



Published in final edited form as:

*Autism Res.* 2019 January ; 12(1): 89–99. doi:10.1002/aur.2004.

## Autism Spectrum Disorder symptoms from 2 to 19: Implications for diagnosing adolescents and young adults.

**Vanessa H. Bal, PhD,**

Department of Psychiatry, Weill Institute for Neurosciences, University of California San Francisco

**So-Hyun Kim, PhD,**

Department of Psychiatry, Weill Cornell Medicine

**Megan Fok, and**

Department of Psychiatry, Weill Institute for Neurosciences, University of California San Francisco

**Catherine Lord, PhD**

Department of Psychiatry, Weill Cornell Medicine

### Abstract

This study explored change in social-communicative symptoms in 140 individuals with childhood Autism Spectrum Disorder (ASD) diagnoses. Trajectories of caregiver-reported social-communicative symptoms were examined for three groups (verbal, delayed speech, minimally verbal) from ages 2 to 19 years. Groups showed comparable levels of social-communicative impairment at 2 and significant decreases in overall symptom levels across the 17-year period ( $p < .001$ ). Across three subdomains, main effects of time and language ( $p < .001$ ) reflected patterns of overall improvement, though children with more impaired language tended to have more caregiver-reported symptoms relative to verbal peers. A significant time-by-language interaction ( $p < .001$ ) reflected that trajectories of socioemotional reciprocity symptoms differed according to patterns of language development. In contrast, improvements in the nonverbal communication domain were seen across language groups, whereas deficits in the development and maintenance of relationships improved for only verbal children. Verbal adults showed significant reductions in the prevalence of several symptoms exhibited during childhood. Improvements suggest that symptoms indicative of ASD in young children may no longer be diagnostic markers in adolescents and adults. Relative stability of several items suggests that impaired facial expression may be a core ASD symptom that warrants more systematic study across the lifespan. Research investigating the manifestation of ASD in older individuals is needed to foster development of appropriate assessment tools and interventions. Differential relationships to developmental factors within the broader social-communication domain underscores a need to focus on more narrowly defined symptom constructs when exploring links between pathophysiology and observable phenotypes.

### Lay Summary

In a sample of 140 participants with Autism Spectrum Disorder (ASD) followed from 2 to 19 years old, this study found that overall social-communicative symptoms improve across childhood and adolescence. However, timing and amount of change varied for different symptom categories and participants with different language abilities. Findings suggest that some older adolescents and

adults with ASD may not exhibit the same difficulties observed in young children with ASD. More research is needed to better understand the strengths and needs of young adults with ASD.

### Keywords

adolescent; adult; symptom; longitudinal

---

### Introduction

Core features of Autism Spectrum Disorder (ASD) vary widely by age, cognitive ability and language level (Gotham et al., 2007; Lord et al., 1999; Sturm et al., 2017). Less is known about how the manifestation of ASD symptoms changes from early childhood to later adolescence and young adulthood. This has important implications for the assessment of ASD in later adolescence and adulthood, as instruments based on what is known about childhood ASD may not be as sensitive to the manifestation of ASD in older individuals. Taking into account expressive language level is a particularly important factor in assessing social-communicative deficits in children with ASD (Klein-Tasman et al., 2007; Gotham et al., 2007). While it may be expected that some symptoms change with ongoing development of expressive language, other symptoms might remain stable regardless of language level. The current study examines change in ASD social-communication symptoms in relation to level of spoken language in a longitudinal cohort of individuals followed from ages 2 to 19 years.

A handful of longitudinal studies have examined ASD symptom trajectories across childhood and early adolescence. For example, in a sample of 345 children followed from 2 to 15 years of age, Gotham and colleagues (2012) identified four latent classes using Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999; Lord et al., 2012) calibrated severity scores (Gotham et al., 2009), which encompass social-communication and repetitive behaviors. The majority of children had persistently high or moderate levels of symptoms, while 9% showed worsening skills and 7% showed improvement. All groups showed a trend toward increasing verbal IQ (VIQ), but the improving class had both earlier and greater overall VIQ gains compared to the other groups, suggesting a role for verbal abilities in ASD symptom changes.

Several other studies have examined change in caregiver-reported symptoms on the Autism Diagnostic Interview-Revised (ADI-R; Rutter et al., 2003) across childhood, adolescence and adulthood. In 48 children diagnosed with ASD between 2–5 years of age, McGovern & Sigman (2005) reported declines in social impairments between approximately 13 and 19 years of age, with larger improvements observed for individuals with IQs above 70. In a subsample of the same cohort assessed again at 26 years of age (n=20), Gillespie-Lynch and colleagues (2012) reported decreases in social-communicative symptoms, with early childhood joint attention and language skills predicting adult impairment. In another cohort, Shattuck and colleagues (2007) reported improvements in the social, but not nonverbal communication, scores over a 4.5 year period. The proportion of items endorsed at follow-up suggested lower “prevalence” of several symptoms. Although this sample was large

(n=241), participant age at initial visit ranged from 10 to 52 years, making it difficult to determine the timing of symptom decreases.

Taken together, the extant literature provides some evidence to suggest that ASD symptoms change across childhood and adolescence. However, many studies have been limited by small samples or large age ranges at each time point, making it difficult to pinpoint the timing of skill development. Moreover, most have focused on broad domains of social-communication, making it difficult to delineate symptoms improving over time from features that remain more stable across the lifespan. Finally, although expressive language is known to be an important factor in the assessment of ASD symptoms, studies have not systematically examined how symptom trajectories relate to language development over time. The current study aimed to: 1) Investigate how trajectories of social-communicative symptoms across childhood and adolescence vary in groups defined by different patterns of language development; and 2) Demonstrate implications of social-communicative symptom change for the assessment of DSM-5 ASD criteria in young adults. While understanding trajectories of restricted and repetitive behaviors (RRBs) will be necessary to inform assessment of adults, the complex nature of this class of behaviors merits individual attention; therefore, RRBs were not included in the present analyses.

## Methods

### Participants

Participants were drawn from a prospective longitudinal sample of 192 children referred for assessment of possible ASD prior to 37 months to clinics in Chicago and North Carolina (Lord et al., 2006). Participants were assessed at approximately ages 2, 3, 9 and 19; children in North Carolina were also seen at age 5. The current study includes 140 children who received a childhood ASD diagnosis and whose caregivers completed an ADI-R at T2, T3 and either T9 or T19. The sample was predominantly male (87%) and White (70%) or African American (29%). Table 1 provides additional participant characteristics.

At T19, 98/140 caregivers were interviewed using the ADI-R; 95 participants also completed an in-person assessment (ADOS, cognitive test). Best estimate clinical diagnoses were made based on all available information regarding current presentation; assessors were blinded to previous diagnoses. At T19, 7 participants were given non-ASD diagnoses (Fragile X, idiopathic Intellectual Disability, Tourette's), 8 individuals no longer met criteria for any DSM-5 clinical diagnosis and 3 were not given diagnoses because they did not complete an in-person assessment. All 140 participants were included in the trajectory analyses, as they were all diagnosed with ASD in childhood, making their inclusion directly relevant to examination of symptom change. Analyses to demonstrate implications of symptom change for assessment of DSM-5 criteria were limited to the 80 participants who received clinical ASD diagnoses at T19, as it was of interest to compare symptom reports in early childhood and young adulthood to identify which symptoms may no longer be markers for ASD in young adults with ASD.

## Measures

The ADI-R (Rutter et al., 2003) is a semi-structured caregiver interview designed to assess current and past ASD symptoms. The Diagnostic Algorithm is organized into three behavioral domains based on DSM-IV and ICD-10 criteria for Autism: Reciprocal Social, Communication and RRBs. A fourth domain indicates whether developmental abnormalities were present before 3 years old. For individuals 4 years and older, domain cut-offs for an instrument classification of “Autism” are based on scores reflecting past behavior. A Current Behavior algorithm comprising scores reflecting behavior in the past three months is also available for treatment planning, but not intended for diagnostic purposes. The Current Behavior domains overlap with the Diagnostic Algorithm, but items vary by age.

Level of language and social-communicative symptoms were derived from the ADI-R as follows:

**Language Group.**—Language was derived from ADI-R item #30, Overall Level of Language (OLL), which captures the caregiver-reported current level of speech. OLL is coded as follows: 2=Fewer than 5 words or speech not used daily; 1=No functional 3-word phrases, but uses at least 5 words daily; 0=Daily, functional use of phrases (3+ words, sometimes including a verb). The majority of children were language delayed at T2 (i.e., 95%>0); however, there was more variability by T3 (i.e., 35%=0). Thus, children were divided into three groups based upon caregiver-reported language at T3 and T19: Verbal-Verbal (V-V) were 49 children using phrases (i.e., OLL=0) at both assessments, Delayed-Verbal (D-V) were 46 children using single words or less (OLL>0) at T3 and phrases by 19 (OLL=0) and Delayed-Minimally Verbal (D-MV) were 45 children using single words or less (OLL>0) at both time points. Language trajectories are shown in Figure S1. Caregiver-reported language level from T9-T19 was stable for 87% of the 83 participants seen at both time points; T9 language level was used for 15 participants not seen at T19. Use of ADI-R OLL to categorize children by broad language abilities shows good agreement with direct measures of language ability (Bal et al., 2016) and may be more appropriate than age estimates from standardized language tests (Lord & Pickles, 1996).

**Social-Communication Symptoms.**—The focus of analyses was 14 items (10 from the Reciprocal Social domain and 4 from the Communication domain) on the Current Behavior Algorithm that are scored for participants of all ages and language levels. Items are scored on a scale of 0–3. Per ADI-R conventions, scores of 2 and 3 were collapsed. Higher scores indicate presence of abnormality, therefore declines in scores over time signify improvement or decrease in symptoms. All 14 items were summed to provide an overall social-communication total. To facilitate interpretation of implications for current assessment of young adults, items were organized into three social-communication subdomains mirroring DSM-5 ASD criteria (deficits in socioemotional reciprocity [Figure 2], nonverbal communicative behaviors [Figure 3] and development/maintenance of social relationships [Figure 4]) recommended by Huerta and colleagues (2012; item-to-subdomain mapping is shown in Table S1 and Figures 2–4). Validity of these groupings was assessed using Confirmatory Factor Analysis in MPlus (Muthén & Muthén, 1998). Comparative Fit Index was .98 at both T3 and T19, indicating a good fit (Skondral & Rabe-Hesketh, 2004).

**Covariates.**—NVIQ and demographics were included as covariates due to differences in T2 symptom levels for children with lower NVIQ and higher maternal education (see statistical analyses below) and the previously documented higher attrition rate of African American families with lower maternal education in this sample (Anderson et al., 2014). Nonverbal IQ (NVIQ) at T3 was derived from the Mullen Scales of Early Learning (Mullen, 1997; n=134) or the Differential Ability Scales (DAS; Elliott, 1990; n=6). Race and maternal education were reported by a caregiver on a demographic questionnaire. Race categories were collapsed to compare white to non-white and maternal education compared Bachelor's degree or higher to some college or less.

**Treatment.**—Amount and type of treatment accessed were examined in post hoc analyses. Information regarding treatment was collected via caregiver logs and interviews at T2-T9. The number of treatment hours was coded for four categories: mentored, parent-implemented structured teaching (MPST; a home program modeled after the TEACCH extended diagnostic services; see (Mesibov et al., 2005), Applied Behavior Analysis, Speech and Other (i.e., occupational, music, etc). In addition, five classroom placements were recorded: special autism class, multi-categorical, regular education with resource room, regular education with aide and regular education without support.

### Statistical Analyses

All analyses were restricted to individuals who were evaluated using the ADI-R at least three times: T2, T3, and either T9 (n=132) or T19 (n=98). Out of 700 interviews, item-level missing data for 112 interviews were computed based on a regression analysis with multiple imputation using a Markov chain Monte Carlo method (Raghunathan et al, 2001; Reynolds, 2000) allowing prediction from the 14 items available at each of the 5 time points, T3 NVIQ and demographic factors (gender, race, maternal education). One hundred imputed datasets using 30 iterations were generated, each with different imputed values to reflect uncertainty of their true value. Estimates of means and other parameters of interest were obtained by averaging results from each dataset and standard errors calculated that accounted for between- and within-imputation variation (Rubin, 1987). There were no differences in gender or level of language for children with and without imputed data. Children for whom data was imputed at T5 were more likely to be white and have higher maternal education; this is consistent with the demographic composition of the Chicago families, who were not seen at T5. In contrast, children for whom data was imputed at T19 were more likely to be non-white and have lower maternal education; reflecting previously reported attrition of families (Anderson et al., 2014). Language outcome groups did not differ on race or maternal education, therefore differences in missing vs. imputed cases are not anticipated to have biased results. Nonetheless, race and maternal education were considered in both the imputation and covaried in the analyses. Results were highly similar when analyses were conducted on the subset of children with complete data (available upon request), therefore findings reported below include imputed data to maximize sample size and increase power.

Trajectories of the overall social-communication total, three DSM-5 social-communicative subdomains and 14 items were examined using Linear Mixed Models (MIXED, SPSS24). Cohort (T3, T5, T9 and T19 relative to T2), Language Group (D-V and D-MV relative to V-

V) and Cohort-by-Language Group were entered as predictors, controlling for NVIQ, gender, race and maternal education. NVIQ and maternal education were significant in models predicting each subdomain. This reflects higher T2 symptom scores for children with lower NVIQ and higher maternal education; exploratory analyses indicated that symptom trajectories did not differ by either characteristic. Results reported below reflect pooled estimates and statistics across the 100 imputations. Estimated marginal means (EMM) were used to plot trajectories by language group; within-group pairwise comparisons are reported to assess level of change between time points. Significance level was set at  $p .005$  to reduce Type I error due to multiple comparisons.

Items were summed in each of the three subdomains to assess the utility of the ADI-R current items for assessing DSM-5 social-communication criteria for ASD. Prevalence of symptoms at T3 and T19 relative to T2 was assessed using the McNemar's  $\chi^2$  test with correction for continuity. When the number of individuals whose score changed (i.e., 0 vs. 1 or 2) was less than 25, the binomial distribution was used (Sheskin, 2004).

Post hoc analyses were conducted to assess whether change in symptoms may be related to difference in treatment access. Chi square analyses were used to compare whether there were group differences in types or amount of treatment accessed at T2-T9 and One-way ANOVA were used to compare the number of hours of each treatment received.

## Results

### Changes in language level

Figure S1 shows changes in caregiver-reported language (i.e., OLL scores) over time. At T2, the majority of children (i.e., 86% V-V group, 100% D-V and D-MV groups) were reportedly not yet using phrase speech. By definition, all V-V children were using phrases and D-V and D-MV children continued to use single words or less at T3 (Figure S1; Table 1). At T5, the average caregiver-reported level of language continued to increase for the D-V group, reflecting an increase in the proportion of children using phrases from T3-T5. By T19, all D-V children reportedly had phrase speech (OLL=0). In contrast, there was limited change in average language level from T3-T19 for the D-MV; only a small number of individuals moved from using few-to-no-words at T3 (OLL=2) to using 5 or more words at T19 (OLL=1; see Table 1). Differences in the timing of phrase speech onset for V-V and D-V groups provides the opportunity to explore symptom changes in relation to phrase speech development, whereas the D-MV group's relative stability in language allows identification of symptom changes associated with general maturation.

### Overall Social-Communication Total (S-C)

For S-C, there were significant main effects of cohort and language group, as well as a cohort-by-language interaction (all  $p .001$ ; Table S1]. As shown in Figure 1A, levels of social-communicative impairments were comparable across language groups at T2. All groups showed significant decreases in symptoms across the 17-year study period ( $p .001$ ), though the D-MV group showed more modest levels of decline ( $EMM_{diff}=4.41$ ) compared to the V-V ( $EMM_{diff}=8.72$ ) and D-V groups ( $EMM_{diff}=7.23$ ). The V-V group showed the

largest decreases in S-C deficits between T2 and T3 ( $EMM_{diff}=6.84$ ,  $p=.001$ ), resulting in significantly fewer T3 symptoms than both the D-V and D-MV groups. In contrast, the D-V group made little change from T2 to T3 ( $EMM_{diff}=1.11$ ,  $p=1.00$ ), but scores decreased from T3 to T5 ( $EMM_{diff}=4.21$ ,  $p=.002$ ). The D-MV group's declines were not significant for any specific period, reflecting slower, steadier improvements; they maintained more symptoms than the V-V group at all time points (Figure 1A).

### Socioemotional Reciprocity (SER)

There were significant main effects of cohort and language group and a significant cohort-by-language group interaction (all  $p<.001$ ; Table S1). All groups showed significant reductions in SER impairments from T2-T19 ( $p<.005$ ; Figure 1B). There was a tendency for the language-delayed groups to exhibit more impaired SER than the V-V group; however, group differences varied over time due to differences in the timing of symptom changes. As noted with overall S-C symptoms, the greatest improvements appeared to be related to the onset of phrase speech. For the V-V group, scores declined the most between T2-T3 ( $EMM_{diff}=3.70$ ,  $p<.001$ ), whereas D-V children exhibited significant change between T3-T5 ( $EMM_{diff}=2.00$ ,  $p<.001$ ). The D-MV groups changes were gradual across the 17-year study and not significant for any given developmental period.

Main effects and interactions are provided in Table S1. Within the SER domain, *Showing* and *Offering Comfort* showed similar patterns to the SER total for the verbal groups; the D-MV group did not show significant declines in either symptom (Figure 2A–B). Other symptoms, however, showed somewhat different trajectories. For example, there were significant effects of time and language group, as well as time-by-language interaction ( $p<.001$ ) for *Use of Other's Body* (Figure 2C). V-V children showed significant reductions in this symptom from T2-T3 and T5-T9 ( $p<.001$ ), whereas significant reductions were only observed from T3-T9 for the D-V ( $p<.001$ ) and T3-T19 for the D-MV group ( $p<.001$ ). Only main effects of language and time were significant ( $p<.001$ ) for *Shared Enjoyment* (Figure 2D), with all language groups showing improvement from T2-T19 ( $p<.005$ ). For *Social Smiling*, only a main effect of language group was significant ( $p<.001$ ; Figure 2E).

### Nonverbal Communication (NVC)

There were significant main effects of cohort and language group ( $p<.001$ ) and a significant cohort-by-language interaction ( $p=.002$ ; Table S1). Each of the three language groups showed significant reductions in symptoms between T2-T19 ( $p<.001$ ). The D-MV group had more impaired NVC at T2 and T3 than the V-V group, but no longer differed by T19 (Figure 1C). The V-V and D-V groups only differed at T3, reflecting the earlier decline in NVC symptoms for the V-V group that corresponded to the onset of phrase speech (T2-T3  $EMM_{diff}=2.94$ ;  $p<.001$ ).

As shown in Table S1, five NVC items showed main effects of cohort and language ( $p<.001$ ). *Quality of Social Overtures* and *Pointing* exhibited marginally significant cohort-by-language interactions ( $p<.05$ ; Figure 3A–B). Closely mimicking the overall pattern of change observed for the NVC subdomain, both verbal groups showing significant changes approximating timing of language onset (V-V T2-T3  $p<.001$ ; D-V T3-T5  $p<.005$ ). The D-

MV group also exhibited significant change between T2-T5 ( $p < .001$ ). *Nodding* followed a similar trajectory, though both verbal groups continued to show improvement until T9 (Figure 3C). Language delayed groups exhibited greater impairment than the V-V group on both *Head Shaking* and *Conventional/Instrumental Gestures* (Figure 3D–E), but similar overall patterns of change over time. Only a significant main effect of language emerged for *Range of Facial Expressions* ( $p < .001$ ); symptom levels were mostly stable over time.

### Development & Maintenance of Relationships (DMR)

For the DMR subdomain, there were significant main effects of cohort and language group ( $p < .001$ ; Table S1). The V-V group showed a reduction in symptoms from T2-T19 ( $p < .001$ ; Figure 1D), but neither language delayed group exhibited significant changes over time. Similar to other subdomains, improvements in the V-V group were observed only between T2-T3 ( $EMM_{diff}=1.04$ ;  $p=.002$ ).

Within the DMR domain, *Offering to Share* showed significant effects of cohort and language ( $p < .001$ ) and a marginal cohort-by-language interaction ( $p=.006$ ; Figure 4A, Table S1). While the V-V group scores declined from T2-T3 and the D-V group from T2-T5 ( $p < .001$ ), the D-MV group remained stable ( $p=0.16$ ). Effects of cohort and language were significant ( $p < .001$ ) for *Appropriateness of Social Responses*; only the V-V group showed significant change over time ( $p=.005$ ; Figure 4B, Table S1). *Inappropriate Facial Expressions* was the only symptom for which there were no significant main effects; T2 levels of impairment were low for all groups and remained stable across time (Figure 4C, Table S1).

### Implications of symptom change for assessment of DSM-5 ASD criteria in young adults

Of the 80 young adults who received best estimate clinical diagnoses of ASD at T19, 11 (13.8%) had no caregiver-reported current symptoms in at least one of the three subdomains. Four of the 11 reported no current SER deficits, seven indicated no NVC symptoms, and four no difficulties in DMR.

As shown in Table S2, prevalence of most symptoms was highest at T2, with 67–100% of children within each language group receiving non-zero scores (indicating presence of abnormality) on 13/14 of the items. For the V-V group, significantly fewer participants demonstrated impairment on *Social Smiling* at T3 (33%) compared to T2 (74%;  $p=.003$ ). Neither language-delayed group exhibited significant reductions in any symptom at T3. From T2-T19, the V-V and D-V groups showed significant improvement on four items (*Showing, Use of Other's Body, Pointing, Nodding and Head Shaking*;  $p < .005$ ; Table S2). The D-V group also showed significant improvement in *Quality of Social Overtures* at T19 ( $p < .005$ ). Several other items showed marginal ( $p < .05$ ) improvement in the V-V group. Endorsement of *Range of Facial Expressions, Inappropriate Facial Expressions* and *Appropriateness of Social Responses* remained stable over time for all three groups. *Social Smiling* was the only item that exhibited a significant increase in endorsement (T3-T19 V-V only,  $p < .005$ ).



## Posthoc Analyses

Comparison of groups indicated no differences in type of treatment or duration at any time point (see Table S3 and online supplement for additional information).

## Discussion

This study demonstrates that children with ASD who have different patterns of language development show different trajectories of social-communicative (S-C) impairments reported by caregivers across childhood and adolescence. Previous studies reported that S-C symptoms decreased or improved between child and young adulthood (Gillespie-Lynch et al., 2012; McGovern & Sigman, 2005; Shattuck et al., 2007). This paper adds to extant literature by demonstrating that some changes may be associated with maturation as children get older, while many improvements appear to be related to phrase speech development. Associations with language were evident from both the levels of symptoms exhibited by each group at a given time point, as well as the timing of changes observed for different language groups. Language-delayed children who remained minimally verbal (D-MV) exhibited the highest overall S-C scores (reflecting more symptoms or greater impairment) across each of the five time points assessed. Symptom levels of children who transitioned from single words to phrase speech between T2 and T3 (V-V) and those who acquired phrases later (D-V) differed only at T3, reflecting that the V-V group's greatest symptom declines (i.e., improvement) occurred during the same period that they developed language. The D-V group, on the other hand, tended to show improvement later (e.g., between T3-T5), also approximating the timing of their phrase speech development.

Each of the three subdomains comprising the overall total exhibited somewhat different trajectories. The Socioemotional Reciprocity (SER) and Nonverbal Communication (NVC) subdomains showed similar trajectories as the overall S-C total, particularly for the two verbal groups. Significant improvements in caregiver-reported symptoms were observed for both the V-V and D-V groups during the same period in which the groups transitioned to phrase speech. Except for the T3 differences resulting from earlier V-V gains, symptom levels were mostly comparable across the study period. D-MV group showed improved SER but maintained higher symptom levels than their verbal counterparts throughout childhood and adolescence, whereas their NVC impairments improved to the point that symptom levels did not differ from verbal young adults by T19. Difficulties with Development and Maintenance of Relationships (DMR) showed greater stability than the other two subdomains. Although the V-V group showed early improvement, both language-delayed groups exhibited very limited change. Taken together, trajectories with differential associations to language provide initial evidence for separable subdomains of nonverbal social-communicative ASD symptoms. Breaking down social-communicative behaviors into these narrower constructs may be important to studies seeking to understand the pathophysiology of ASD symptoms. Nonetheless, it is important to remember that these symptom subdomains encompass multiple complex behaviors that also show somewhat different trajectories. Isolating biological contributions to such broad constructs will require identification of the specific mechanisms (e.g., emotion or cognitive processes) underlying observable impairments.

Individuals who remained MV tended to show the least caregiver-reported change across subdomains; however, improvements were noted in individual symptoms in both SER (*Shared Enjoyment, Use of Others' Body*) and NVC (*Quality of Social Overtures*, which reflects integration of eye contact during interactions, *Pointing, Nodding*). Although MV young adults remained significantly more impaired in most symptoms than their verbal peers, these changes suggest some ongoing skill development/symptom abatement despite relatively little change in language.

A few symptoms (*Social Smiling, Range of Facial Expressions, Inappropriate Facial Expressions*) did not show main effects of time. Although these items were categorized under different subdomains of social-communicative behaviors by Huerta and colleagues (2012), each relates to facial expressions (Huerta, Bishop, Duncan, Hus, & Lord, 2012). Despite some variation across childhood, the overall tendency for these symptoms to persist at similar levels in toddlerhood and young adulthood suggests that emotional expression through facial behavior may be a core deficit in ASD that warrants more systematic study across the lifespan. Considering this finding is based upon caregiver report, it will be particularly important for future studies to explore change using methods that quantify facial behavior in an objective manner (Yirmiya et al., 1989; Mathersul et al., 2013).

The observed decreases in the prevalence of many symptoms have important implications for diagnosis of verbal adolescents and adults, particularly those with a history of less significant language delay. Currently, diagnostic classification on instruments such as the ADI-R is based on past behaviors; there are not diagnostic cutoffs for the Current Behavior Algorithm. For older individuals presenting for a first-time diagnosis, reliance on retrospective recall may be affected by caregiver memory, particularly if the caregiver was not concerned about their child's behavior in earlier childhood (Havdahl et al., 2017), as may be the case in later-life diagnoses. It is critical that researchers and clinicians are aware that the longitudinal data presented here, based on caregiver reports of *current* behaviors, suggest that current ADI-R scores cannot simply be applied to the Diagnostic Algorithm to determine whether an older adolescent or adult currently meets ASD criteria.

The present results should be replicated using other instruments. Nonetheless, findings highlight a critical need for future research to characterize the manifestation of ASD symptoms in adults and assess the utility of caregiver-reported current behaviors for diagnostic and descriptive purposes in older individuals of varying language abilities. While history is an important factor in establishing an ASD diagnosis, mapping the current constellation of symptoms is critical to the development of instruments that are sensitive to ASD symptomatology in adults. It will be important that such instruments are designed to collect information using both self-report and observations from other informants, as different perspectives are likely to provide important insights into treatment targets and our understanding of ongoing social-communicative development.

The present findings also have implications for the conceptualization of ASD as a lifelong developmental disorder. While it is tempting to interpret symptom changes as representative of a decrease in autism severity it is perhaps more appropriate to consider improvements as evidence of ongoing development in basic social-communicative skills. While we observed

improvement in many symptoms considered to be core deficits in childhood ASD, most participants continued to meet diagnostic criteria for ASD at T19. Given age and language-related associations, summations of current ADI-R items are unlikely to be a meaningful reflection of social-communicative severity and changes should not be interpreted as adults having milder or less severe ASD (Seltzer et al., 2003). Rather, the decrease in scores over time likely reflects a qualitative shift in the types of symptoms defining ASD at different periods of development and suggests that childhood symptoms of ASD may not be the best markers for ASD in young adulthood. As Rogers and Pennington (1991) observed, “We expect to see some signs of a deeper underlying deficit specific to autism stand out during a specific developmental stage only to be accomplished to some degree at a later developmental stage and replaced by other symptoms of the underlying deficit” (p. 146). As we increase our understanding of mechanisms underlying the observable symptoms on which we base ASD diagnoses, it will be important to develop instruments that incorporate the qualitative changes across development in order to estimate symptom severity in adults and track age- and treatment-related change.

This study focused on changes in nonverbal social-communicative symptoms on the ADI-R, a caregiver-report diagnostic interview. It is possible that symptom trajectories observed in this study may be influenced by a number of caregiver factors, such as expectations of behavior and accuracy of observation, rather than the “true” course of symptoms across childhood and adolescence. Readers should also be reminded that ADI-R items are ordinal, not continuous, therefore changes encompass both qualitative and quantitative changes in symptoms.

Very few participants in this sample acquired phrase speech at an early age, thereby limiting our ability to assess changes in verbal communication items (e.g., conversation, social chat, stereotyped speech). Future studies with larger cohorts of non-language delayed children should assess how language-based social-communicative symptoms might change in relation to age and ongoing language development (e.g., phrases to sentences), as these are likely to be important diagnostic markers for verbal adults. In addition, few participants in this sample received ABA or MPST intervention and the specific intervention targets for each individual were not known. Clinical trials with comprehensive data regarding types, durations and foci of interventions for both treatment and control groups will continue to inform our understanding of the effects of different interventions on ASD symptom trajectories.

It is also necessary to highlight that these results are based upon a single cohort of children referred for ASD prior to 3 years of age in the early 1990s, which may not be a representative sample of children with ASD referred today. Methods used to identify clusters or classes of children based on their symptom trajectories (e.g., discrete mixture models, such as Proc Traj; Nagin, 2005) could have been applied to identify the range of different symptom trajectories (and how language predicts group membership). However, the mixed model approach using pre-defined language groups was chosen to explicitly examine the course of symptoms in relation to language as it was reported on the ADI-R. We hoped this would facilitate application of results to clinical cases and current samples that may vary with respect to proportions of language delayed or MV individuals.

Variability in age of interview at each time point and less dense sampling in adolescence warrants caution in interpreting the precise age of observed changes in symptom presentations. Moreover, although a more conservative alpha level of .005 was used to reduce Type I error, analysis of individual items resulted in a large number of comparisons. Replication of results in independent cohorts using symptom measures based on direct observational measures designed to capture change is clearly needed, particularly to assess which skill development may be attributable to language vs. more general maturation. Nonetheless, that three separate items capturing facial behavior seem to remain relatively stable suggest a promising area for more focused future research.

Finally, data were not missing completely at random; children whose data were imputed at T5 and T19 differed on demographic characteristics due to differences in study protocol (Chicago not seen at T5) and attrition at T19 (lost participants were more likely to be non-white and have lower maternal education). Language outcome groups did not differ on race or maternal education, therefore these differences in missing vs. imputed cases are not anticipated to have biased results. Furthermore, results were highly similar when analyses were limited to cases with complete data.

The present study demonstrates variability in the trajectories of caregiver-reported social-communicative impairments used to characterize ASD across childhood and adolescence. Results highlight the significance of both age and changes in language abilities in symptom presentation: declines in symptom levels (i.e., “improvements”) were often related to language development, though some symptoms seemed to improve regardless of language change. Abnormalities in facial expressions appeared to persist into young adulthood. The present findings demonstrate the need to exercise caution when assessing current symptomatology of young adults with instruments developed for use with younger children (such as the ADI-R); items may not capture impairments relevant to diagnosis in older individuals. Thus, while it is encouraging to highlight ongoing development demonstrated by an improvement in childhood symptoms, interpreting changes as reflective of a decrease in autism severity fails to acknowledge that many young adults with ASD continue to exhibit significant social-communicative impairments that require support across the lifespan. Future research is needed to characterize ASD in later adolescence and adulthood and better understand the observable manifestation of underlying impairments that support their clinical diagnoses. Such knowledge is critical to develop diagnostic instruments and metrics of severity for use with adults, as well as to inform development of better services and supports.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

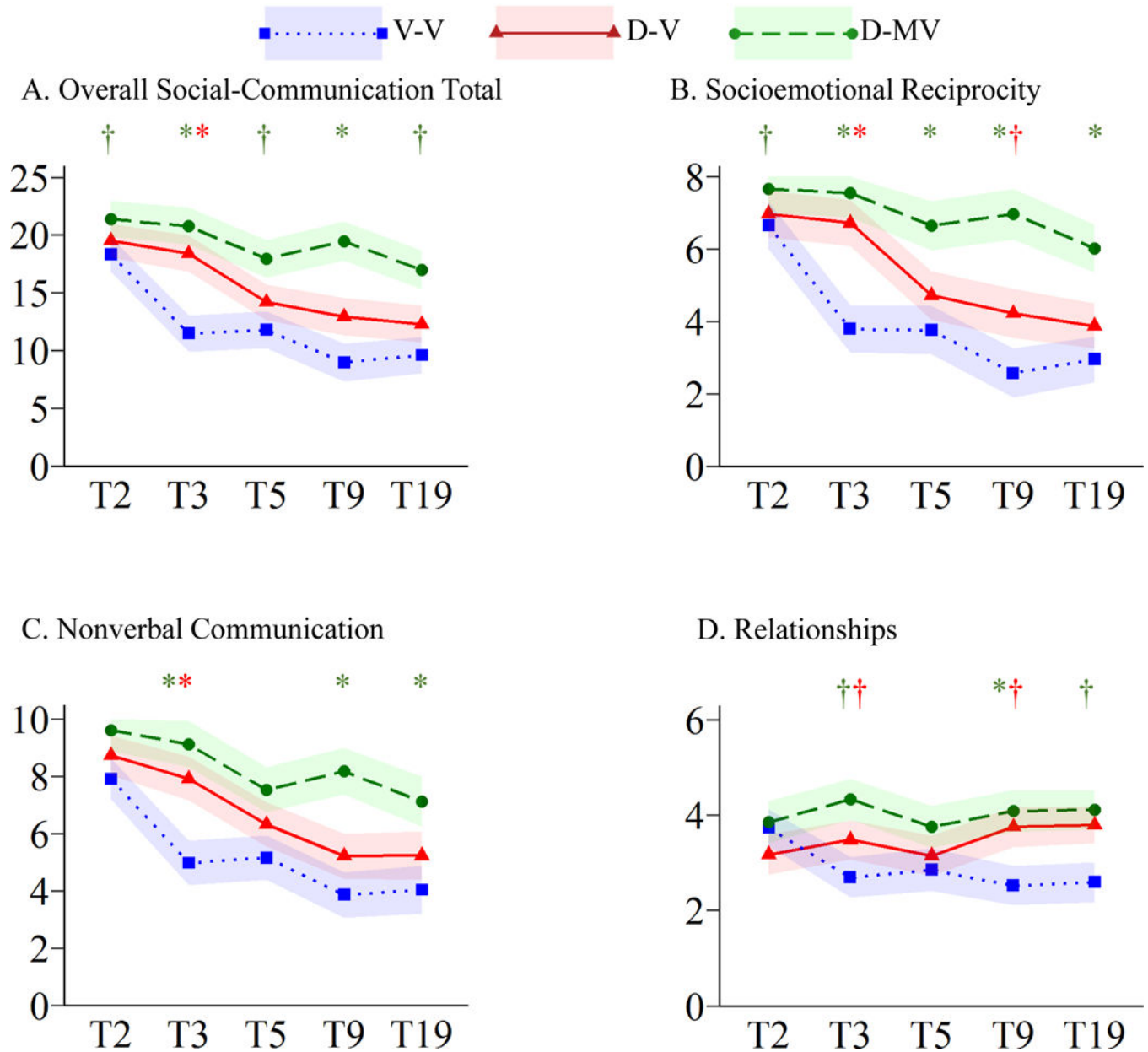
## Acknowledgements

We gratefully acknowledge the participants and their families for their ongoing participation. This study was funded by NIMH R01- MH081873 to CL. CL acknowledges receipt of royalties for the Autism Diagnostic Interview–Revised (ADI-R); profits from this study were donated to charity.

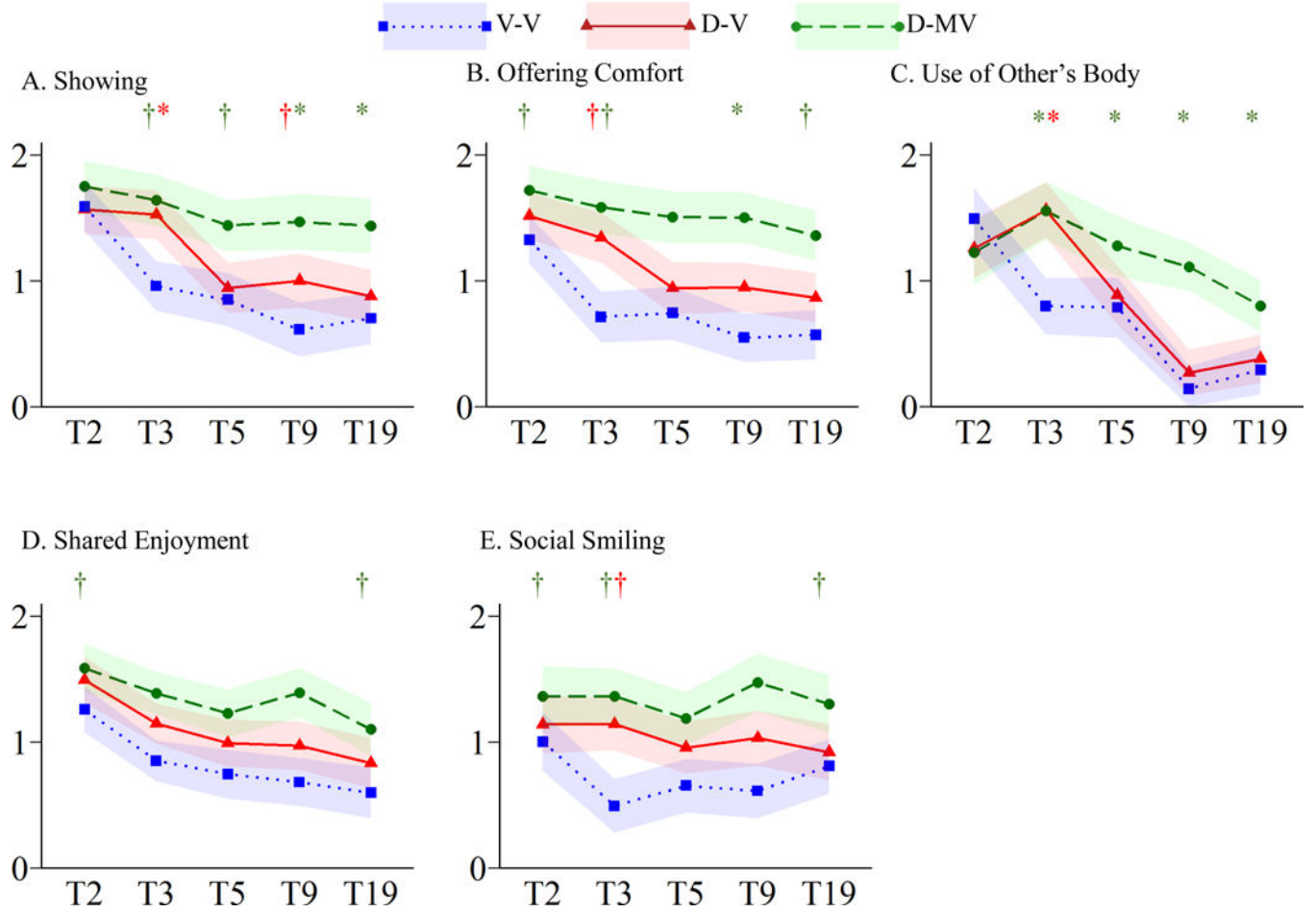
## References

- Anderson DK, Liang JW, & Lord C (2014). Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 55(5), 485–494. 10.1111/jcpp.12178 [PubMed: 24313878]
- Bal VH, Katz T, Bishop SL, & Krasileva K (2016). Understanding definitions of minimally verbal across instruments: evidence for subgroups within minimally verbal children and adolescents with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 57(12), 1424–1433. 10.1111/jcpp.12609
- Elliott CD (1990). *Differential Ability Scales*. San Antonio, TX: Psychological Corporation.
- Gillespie-Lynch K, Sepeta L, Wang Y, Marshall S, Gomez L, Sigman M, & Hutman T (2012). Early childhood predictors of the social competence of adults with autism. *Journal of Autism and Developmental Disorders*, 42(2), 161–174. 10.1007/s10803-011-1222-0 [PubMed: 22187106]
- Gotham K, Pickles A, & Lord C (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39(5), 693–705. 10.1007/s10803-008-0674-3 [PubMed: 19082876]
- Gotham K, Risi S, Pickles A, & Lord C (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, 37(4), 613–627. 10.1007/s10803-006-0280-1 [PubMed: 17180459]
- Havdahl KA, Bishop S, Surén P, Øyen A, Lord C, Pickles A, ... Stoltenberg C (2017). The influence of parental concern on the utility of autism diagnostic instruments, 10(1), 1672–1686. 10.1002/aur.1817
- Huerta M, Bishop S, Duncan A, Hus V, & Lord C (2012). Application of DSM-5 criteria for Autism Spectrum Disorder to three samples of children with DSM-IV diagnoses of Pervasive Developmental Disorders, 169(10), 1056–1064.
- Klein-Tasman BP, Risi S, & Lord CE (2007). Effect of language and task demands on the diagnostic effectiveness of the autism diagnostic observation schedule: the impact of module choice. *Journal of Autism and Developmental Disorders*, 37(7), 1224–1234. 10.1007/s10803-006-0266-z [PubMed: 17149669]
- Lord C, & Pickles A (1996). Language level and nonverbal social-communicative behaviors in autistic and language-delayed children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(11), 1542–1550. 10.1097/00004583-199611000-00024 [PubMed: 8936922]
- Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, & Pickles A (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry*, 63(6), 694–701. 10.1001/archpsyc.63.6.694 [PubMed: 16754843]
- Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, & Bishop S (2012). *Autism Diagnostic Observation Schedule: ADOS-2*. Los Angeles, CA: Western Psychological Services.
- Lord C, Rutter M, DiLavore PS, & Risi S (1999). *Autism Diagnostic Observation Schedule (ADOS)*. Los Angeles, CA: Western Psychological Services.
- Mathersul D, McDonald S, & Rushby JA (2013). Automatic facial responses to briefly presented emotional stimuli in autism spectrum disorder, 94, 397–407.
- McGovern CW, & Sigman M (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry*, 46(4), 401–408. 10.1111/j.1469-7610.2004.00361.x [PubMed: 15819649]
- Mesibov G, Shea V, & Schopler E (2005). *The Teacch Approach to Autism Spectrum Disorders*. New York: Kluwer Academic/Plenum 10.1007/978-0-306-48647-0
- Mullen E (1997). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service, Inc.
- Muthén LK, & Muthén BO (1998). *Mplus User's Guide*. (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Nagin DS (2005). *Group-based modeling of development*. Cambridge, Mass: Harvard University Press.
- Raghunathan TE, Lepkowski JM, Van Hoewyk J, & Solenberger P (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models., 27(1), 85–95.

- Reynolds AJ (n.d.). Success in early intervention: the Chicago child parent centers. U of Nebraska Press.
- Rogers SJ, & Pennington BF (1991). A theoretical approach to the deficits in infantile autism. *Development and Psychopathology*, 3(02), 137–162. 10.1017/S0954579400000043
- Rubin DB (1987). Multiple imputation for nonresponse in surveys. John Wiley and Sons.
- Rutter M, Le Couteur A, & Lord C (2003). *Autism Diagnostic Interview-Revised*. Los Angeles, CA: Western Psychological Services.
- Seltzer MM, Krauss MW, Shattuck PT, Orsmond G, Swe A, & Lord C (2003). The symptoms of autism spectrum disorders in adolescence and adulthood. *Journal of Autism and Developmental Disorders*, 33(6), 565–581. [PubMed: 14714927]
- Shattuck PT, Seltzer MM, Greenberg JS, Orsmond GI, Bolt D, Kring S, ... Lord C (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 37(9), 1735–1747. 10.1007/s10803-006-0307-7 [PubMed: 17146700]
- Sheskin D (2004). *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, FL: Chapman & Hall/CRC Press.
- Skondral A, & Rabe-Hesketh S (n.d.). *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*. Boca Raton, FL: Chapman & Hall/CRC Press.
- Sturm A, Kuhfeld M, Kasari C, & McCracken J (2017). Development and validation of an item response theory-based Social Responsiveness Scale short form, 58(9), 1053–1061. 10.1111/jcpp.12731
- Yirmiya N, Kasari C, Sigman M, & Mundy P (1989). Facial expressions of affect in autistic, mentally retarded and normal children. *Journal of Child Psychology and Psychiatry*, 30(5), 725–735. [PubMed: 2793960]

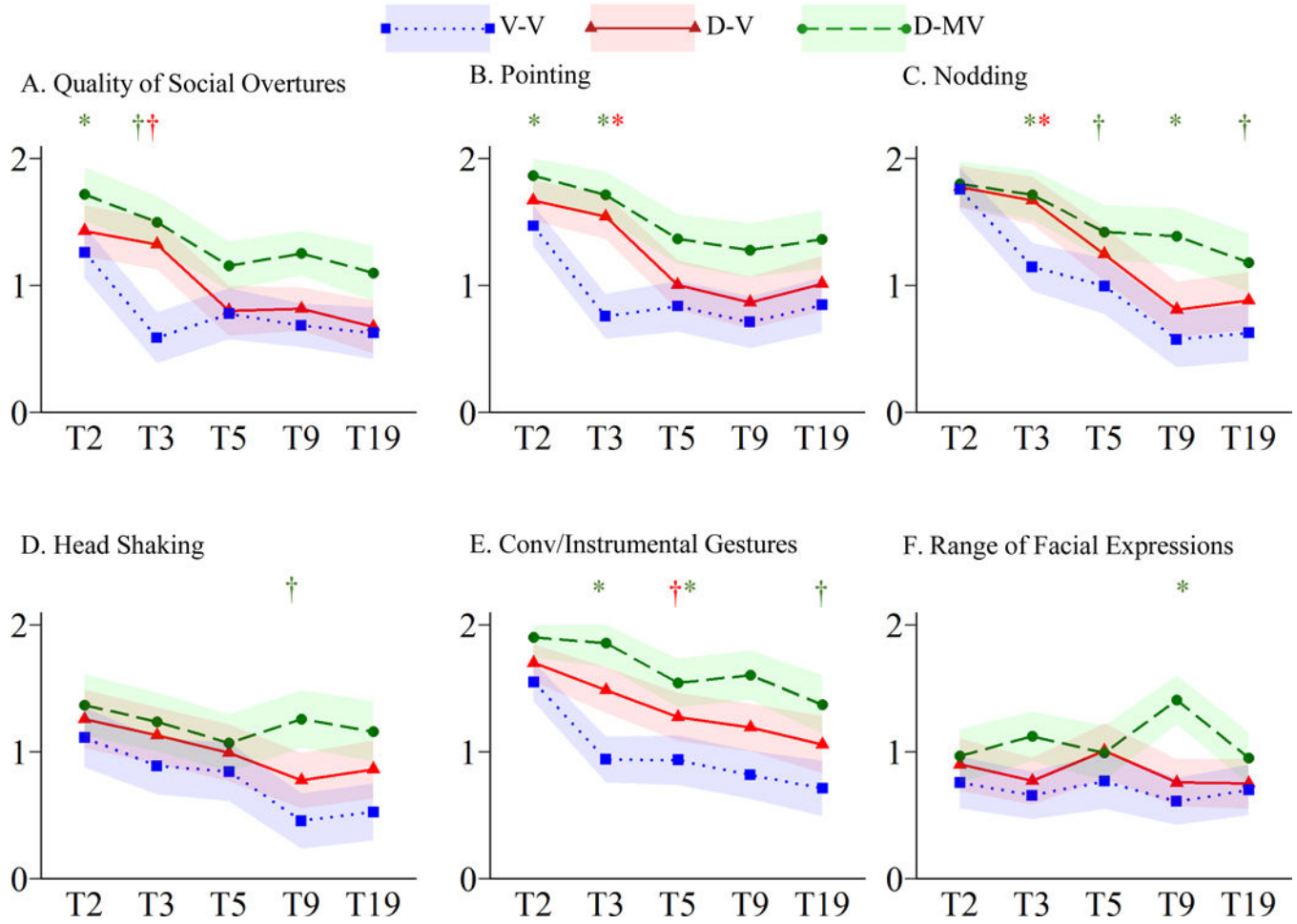


**Figure 1.**  
 Trajectories of Overall Social-Communicative Total and Subdomains Symptoms  
 †p .05 \*p .005; V-V=Verbal at T3 & T19; D-V=Delayed at T3, Verbal at T19; D-MV=Delayed at T3, Minimally Verbal at T19

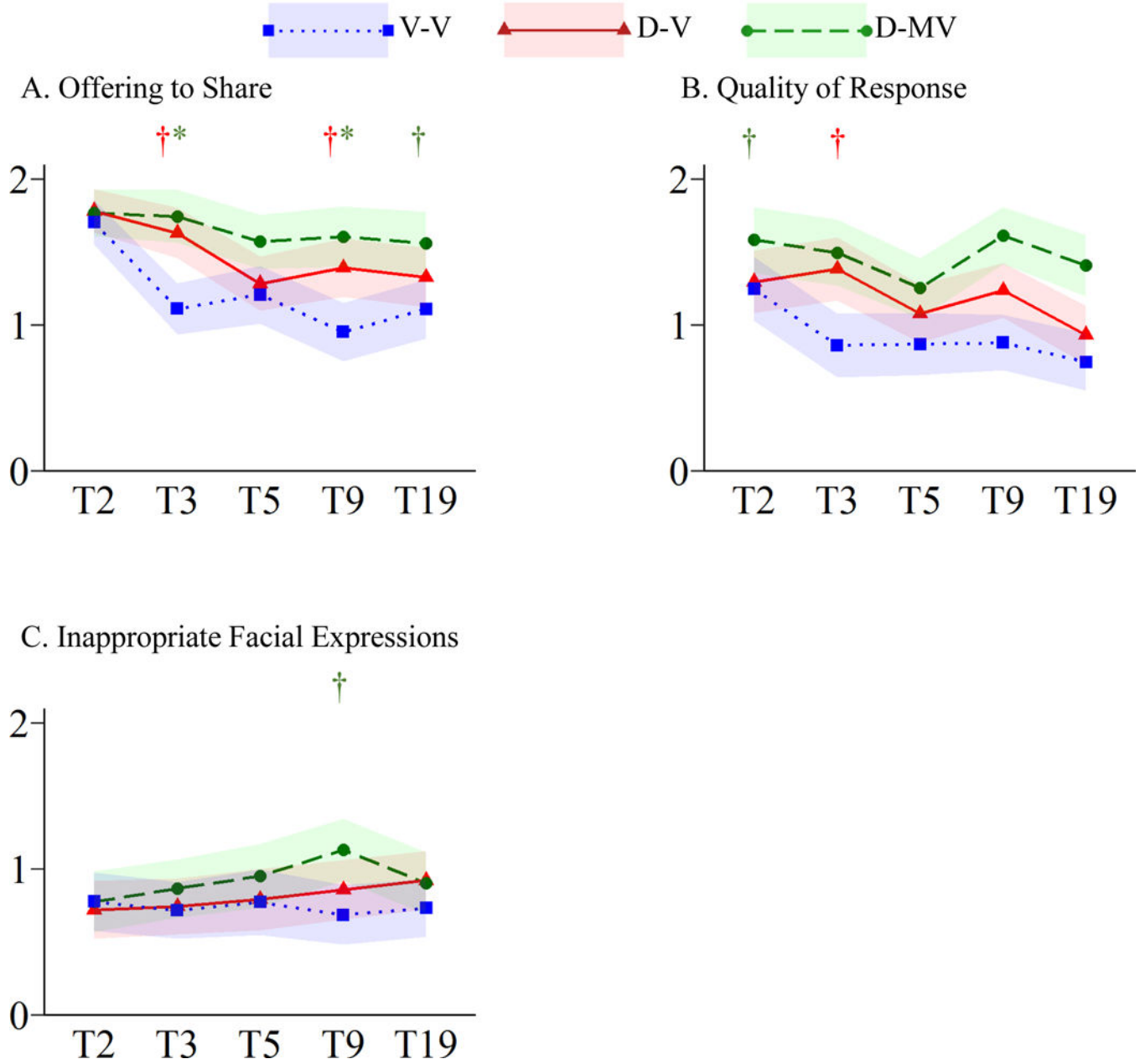


**Figure 2.**  
Trajectories of Socioemotional Reciprocity Symptoms  
†p .05 \*p .005





**Figure 3.**  
 Trajectories of Nonverbal Communication Symptoms  
 †p .05 \*p .005



**Figure 4.**  
Trajectories of Development and Maintenance of Relationships Symptoms  
†p .05 \*p .005

TABLE 1

Participant demographics at T3 and T19

Lang Group	T3				T19*			
	V-V (N=49)	D-V (N=46)	D-MV (N=45)	V-V (N=37)	D-V (N=31)	D-MV (N=30)		
	M (SEM)	M (SEM)	M (SEM)	M (SEM)	M (SEM)	M (SEM)	M (SEM)	M (SEM)
Age (yrs)	3.69 (0.07)	3.41 (0.07)	3.47 (0.07)	18.90 (0.20)	19.20 (0.07)	19.20 (0.22)		
NVIQ	82.00 (2.23)	63.10 (2.19)	47.50 (2.49)	93.50 (5.33)	57.90 (4.97)	16.50 (1.62)		
Social-Comm	10.24 <sup>††</sup> (0.92)	18.63 <sup>†§</sup> (0.89)	21.87 <sup>†§</sup> (0.64)	8.37 <sup>††</sup> (0.87)	12.53 <sup>†§</sup> (0.86)	18.06 <sup>†§</sup> (0.73)		
SER	3.49 <sup>††</sup> (0.33)	6.78 <sup>†</sup> (0.33)	7.82 <sup>†</sup> (0.33)	2.65 <sup>††</sup> (0.32)	3.94 <sup>†§</sup> (0.34)	6.00 <sup>†§</sup> (0.30)		
NVC	4.41 <sup>††</sup> (0.47)	8.04 <sup>†§</sup> (0.47)	9.62 <sup>†§</sup> (0.26)	3.47 <sup>††</sup> (0.44)	5.37 <sup>†§</sup> (0.49)	7.59 <sup>†§</sup> (0.39)		
DMR	2.35 <sup>††</sup> (0.24)	3.80 <sup>†</sup> (0.20)	4.42b (0.21)	2.25 <sup>††</sup> (0.24)	3.23 <sup>†§</sup> (0.22)	4.18 <sup>†§</sup> (0.18)		
Lang Level (%)								
Phrases	100%	0%	0%	100%	100%	0%		
5+ words	0%	54%	16%	0%	0%	40%		
<5 words	0%	46%	84%	0%	0%	60%		
Female	10.2%	15.2%	13.3%	8.1%	19.4%	10.0%		

Note: Superscripts indicate significant difference from same symbol group ( $p < .05$ ). V-V = Verbal at T3 & T19; D-V = Delayed at T3, Minimally Verbal at T19; D-MV = Delayed at T3, Minimally Verbal at T19; NVIQ = nonverbal IQ; Exp AE = expressive language age equivalent; Social-Comm, Overall Social-Communication Total; SER = Socioemotional Reciprocity; NVC = Nonverbal Communication; DMR = Development and Maintenance of Relationships.

\* Means and standard errors are based on imputed data for T19.