattentive, ADHD-Hyperactive). Effect sizes (ES) for these binary diagnostic status variables will represent Cohen's d, and Student's $t*\sqrt{n}$ for continuous variables such as Age and IQ (see Methods). The utility of this approach is clearly illustrated by its application to *externalizing* symptoms, defining outward-directed behaviors such as aggression, rule-breaking, and impulsivity. While in this cohort ASD traits (ASSQ) moderately

correlate with externalizing problems (CBCL [58]) (ρ(2119) = 0.38, p<0.001, Fig. 2B), the regression model reveals that this is an artifact of comorbidity. After accounting for all factors simultaneously, externalizing behaviors were predicted by ADHD-Hyperactivity (Effect Size (ES)=0.3), but not ASD (Fig. 2E). On the other hand, we found internalizing problems (inward-directed difficulties like anxiety, depression, and withdrawal) were associated with older ages (ES=0.195), female sex (ES=0.081), and ASD diagnosis and/or ADHD-Inattentive status (ES=0.125 and 0.072, respectively) (Fig. 2E, F). This highlights a shared burden of emotional and behavioral dysregulation in both conditions [59] and demonstrates the model's power to parse complex clinical presentations into more specific, disorder-related patterns. Such a model is useful for measures like speech, language, and movement, which are also affected by development [60], and are the focus of this study.

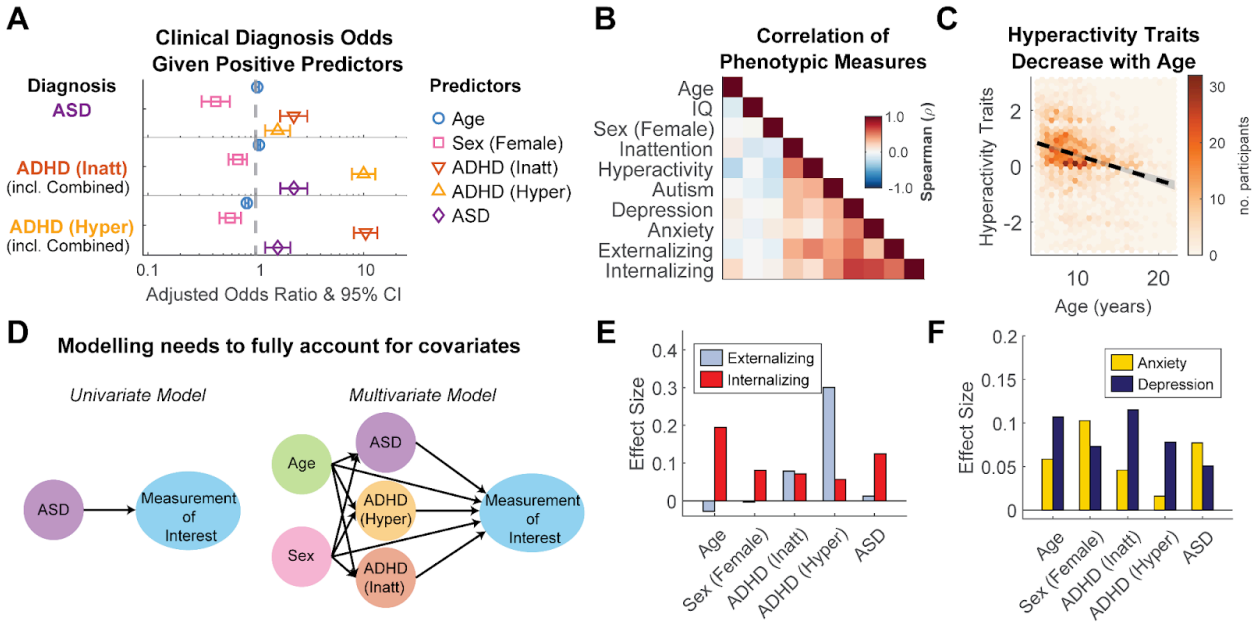


Figure 2. Clinical comorbidity and the necessity of multivariate modeling. A) Adjusted odds ratios showing the likelihood of a co-occurring positive diagnosis in individuals with positive predictors (or one-year change in the case of age). Females are less likely than males to receive an ADHD or ASD diagnosis (OR < 1). **B)** Correlation matrix illustrating how clinical instruments aimed at measuring specific neurodevelopmental disorder traits, like ASSQ for autism and SWAN for ADHD, are highly correlated with other measurements like depression (MFQ), anxiety (SCARED), and internalizing and externalizing issues (CBCL), but also with age, IQ, and sex, which makes it difficult to link observed behavior to each of the instrument scales independently. **C)** Two-dimensional histogram showing the negative correlation between age and hyperactivity traits (SWAN scale) (ρ(2194) = -0.28, p<0.001). **D)** Conceptual diagram of the multivariate regression model used to isolate the unique effects of each predictor. A "Univariate Model" ignores covariation, while the "Multivariate Model" controls for shared variance, with Age and Sex as a common source of variance. **E)** Effect sizes from the multivariate model predicting internalizing and externalizing behaviors (CBCL) and **F)** symptoms of anxiety (SCARED) and depression (MFQ). Effect sizes represent the unique contribution of each factor (Sex coded as Female = 1, Male = 0). Full names for clinical instruments (SWAN, ASSQ, MFQ, SCARED, CBCL) are provided in the 'Clinical Instruments' section in the Methods. Detailed model statistics for panels A, E, and F are available in Supplementary Materials Sections S1.1 and S1.2.

Developmental factors dominate over diagnostic effects in structural language

Structural and pragmatic language abilities are typically assessed using standardized psychometric tests, which place heavy demands on non-linguistic abilities like attention and impulse control, potentially confounding results [61]. Here, we used AI-driven tools to extract objective metrics of structural (amount of speech, speech rate, lexical diversity, filler-word usage, etc.) and pragmatic (conversation presence, question-answer coherence, self-referential focus) language abilities from naturalistic conversational responses.

Our analysis revealed that the structural language abilities measured here were mostly unaffected by either an ADHD or ASD diagnosis (Fig. 2E). Across a range of metrics, our multivariate model, which simultaneously accounts for age, sex, IQ, and diagnosis, consistently confirmed that development (age) was the primary driver of the observed linguistic abilities (Fig. 2E). This confound is illustrated by the apparent relationship between lexical diversity and hyperactivity: while a simple correlation suggested a significant negative association (Fig. 3B, $\rho(2160)=-0.16$, $p<0.001$), implying that hyperactivity impairs vocabulary use, our multivariate model (Fig. 3C-D) revealed this variation could be attributed to the stronger effects of age (ES=0.489) and IQ (ES=0.273).

This pattern of a dominant age effect and smaller effects otherwise held across all features we analyzed. For structural language metrics, in addition to age, IQ consistently emerged as a reliable predictor (Fig. 3E). Older participants spoke more (ES=0.398), faster (ES=0.563), and more coherently (ES=0.071). The only significant diagnostic marker was a subtle increase in self-referential speech (i.e., use of first-person pronouns) among participants with ASD (ES=0.1). We also observed secondary demographic effects, with females speaking faster (ES=0.078) and answering more coherently (ES=0.089) than males, and higher IQ and older ages being associated with more frequent use of filler words (ES=0.11 & 0.21, respectively), potentially reflecting the cognitive load of formulating more complex sentences.

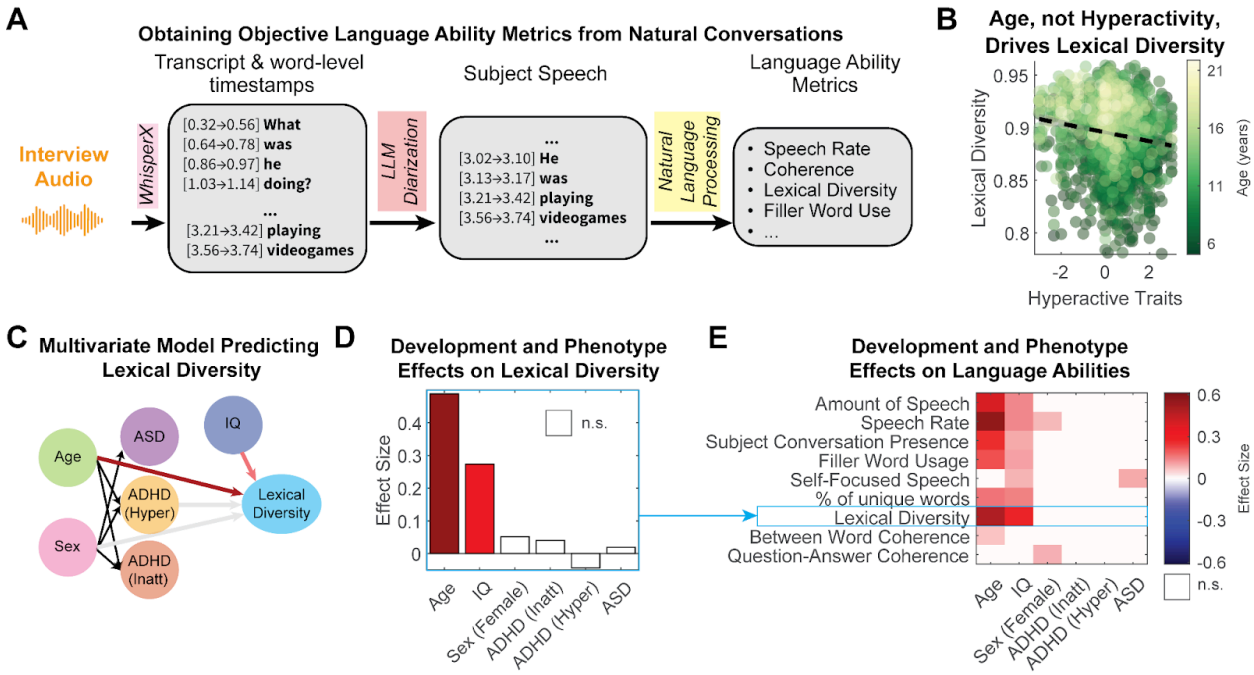


Figure 3. Language abilities are primarily driven by development, not by ADHD or ASD. A) The computational pipeline used to extract objective structural and pragmatic language metrics from transcribed audio. **B)** Hyperactive traits (SWAN) and lexical diversity are correlated ($\rho(2160)=-0.16$, $p<0.001$). However, the color gradient, representing age, reveals that this association is confounded by development. **C)** Diagram of the multivariate model used to disentangle the effects of diagnosis and developmental covariates on lexical diversity. Arrows are color-coded to the effect size magnitude in the multivariate regression model (panel D). Univariate associations between DSM-5 diagnoses and demographics with lexical diversity (grayed arrows) are instead better explained by age and IQ in the multivariate model. Black arrows indicate effects tested in Fig. 2A. (See Supplementary Materials Section S1.1) **D)** Effect sizes from the multivariate model. **E)** A summary matrix of effect sizes from the multivariate model applied to a range of language metrics. The results show the consistent and strong effects of age and IQ and the general lack of diagnostic effects across all measured aspects of language abilities. Significance cut-off at $p < 0.01$ (white otherwise; Bonferroni-corrected across the 9 models tested). Detailed multivariate model reports are available in Supplementary Materials Section S1.3.

ASD, but not ADHD, is associated with atypical narrative, story comprehension, and perspective taking

While our analysis revealed that the measured structural and pragmatic language abilities are primarily shaped by development rather than diagnosis, impairments in social communication often stem from difficulties with higher-order functions. These include factual memory, comprehending the central theme of the story, and inferring the mental and emotional states of others, a perspective-taking ability also known as Theory of Mind (ToM). The interview questions were designed to tap into this higher-order functions (see Extended Data Table S1 and Extended Data Table S2). To objectively quantify these functions at scale, we developed a computational pipeline to analyze the semantic content of participants' interview responses. This approach measured how similar each answer was to that expected from a typical adult (Fig. 4 A-C, assessed here from the older typically developing youths).

The multivariate analysis revealed a distinct pattern of impairment specific to ASD (Fig. 4E). After controlling for the strong associations with age (e.g. Fig. 4D), an ASD diagnosis was uniquely associated with producing less typical and detailed narratives (ES=-0.117), poorer comprehension of the story's central theme—that both the boy and the puppy in the story were similar because they were missing a leg— (ES=-0.097), and provided more idiosyncratic, or unusual, descriptions of their own emotions (ES=-0.104).

These results provide objective, quantitative evidence for challenges in narrative construction and comprension [62] ('Narrative Detail') and perspective taking [30] ('Thematic Understanding') that are a core feature of autism. In contrast, after accounting for all covariates and multiple-comparison correction, neither the ADHD-Inattentive nor the ADHD-Hyperactive status were significantly associated with any of these social-cognitive measures. Recall that ADHD-Combined was coded as both; therefore, this null result applies to the Combined presentation as well. This suggests that the communication difficulties sometimes attributed to ADHD [63] do not extend to these functions.

Age remained a powerful predictor across most semantic language use measures (Fig. 4E). Consistent with typical social-cognitive maturation [44], older participants provided more detailed narratives (ES=0.552), demonstrated a better understanding of the movie's main message

(ES=0.383), and showed more typical positive and negative preferences (ES=0.21 & 0.192). Similarly, higher IQ was associated with more typical narrative abilities (ES=0.223). We also observed a secondary finding related to sex, with females tending to produce more typical narratives (ES=0.103) and less divergent descriptions of their emotions (ES=0.119) and positive preferences (ES=0.137) compared to male participants (Fig. 4E). In none of our analyses did we see a marked interaction of these effects with diagnostic status. One exception was the age effect ASD on 'Narrative Detail' which tended to increase with age (Fig. 4D; see follow-up analysis under Statistical Methods).

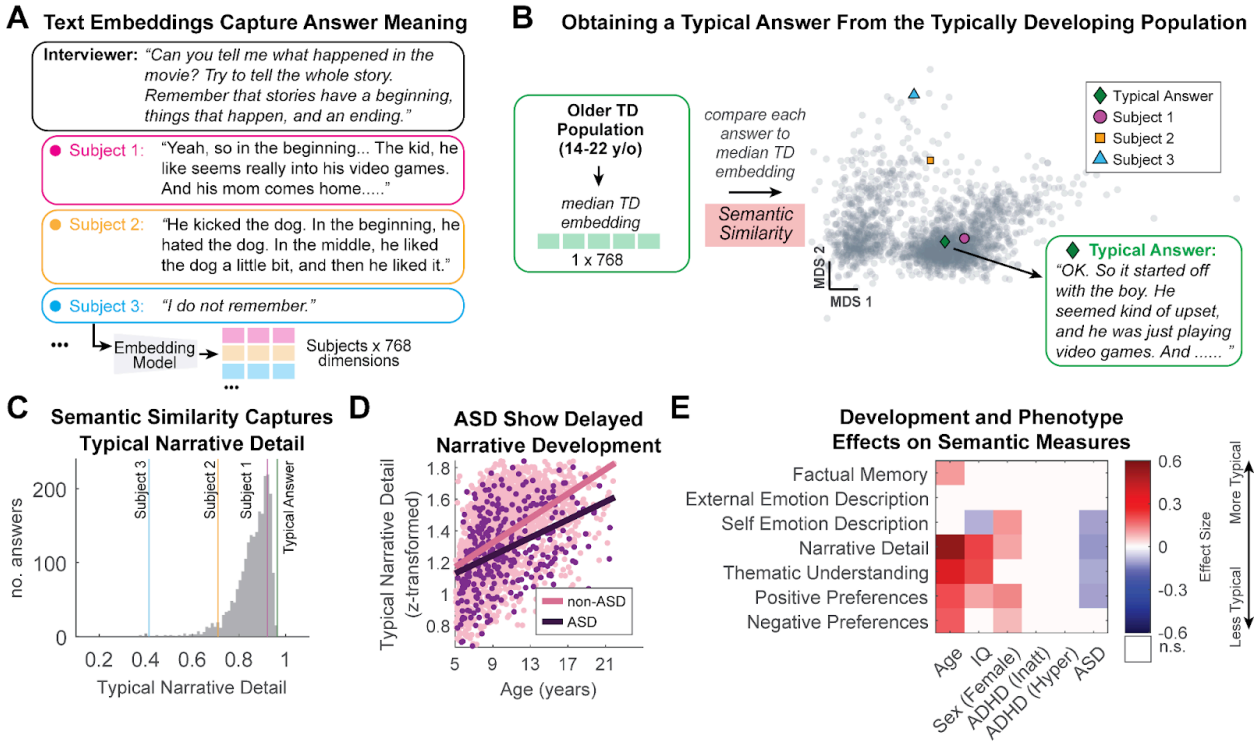


Figure 4. Semantic analysis reveals ASD-specific effects on narrative and social-cognitive abilities. A) The computational pipeline for semantic analysis. Participants' answers to 23 predefined interview questions were parsed and converted into numerical semantic representations (embeddings). **B)** A "typical answer" is obtained from the semantic similarity to the median embedding from the older TD participants (14-22 y/o). **C)** Semantic similarity to the TD prototype is used as a measure of 'typicality'. For the narrative recall question in panel A, higher similarity reflects a more detailed and typical recount of the narrative. Subject 1 provides a detailed answer (high similarity), whereas Subject 3 does not recall the story (low similarity). **D)** Joint distribution illustrating the positive association between age and the production of more typical narratives, but the diminished abilities in the ASD group compared to the non-ASD group. **E)** Matrix of effect sizes from the multivariate model, assessing the impact of diagnosis and demographic factors on various social-cognitive domains probed by the interview questions (e.g., factual memory, emotional description, and thematic understanding). Only effect sizes with $p<0.01$, after Bonferroni correction, are shown in color. See Extended Data Table S1 for a full list of questions and Supplementary Section S1.4 for detailed multivariate model reports.

Autism Spectrum Disorder has a distinct vocal profile

The acoustic patterns of pitch, intensity, and rhythm are critical for conveying the intent of an utterance, and atypical prosody is a widely reported clinical feature of ASD [32]. To this end, we employed an automated pipeline to extract a comprehensive set of acoustic features from the segmented speech of each child, including measures of pitch, intensity (loudness), and voice quality (Fig. 5A).

Our multivariate analysis revealed a distinct acoustic signature specific to ASD. After controlling for age, sex, and IQ, an ASD diagnosis was significantly associated with increased mean pitch (ES=0.079), greater pitch variability (ES=0.094), and higher vocal loudness (ES=0.083) (Fig. 5C). Furthermore, we identified significant markers related to voice quality: participants with ASD exhibited significantly higher dysphonia severity (ES=0.133, Fig. 5B), as measured by cepstral peak prominence (CPP) [64], and increased breathiness (ES=0.117), measured by the glottal-to-noise excitation ratio [65]. These quantitative findings align with clinical descriptions of an "unusual" voice in ASD [66]. Notably, this vocal profile was specific to ASD. Consistent with our findings for language content, narratives, and perspective taking, neither an ADHD-Inattentive nor an ADHD-Hyperactive status (which includes ADHD-Combined) was associated with any of the measured vocal prosody features. Developmental effects were also prominent, with age being a significant predictor of reduced pitch (ES=-0.419), pitch and loudness variability (ES=-0.415 & -0.14), and to a lesser degree breathiness (ES=-0.104), and dysphonia (ES=0.086). Together, these results identify an objective vocal signature for ASD that is independent of ADHD status and comorbidity and follows a clear developmental course.

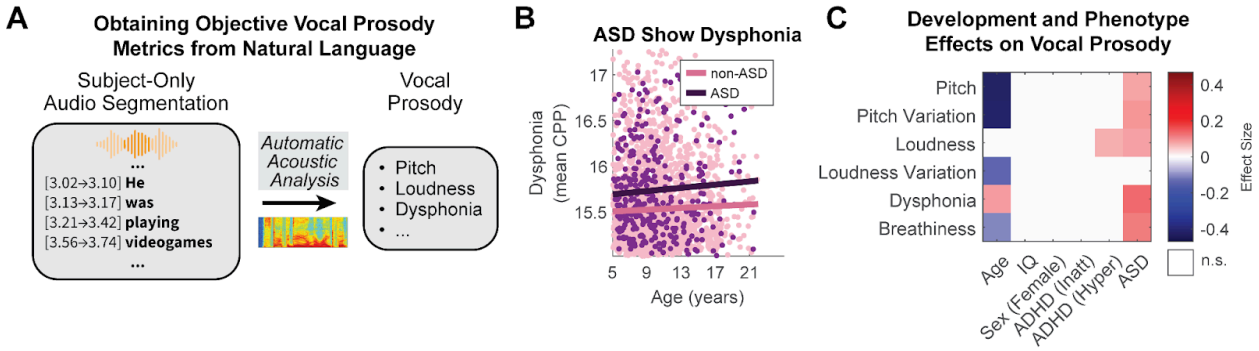


Figure 5. Automatic vocal prosody metrics reveal an ASD-specific acoustic profile. A) Automatic extraction of vocal prosody features involved transcribing and diarizing speech, thus isolating the children's speech from that of the adult interviewer, and extracting audio-based features using an open-source toolbox (OpenWillis [67]). **B)** ASD participants show increased dysphonia compared to non-ASD participants, which may be related to an "unusual" voice in ASD. **C)** Effect Size matrix illustrating the effects of development and ASD on vocal prosody. Effect Sizes in this matrix are estimated again using a multivariate regression model, with $p<0.01$, after Bonferroni correction. See Supplementary Materials Section S1.5 for detailed multivariate model reports.

Increased motor activity is uniquely associated with ADHD-Hyperactive diagnostic status

Motor behaviors are relevant diagnostic features of ASD and ADHD, with hyperactivity being a key diagnostic criterion in ADHD, and stereotypical and repetitive motor behaviors being among the defining criteria for ASD. Yet motor behaviors are rarely quantified objectively in naturalistic

settings. We applied computer vision models to the interview videos to extract objective metrics of upper-body and facial movement. Using pose and facial landmark estimation models, we computed the average frame-to-frame displacement for several regions of interest, including the upper body, head, eyes, and face (Fig. 6A).

Consistent with a typical developmental trajectory, age was a strong negative predictor of movement across all measures (ES ranging from -0.31 to -0.46). After accounting for these developmental effects, we identified a specific motor signature. ADHD-Hyperactive status was uniquely and robustly associated with increased movement across all measured body and facial regions (ES=0.10-0.14) (Fig. 6B-C). Conversely, neither ASD nor ADHD-Inattentive were linked to any significant changes in motor activity. We also observed a significant association between female sex and increased overall body movement (ES=0.105). These results demonstrate that while increased motor activity is a powerful, objective marker, it is specific to the ADHD-Hyperactive diagnosis and does not generalize to ASD or the Inattentive diagnosis in our population.

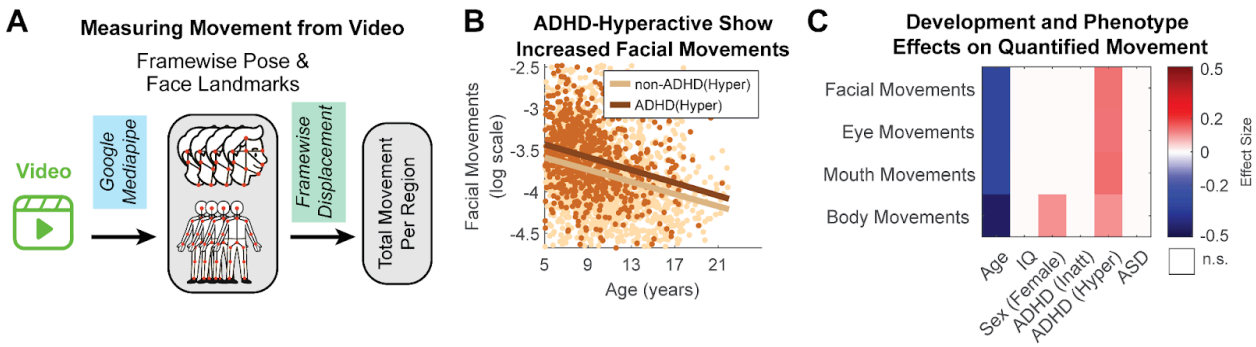


Figure 6. Associations of demographics and phenotype with body and face movements. A) We use Google's Mediapipe computer vision models to obtain framewise locations for anatomical and facial landmarks. We then calculate the average framewise displacement for the different regions of interest. **B)** ADHD-Hyperactive youth show increased facial movements across the whole developmental spectrum, compared to non-ADHD Hyperactive. **C)** Effect Size matrix illustrating the effects of development and an ADHD-Hyperactive status on all measures of face and body movements. Effect Sizes in this matrix are estimated using multivariate regression, with $p<0.01$, after Bonferroni correction. See Supplementary Materials Section S1.6 for detailed multivariate model reports.

Discussion

This study used digital phenotyping to objectively analyze communication and motor behavior across a wide developmental range in youths with ADHD and/or ASD. We used relatively short videos from a scripted interview to identify objective markers of these domains using artificial intelligence. By using a large, community-recruited cohort, this study captured the variability seen in real-world clinical settings.

The dominant driver of variance in our measures of language, speech, and motor behavior was age. Language deficits in ADHD were fully explained by the confounding effect of age and

co-occurring ASD. Conversely, ASD showed structurally intact language but significant impairments in narrative ability and perspective-taking, coupled with a unique vocal profile.

Previous research, often using standardized assessments, has often linked ADHD to language disturbances affecting both structural and pragmatic, or social, aspects of communication [63,68]. This view is supported by community-based studies using standardized language assessments, which have found that children with ADHD are nearly three times more likely to have co-occurring language problems than their typically developing peers, with prevalence rates of impairment estimated between 40-45% [69,70]. Our findings, derived from objective quantification of spontaneous conversation, challenge this view. After accounting for developmental factors such as age and IQ, we did not detect a consistent association between an ADHD diagnosis and structural or pragmatic language abilities, aligning with findings from other naturalistic studies [61]. In other words, the changing prevalence of ADHD with age and its association with IQ may have confounded previous language studies that do not take these factors into account. Additionally, the discrepancy between our findings and studies reporting high rates of language impairment may be explained by methodological differences. Standardized psycholinguistic tests often place significant demands on non-linguistic abilities such as sustained attention, impulse control, and working memory capacity [71]. Poor performance on these assessments by children with ADHD may therefore reflect these core executive function deficits rather than a primary limitation in language abilities [61].

The finding that apparent language deficits are better explained by executive dysfunction extends from basic structural language abilities to higher-order social cognition. Our investigation of narrative ability, story comprehension, and perspective-taking (ToM) found that ADHD was similarly unassociated with any of these domains, in opposition to some previous studies [72]. These social-cognitive functions were, like language abilities, strongly moderated by age and IQ, but were uniquely and significantly impaired in participants with ASD. This suggests that the pragmatic communication challenges in ADHD are, in a similar fashion, more likely linked to underlying executive function deficits, such as poor inhibition and planning, which impact the organization and delivery of speech, rather than a fundamental deficit in understanding language or the mental states of others [73,74]. Our results reinforce the view that impaired perspective taking (ToM) is a critical differentiator between ASD and ADHD [75] and suggest its presence may serve as a strong indicator of comorbid ASD in individuals with ADHD.

In contrast to these findings for ADHD, our results for ASD provide quantitative support for the classic conceptualization of the disorder, defined by the core challenges in social communication [5]. In this study, participants with ASD performed similarly to their peers on structural language measures like lexical diversity and coherence, with their language abilities following a typical developmental trajectory. The one discernible feature at this level was a greater tendency for self-referential speech, marked by an increased use of first-person pronouns. This subtle linguistic pattern may serve as an objective marker of the self-focused perspective often described in autism, and echoes findings of pronoun use abnormalities that suggest an asynchrony between linguistic and social development [76].

The most prominent communication challenges in our ASD cohort emerged clearly in these higher-order domains of narrative and social cognition. Our semantic analysis revealed that participants with ASD produced less "typical" narratives, suggesting a divergence in identifying and recounting the most socially salient aspects of the story. They also were less likely to grasp the film's central theme, pointing to a challenge in moving beyond literal events to understand the overarching message or 'gist' of the story [62]. Most notably, their idiosyncratic emotion descriptions provide quantitative evidence of the ToM-related challenges that are characteristic of the condition [77]. These results, obtained at an unprecedented scale and in a naturalistic, conversational context, both validate and extend foundational work that used computational natural language processing to objectively measure narrative skills in autism. Previous studies, which primarily employed Latent Semantic Analysis, had already demonstrated that individuals with ASD show diminished narrative quality in demanding recall tasks [78] and in response to emotionally evocative scenes [79]. Our findings confirm that these core deficits are promising candidate digital markers for ASD.

This study further provides a quantitative basis for the long-observed atypical prosody in ASD, which, despite being clinically apparent, has been difficult to define with objective, replicable metrics [35]. Our large-scale analysis robustly confirmed the findings of increased pitch and pitch variability [34], but also clarified these prior inconsistencies by demonstrating a significant and consistent increase in vocal intensity (loudness) in our ASD cohort. Furthermore, we identified a distinct acoustic profile characterized by increased dysphonia and breathiness as markers of ASD, providing objective validation for long-standing clinical descriptions of autistic voices as "hoarse" or "creaky" [66]. Most importantly, these associations were absent in ADHD, revealing yet another distinct marker in both disorders.

Complementing the analysis of communication, our objective quantification of motor behavior revealed a particular signature that further differentiates the clinical presentations of ADHD and ASD. Motor difficulties are a well-documented feature of ADHD, but are typically measured with subjective rating scales or specialized sensors like actigraphy [43]. Our study demonstrates the feasibility of using unobtrusive, scalable computer vision methods to quantify motor activity in a naturalistic conversational setting, a technique with significant potential for both clinical and in-home monitoring [80]. We found that, consistent with typical development, overall movement decreased with age. After accounting for this maturational trend, increased movement of the head, body, and face was a specific marker of an ADHD-Hyperactive diagnostic status.

Conversely, our analysis did not find an association between increased movement and diagnoses of either ADHD-Inattentive or ASD. This null finding reinforces the specificity of our measure to the hyperactive-impulsive domain of ADHD. The absence of a detectable general motor signature for ASD in the present work, however, warrants careful interpretation. Motor challenges are increasingly recognized as a core feature of ASD [42]. It is likely that the global displacement metrics used in our study, while effective at capturing the overt restlessness of hyperactivity, are not sensitive to the repetitive motor patterns characteristic of autism [41]. The moderate levels of autism in our sample may also mean that stereotypical motor mannerisms

were less apparent in this context. Future work could employ more granular computer vision techniques, such as those that classify specific gesture types or postural patterns [81–85], to better characterize the nuanced motor profile of ASD [41,86]. Nevertheless, our findings establish objective, video-derived hyperactivity as a robust and specific digital biomarker for the ADHD-Hyperactive diagnostic status.

Finally, our results consistently highlight the influence of age and sex across all measured behavioral domains. Age was a powerful predictor of maturation, driving changes in everything from fundamental motor control and vocal pitch to the complexity of socio-cognitive and linguistic abilities. Similarly, distinct sex-specific profiles emerged in vocal patterns, language, and body movement. Developmental trajectories and demographic factors are not merely noise to be controlled for, but essential elements of the behavioral phenotype. Any search for valid, disorder-specific digital biomarkers must therefore be built upon developmental models that can successfully disentangle the effects of pathology from the natural course of maturation and sex-based differences.

Beyond these demographic and developmental factors, our models also included general cognitive ability by using Full-Scale IQ as a covariate. We acknowledge that IQ is not a pure independent measure of aptitude but partially an outcome that is affected by the disorder itself [87]. Statistically, it can behave similarly to sex, which is associated with the disorder due to varying prevalence. In the present cohort, there is a mean drop in IQ of approximately 10 points from typically developing kids with a positive diagnosis of ASD or ADHD (Table 1). However, IQ also varies substantially within each diagnostic status, with a variance that is no different than in typically developing youths (Table 1). In particular, it is widely recognized that cognitive ability can vary widely within ASD [48,49]. It is therefore expected to serve as an independent predictor, for instance, of language abilities in ASD [32,88]. Nevertheless, to ensure the robustness of our core findings, we repeated our primary analyses without including IQ as a covariate. The significant associations between diagnostic status and our key behavioral markers for ASD and ADHD remained consistent (see Extended Data Fig. S3). Similarly, the results did not change if using verbal IQ scores as a covariate instead.

Current clinical practice includes the use of established clinical instruments that measure behavioral presentations. As one may expect, our measures of movement, language, and speech correlate strongly with these established metrics (see Extended Data Fig. S4). However, it appears that our automated features are at times more specific or more sensitive to the ADHD and ASD presentations. However, we want to clarify that we are not suggesting these automated behavioral features should replace current diagnostic tools. Traditional diagnostic clinical instruments appear to be better matched to the clinical categories as currently defined in the DSM-5 (see Extended Data Fig. S5).

Several limitations should be noted. Our sample primarily consisted of mild to moderate ASD, which might explain the absence of behavioral stereotypes more commonly observed in more severe cases. Although our models accounted for the significant overlap between ADHD and ASD, other co-occurring conditions could have influenced the results. However, we ruled out the

impact of child anxiety on our behavioral measures by repeating all analyses while modeling for co-occurring positive anxiety diagnoses, and the associations with ADHD or ASD remained consistent (Extended Data Fig. S6). To unambiguously assign variance to the Hyperactive and Inattentive status, we chose not to code for Combined presentation, as in some previous works [89–91], however, when coding in the regression analysis instead for Inattentive and Combined presentation, the results are identical (Extended Data Fig S7). Furthermore, as a cross-sectional study, we could identify developmental trends across different ages but could not track changes within a single child over time or establish causality. Future work should validate our brief, video-based measures against long-term behavioral data, such as from wearable sensors, to ensure they accurately capture a child's typical behavior. Applying these computational tools in other critical settings, including parent-child interactions and standardized assessments like the ADOS, will also be essential to confirm their broader clinical utility.

In summary, our multivariate analyses, accounting for developmental effects, identified distinct behavioral tendencies that differentiate ADHD and ASD at the group level. ADHD, particularly the hyperactive-impulsive and combined presentations, were characterized by increased motor activity, while language and narrative difficulties did not emerge as robust markers once development and comorbidity were considered. Conversely, ASD exhibited a unique social-communication signature, with mostly typical structural language but notable impairments in narrative skills and perspective taking, along with a distinctive vocal profile characterized by higher pitch, greater intensity, and altered voice quality. These results demonstrate that digital phenotyping can provide objective, scalable, and developmentally sensitive markers to help differentiate overlapping neurodevelopmental disorders.

Methods

Clinical Population

This study used video-recorded conversations with 2,341 children (aged 5-22, M=10.1, SD = 3.39 years) as part of the Healthy Brain Network (HBN) [50]. The cohort included 1,455 participants with clinician-confirmed ADHD: 653 with inattentive presentation, 93 with hyperactive-impulsive presentation, and 625 with combined presentation. An additional 84 showed ADHD traits without meeting diagnostic criteria for a specific ADHD presentation (DSM-5 ADHD-other/unspecified). The sample also comprised 349 individuals with mild to moderate ASD, with 81% of ASD cases also having a comorbid ADHD diagnosis of some type. The cohort included 168 typically developing individuals and 653 with other clinical diagnoses, mostly anxiety and learning disorders. The overall male-to-female ratio was 1.98:1 (1556/785). Detailed demographic and clinical characteristics per diagnosis are available in Table 1.

Of the 2,341 participants, 1143 (48.8%) were white/caucassian, 311 (13.3%) black/african american, 226 hispanic (9.7%), 55 (2.3%) asian, and 42 (1.8%) identified as other races. 382 (16.3%) identified as two or more races. 182 (7.8%) participants did not specify, or their race was unknown.

Behavioral Task and Data Acquisition

Participants were recorded during a semi-structured interview conducted in a controlled laboratory setting. As part of the HBN protocol, participants watched "The Present," a short, emotive animated film, and were then prompted by a clinician to narrate the story and answer perspective-taking questions in the form of a scripted interview. High-definition video of the participant's face and upper body was captured using a Canon XC15 camcorder, while high-fidelity audio was recorded simultaneously with a Røde NT1 microphone. This scripted procedure provided the raw data for subsequent automated analysis of motor, prosodic, and linguistic features.

Diagnostic Procedure

The HBN is a large-scale biobank that utilizes a community-referred recruitment model to capture the wide heterogeneity inherent in developmental psychopathology. As such, the HBN sample is enriched for individuals meeting DSM-5 criteria, though cases are not typically severe. Diagnoses within the HBN are determined by licensed clinicians at the Child Mind Institute based on a consensus model that integrates multiple sources of information. The primary diagnostic instrument is the semi-structured Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) [92] for DSM-5, which is administered to both participants and their parents. The final consensus diagnosis also incorporates behavioral observations during testing, parental and self-reported clinical questionnaires, and the participant's developmental and educational history. For ASD, the primary evaluation relies on the KSADS autism module and questionnaires (e.g., SRS-2, SCQ, ASSQ), with only a limited subset of participants receiving supplemental "gold-standard" assessments like the ADOS or ADI-R.

For our analysis, clinician-confirmed diagnoses were modeled to assess the impact of specific disorders. ADHD diagnoses were categorized as either primarily inattentive or hyperactive/impulsive diagnoses. Participants with a clinician-confirmed diagnosis of ADHD-Combined presentation were considered positive for both hyperactive and inattentive status. ASD diagnoses were modeled as a single binary variable (1 for positive, 0 for negative). The vast majority of these individuals (>95%) were mild to moderate (only a few requiring maximum support), consistent with the HBN's inclusion criteria, which require participants to be verbal (see Extended Data Fig. S2 for distribution of autism clinical instrument scores). The control group, designated as the Typically Developing Population, consisted of individuals with a confirmed negative diagnostic clinician decision in all DSM-5 categories.

Clinical Instruments

Participants in the HBN are deeply phenotyped using a comprehensive battery of clinical instruments. For this study, we defined ADHD traits using the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale (SWAN) [55], and ASD traits using the Autism Spectrum Screening Questionnaire (ASSQ) [54]. These scales were selected as they showed the strongest unique association with their respective clinician-confirmed diagnoses compared to other available scales (see Extended Data Fig. S5). Other clinical instruments used in this study include the Mood and Feelings Questionnaire (MFQ, depressive traits) [57], Screen for Child

Anxiety Related Disorders (SCARED-P, anxiety traits) [56], and the Child Behavior Checklist (CBCL) [58], with two factors: internalizing and externalizing traits. A full description of age, sex, IQ, and clinical instrument scale scores for each diagnostic group is provided in Table 1.

IQ score

Cognitive ability was assessed using a composite IQ score. Full-Scale IQ scores were primarily measured using the Wechsler Intelligence Scale for Children (WISC-V) [51] (N=1984, mean=99.2, SD=16.7). Exceptions included early participants who were administered the Wechsler Abbreviated Scale of Intelligence (WASI) [52] (N=36, mean=97.8, SD=14.7), and children under age 6 or with known IQ below 70, for whom the Kaufman Brief Intelligence Test (KBIT) [53] was used (N=166, mean=100.4, SD=15.8). Full-Scale IQ Composite scores were available for 2186 out of the 2341 participants (93.3%). Similarly, verbal IQ scores were obtained by combining the WISC-VCI (N=1986, mean=103.6 SD=16.6), WASI-VCI (N=36, mean=99.3 SD=14.4), and KBIT-verbal (N=166, mean=100.6 SD=15.9) scores. Verbal IQ scores were available for 2188/2341 participants (93.5%). Main analysis and results were computed using Full-Scale IQ as covariate. However, main associations between diagnoses and measured behavior were independent of including IQ as covariate or not (Extended Data Fig. S3). Similarly, the results did not change significantly if using Verbal IQ scores instead.

Transcription, Diarization, & Question-Answer extraction

The audio recordings were first transcribed into text using the WhisperX model (v.3.3.0, Whisper model *large-v2*) [93] in a local workstation, which generated a raw transcript with timestamps for each word. Following transcription, speaker diarization was performed to separate the raw text into words spoken by the interviewer and the participant. This was achieved using a custom prompt with Google's large language model (*gemini-2.0-flash*). The diarized transcript was then used for a final extraction step, where a separate LLM prompt isolated the specific questions from the protocol and the full corresponding answers provided by the participant. A list of the scripted questions, as well as the "typical answer", i.e. the answer semantically closest to the median embedding of the older typical population (14-22 y/o), for each question, the typical answer is given in Extended Data Table S1 (see 'Semantic Analysis' method section for details). The complete prompts for LLM-based diarization and question-answer extraction are available in the Supplementary Materials Sections S2 and S3.

Description of Behavioral Variables

Various open-source tools were used to obtain objective measures of language, speech prosody, semantic content, and movement from the interview data. A summary of the measures is provided in Extended Data Table S2.

Language and Speech Prosody Measures. We extracted a set of acoustic and linguistic features using the Openwillis toolkit (v.3.0.5) [67]. Openwillis is a Python wrapper developed by Brooklyn Health that automates feature extraction by integrating established open-source speech and natural language processing tools for digital phenotyping, and can operate locally.

Semantic Measures. To characterize the semantic content of participants' answers, we utilized Google's 'Gecko' text embedding model (*text-embedding-004*) to generate 768 dimensional

numerical representations of each response [94]. First, we established a normative semantic baseline by calculating the median embedding for each of the 23 questions across all answers from the old participants in the typically developing (TD) population (14-22 y/o). Then, for each participant's answer, we computed its cosine similarity to the corresponding TD median embedding. This yielded a score representing how semantically ‘typical’ the response was. Participants with missing answers were assigned a missing score for that question. The interpretation of this score depends on the question's content; for narrative-recount questions, higher similarity reflects greater "typical narrative detail," whereas for questions about emotion, lower similarity may indicate more nuanced or idiosyncratic descriptions. Finally, these 23 individual typicality scores were averaged into conceptually meaningful domains based on the function of the questions. These domains included narrative detail, factual memory, thematic and emotional understanding, and perspective taking (e.g., describing one's own or a character's emotions, or the similarity between the dog and the kid and how this relates to the kid’s behavior). For example, the 'self-emotion description' score was the average semantic similarity score across answers to questions 14, 17, 20, and 23 (see Extended Data Table S1 for details on these questions and examples of the ‘Typical Answers’). The Fisher-Z transform (inverse hyperbolic tangent, artanh) was applied to the semantic (cosine) similarity measures prior to regression analysis to approximate normal distributions.

Movement measures. We extracted anatomical and facial landmarks from video recordings at the native 30Hz sampling rate of the video recordings using Google’s Mediapipe Holistic model (v.0.10.11) in a local machine. The raw 3D landmark coordinates were filtered to remove low-confidence data based on two criteria applied over a sliding window. First, segments with more than 10% missing values in a 2-second window were excluded. Second, segments exhibiting excessive jitter, defined as a standard deviation of the x-coordinate exceeding empirically-defined thresholds (0.1 for face; 0.01 for pose and landmarks) in 10-frame windows, were also excluded. A landmark point was considered valid for analysis only if it passed both of these checks. To ensure comparability across participants, the filtered data were normalized. Facial landmarks were aligned to a canonical model using an affine transformation [95] to correct for head rotation, translation, and scale. Body landmarks were normalized by the frame-by-frame inter-shoulder distance to render movements invariant to body size and participant's distance from the camera. Finally, movement was quantified as the frame-to-frame 3D Euclidean displacement for each landmark. These displacement values were then averaged within predefined anatomical regions (e.g., face, mouth, eyes, and body) and across all valid frames to yield a single mean movement score per region for each participant. Log of the average movement was used to model the effects of diagnosis on such measures.

Analytical Approach & Statistical Modeling

To test our primary research questions, we modeled the relationship between the measurements of interest and clinical diagnoses using multivariate linear regression. The model isolated the unique contribution of ASD and ADHD diagnostic status while controlling for key developmental and cognitive confounders. For each behavioral metric, the statistical model took the following form:

$$\text{Metric} \sim \text{Age} + \text{IQ} + \text{Sex} + \text{ADHD-Inatt} + \text{ADHD-Hyper} + \text{ASD} \quad (1)$$

with one model for each dependent variable (e.g., speech rate, lexical diversity), and the predictors were age (in years), a composite IQ score (Full-Scale), Sex (coded binary female=1, male=0), and binary variables for clinician-confirmed diagnoses of ASD, ADHD-Inattentive, and ADHD-Hyperactive (ADHD-Combined presentation was modeled as positive in both).
Every participant counts as n=1, with a set of numbers describing them: Age, IQ, ASD, ADHD-Hyper, ADHD-Inatt. For instance, an 8 year old boy with an IQ of 90 who presents with ASD and ADHD-Combined is coded as: 8, 90, 0, 1, 1, 1. Models were fitted using the bisquare 'robust' linear fitting option in MATLAB (2024b). To quantify and compare the effect of each predictor, we estimated its effect size from the t-statistic and error degrees of freedom DFE:

$$ES = \text{t-stat} / \text{sqrt}(DFE) \tag{2}$$

This normalization allows for the comparison of effects across different variables, even when the number of observations varies due to missing data for a specific behavioral metric. This measure is the traditional Cohen's d, in the case of binary variables (Sex and Disorders), and a sensible measure of effect size (mean over standard deviation of the beta coefficient) in the case of regression with continuous variables (Age and IQ). Statistical tables for each model of the behavioral outcome measures, including all beta coefficients, standard errors, t-statistics, effect sizes, and p-values, are provided in Supplementary Materials Sections S1.3-S1.6.

For continuous outcome measures such as the trait measures (SCARED, MFQ, CBCL-Internalizing, CBCL-Externalizing), we use the same model as with the behavioral outcome measures above (Fig. 2E-F, Supplement Materials Section S1.2).

For binary outcome measures (clinical diagnosis of ASD, ADHD-inattentive, and ADHD-hyperactive), we use a logistic binomial model and report odds ratios (OR) as the exponents of the beta coefficients (Fig. 2A and black arrows in Fig. 3C, Supplementary Materials Section S1.1).

$$OR = \exp(\beta). \tag{3}$$

Note that the linear model (1) does not include interaction terms. For age in particular, this means that the effects of the disorders are modeled as a constant gap between a positive and a negative diagnosis. However, for some behavioral features, the gap appears to increase with age (e.g., Fig. 4D, 5B). Such an overall gap can be detected by model (1) even if it increases with age. Given the reduced statistical power of larger models, we have therefore avoided interaction terms. Nevertheless, we performed a post-hoc analysis including an Age*Disorder interaction whenever there was a significant finding of age and ASD, ADHD-Hyperactivity, or ADHD-Inattention diagnostic status. This follow-up analysis suggests that there is an increasing gap with age for narrative detail in ASD (Fig. 4D, p=0.014 uncorrected), and for body movement in ADHD-Hyperactivity (Fig. 6C, p=0.044 uncorrected). Conversely, differences appear to diminish with age for pitch, pitch variation, and breathiness in ASD (Fig. 5C, p=0.004, 0.022 & 0.029, uncorrected).

Ethics Statement

The data used in this study were collected between 2016 and 2022. The HBN protocol was approved by the Chesapeake Institutional Review Board. Written informed consent was obtained from participants aged 18 or older, and from legal guardians for participants younger than 18. Access to the data was governed by a Data Transfer and Use Agreement (dated 6/27/2024) between the Child Mind Institute and The City University of New York.

Data Availability

Phenotypical, diagnostic, and interview data are available upon request and with the establishment of a Data Use Agreement with the Child Mind Institute at **https://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/.**

Code Availability

The full code pipeline used for speech transcription and diarization, as well as to extract language, vocal prosody, semantic embeddings, and body and face landmarks, is publicly available at **https://github.com/asortubay/AI_behavioral_analysis**.

Acknowledgments & Funding

NIH-NIMH P50 MH109429, CA Department of Health, grant to Nathan Klein Institute, Milham (PI).

References

1. Kogan, M. D. *et al.* Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007. *Pediatrics* **124**, 1395–1403 (2009).
2. Polanczyk, G., De Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am. J. Psychiatry* **164**, 942–948 (2007).
3. Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C. & Rohde, L. A. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *Int. J. Epidemiol.* **43**, 434–442 (2014).
4. Yang, Y. *et al.* Prevalence of neurodevelopmental disorders among US children and adolescents in 2019 and 2020. *Front. Psychol.* **13**, 997648 (2022).

5. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. (American psychiatric association, 2013).
6. Grzadzinski, R. *et al.* Examining Autistic Traits in Children with ADHD: Does the Autism Spectrum Extend to ADHD? *J. Autism Dev. Disord.* **41**, 1178–1191 (2011).
7. Hattori, J. *et al.* Are pervasive developmental disorders and attention-deficit/hyperactivity disorder distinct disorders? *Brain Dev.* **28**, 371–374 (2006).
8. Hollingdale, J., Woodhouse, E., Young, S., Fridman, A. & Mandy, W. Autistic spectrum disorder symptoms in children and adolescents with attention-deficit/hyperactivity disorder: a meta-analytical review. *Psychol. Med.* **50**, 2240–2253 (2020).
9. Holtmann, M., Bölte, S. & Poustka, F. Attention deficit hyperactivity disorder symptoms in pervasive developmental disorders: association with autistic behavior domains and coexisting psychopathology. *Psychopathology* **40**, 172–177 (2007).
10. Nijmeijer, J. S. *et al.* Attention-deficit/hyperactivity disorder and social dysfunctioning. *Clin. Psychol. Rev.* **28**, 692–708 (2008).
11. Reiersen, A. M., Constantino, J. N., Volk, H. E. & Todd, R. D. Autistic traits in a population-based ADHD twin sample. *J. Child Psychol. Psychiatry* **48**, 464–472 (2007).
12. Simonoff, E. *et al.* Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *J. Am. Acad. Child Adolesc. Psychiatry* **47**, 921–929 (2008).
13. Mulligan, A. *et al.* Autism symptoms in Attention-Deficit/Hyperactivity Disorder: A Familial trait which Correlates with Conduct, Oppositional Defiant, Language and Motor Disorders. *J. Autism Dev. Disord.* **39**, 197–209 (2009).
14. Reiersen, A. M., Constantino, J. N., Grimmer, M., Martin, N. G. & Todd, R. D. Evidence for shared genetic influences on self-reported ADHD and autistic symptoms in young adult Australian twins. *Twin Res. Hum. Genet.* **11**, 579–585 (2008).
15. Rommelse, N. N., Franke, B., Geurts, H. M., Hartman, C. A. & Buitelaar, J. K. Shared

heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *Eur. Child Adolesc. Psychiatry* **19**, 281–295 (2010).

16. Ronald, A., Simonoff, E., Kuntsi, J., Asherson, P. & Plomin, R. Evidence for overlapping genetic influences on autistic and ADHD behaviours in a community twin sample. *J. Child Psychol. Psychiatry* **49**, 535–542 (2008).
17. Aoki, Y. *et al.* Association of White Matter Structure With Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder. *JAMA Psychiatry* **74**, 1120–1128 (2017).
18. Brieber, S. *et al.* Structural brain abnormalities in adolescents with autism spectrum disorder and patients with attention deficit/hyperactivity disorder. *J. Child Psychol. Psychiatry* **48**, 1251–1258 (2007).
19. Christakou, A. *et al.* Disorder-specific functional abnormalities during sustained attention in youth with attention deficit hyperactivity disorder (ADHD) and with autism. *Mol. Psychiatry* **18**, 236–244 (2013).
20. Di Martino, A. *et al.* Shared and Distinct Intrinsic Functional Network Centrality in Autism and Attention-Deficit/Hyperactivity Disorder. *Biol. Psychiatry* **74**, 623–632 (2013).
21. Bedford, S. A. *et al.* Brain-Charting Autism and Attention-Deficit/Hyperactivity Disorder Reveals Distinct and Overlapping Neurobiology. *Biol. Psychiatry* **97**, 517–530 (2025).
22. Norman, L. J. *et al.* Cross-sectional mega-analysis of resting-state alterations associated with autism and attention-deficit/hyperactivity disorder in children and adolescents. *Nat. Ment. Health* **3**, 709–723 (2025).
23. Grzadzinski, R., Dick, C., Lord, C. & Bishop, S. Parent-reported and clinician-observed autism spectrum disorder (ASD) symptoms in children with attention deficit/hyperactivity disorder (ADHD): implications for practice under DSM-5. *Mol. Autism* **7**, 7 (2016).
24. Hayashi, W. *et al.* ASD symptoms in adults with ADHD: a preliminary study using the ADOS-2. *Eur. Arch. Psychiatry Clin. Neurosci.* **272**, 217–232 (2022).
25. Rosen, T. E., Mazefsky, C. A., Vasa, R. A. & Lerner, M. D. Co-occurring psychiatric

conditions in autism spectrum disorder. *Int. Rev. Psychiatry* **30**, 40–61 (2018).

26. Geurts, H. M. & Embrechts, M. Language Profiles in ASD, SLI, and ADHD. *J. Autism Dev. Disord.* **38**, 1931–1943 (2008).
27. Bishop, D. V. & Baird, G. Parent and teacher report of pragmatic aspects of communication: use of the Children's Communication Checklist in a clinical setting. *Dev. Med. Child Neurol.* **43**, 809–818 (2001).
28. Geurts, H. M. *et al.* Can the Children's Communication Checklist differentiate between children with autism, children with ADHD, and normal controls? *J. Child Psychol. Psychiatry* **45**, 1437–1453 (2004).
29. Geurts, H. M. & Embrechts, M. Language profiles in ASD, SLI, and ADHD. *J. Autism Dev. Disord.* **38**, 1931–1943 (2008).
30. Baron-Cohen, S., Leslie, A. M. & Frith, U. Does the autistic child have a "theory of mind" ? *Cognition* **21**, 37–46 (1985).
31. Kim, Y.-S. G., Dore, R., Cho, M., Golinkoff, R. & Amendum, S. J. Theory of mind, mental state talk, and discourse comprehension: Theory of mind process is more important for narrative comprehension than for informational text comprehension. *J. Exp. Child Psychol.* **209**, 105181 (2021).
32. Tager-Flusberg, H., Paul, R. & Lord, C. Language and communication in autism. *Handb. Autism Pervasive Dev. Disord.* **1**, 335–364 (2005).
33. McCann, J. & Peppé, S. Prosody in autism spectrum disorders: a critical review. *Int. J. Lang. Commun. Disord.* **38**, 325–350 (2003).
34. Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M. & Gaigg, S. B. "Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis". *Autism Res.* **10**, 384–407 (2017).
35. Trayvick, J. *et al.* Speech and language patterns in autism: Towards natural language processing as a research and clinical tool. *Psychiatry Res.* **340**, 116109 (2024).

36. Hamdan, A.-L. *et al.* Vocal Characteristics in Children With Attention Deficit Hyperactivity Disorder. *J. Voice* **23**, 190–194 (2009).
37. Barona-Lleo, L. & Fernandez, S. Hyperfunctional Voice Disorder in Children With Attention Deficit Hyperactivity Disorder (ADHD). A Phenotypic Characteristic? *J. Voice* **30**, 114–119 (2016).
38. Briend, F. *et al.* Voice acoustics allow classifying autism spectrum disorder with high accuracy. *Transl. Psychiatry* **13**, 250 (2023).
39. Eni, M. *et al.* Reliably quantifying the severity of social symptoms in children with autism using ASDSpeech. *Transl. Psychiatry* **15**, 14 (2025).
40. Godel, M. *et al.* Prosodic signatures of ASD severity and developmental delay in preschoolers. *Npj Digit. Med.* **6**, 99 (2023).
41. de Belen, R. A. J., Bednarz, T., Sowmya, A. & Del Favero, D. Computer vision in autism spectrum disorder research: a systematic review of published studies from 2009 to 2019. *Transl. Psychiatry* **10**, 1–20 (2020).
42. Zampella, C. J., Wang, L. A. L., Haley, M., Hutchinson, A. G. & de Marchena, A. Motor Skill Differences in Autism Spectrum Disorder: a Clinically Focused Review. *Curr. Psychiatry Rep.* **23**, 64 (2021).
43. De Crescenzo, F. *et al.* The use of actigraphy in the monitoring of sleep and activity in ADHD: A meta-analysis. *Sleep Med. Rev.* **26**, 9–20 (2016).
44. Hartman, C. A., Geurts, H. M., Franke, B., Buitelaar, J. K. & Rommelse, N. N. J. Changing ASD-ADHD symptom co-occurrence across the lifespan with adolescence as crucial time window: Illustrating the need to go beyond childhood. *Neurosci. Biobehav. Rev.* **71**, 529–541 (2016).
45. Larsson, H., Dilshad, R., Lichtenstein, P. & Barker, E. D. Developmental trajectories of DSM-IV symptoms of attention-deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology. *J. Child Psychol. Psychiatry* **52**, 954–963 (2011).

46. Loomes, R., Hull, L. & Mandy, W. P. L. What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 466–474 (2017).
47. Ramtekkar, U. P., Reiersen, A. M., Todorov, A. A. & Todd, R. D. Sex and age differences in Attention-Deficit/Hyperactivity Disorder symptoms and diagnoses: Implications for DSM-V and ICD-11. *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 217-28.e1–3 (2010).
48. Joseph, R. M. & Fein, D. The significance of IQ and differential cognitive abilities for understanding ASD. *Neuropsychol. Autism* **281**, 294 (2011).
49. Charman, T. *et al.* Defining the cognitive phenotype of autism. *Brain Res.* **1380**, 10–21 (2011).
50. Alexander, L. M. *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* **4**, 170181 (2017).
51. Wechsler, D. *Wechsler Intelligence Scale for Children-Fifth Edition*. (Pearson, Bloomington, MN).
52. Wechsler, D. Wechsler abbreviated scale of intelligence--. (1999).
53. Kaufman, A. S. & Kaufman, N. L. Kaufman brief intelligence test. *Encycl. Spec. Educ. Ref. Educ. Child. Adolesc. Adults Disabil. Except. Individ.* (2013).
54. Ehlers, S., Gillberg, C. & Wing, L. A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. *J. Autism Dev. Disord.* **29**, 129–141 (1999).
55. Swanson, J. M. *et al.* Categorical and Dimensional Definitions and Evaluations of Symptoms of ADHD: History of the SNAP and the SWAN Rating Scales. *Int. J. Educ. Psychol. Assess.* **10**, 51–70 (2012).
56. Birmaher, B. *et al.* Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. *J. Am. Acad. Child Adolesc. Psychiatry* **38**, 1230–1236 (1999).

57. Angold, A., Costello, E. J., Messer, S. C. & Pickles, A. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Int. J. Methods Psychiatr. Res.* (1995).

58. Achenbach, T. M. *Integrative Guide for the 1991 CBCL/4-18, YSR, and TRF Profiles*. (Department of Psychiatry, University of Vermont, 1991).

59. Carta, A. *et al.* Characterization of Clinical Manifestations in the Co-occurring Phenotype of Attention Deficit/Hyperactivity Disorder and Autism Spectrum Disorder. *Front. Psychol.* **11**, (2020).

60. Menyuk, P. & Brisk, M. E. Language Development in Middle Childhood: Ages 9–13. in *Language development and education: Children with varying language experiences* 118–141 (Springer, 2005).

61. Redmond, S. M. Conversational profiles of children with ADHD, SLI and typical development. *Clin. Linguist. Phon.* **18**, 107–125 (2004).

62. Diehl, J. J., Bennetto, L. & Young, E. C. Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *J. Abnorm. Child Psychol.* **34**, 83–98 (2006).

63. Korrel, H., Mueller, K. L., Silk, T., Anderson, V. & Sciberras, E. Research Review: Language problems in children with Attention-Deficit Hyperactivity Disorder–a systematic meta-analytic review. *J. Child Psychol. Psychiatry* **58**, 640–654 (2017).

64. Fraile, R. & Godino-Llorente, J. I. Cepstral peak prominence: A comprehensive analysis. *Biomed. Signal Process. Control* **14**, 42–54 (2014).

65. Godino-Llorente, J. I. *et al.* The Effectiveness of the Glottal to Noise Excitation Ratio for the Screening of Voice Disorders. *J. Voice* **24**, 47–56 (2010).

66. Fusaroli, R. *et al.* Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in American and Danish autistic children. *Autism Res.* **15**, 653–664 (2022).

67. Worthington, M., Efstathiadis, G., Yadav, V. & Abbas, A. 172. OpenWillis: An Open-Source Python Library for Digital Health Measurement. *Biol. Psychiatry* **95**, S169–S170 (2024).
68. Bellani, M., Moretti, A., Perlini, C. & Brambilla, P. Language disturbances in ADHD. *Epidemiol. Psychiatr. Sci.* **20**, 311–315 (2011).
69. Sciberras, E. *et al.* Language problems in children with ADHD: A community-based study. *Pediatrics* **133**, 793–800 (2014).
70. Tirosh, E. & Cohen, A. Language deficit with attention-deficit disorder: A prevalent comorbidity. *J. Child Neurol.* **13**, 493–497 (1998).
71. Oram, J., Fine, J., Okamoto, C. & Tannock, R. Assessing the language of children with attention deficit hyperactivity disorder. *Am. J. Speech Lang. Pathol.* **8**, 72–80 (1999).
72. Jepsen, I. B., Hougaard, E., Matthiesen, S. T. & Lambek, R. A systematic review and meta-analysis of narrative language abilities in children with Attention-Deficit/Hyperactivity Disorder. *Res. Child Adolesc. Psychopathol.* **50**, 737–751 (2022).
73. Geurts, H. M., Broeders, M. & Nieuwland, M. S. Thinking outside the executive functions box: Theory of mind and pragmatic abilities in attention deficit/hyperactivity disorder. *Eur. J. Dev. Psychol.* **7**, 135–151 (2010).
74. Tannock, R. & Schachar, R. Executive dysfunction as an underlying mechanism of behavior and language problems in attention deficit hyperactivity disorder. in *Language, learning, and behavior disorders: Developmental, biological, and clinical perspectives* 128–155 (Cambridge University Press, New York, NY, US, 1996).
75. Taurines, R. *et al.* ADHD and autism: differential diagnosis or overlapping traits? A selective review. *ADHD Atten. Deficit Hyperact. Disord.* **4**, 115–139 (2012).
76. Naigles, L. R. *et al.* “You’re telling me!” The prevalence and predictors of pronoun reversals in children with autism spectrum disorders and typical development. *Res. Autism Spectr. Disord.* **27**, 11–20 (2016).
77. Baron-Cohen, S., Leslie, A. M. & Frith, U. Does the autistic child have a “theory of mind” ?

Cognition **21**, 37–46 (1985).

78. Losh, M. & Gordon, P. C. Quantifying Narrative Ability in Autism Spectrum Disorder: A Computational Linguistic Analysis of Narrative Coherence. *J. Autism Dev. Disord.* **44**, 3016–3025 (2014).
79. Lee, M. *et al.* What's the story? A computational analysis of narrative competence in autism. *Autism* **22**, 335–344 (2018).
80. Pecoraro, P. M., Marsili, L., Espay, A. J., Bologna, M. & di Biase, L. Computer Vision Technologies in Movement Disorders: A Systematic Review. *Mov. Disord. Clin. Pract.* (2025).
81. Cook, A., Mandal, B., Berry, D. & Johnson, M. Towards Automatic Screening of Typical and Atypical Behaviors in Children With Autism. in *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* 504–510 (2019). doi:10.1109/DSAA.2019.00065.
82. Kojovic, N., Natraj, S., Mohanty, S. P., Maillart, T. & Schaer, M. Using 2D video-based pose estimation for automated prediction of autism spectrum disorders in young children. *Sci. Rep.* **11**, 15069 (2021).
83. Lemler, C. *et al.* Semi-Automated Multi-Label Classification of Autistic Mannerisms by Machine Learning on Post Hoc Skeletal Tracking. *Autism Res.* **18**, 833–844 (2025).
84. Vyas, K. *et al.* Recognition Of Atypical Behavior In Autism Diagnosis From Video Using Pose Estimation Over Time. in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)* 1–6 (2019). doi:10.1109/MLSP.2019.8918863.
85. Zhang, Y., Tian, Y., Wu, P. & Chen, D. Application of Skeleton Data and Long Short-Term Memory in Action Recognition of Children with Autism Spectrum Disorder. *Sensors* **21**, 411 (2021).
86. Koehler, J. C. *et al.* Machine learning classification of autism spectrum disorder based on reciprocity in naturalistic social interactions. *Transl. Psychiatry* **14**, 76 (2024).
87. Dennis, M. *et al.* Why IQ is not a covariate in cognitive studies of neurodevelopmental

disorders. *J. Int. Neuropsychol. Soc.* **15**, 331–343 (2009).

88. Faerman, A. *et al.* Language Abilities are Associated with Both Verbal and Nonverbal Intelligence in Children on the Autism Spectrum. *Dev. Neuropsychol.* **48**, 248–257 (2023).
89. Conway, K. P. *et al.* Association between COVID-19 risk-mitigation behaviors and specific mental disorders in youth. *Child Adolesc. Psychiatry Ment. Health* **17**, 14 (2023).
90. Restrepo, A. *et al.* Problematic internet use in children and adolescents: associations with psychiatric disorders and impairment. *BMC Psychiatry* **20**, 252 (2020).
91. Kramer, E. *et al.* Diagnostic Associations of Processing Speed in a Transdiagnostic, Pediatric Sample. *Sci. Rep.* **10**, 10114 (2020).
92. Kaufman, J. *et al.* Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 980–988 (1997).
93. Bain, M., Huh, J., Han, T. & Zisserman, A. WhisperX: Time-Accurate Speech Transcription of Long-Form Audio. in *INTERSPEECH 2023* 4489–4493 (ISCA, 2023). doi:10.21437/Interspeech.2023-78.
94. Lee, J. *et al.* Gecko: Versatile Text Embeddings Distilled from Large Language Models. Preprint at https://doi.org/10.48550/arXiv.2403.20327 (2024).
95. Ortubay, A. S., Parra, L. C. & Madsen, J. Real-time estimation of overt attention from dynamic features of the face using deep-learning. Preprint at https://doi.org/10.48550/arXiv.2409.13084 (2024).

Extended Data

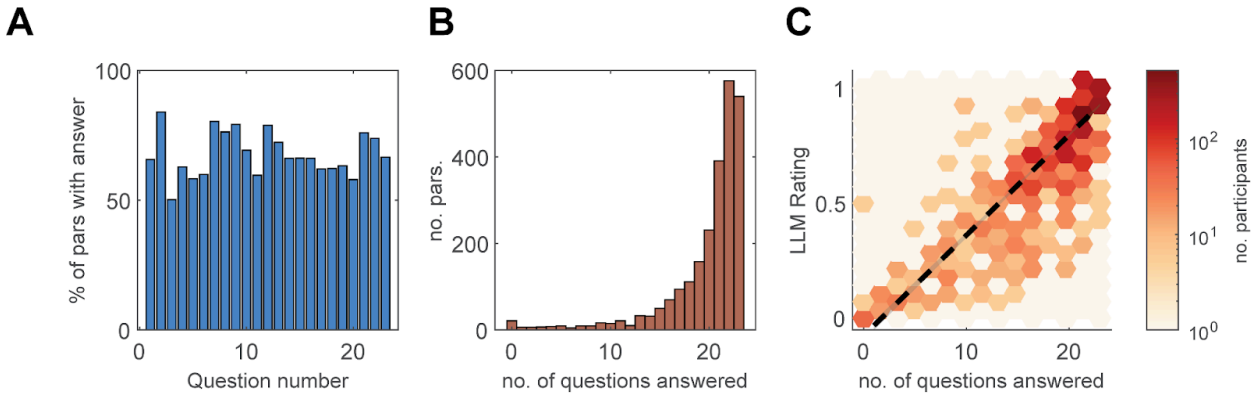
Extended Data Table S1. Interview Questions & Typical Answer

Semantic Measure		Question	Typical Answer, Typically Developing Children
—	1	"I hope you enjoyed the last movie. Have you seen it before?"	"Yes. Yeah."
Narrative Detail	2	"Can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning, things that happen, and an ending."	"OK. So it started off with the boy. He seemed kind of upset, and he was just playing video games. And it was kind of dark in their house and stuff. And his mom comes home, and she like turns on the lights and everything and she gives him, she gives him the gift in the box so he stops playing and then he opens the box and he like sees a puppy and then he kind of gets excited and then he sees that the puppy doesn't have a leg, like is missing one leg. So then he's like not interested and he kind of just puts it off to the side but the dog doesn't really know what's going on. It's just like excited to be with someone like an owner or something. And then he goes back to playing video games, and the dog is kind of just searching around the house and finds the ball. And you can tell that the dog's kind of having trouble walking and playing with it and stuff. And so then he just kind of watches the owner. He kind of watches as the dog's kind of struggling. but he's playing with the ball and it's kind of making him kind of happy watching him just like play around and he's like still happy. And then he eventually turns off his video game and then it ends when they go outside to play. And you realize that he has some, like he, his leg also is missing."
—	3	"Do you remember anything else from the story?"	"Um... No. Okay."
Positive Preferences	4	"What are some of the things you liked about the movie?"	"I liked the dog. And I also liked how he liked him at the end and started playing with him. Yeah, that's nice."
Negative Preferences	5	"Is there something you didn't like about the movie?"	"There wasn't a whole lot that I didn't like. I mean, like, honestly, I didn't like the fact that he was being mean in the beginning, kicking the dog."
Factual Memory	6	"Who gave the boy a box?"	"The mom."
Factual Memory	7	"What was in the box?"	"A dog."
Factual Memory	8	"What was the boy doing before he got the box?"	"Playing video games."
Factual Memory	9	"What was the puppy playing with?"	"A ball, red ball."
Thematic Understanding	10	"How are the puppy and the boy the same?"	"They both are missing a leg."
—	11	"In the movie, who is missing a leg? The boy, the puppy, or no one?"	"Both."
		Kids rewatch a short clip from "The Present" (00:53-01:02)	
External Emotion Description	12	"How do you think the puppy was feeling?"	"Happy. Really happy."
External Emotion Description	13	"How do you think the boy was feeling?"	"Happy. A little excited."
Self Emotion Description	14	"And how did you feel while you were watching that part?"	"Happy."
		Kids rewatch a short clip from "The Present" (01:03-01:09)	
External Emotion Description	15	"How do you think the puppy was feeling?"	"At first he was okay and then he was probably upset."
External Emotion Description	16	"How do you think the boy was feeling?"	"Disappointed, a little angry, maybe."
Self Emotion Description	17	"And how did you feel while you were watching that part?"	"A little bit upset, because it was pretty sad and all that."
		Kids rewatch a short clip from "The Present" (01:14-01:23)	
External Emotion Description	18	"How do you think the puppy was feeling?"	"Sad. Like he tried again to like play and all that and then he just kicked him over to get lost."
External Emotion Description	19	"How do you think the boy was feeling?"	"Angry. Still mad."
Self Emotion Description	20	"And how did you feel while you were watching that part?"	"A little bit sadder because he was actually kind of unleashing his emotions at the puppy.."
		Kids rewatch a short clip from "The Present" (03:03-03:21)	
External Emotion Description	21	"How do you think the puppy was feeling?"	"Happy."
External Emotion Description	22	"How do you think the boy was feeling?"	"Happy, too."
Self Emotion Description	23	"And how did you feel while you were watching that part?"	"Good. Happy."

Extended Data Table S2. Description of Behavioral Variables.

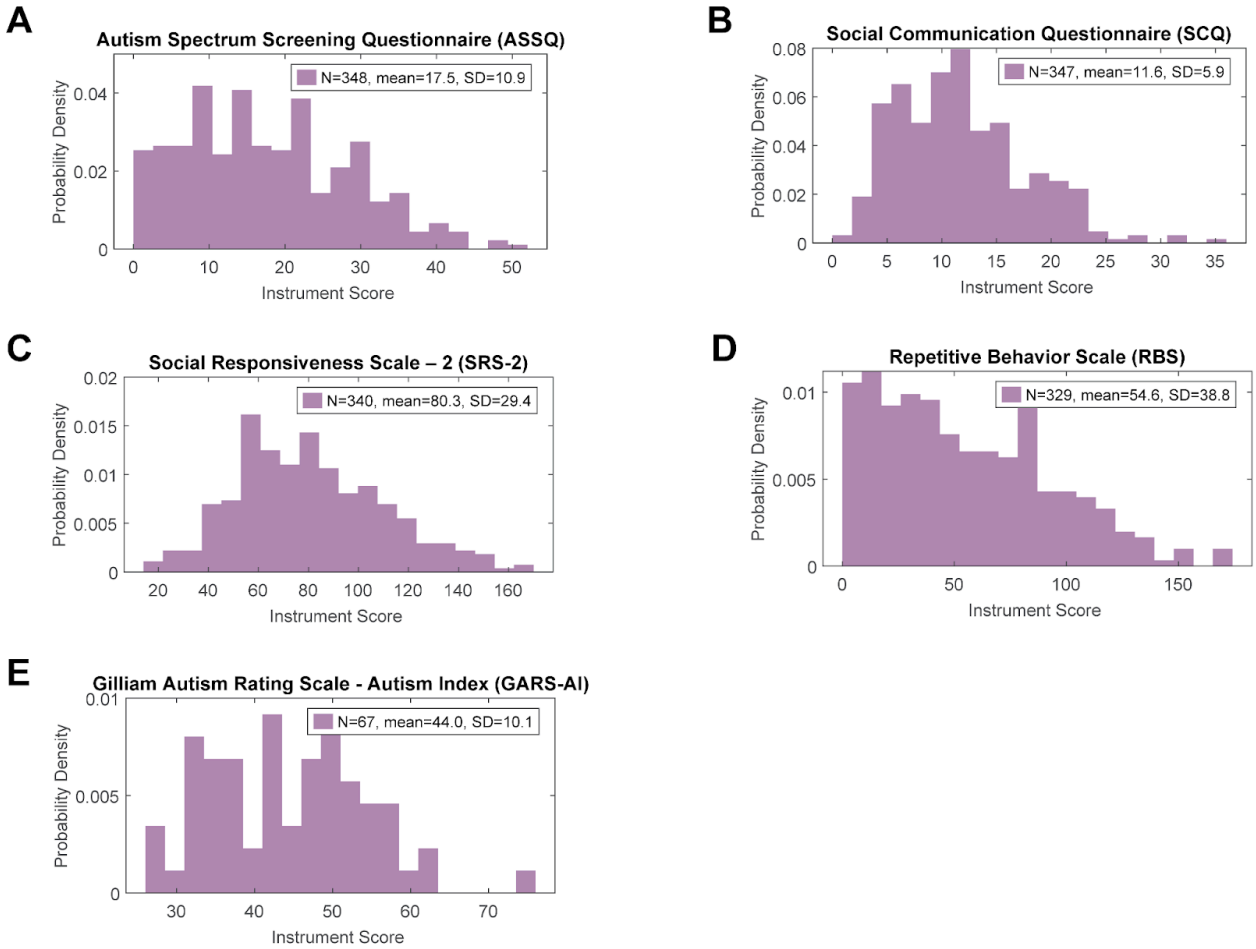
Group	Measured Aspect	Variable	Measurement
Language Abilitiy Measures	Structural Language	*Amount of Speech*	Total subject speech length in words
		Speech Rate	Subject speech rate, in words per minute
		Filler Word Usage	Number of filler words ('um', 'uh', 'er', 'ah', 'like', 'you know', 'so', 'okay', 'right') per minute of subject speech
		% of unique words	Percentage of unique words in total subject speech
		Lexical Diversity	Average Moving Average Type Token Ratio with a window size of 10
		Between Word Coherence	Average measure of semantic similarity of each word to the preceding word using LLMs
	Pragmatic Language	*Subject Conversation Presence*	Subject speech percentage from total conversation speech
		Self-Focused Speech	Percentage of pronouns in the utterance that are first-person singular pronouns
		Question-Answer Coherence	Average semantic similarity of current speaker turn to previous turn of the other speaker
Semantic Measures (Socio-Cognitive Abilities)	Memory	*Factual Memory*	Answer Typicality when asked about facts in the movie
	Social Judgment & Salience	*External Emotion Description*	Answer Typicality when asked about dog/child emotions in the movie
	Theory of Mind	*Self Emotion Description*	Answer Typicality when asked about self emotions in the movie
	Narrative Production, Comprehension & Theory of Mind	*Narrative Detail*	Answer Typicality when asked to give a detailed recall of the movie
	Theory of Mind	*Thematic Understanding*	Answer Typicality when asked about the similarities of the dog and kid, and their shared experience
	Social Judgment & Salience	*Positive Preferences*	Answer Typicality when asked about what they liked in the movie
	Social Judgment & Salience	*Negative Preferences*	Answer Typicality when asked about what they did not like in the movie
Vocal Prosody Measures		*Pitch*	Vocal Pitch (f0)
		Pitch Variation	Pitch variation (std(f0))
		Loudness	Vocal intensity in dB
		Loudness Variation	Intensity variation (std(dB))
		Dysphonia	Mean cepstral peak prominence
		Breathiness	Mean glottal-to-noise excitation ratio
Movement Measures		*Facial Movements*	Total movement in the video, where landmarks of interest were visible
		Eye Movements	Total movement in the video, where landmarks of interest were visible
		Mouth Movements	Total movement in the video, where landmarks of interest were visible
		Body Movements	Total movement in the video, where landmarks of interest were visible

Extended Data Figure S1. Validity of interview questions.



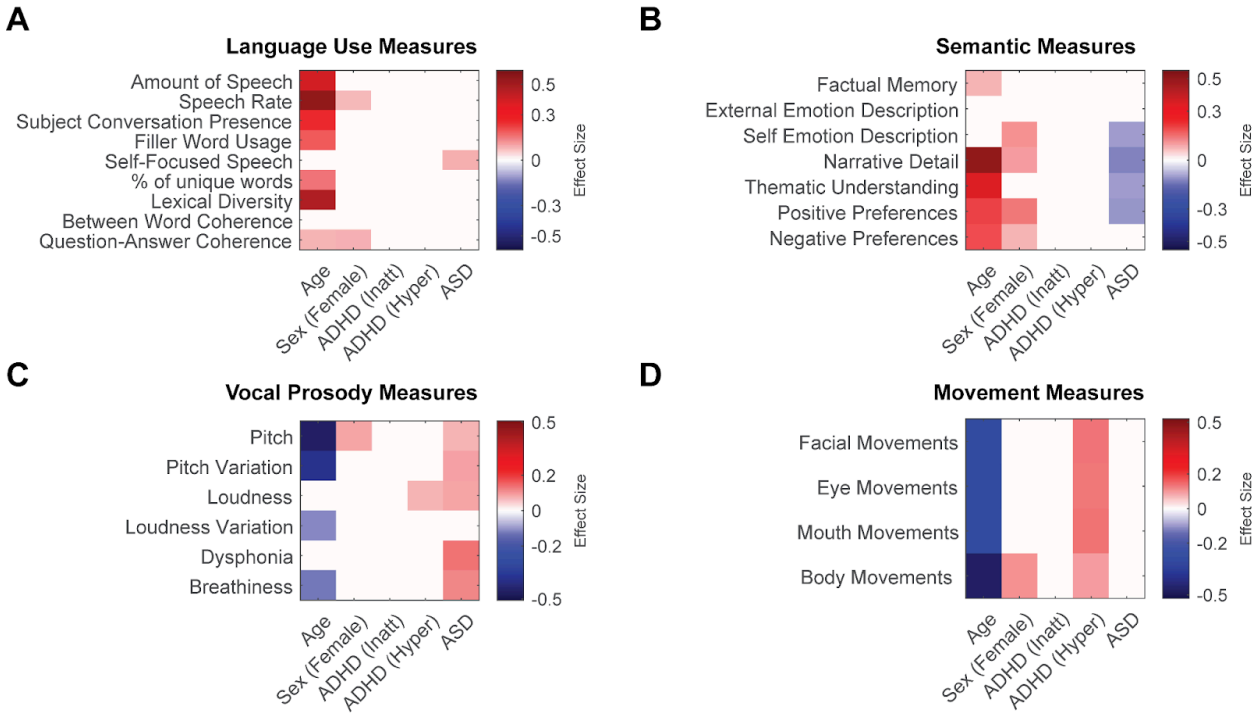
Extended Data Figure S1. Validity of interview questions. A) Not all questions were asked to every participant, or the answer was not available. In average, 67.7% of participants answered each question. B) Each participant answered 19.9 questions in average. C) The number of questions answered by each participant is correlated with the LLM's rating of the conversation completion ($\rho(2339) = 0.79$, $p<0.001$), see Supplementary Materials Section S3 for rating prompt.

Extended Data Figure S2. Distribution of Autism Scores in the ASD-positive population



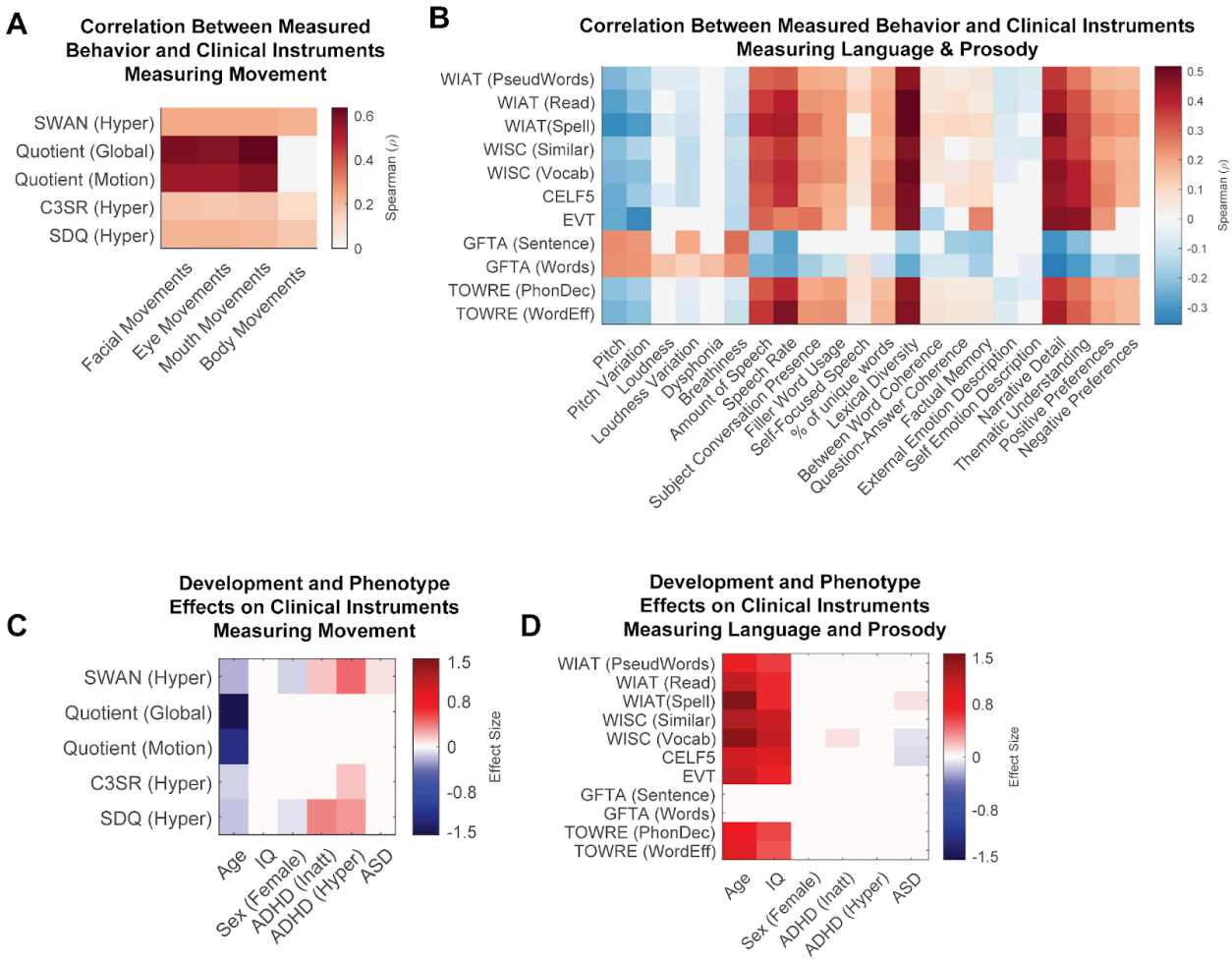
Extended Data Figure S2. Distribution of Autism Scores in the ASD-positive population. The Healthy Brain Network's deep phenotyping protocol provides multiple parent- and self-report questionnaires to measure traits associated with ASD. These included the Autism Spectrum Screening Questionnaire (ASSQ), the Social Communication Questionnaire (SCQ), the Social Responsiveness Scale-2 (SRS-2). The ASD population is comprised primarily of mild-to-moderate cases.

Extended Data Figure S3. Effect size matrices for primary behavioral outcomes after excluding IQ as covariate.



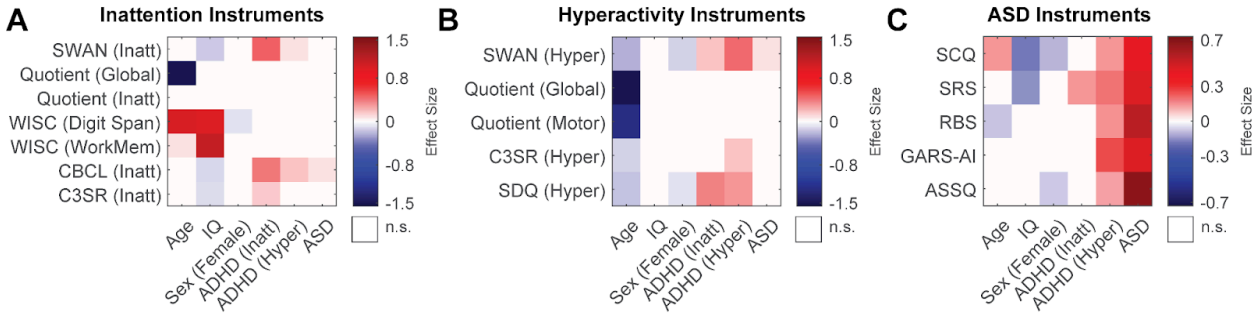
Extended Data Figure S3. Effect size matrices for primary behavioral outcomes after excluding IQ as covariate. The models (A through D) are identical to the primary multivariate models presented in the main text, but without modelling IQ as covariate. Results are mainly consistent. Effect sizes are estimated using multivariate regression, with $p<0.01$ after Bonferroni correction.

Extended Data Figure S4. Comparing Clinical Instruments to Measured Behavior.



Extended Data Figure S4. Comparing Clinical Instruments to Measured Behavior. Spearman correlation between automatically measured behaviors and the clinical instrument scores measuring movement (A), and language use/prosody (B). C-D) Effects of development and diagnosis on such clinical instrument scores. As one may expect, our measures of movement, language and speech, correlate strongly with these established metrics (A, B). When using standard instruments it appears that ADHD-Inattention is also associated with elevated motor activity (C). Thus, our movement measures appear to be more specific to the ADHD-Hyperactive status (cf. main text Fig. 6C with panel C above). On the other hand, when using standard instruments, it appears that ASD has a lesser effect on language use (D). Therefore our features related to social communication appear to be more sensitive to the ASD presentation (cf. main text Fig. 4E with panel D above). Effect sizes in C, D measured using multivariate models, significance at $p<0.01$ after Bonferroni correction, white otherwise. For A and B, significant at a cutoff of $p<0.05$, uncorrected, white otherwise.

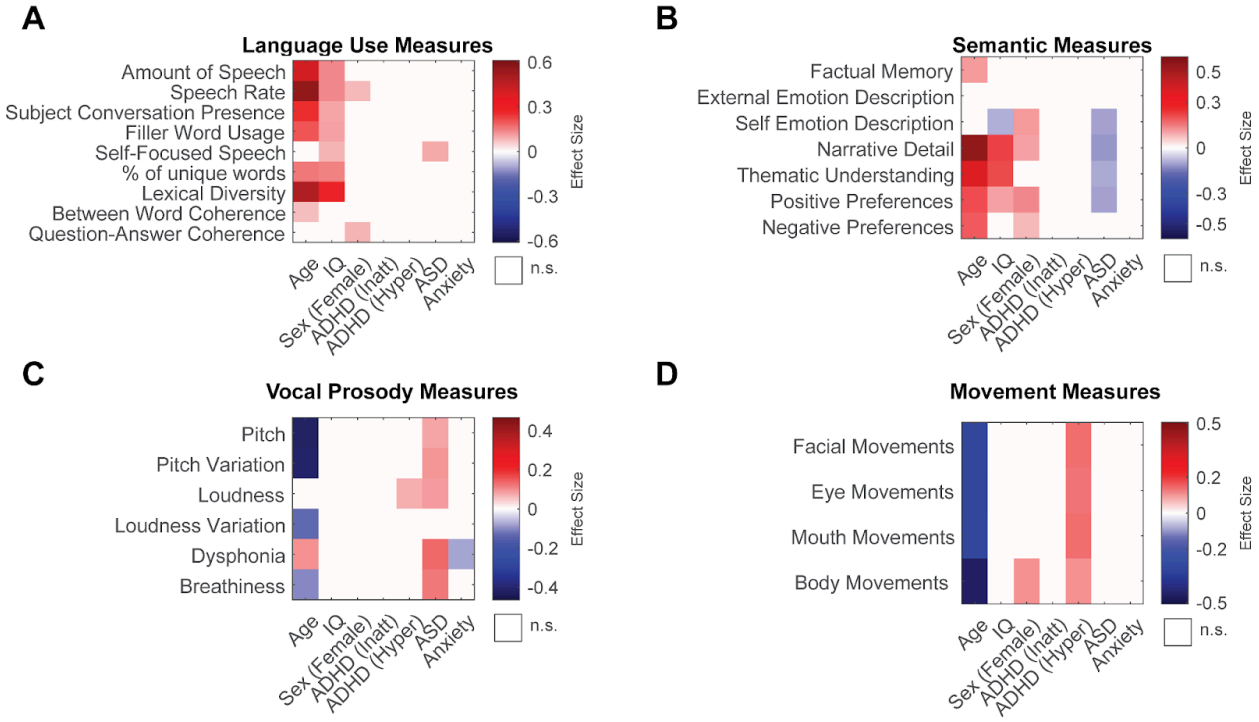
Extended Data Figure S5. Unique association of clinical questionnaire scales with clinician-confirmed diagnoses.



Extended Data Figure S5. Unique association of clinical questionnaire scales with clinician-confirmed diagnoses. The Healthy Brain Network's deep phenotyping protocol provides multiple parent- and self-report questionnaires to measure traits associated with ADHD and ASD.The effect size matrices reveal which clinical instruments are more uniquely explained by specific positive diagnoses while accounting for development and demographic effects. To select the most specific and robust trait measures for our analyses in main text Figures 2-3, we evaluated which of these scales best and most uniquely reflected the corresponding gold-standard, clinician-confirmed diagnosis. The SWAN Inattention and Hyperactivity subscales and the ASSQ emerged as the strongest and most specific associations for their corresponding clinician-confirmed diagnoses. Therefore, we selected these scales for all subsequent analyses where continuous measures of ADHD and ASD traits were required. Effect sizes measured using multivariate models, significance at $p<0.01$ after Bonferroni correction, white otherwise.

Extended Data Figure S6. Effect size matrices for primary behavioral outcomes after including anxiety as covariate.

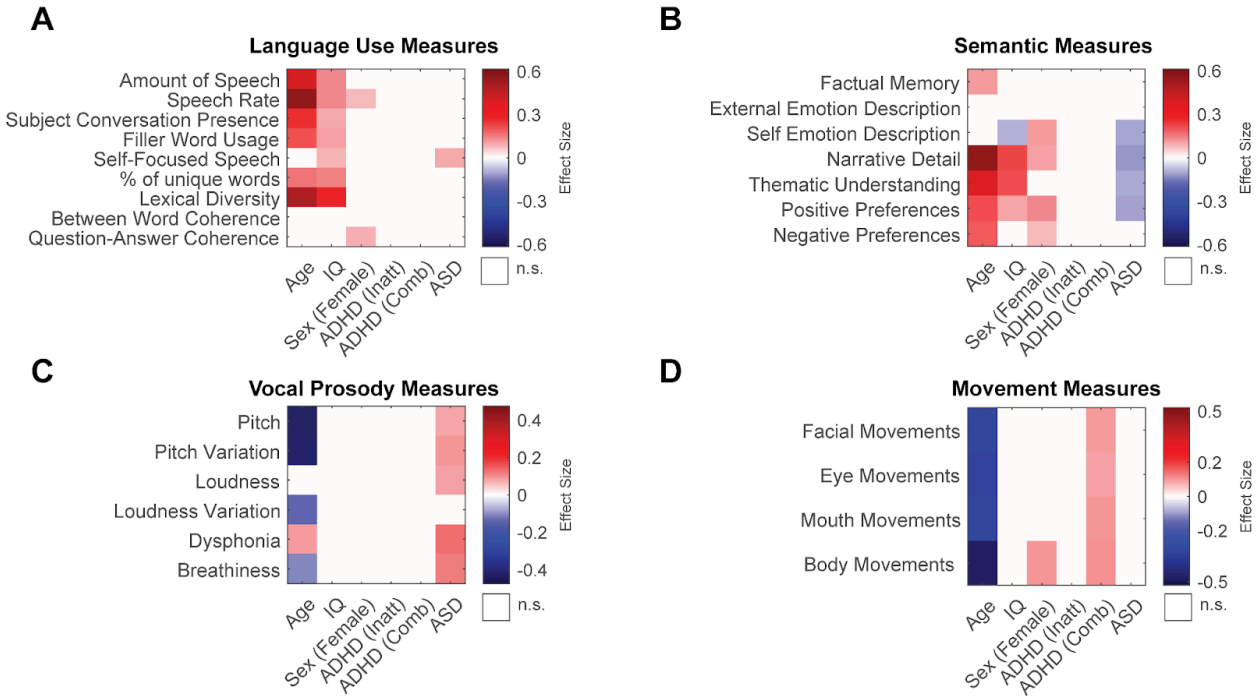
(Figure S4).



Extended Data Figure S6. Effect size matrices for primary behavioral outcomes after including anxiety as covariate. Given the high rates of comorbid anxiety in both ADHD and ASD, and the potential for a semi-structured interview with an unfamiliar clinician to elicit anxious behaviors, we modeled anxiety as a key potential confound. Behaviors associated with anxiety such as motor restlessness, social reticence, or altered vocal tone could overlap with the objective behavioral markers under investigation and potentially obscure the specific effects of ADHD and ASD. To address this and confirm the robustness of our primary findings, we repeated our primary multivariate regressions, adding a binary variable for any clinician-confirmed DSM-5 anxiety diagnosis as an additional covariate (N = 789). This allowed us to test whether the observed associations between our objective behavioral markers and ADHD/ASD diagnoses remained significant after statistically controlling for the influence of anxiety. We find that all the associations between ASD and/or ADHD with our behavioral measures are independent of whether we model anxiety or not. Effect sizes measured using multivariate models, significance at $p<0.01$ after Bonferroni correction, white otherwise.

Extended Data Figure S7. Results are independent of how the ADHD presentations are coded.

.



Extended Data Figure S7. Results are independent of how the ADHD presentations are coded. Throughout the work presented in this manuscript, clinician-diagnosed ADHD-Combined presentations were modelled as positive in both ADHD-Hyperactive and ADHD-Inattentive status. This was done to maximize the number of participants in the modelled group, while maintaining the distinction of the two main inattentive and hyperactive diagnostic status. Previous studies on the HBN population have modelled this overlap by considering the ADHD-Combined separate from the ADHD-Inattentive presentations [89–91]. We repeated the analyses on the behavioral features by following this approach, and the results were mostly consistent with our associations with the ADHD-Hyperactive presentation. The models (A through D) are identical to the primary multivariate models presented in the main text, but

ADHD modelled separately as the ADHD-Combined (N=625) and ADHD-Inattentive (N=653) presentations. Effect sizes are estimated using multivariate regression, with $p<0.01$ after Bonferroni correction.

Supplementary Material

Contents

Supplementary Materials Section S1. Multivariate Model Reports.....39
Supplementary Materials Section S1.1. Diagnoses.....39
Supplementary Materials Section S1.2. Clinical Instruments..... 39
Supplementary Materials Section S1.3. Language Use Measures..... 41
Supplementary Materials Section S1.4. Semantic Measures.....44
Supplementary Materials Section S1.5. Vocal Prosody Measures.....46
Supplementary Materials Section S1.6. Movement Measures..... 48
Supplementary Materials Section S2. Diarizing Prompt for LLM.....50
Supplementary Materials Section S2. Question-Answer Extraction Prompt for LLM..... 51

Supplementary Materials Section S1. Multivariate Model Reports.

p-values reported here appear uncorrected for multiple comparisons. For Figures 3-6, Effect Size matrices show significant effect sizes with a cutoff of p<0.01, white otherwise. p-values for the displays and analyses are Bonferroni-corrected for multiple comparisons by multiplying the model p-values (reported below) by the number of tested behavioral features.

Supplementary Materials Section S1.1. Diagnoses

Detailed model reports for panel A in Figure 2 and black arrows in panel C in Figure 3.

--- Results for Outcome: ASD ---

Predictor	Odds Ratio (OR)	95% CI	p-value
Age	1.033	[0.997, 1.07]	0.0746
Sex (Female)	0.426	[0.315, 0.575]	2.57E-08
ADHD (Inatt)	2.251	[1.69, 3]	3.55E-08
ADHD (Hyper)	1.597	[1.22, 2.08]	0.000548

--- Results for Outcome: ADHD (Inatt) ---

Predictor	Odds Ratio (OR)	95% CI	p-value
Age	1.071	[1.04, 1.1]	1.03E-06
Sex (Female)	0.681	[0.56, 0.828]	0.000118
ADHD (Hyper)	10.008	[7.8, 12.8]	1.38E-73
ASD	2.256	[1.69, 3.01]	3.62E-08

--- Results for Outcome: ADHD (Hyper) ---

Predictor	Odds Ratio (OR)	95% CI	p-value
Age	0.826	[0.797, 0.856]	7.84E-26
Sex (Female)	0.58	[0.46, 0.731]	3.90E-06
ADHD (Inatt)	10.46	[8.13, 13.5]	1.44E-74
ASD	1.601	[1.22, 2.09]	0.000593

Supplementary Materials Section S1.2. Clinical Instruments

Detailed model reports for panels E & F in Figure 2.

--- Results for Feature: Externalizing (CBCL) ---

Model Fit: $R^2 = 0.146$, Adjusted $R^2 = 0.144$

Overall Model: $F(5, 2126) = 72.513$, $p = 2.96e-70$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.0753	0.0584	-1.29	-0.028	0.197
Sex (Female)	-0.0637	0.39	-0.163	-0.00354	0.87
ADHD (Inatt)	1.48	0.407	3.63	0.0787	0.000292
ADHD (Hyper)	6.06	0.438	13.8	0.3	9.10E-42
ASD	0.298	0.517	0.577	0.0125	0.564

--- Results for Feature: Internalizing (CBCL) ---

Model Fit: $R^2 = 0.087$, Adjusted $R^2 = 0.085$

Overall Model: $F(5, 2126) = 40.583$, $p = 5.62e-40$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.451	0.0501	8.98	0.195	5.57E-19
Sex (Female)	1.25	0.335	3.74	0.0811	1.91E-04
ADHD (Inatt)	1.16	0.35	3.31	0.0718	0.000951
ADHD (Hyper)	0.991	0.376	2.63	0.0571	0.00851
ASD	2.55	0.444	5.75	0.125	1.05E-08

--- Results for Feature: Anxiety Traits (SCARED) ---

Model Fit: $R^2 = 0.040$, Adjusted $R^2 = 0.038$

Overall Model: $F(5, 2142) = 17.822$, $p = 2.4e-17$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value

Age	0.19	0.0701	2.71	0.0586	0.00677
Sex (Female)	2.36	0.496	4.76	0.103	2.04E-06
ADHD (Inatt)	1.1	0.515	2.13	0.046	0.0333
ADHD (Hyper)	0.411	0.556	0.739	0.016	0.46
ASD	2.32	0.648	3.58	0.0773	3.54E-04

--- Results for Feature: Depression Traits (MFQ) ---

Model Fit: $R^2 = 0.078$, Adjusted $R^2 = 0.076$

Overall Model: $F(5, 2151) = 36.627$, $p = 4.07e-36$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.23	0.0464	4.97	0.107	7.33E-07
Sex (Female)	1.11	0.328	3.38	0.0729	0.000729
ADHD (Inatt)	1.82	0.341	5.34	0.115	1.00E-07
ADHD (Hyper)	1.33	0.368	3.62	0.0781	0.000299
ASD	1.01	0.429	2.35	0.0507	0.0187

Supplementary Materials Section S1.3. Language Use Measures

Detailed model reports for panels D & E in Figure 3.

--- Results for Feature: Amount of Speech ---

Model Fit: $R^2 = 0.154$, Adjusted $R^2 = 0.151$

Overall Model: $F(6, 2104) = 63.662$, $p = 8.01e-73$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0881	0.00483	18.3	0.398	2.98E-69
IQ	0.00529	0.000845	6.25	0.136	4.87E-10
Sex (Female)	0.0646	0.0302	2.14	0.0466	0.0327
ADHD (Inatt)	0.0435	0.0316	1.38	0.03	0.169
ADHD (Hyper)	0.00959	0.0341	0.281	0.00613	0.779
ASD	0.00618	0.04	0.154	0.00337	0.877

--- Results for Feature: Speech Rate ---

Model Fit: $R^2 = 0.251$, Adjusted $R^2 = 0.249$

Overall Model: $F(6, 2104) = 117.372$, $p = 4.34e-128$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	5.49	0.212	25.8	0.563	6.74E-128

IQ	0.231	0.0372	6.21	0.135	6.32E-10
Sex (Female)	4.77	1.33	3.58	0.0781	0.000349
ADHD (Inatt)	-1.45	1.39	-1.04	-0.0228	0.297
ADHD (Hyper)	4.02	1.5	2.68	0.0583	0.0075
ASD	-3.28	1.76	-1.86	-0.0406	0.063

--- Results for Feature: Subject Conversation Presence ---

Model Fit: $R^2 = 0.072$, Adjusted $R^2 = 0.069$

Overall Model: $F(6, 2104) = 27.199$, $p = 2.19e-31$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	1.33	0.113	11.7	0.256	8.63E-31
IQ	0.0914	0.0199	4.6	0.1	4.51E-06
Sex (Female)	1.11	0.711	1.56	0.034	0.119
ADHD (Inatt)	1.55	0.743	2.08	0.0454	0.0376
ADHD (Hyper)	0.0615	0.801	0.0768	0.00167	0.939
ASD	0.225	0.941	0.24	0.00522	0.811

--- Results for Feature: Filler Word Usage ---

Model Fit: $R^2 = 0.062$, Adjusted $R^2 = 0.060$

Overall Model: $F(6, 2122) = 23.512$, $p = 4.96e-27$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.272	0.0282	9.65	0.21	1.33E-21
IQ	0.0251	0.00496	5.06	0.11	4.60E-07
Sex (Female)	0.0974	0.177	0.552	0.012	0.581
ADHD (Inatt)	0.0397	0.185	0.215	0.00466	0.83
ADHD (Hyper)	0.111	0.199	0.559	0.0121	0.576
ASD	-0.388	0.233	-1.66	-0.0361	0.0968

--- Results for Feature: Self-Focused Speech ---

Model Fit: $R^2 = 0.031$, Adjusted $R^2 = 0.029$

Overall Model: $F(6, 2100) = 11.306$, $p = 1.83e-12$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.0402	0.0977	-0.411	-0.00897	0.681

IQ	0.0669	0.0171	3.9	0.0852	9.78E-05
Sex (Female)	-1.23	0.613	-2	-0.0437	0.0456
ADHD (Inatt)	0.711	0.641	1.11	0.0242	0.267
ADHD (Hyper)	-0.313	0.691	-0.454	-0.0099	0.65
ASD	3.7	0.81	4.57	0.0996	5.28E-06

--- Results for Feature: % of unique words ---

Model Fit: $R^2 = 0.065$, Adjusted $R^2 = 0.062$

Overall Model: $F(6, 2104) = 24.246$, $p = 6.8e-28$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.188	0.0255	7.35	0.16	2.93E-13
IQ	0.0298	0.00447	6.65	0.145	3.62E-11
Sex (Female)	0.446	0.16	2.79	0.0607	0.00538
ADHD (Inatt)	-0.294	0.167	-1.76	-0.0383	0.0788
ADHD (Hyper)	-0.243	0.18	-1.35	-0.0294	0.177
ASD	0.0827	0.212	0.39	0.00851	0.696

--- Results for Feature: Lexical Diversity ---

Model Fit: $R^2 = 0.246$, Adjusted $R^2 = 0.244$

Overall Model: $F(6, 2104) = 114.224$, $p = 5.05e-125$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.00446	0.000199	22.4	0.489	6.63E-100
IQ	0.000437	3.49E-05	12.5	0.273	9.25E-35
Sex (Female)	0.00295	0.00125	2.37	0.0516	0.0181
ADHD (Inatt)	0.00242	0.00131	1.86	0.0405	0.0635
ADHD (Hyper)	-0.00282	0.00141	-2	-0.0437	0.0453
ASD	0.00149	0.00165	0.899	0.0196	0.369

--- Results for Feature: Between Word Coherence ---

Model Fit: $R^2 = 0.013$, Adjusted $R^2 = 0.010$

Overall Model: $F(6, 2104) = 4.709$, $p = 9.03e-05$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.000141	4.30E-05	3.28	0.0714	0.00107
IQ	1.99E-06	7.53E-06	0.264	0.00576	0.791

Sex (Female)	0.000214	0.000269	0.793	0.0173	0.428
ADHD (Inatt)	-0.000485	0.000282	-1.72	-0.0375	0.0858
ADHD (Hyper)	1.73E-05	0.000304	0.0568	0.00124	0.955
ASD	-0.00104	0.000357	-2.92	-0.0637	0.00353

--- Results for Feature: Question-Answer Coherence ---

Model Fit: $R^2 = 0.014$, Adjusted $R^2 = 0.011$

Overall Model: $F(6, 2087) = 5.029$, $p = 3.96e-05$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.00151	0.000526	2.87	0.0628	0.00416
IQ	-9.45E-06	9.20E-05	-0.103	-0.00225	0.918
Sex (Female)	0.0134	0.00329	4.08	0.0892	4.73E-05
ADHD (Inatt)	-0.00316	0.00344	-0.92	-0.0201	0.358
ADHD (Hyper)	0.00374	0.00371	1.01	0.0221	0.314
ASD	-0.0047	0.00435	-1.08	-0.0236	0.281

Supplementary Materials Section S1.4. Semantic Measures

Detailed model reports for panel E in Figure 4.

--- Results for Feature: Factual Memory ---

Model Fit: $R^2 = 0.018$, Adjusted $R^2 = 0.015$

Overall Model: $F(6, 2018) = 6.284$, $p = 1.48e-06$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0108	0.0021	5.13	0.114	3.11E-07
IQ	0.000716	0.00037	1.94	0.0431	0.0528
Sex (Female)	-0.0164	0.0132	-1.24	-0.0276	0.214
ADHD (Inatt)	0.00958	0.0138	0.694	0.0155	0.488
ADHD (Hyper)	-0.00847	0.0149	-0.57	-0.0127	0.569
ASD	-0.0435	0.0173	-2.51	-0.0559	0.0122

--- Results for Feature: External Emotion Description---

Model Fit: $R^2 = 0.005$, Adjusted $R^2 = 0.002$

Overall Model: $F(6, 2096) = 1.729$, $p = 0.11$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.00207	0.000879	-2.35	-0.0513	1.89E-02

IQ	5.99E-05	0.000155	0.387	0.00845	6.99E-01
Sex (Female)	0.00226	0.00551	0.41	0.00895	0.682
ADHD (Inatt)	0.0126	0.00577	2.19	0.0478	0.0289
ADHD (Hyper)	-0.00409	0.00622	-0.657	-0.0144	0.511
ASD	0.00165	0.00732	0.226	0.00494	0.821

--- Results for Feature: Self Emotion Description ---

Model Fit: $R^2 = 0.035$, Adjusted $R^2 = 0.032$

Overall Model: $F(6, 2080) = 12.671$, $p = 4.34e-14$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.00156	0.00139	1.13	0.0247	2.60E-01
IQ	-0.000989	0.000244	-4.05	-0.0888	5.36E-05
Sex (Female)	0.0475	0.00872	5.45	0.119	5.67E-08
ADHD (Inatt)	0.000366	0.00911	0.0401	0.00088	0.968
ADHD (Hyper)	0.0197	0.00981	2	0.0439	0.0452
ASD	-0.0547	0.0116	-4.74	-0.104	2.32E-06

--- Results for Feature: Narrative Detail ---

Model Fit: $R^2 = 0.282$, Adjusted $R^2 = 0.280$

Overall Model: $F(6, 2084) = 136.422$, $p = 5.2e-146$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0439	0.00174	25.2	0.552	1.53E-122
IQ	0.0031	0.000305	10.2	0.223	8.02E-24
Sex (Female)	0.0509	0.0108	4.69	0.103	2.89E-06
ADHD (Inatt)	0.0324	0.0113	2.86	0.0626	0.00428
ADHD (Hyper)	-0.0327	0.0122	-2.67	-0.0584	0.0077
ASD	-0.0765	0.0144	-5.32	-0.117	1.13E-07

--- Results for Feature: Thematic Understanding ---

Model Fit: $R^2 = 0.164$, Adjusted $R^2 = 0.161$

Overall Model: $F(6, 1859) = 60.793$, $p = 5.6e-69$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0647	0.00392	16.5	0.383	3.29E-57

IQ	0.00645	0.000699	9.23	0.214	7.18E-20
Sex (Female)	-0.0216	0.0244	-0.883	-0.0205	0.377
ADHD (Inatt)	0.0368	0.0257	1.43	0.0331	0.153
ADHD (Hyper)	-0.039	0.0278	-1.4	-0.0325	0.161
ASD	-0.136	0.0325	-4.2	-0.0973	2.84E-05

--- Results for Feature: Positive Preferences ---

Model Fit: $R^2 = 0.087$, Adjusted $R^2 = 0.084$

Overall Model: $F(6, 1867) = 29.728$, $p = 3.47e-34$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0144	0.00158	9.09	0.21	2.51E-19
IQ	1.25E-03	0.00028	4.47	0.103	8.30E-06
Sex (Female)	0.0596	0.01	5.94	0.137	3.40E-09
ADHD (Inatt)	0.00264	0.0106	0.25	0.00579	0.803
ADHD (Hyper)	-0.00982	0.0114	-0.863	-0.02	0.388
ASD	-0.061	0.0133	-4.6	-0.106	4.57E-06

--- Results for Feature: Negative Preferences---

Model Fit: $R^2 = 0.050$, Adjusted $R^2 = 0.047$

Overall Model: $F(6, 1798) = 15.771$, $p = 9.98e-18$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0108	0.00132	8.14	0.192	7.32E-16
IQ	0.000406	0.000234	1.73	0.0408	0.0836
Sex (Female)	0.0272	0.00839	3.24	0.0764	0.00121
ADHD (Inatt)	-0.00306	0.00879	-0.348	-0.00821	0.728
ADHD (Hyper)	-0.0106	0.00948	-1.11	-0.0263	0.266
ASD	-0.0158	0.011	-1.43	-0.0338	0.152

Supplementary Materials Section S1.5. Vocal Prosody Measures

Detailed model reports for panel B in Figure 5.

--- Results for Feature: Pitch ---

Model Fit: $R^2 = 0.172$, Adjusted $R^2 = 0.170$

Overall Model: $F(6, 2109) = 72.945$, $p = 7.27e-83$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-5.99	0.312	-19.2	-0.419	4.06E-76
IQ	-0.016	0.0547	-0.292	-0.00636	0.77
Sex (Female)	4.33	1.95	2.22	0.0483	0.0266
ADHD (Inatt)	0.198	2.04	0.0971	0.00211	0.923
ADHD (Hyper)	6.1	2.2	2.78	0.0605	0.00554
ASD	9.41	2.58	3.65	0.0794	0.000271

--- Results for Feature: Pitch Variation---

Model Fit: $R^2 = 0.177$, Adjusted $R^2 = 0.174$

Overall Model: $F(6, 2109) = 75.459$, $p = 1.52e-85$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-2.41	0.126	-19.1	-0.415	7.72E-75
IQ	0.0318	0.0222	1.44	0.0313	1.51E-01
Sex (Female)	1.47	0.791	1.86	0.0405	0.0633
ADHD (Inatt)	0.479	0.827	0.58	0.0126	0.562
ADHD (Hyper)	2.52	0.891	2.83	0.0616	4.72E-03
ASD	4.5	1.05	4.3	0.0937	1.77E-05

--- Results for Feature: Loudness---

Model Fit: $R^2 = 0.021$, Adjusted $R^2 = 0.018$

Overall Model: $F(6, 2109) = 7.474$, $p = 6.11e-08$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0745	0.0461	1.62	0.0352	0.106
IQ	-0.0169	0.00808	-2.09	-0.0456	0.0364
Sex (Female)	-0.684	0.288	-2.37	-0.0516	0.0178
ADHD (Inatt)	-0.767	0.302	-2.54	-0.0554	0.0111
ADHD (Hyper)	1.05	0.325	3.23	0.0702	1.28E-03
ASD	1.45	0.381	3.8	0.0828	1.49E-04

--- Results for Feature: Loudness Variation---

Model Fit: $R^2 = 0.029$, Adjusted $R^2 = 0.026$

Overall Model: $F(6, 2109) = 10.384$, $p = 2.27e-11$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.0922	0.0143	-6.44	-0.14	1.45E-10
IQ	-0.00428	0.00251	-1.7	-0.0371	0.0887
Sex (Female)	-0.274	0.0896	-3.06	-0.0666	0.00226
ADHD (Inatt)	0.0185	0.0937	0.197	0.00429	0.844
ADHD (Hyper)	-0.00554	0.101	-0.0549	-0.00119	0.956
ASD	0.26	0.118	2.2	0.0479	0.0281

--- Results for Feature: Dysphonia---

Model Fit: $R^2 = 0.052$, Adjusted $R^2 = 0.050$

Overall Model: $F(6, 2109) = 19.360$, $p = 4.39e-22$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	0.0134	0.00341	3.94	0.0857	8.54E-05
IQ	-0.00104	0.000598	-1.74	-0.0378	0.0824
Sex (Female)	-0.0243	0.0213	-1.14	-0.0248	0.255
ADHD (Inatt)	-0.0284	0.0223	-1.27	-0.0278	0.203
ADHD (Hyper)	0.0175	0.024	0.729	0.0159	4.66E-01
ASD	0.172	0.0282	6.1	0.133	1.26E-09

--- Results for Feature: Breathiness---

Model Fit: $R^2 = 0.028$, Adjusted $R^2 = 0.025$

Overall Model: $F(6, 2109) = 10.032$, $p = 5.94e-11$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.00235	0.000495	-4.76	-0.104	2.11E-06
IQ	-6.02E-05	8.68E-05	-0.694	-0.0151	0.488
Sex (Female)	0.00597	0.0031	1.93	0.042	0.0541
ADHD (Inatt)	-0.00432	0.00324	-1.33	-0.029	0.183
ADHD (Hyper)	0.00247	0.00349	0.707	0.0154	4.80E-01
ASD	0.0219	0.00409	5.36	0.117	9.46E-08

Supplementary Materials Section S1.6. Movement Measures

Detailed model reports for panel B in Figure 6.

--- Results for Feature: overall_mean ---

Model Fit: $R^2 = 0.136$, Adjusted $R^2 = 0.134$

Overall Model: $F(6, 2141) = 56.376$, $p = 6.9e-65$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.0474	0.00323	-14.7	-0.317	1.93E-46
IQ	-0.000256	0.000568	-0.452	-0.00976	0.652
Sex (Female)	-0.026	0.0202	-1.29	-0.0278	0.199
ADHD (Inatt)	-0.0177	0.0212	-0.835	-0.018	4.04E-01
ADHD (Hyper)	0.147	0.0229	6.41	0.139	1.80E-10
ASD	0.0739	0.0267	2.76	0.0598	0.00575

--- Results for Feature: eyes_mean ---

Model Fit: $R^2 = 0.142$, Adjusted $R^2 = 0.140$

Overall Model: $F(6, 2141) = 59.035$, $p = 7.78e-68$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.0489	0.00328	-14.9	-0.322	7.28E-48
IQ	-0.000178	0.000576	-0.309	-0.00669	7.57E-01
Sex (Female)	-0.0551	0.0205	-2.69	-0.058	0.0073
ADHD (Inatt)	-0.0224	0.0215	-1.04	-0.0225	2.97E-01
ADHD (Hyper)	0.143	0.0232	6.14	0.133	1.01E-09
ASD	0.0794	0.0271	2.93	0.0633	3.46E-03

--- Results for Feature: lips_mean ---

Model Fit: $R^2 = 0.130$, Adjusted $R^2 = 0.127$

Overall Model: $F(6, 2141) = 53.259$, $p = 2.07e-61$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.0449	0.00315	-14.3	-0.308	3.53E-44
IQ	-0.000272	0.000553	-0.492	-0.0106	0.623
Sex (Female)	-0.00655	0.0197	-0.333	-0.00719	7.39E-01
ADHD (Inatt)	-0.0135	0.0206	-0.654	-0.0141	5.13E-01
ADHD (Hyper)	0.145	0.0223	6.49	0.14	1.03E-10
ASD	0.0686	0.026	2.64	0.057	0.00845

--- Results for Feature: overall_mean_pose ---

Model Fit: $R^2 = 0.209$, Adjusted $R^2 = 0.207$

Overall Model: $F(6, 2155) = 95.086$, $p = 3.18e-106$

Predictor	Coefficient (β)	Std. Error	t-stat	Effect Size	p-value
Age	-0.06	0.0028	-21.4	-0.461	2.65E-92
IQ	-0.0013	0.000492	-2.65	-0.057	8.18E-03
Sex (Female)	0.0855	0.0175	4.88	0.105	1.14E-06
ADHD (Inatt)	0.0139	0.0184	0.756	0.0163	4.50E-01
ADHD (Hyper)	0.0967	0.0198	4.88	0.105	1.13E-06

Supplementary Materials Section S2. Diarizing Prompt for LLM

```
""You are given a transcript of an interview between a clinician (interviewer) and a subject. The interviewer asks
questions from a predefined list, though slight variations may occur. Your task is to label the dialogue as either
"Interviewer:" or "Subject:", removing timestamps while preserving the dialogue structure.

### Instructions:
1. Identify lines spoken by the **interviewer** based on a predefined question list. These questions may have slight
variations but follow the same general intent.
2. Label all other responses as **"Subject:"**
3. Remove timestamps but keep the text structure intact.
4. Ensure proper formatting with a new line between each labeled turn.

### Example Input:

0.709 -> 2.83: I hope you enjoyed the last movie about the puppy.
2.95 -> 3.851: Have you seen it before?
4.451 -> 4.751: No.
5.231 -> 6.792: Can you tell me what happened in the movie?
6.852 -> 8.013: Try to tell the whole story.
8.093 -> 11.515: Remember that stories have a beginning, things that happen, and an ending.
12.895 -> 15.877: I didn't watch the whole thing though.
```

```
16.637 -> 18.478: Can you tell me about the part that you did watch?
19.139 -> 19.639: Okay.

### Example Output:

Interviewer: I hope you enjoyed the last movie about the puppy.
Interviewer: Have you seen it before?
Subject: No.
Interviewer: Can you tell me what happened in the movie?
Interviewer: Try to tell the whole story.
Interviewer: Remember that stories have a beginning, things that happen, and an ending.
Subject: I didn't watch the whole thing though.
Interviewer: Can you tell me about the part that you did watch?
Subject: Okay.

### If the transcript contains these specific sentences, you will label them as "clip" instead of interviewer or subject:

Whoa! Cool!
You gotta be kidding me
Get lost
Mom, I'll/we'll be outside

### The interviewer question list:

So I hope you enjoyed the last movie. Have you seen it before?
So can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning,
things that happen, and an ending.
Do you remember anything else from the story?
So what are some of the things you liked about the movie?
What are some of the things you didn't like about the movie?
Who gave the boy a box?
What was in the box?
What was the boy doing before he got the box?
What was the puppy playing with?
How are the puppy and the boy the same?
So in the movie, who is missing a leg? The boy, the puppy, both the boy and the puppy, or no one?
So we're going to watch a short clip from the movie and then we'll talk about it.
How do you think the puppy was feeling?
How do you think the boy was feeling?
And how did you feel while you were watching that part?
How do you think the puppy was feeling?
How do you think the boy was feeling?
And how did you feel while you were watching that part?
How do you think the puppy was feeling?
How do you think the boy was feeling?
And how did you feel while you were watching that part?
How do you think the puppy was feeling?
How do you think the boy was feeling?
And how did you feel while you were watching this part?
Great, thank you

### In some cases, the interviewer may ask additional questions or make comments. You should label these as
"Interviewer:" as well using their specific context.

### The transcript:
<<transcript>>
```

Supplementary Materials Section S2. Question-Answer Extraction Prompt for LLM

```
"""Task: Extract Answers from Interview Transcript

You are tasked with extracting the participant's answers to a specific set of predefined questions from an interview
transcript. The questions relate to the short film "The Present." The interviewer *may* have asked these questions
verbatim, asked them in a slightly different way, or not asked them at all. The subject may have answered directly,
indirectly, partially, over multiple turns, or not at all.
In the transcript the interviewer asks these eleven questions in order:
1.  "I hope you enjoyed the last movie. Have you seen it before?"
2.  "Can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning,
things that happen, and an ending."
3.  "Do you remember anything else from the story?"
4.  "What are some of the things you liked about the movie?"
5.  "What are some of the things you didn't like about the movie?"
6.  "Who gave the boy a box?"
7.  "What was in the box?"
8.  "What was the boy doing before he got the box?"
9.  "What was the puppy playing with?"
10. "How are the puppy and the boy the same?"
11. "In the movie, who is missing a leg? The boy, the puppy, both the boy and the puppy, or no one?"

After these questions, the interviewer then shows the subject four short clips from the movie and asks the subject to
describe how the characters are feeling in each clip. The interviewer may say something like:
- "Let's watch a short clip from the movie, and then we'll talk about it."
- "Now let's watch another clip"
- "One more clip"
- "One last clip"

After watching each clip, the interviewer asks the subject how they think the characters are feeling and how they
themselves feel while watching the clip. The interviewer may ask these questions in various ways, such as:
- "How do you think the puppy was feeling?"
- "How do you think the boy was feeling?"
- "And how did you feel while you were watching that part?"


Input:

*   A transcript of a conversation between an interviewer and a subject discussing the short film "The Present."

Output:

Your output will consist of two parts:

**Part 1: Question-Answer Extraction**

A series of lines, one for each predefined question. Each line will follow this format:

`"Question" -> "Answer"`

Where:

*   `"Question"` is one of the predefined questions listed below (use the *exact* wording provided here, even if the
interviewer phrased it differently).
*   `"Answer"` is the subject's answer to the question.
```

```
    *   If the question was answered (directly or indirectly), provide the subject's complete answer, concatenating all
relevant parts of their response, even if it spans multiple turns or requires piecing together information from different
parts of the transcript.
    *   If the question was *not* asked, or if the subject provided *no* discernible answer (even after considering the
surrounding context), leave the "Answer" field *blank*. Do *not* write "N/A," "No answer," or any other placeholder. Just
leave it blank.
    * If parts of different answers are mixed, extract only the answer to the intended question.

Predefined Questions:

1.  "I hope you enjoyed the last movie. Have you seen it before?"
2.  "Can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning,
things that happen, and an ending."
3.  "Do you remember anything else from the story?"
4.  "What are some of the things you liked about the movie?"
5.  "What are some of the things you didn't like about the movie?"
6.  "Who gave the boy a box?"
7.  "What was in the box?"
8.  "What was the boy doing before he got the box?"
9.  "What was the puppy playing with?"
10. "How are the puppy and the boy the same?"
11. "In the movie, who is missing a leg? The boy, the puppy, both the boy and the puppy, or no one?"
12. "How do you think the puppy was feeling? (after watching first clip)"
13. "How do you think the boy was feeling? (after watching first clip)"
14. "And how did you feel while you were watching that part? (after watching first clip)"
15. "How do you think the puppy was feeling? (after watching second clip)"
16. "How do you think the boy was feeling? (after watching second clip)"
17. "And how did you feel while you were watching that part? (after watching second clip)"
18. "How do you think the puppy was feeling? (after watching third clip)"
19. "How do you think the boy was feeling? (after watching third clip)"
20. "And how did you feel while you were watching that part? (after watching third clip)"
21. "How do you think the puppy was feeling? (after watching fourth clip)"
22. "How do you think the boy was feeling? (after watching fourth clip)"
23. "And how did you feel while you were watching that part? (after watching fourth clip)"

**Part 2: Conversation Completion Rating**

A single line in the following format:

`"Conversation Completion Rating" -> "Rating"`

Where:

*   `"Rating"` is a numerical score between 0 and 1 (inclusive) representing the overall completion of the conversation,
*specifically with respect to the predefined questions*.
    *   **1:** Represents a perfect completion, where all of the predefined questions were asked (or closely
paraphrased), and the subject provided clear and complete answers to each.
    *   **0:** Represents a very poor completion, where many questions were not asked, or the subject consistently failed
to answer the questions that were asked.
    *   **Intermediate values:** Represent varying degrees of completion quality.  A score of 0.5, for example, would
suggest that roughly half of the questions were asked and answered reasonably well. Use your judgment to assign a score
that reflects the overall completeness and clarity of the question-answer exchange related to the *predefined questions*.
Do *not* assess the overall quality of the conversation beyond its adherence to addressing the core content of these
questions.

Instructions:
```

1. **Read the Entire Transcript:** Carefully read the entire transcript to understand the completion of the conversation and the context of each statement.
2. **Identify Questions and Answers:** For each predefined question:
    *   **Search for the Question:** Look for the exact question, or a close paraphrase, in the interviewer's speech. Note that the interviewer might rephrase, prompt, or interrupt the subject.
    *   **Locate the Answer (If Present):** If the question (or a close variant) was asked, carefully examine the subject's subsequent responses. The answer might be:
        *   **Direct and Immediate:** Right after the question.
        *   **Indirect:** Implied by their statements, requiring inference.
        *   **Partial:** Only part of the answer is given directly.
        *   **Scattered:** Pieces of the answer are given across multiple turns, possibly with interviewer prompts or interruptions.
        *   **Non-existent:** The subject might not answer at all.
    * **Use Context:** Consider surrounding conversational turns from both interviewer and subject to determine if a particular utterance constitutes an answer, even an indirect or partial one.
    * **Combine Response Parts:** If the answer is spread over multiple turns, concatenate *all relevant parts* of the subject's response into a single, coherent answer. Remove any interviewer interjections or prompts from within the concatenated answer. Only include the *subject's* words in the "Answer" field.
4. * **For the questions after the clips:** Because there are multiple clips, you will need to extract the answers to these questions from the conversation after each clip is shown. The interviewer will ask the subject how they think the characters are feeling and how they themselves feel while watching the clip. However, the order of the clips is always the same. Sometimes part of the dialogue is present on the transcript, so in those cases, use that as clue to know to what clip the question is referring to. If you can't clearly now what question/answer corresponds to what clip, restrain from returning and answer for that specific question, leave the "Answer" field blank.:
    * **Clip 1:**: the kid opens the present and sees the puppy, the transcript might say "Whoa, cool."
    * **Clip 2:**: the kid throws the puppy away, the transcript might say "You've got to be kidding me."
    * **Clip 3:**: the kid kicks the puppy, the transcript might say "Get lost!"
    * **Clip 4:**: the kid plays with the puppy, the transcript might say "Mom, I'll / We'll be outside."
3. **Output (Part 1):** Generate the question-answer extraction in the specified `"Question" -> "Answer"` format. Leave the "Answer" field blank if the question was not asked or if no answer can be found.
4. **Assess Conversation Completion (Part 2):** After extracting the question-answer pairs, evaluate the *overall completion* of the conversation *with respect to the predefined questions*. Consider:
    * How many of the predefined questions (or close paraphrases) were asked by the interviewer?
    * Did the conversation stay focused on the topics covered by the predefined questions, or did it frequently deviate?
5. **Output (Part 2):** Provide the "Conversation Completion Rating" on a separate line, using a number between 0 and 1.

Example 1:

    (example 1 transcript):
        Interviewer: So I hope you enjoyed the last movie.
        Interviewer: Have you seen it before?
        Subject: I have at the mug MRI.
        Interviewer: Oh, cool.
        Subject: I liked it, though.
        Interviewer: Great.
        Interviewer: So can you tell me what happened in the movie?
        Interviewer: Try to tell the whole story.
        Interviewer: Remember that stories have a beginning, things that happen, and an ending.
        Subject: So there's a boy who was disabled, and he didn't want to go outside.
        Subject: He wasn't in the mood for anything except for his video games.
        Subject: and then like his mom was trying to like help him go like get some fresh air and then but like everything like his mom tried to do didn't work so then he then she bought um him a puppy so then he started like he didn't want like to see the puppy like anymore because then he realized he realized that he was also like disabled
        Subject: But then he started having a liking to him because he felt like the puppy and him were like the same.
        Subject: So he started playing with the puppy more.
        Subject: So then they went outside and they started playing with the ball.
        Interviewer: Great.
        Interviewer: Do you remember anything else from the story?

```
        Subject: I think that's much it.
        Interviewer: Great.
        Interviewer: So what are some of the things you liked about the movie?
        Subject: I liked how he had a good relationship with his doll.
        Interviewer: What are some of the things you didn't like about the movie?
        Subject: I liked everything about the movie, but
        Subject: I was hoping it would be longer.
        Interviewer: Who gave the boy a box?
        Subject: His mother.
        Interviewer: What was in the box?
        Subject: The puppy.
        Interviewer: What was the boy doing before he got the box?
        Subject: He wasn't really paying attention to anything.
        Subject: He was glued to his video games.
        Interviewer: What was the puppy playing with?
        Subject: He was playing with the red rubber ball.
        Interviewer: How are the puppy and the boy the same?
        Subject: Both of them each have a disabled part of their body.
        Interviewer: So in the movie, who is missing a leg?
        Interviewer: The boy, the puppy, both the boy and the puppy, or no one?
        Subject: The boy.
        Interviewer: Great.
        Interviewer: So we're going to watch a short clip from the movie and then we'll talk about it.
        Subject: Okay.
        clip: Whoa, cool.
        Interviewer: How do you think the puppy was feeling?
        Subject: Happy.
        Interviewer: How do you think the boy was feeling?
        Subject: He's feeling happy too.
        Interviewer: And how did you feel while you were watching that pot?
        Subject: I guess happy.
        Interviewer: Right.
        Interviewer: How do you think the puppy was feeling?
        Subject: Sad.
        Interviewer: How do you think the boy was feeling?
        Subject: He was feeling angry.
        Interviewer: And how did you feel while you were watching that mom?
        Subject: I didn't like it.
        Interviewer: How do you think the puppy was feeling?
        Subject: He didn't mind.
        Interviewer: How do you think the boy was feeling?
        Subject: He was still angry.
        Interviewer: And how did you feel while you were watching that pup?
        Subject: I feel bad for the puppy.
        Interviewer: Great.
        clip: Mom, I'll be outside.
        Interviewer: How do you think the puppy was feeling?
        Subject: Um, happy.
        Interviewer: How do you think the boy was feeling?
        Subject: He was feeling happy too.
        Interviewer: And how did you feel while you were watching this pot?
        Subject: I was happy.
        Interviewer: Great, thank you.

    (example 1 output):
        "I hope you enjoyed the last movie. Have you seen it before?" -> "I have at the mug MRI."
        "Can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning,
things that happen, and an ending." -> "So there='s a boy who was disabled, and he didn't want to go outside. He wasn't
```

in the mood for anything except for his video games. And then like his mom was trying to like help him go like get some fresh air and then but like everything like his mom tried to do didn't work so then he then she bought um him a puppy so then he started like he didn't want like to see the puppy like anymore because then he realized he realized that he was also like disabled. But then he started having a liking to him because he felt like the puppy and him were like the same. So he started playing with the puppy more. So then they went outside and they started playing with the ball."

"Do you remember anything else from the story?" -> "I think that's much it."

"What are some of the things you liked about the movie?" -> "I liked how he had a good relationship with his doll."

"What are some of the things you didn't like about the movie?" -> "I liked everything about the movie, but I was hoping it would be longer."

"Who gave the boy a box" -> "His mother."

"What was in the box?" -> "The puppy."

"What was the boy doing before he got the box?" -> "He wasn't really paying attention to anything. He was glued to his video games."

"What was the puppy playing with?" -> "He was playing with the red rubber ball."

"How are the puppy and the boy the same?" -> "Both of them each have a disabled part of their body."

"In the movie, who is missing a leg? The boy, the puppy, both the boy and the puppy, or no one?" -> "The boy."

"How do you think the puppy was feeling? (after watching first clip)" -> "Happy."

"How do you think the boy was feeling? (after watching first clip)" -> "He's feeling happy too."

"And how did you feel while you were watching that part? (after watching first clip)" -> "I guess happy."

"How do you think the puppy was feeling? (after watching second clip)" -> "Sad."

"How do you think the boy was feeling? (after watching second clip)" -> "He was feeling angry."

"And how did you feel while you were watching that part? (after watching second clip)" -> "I didn't like it."

"How do you think the puppy was feeling? (after watching third clip)" -> "He didn't mind."

"How do you think the boy was feeling? (after watching third clip)" -> "He was still angry."

"And how did you feel while you were watching that part? (after watching third clip)" -> "I feel bad for the puppy"

"How do you think the puppy was feeling? (after watching fourth clip)" -> "Um, happy."

"How do you think the boy was feeling? (after watching fourth clip)" -> "He was feeling happy too."

"And how did you feel while you were watching that part? (after watching fourth clip)" -> "I was happy."

"Rating" -> "1"

Example 2:

(example 2 transcript):

Interviewer: Okay, so I hope you enjoyed the last movie.

Interviewer: Have you seen it before?

Subject: No

Interviewer: Um, so I'm going to need you to talk a little bit.

Interviewer: Is that okay?

Subject: Okay.

Interviewer: Can you tell me what happened in the movie?

Interviewer: Try to tell the whole story.

Interviewer: Remember that stories have a beginning, things that happen, and an ending.

Interviewer: Great.

Interviewer: Do you remember anything else from the story?

Interviewer: Okay.

Interviewer: What are some of the things you liked about the movie?

Interviewer: And what are some of the things you didn't like about the movie?

Interviewer: So who gave the boy a box?

Interviewer: What was in the box?

Interviewer: What was the boy doing before he got the box?

Interviewer: What was the puppy playing with?

Interviewer: How are the puppy and the boy the same?

Interviewer: Perfect.

Interviewer: So in the movie, who is missing a leg?

Interviewer: The boy, the puppy, both the boy and the puppy, or no one?

Subject: Both the boy and the puppy.

```
        Interviewer: Great.
        Interviewer: So we're going to watch a couple of short clips from the movie and talk about them, OK?
        Subject: OK.
        clip: Whoa.
        clip: Cool.
        Interviewer: How do you think the puppy was feeling?
        Subject: Happy.
        Interviewer: How do you think the boy was feeling?
        Subject: Happy.
        Interviewer: Happy?
        Interviewer: And how did you feel while you were watching this part?
        Subject: Good.
        Interviewer: Good?
        Interviewer: Great.
        clip: You've got to be kidding me.
        Interviewer: How do you think the puppy was feeling?
        Interviewer: And how do you think the boy was feeling?
        Interviewer: And how did you feel while you were watching this part?
        clip: Get lost!
        Interviewer: How do you think the puppy was feeling?
        Interviewer: And how do you think the boy was feeling?
        Interviewer: And how did you feel while you were watching this pot?
        Interviewer: Okay, one more.
        clip: Mom!
        clip: We'll be outside!
        Interviewer: How do you think the puppy was feeling?
        Subject: Happy.
        Interviewer: How do you think the boy was feeling?
        Subject: Happy.
        Interviewer: And how did you feel while you were watching his paw?
        Subject: Happy.
        Interviewer: Great, thank you.


    (example 2 output):
        "I hope you enjoyed the last movie. Have you seen it before?" -> "No."
        "Can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning,
things that happen, and an ending." -> ""
        "Do you remember anything else from the story?" -> ""
        "What are some of the things you liked about the movie?" -> ""
        "What are some of the things you didn't like about the movie?" -> ""
        "Who gave the boy a box" -> ""
        "What was in the box?" -> ""
        "What was the boy doing before he got the box?" -> ""
        "What was the puppy playing with?" -> ""
        "How are the puppy and the boy the same?" -> ""
        "In the movie, who is missing a leg? The boy, the puppy, both the boy and the puppy, or no one?" ->  "Both the
boy and the puppy."
        "How do you think the puppy was feeling? (after watching first clip)" -> "Happy."
        "How do you think the boy was feeling? (after watching first clip)" -> "Happy."
        "And how did you feel while you were watching that part? (after watching first clip)" -> "Good."
        "How do you think the puppy was feeling? (after watching second clip)" -> ""
        "How do you think the boy was feeling? (after watching second clip)" -> ""
        "And how did you feel while you were watching that part? (after watching second clip)" -> ""
        "How do you think the puppy was feeling? (after watching third clip)" -> ""
        "How do you think the boy was feeling? (after watching third clip)" -> ""
        "And how did you feel while you were watching that part? (after watching third clip)" -> ""
        "How do you think the puppy was feeling? (after watching fourth clip)" -> "Happy."
```

```
        "How do you think the boy was feeling? (after watching fourth clip)" -> "Happy."
        "And how did you feel while you were watching that part? (after watching fourth clip)" -> "Happy."
        "Rating" -> "0.1"

    (example 3 transcript):
        Interviewer: Have you seen that last movie before, the puppy one?
        Subject: I don't think I have.
        Interviewer: Can you try and tell me what happens in the movie?
        Interviewer: The full beginning, middle, and end.
        Subject: Um, the beginning, the beginning of the movie is, um, so it's the middle, mom gives him the puppy, but
he doesn't like it on first because it has three legs, and then towards the end,
        Subject: The last time I played with them.
        Subject: I don't know about it.
        Subject: All right, we both have the same disability.
        Subject: Okay.
        Subject: Didn't go home.
        Subject: How deep?
        Interviewer: What was the boy doing before he got the box?
        Interviewer: What was the puppy playing with?
        Interviewer: How are the puppy and the boy the same?
        Interviewer: In the movie, who's missing a leg?
        Interviewer: The boy, the puppy, both the boy and the puppy or no one?
        Interviewer: Okay, now we're going to watch a short clip and then we'll talk about it.
        Interviewer: How's the puppy feeling here?
        Interviewer: How's the boy feeling?
        Interviewer: And how do you feel when you watch his part?
        Interviewer: How's the puppy feeling here?
        Interviewer: How's the boy feeling?
        Interviewer: And how do you feel when you watch this part?
        Interviewer: How's the puppy feeling here?
        Interviewer: How's the boy feeling?
        Interviewer: How do you feel when you watch his fart?
        Interviewer: How's the puppy feeling here?
        Interviewer: How's the boy feeling?
        Interviewer: And how do you feeling off that part?
        Subject: Good.
        Interviewer: Okay, we're all done.

    (example 3 output):
        "I hope you enjoyed the last movie. Have you seen it before?" -> "I don't think I have."
        "Can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning,
things that happen, and an ending." -> "Um, the beginning, the beginning of the movie is, um, so it's the middle, mom
gives him the puppy, but he doesn't like it on first because it has three legs, and then towards the end, The last time I
played with them. I don't know about it. All right, we both have the same disability. Okay. Didn't go home. How deep?"
        "Do you remember anything else from the story?" -> ""
        "What are some of the things you liked about the movie?" -> ""
        "What are some of the things you didn't like about the movie?" -> ""
        "Who gave the boy a box" -> ""
        "What was in the box?" -> ""
        "What was the boy doing before he got the box?" -> ""
        "What was the puppy playing with?" -> ""
        "How are the puppy and the boy the same?" -> ""
        "In the movie, who is missing a leg? The boy, the puppy, both the boy and the puppy, or no one?" ->  ""
        "How do you think the puppy was feeling? (after watching first clip)" -> ""
        "How do you think the boy was feeling? (after watching first clip)" -> ""
        "And how did you feel while you were watching that part? (after watching first clip)" -> ""
        "How do you think the puppy was feeling? (after watching second clip)" -> ""
        "How do you think the boy was feeling? (after watching second clip)" -> ""
```

```
    "And how did you feel while you were watching that part? (after watching second clip)" -> ""
    "How do you think the puppy was feeling? (after watching third clip)" -> ""
    "How do you think the boy was feeling? (after watching third clip)" -> ""
    "And how did you feel while you were watching that part? (after watching third clip)" -> ""
    "How do you think the puppy was feeling? (after watching fourth clip)" -> ""
    "How do you think the boy was feeling? (after watching fourth clip)" -> ""
    "And how did you feel while you were watching that part? (after watching fourth clip)" -> "Good."
    "Rating" -> "0"

(example 4 transcript):
    Interviewer: I hope you enjoyed the last movie.
    Interviewer: Have you seen it before?
    Subject: Yeah.
    Subject: The puppy cartoon?
    Subject: Yeah.
    Interviewer: Can you tell me what happened in the movie and try to tell the whole story?
    Interviewer: Remember that stories have a beginning, things that happen, and an ending.
    Subject: So first there was a boy that was playing a game.
    Subject: So when his mother came in the door was a box.
    Subject: She put it down and then a puppy was in the box.
    Subject: But the boy didn't want to play.
    Subject: Instead he wanted to play on his TV.
    Subject: So first he just kicked the puppy, so then he just came back and he did it again.
    Subject: So then he just went crazy and then got a ball.
    Subject: And then he wanted to play, but then the boy just kicked the ball in the box and the puppy ran with it.
    Subject: And then
    Subject: he dropped the ball back.
    Subject: And then the boy stopped when he was playing, and then he went to lose the puppy.
    Interviewer: Do you remember anything else from the story?
    Subject: That's all I can remember.
    Interviewer: What are some of the things you liked about the movie?
    Subject: When you like
    Subject: And then all of the puppies ran with it and then he got trapped in the box.
    Subject: When he kicked the puppy.
    Subject: His mom.
    Subject: A puppy.
    Subject: Playing on his TV.
    Subject: A ball.
    Interviewer: How are the puppy and the boy the same?
    Subject: They're both playing with them.
    Interviewer: In the movie, who is missing a leg?
    Interviewer: The boy, the puppy, both the boy and the puppy, or no one?
    Subject: The boy and the puppy.
    Interviewer: Let's watch a short clip from the movie, and then we'll talk about it.
    Interviewer: Give me one second, bud.
    Subject: Can we do an interview on it?
    Interviewer: This is the interview.
    Interviewer: Okay, here we go.
    Interviewer: You can sit back.
    Interviewer: Can you sit in your chair?
    Interviewer: How was the puppy feeling?
    Subject: Happy.
    Interviewer: How was the boy feeling?
    Subject: Happy.
    Interviewer: How did you feel when you watched that part?
    Subject: Happy.
    Subject: Ugh.
    clip: You've just got to be kidding me.
```

```
        Interviewer: How was the puppy feeling?
        Subject: He was feeling bad.
        Interviewer: How was the boy feeling?
        Subject: He was feeling, um, angry.
        Interviewer: How did you feel when you watched this part?
        Subject: Bad.
        Subject: Sad.
        Subject: Angry.
        Subject: Boys.
        Subject: Something.
        Subject: Get the boy and the puppy.
        Subject: I miss the boy.
        Interviewer: How's the puppy feeling?
        Subject: Very happy.
        Interviewer: How's the boy feeling?
        Subject: Super happy.
        Interviewer: How did you feel while you were watching that part?
        Subject: Happy.
        Interviewer: That's all, thank you.

    (example 4 output):
        "I hope you enjoyed the last movie. Have you seen it before?" -> "Yeah, The puppy cartoon? Yeah."
        "Can you tell me what happened in the movie? Try to tell the whole story. Remember that stories have a beginning,
things that happen, and an ending." -> "So first there was a boy that was playing a game. So when his mother came in the
door was a box. She put it down and then a puppy was in the box. But the boy didn't want to play. Instead he wanted to
play on his TV. So first he just kicked the puppy, so then he just came back and he did it again. So then he just went
crazy and then got a ball. And then he wanted to play, but then the boy just kicked the ball in the box and the puppy ran
with it. And then he dropped the ball back. And then the boy stopped when he was playing, and then he went to lose the
puppy."
        "Do you remember anything else from the story?" -> "That's all I can remember."
        "What are some of the things you liked about the movie?" -> "When you like. And then all of the puppies ran with
it and then he got trapped in the box."
        "What are some of the things you didn't like about the movie?" -> "When he kicked the puppy."
        "Who gave the boy a box" -> "His mom"
        "What was in the box?" -> "A puppy"
        "What was the boy doing before he got the box?" -> ""Playing on his TV."
        "What was the puppy playing with?" -> "A ball."
        "How are the puppy and the boy the same?" -> "They're both playing with them."
        "In the movie, who is missing a leg? The boy, the puppy, both the boy and the puppy, or no one?" ->  "The boy and
the puppy."
        "How do you think the puppy was feeling? (after watching first clip)" -> "Happy."
        "How do you think the boy was feeling? (after watching first clip)" -> "Happy."
        "And how did you feel while you were watching that part? (after watching first clip)" -> "Happy. Ugh."
        "How do you think the puppy was feeling? (after watching second clip)" -> "He was feeling bad."
        "How do you think the boy was feeling? (after watching second clip)" -> "He was feeling, um, angry."
        "And how did you feel while you were watching that part? (after watching second clip)" -> "Bad."
        "How do you think the puppy was feeling? (after watching third clip)" -> "Sad."
        "How do you think the boy was feeling? (after watching third clip)" -> "Angry."
        "And how did you feel while you were watching that part? (after watching third clip)" -> ""
        "How do you think the puppy was feeling? (after watching fourth clip)" -> "Very happy."
        "How do you think the boy was feeling? (after watching fourth clip)" -> "Super happy."
        "And how did you feel while you were watching that part? (after watching fourth clip)" -> "Happy."
        "Rating" -> "0.8"


### This is the transcript you have to process:
<<<transcript>>>
"""
```