Original Article

autism

Autism
2019, Vol. 23(8) 1911–1926

DOI: 10.1177/1362361318812423
journals.sagepub.com/home/aut
SAGE

Longitudinal stability of reading profiles in individuals with higher functioning autism

Emily J Solari[1], Ryan P Grimm[1], Nancy S McIntyre[2], Matthew Zajic[1] and Peter C Mundy[3]

Abstract
The reading difficulties of individuals with autism spectrum disorders have been established in the literature, with particular attention drawn toward reading comprehension difficulties. Recent papers have highlighted the heterogeneous nature of reading abilities in this population by utilizing statistical methods that allow for investigations of unique reading profiles. This article extends this literature by investigating reading profiles longitudinally, to investigate the stability of reader profiles across time. Latent profile and transition analyses were conducted to establish categorically distinct reading profiles at two time points, 30 months apart. This study also examined whether age and autism symptom severity were related to the profiles at each time point. Finally, transitions between profiles at each time point were identified. Age did not predict profile membership, but there were significant differences in symptom severity that were largely stable over time. Results indicate that heterogeneous reading profiles exist within the autism population, ranging from average reading ability to severe difficulties across different reading subskills. The data from this study demonstrate that reading profiles of children and adolescents with autism spectrum disorders shift when examined across time.

Keywords
autism spectrum disorders, reading profiles

Autism spectrum disorders (ASD) refer to a group of neurodevelopmental disorders characterized by impairments in social interaction, communication, and repetitive and restricted behaviors and interests; prevalence rates indicate that 1 in 59 individuals in the United States are diagnosed with ASD (Baio et al., 2018). In recent years, attention has been drawn toward the academic achievement of individuals with ASD during the school-age years. Approximately 68% of school-age children with ASD exhibit cognitive abilities outside the range of intellectual disability (IQ > 70; Centers for Disease Control and Prevention (CDC, 2016). These individuals diagnosed with ASD, with a normative range of IQ, typically exhibit sentence-level verbal abilities; however, their academic and social outcomes are significantly less than optimal (Seltzer et al., 2004). Many children with ASD, who have normative range IQ, exhibit a discrepancy between intellectual functioning and academic achievement that appears to be consistent with specific learning disabilities; particular attention has been drawn toward discrepancies in reading comprehension, the ability to make meaning from written text, even when students have intact word reading abilities (Huemer and Mann, 2010; Jones et al., 2009; Mayes and Calhoun, 2006; Whitby and Mancil, 2009).

Reading comprehension development

According to the Simple View of Reading (Gough and Tunmer, 1986), reading comprehension is the product of efficient decoding, or word reading, and oral language skills. As such, there are three potential subgroups who may exhibit difficulties with reading comprehension: (a) individuals with intact decoding skills but difficulties in oral language skills, (b) intact oral language coupled with decoding deficiencies, and (c) individuals who have difficulties in both oral language and decoding (Catts et al., 2003). Previous research has applied the Simple View to samples of children diagnosed with ASD (McIntyre et al., 2017a; Ricketts et al., 2013).

The main focus of many previous studies of reading with children diagnosed with ASD has been on a particular

[1]University of Virginia, USA
[2]University of North Carolina at Chapel Hill Frank Porter Graham Child Development Institute, USA
[3]University of California, Davis, USA

Corresponding author:
Emily J Solari, University of Virginia, 417 Emmet St. South, PO Box 400273, Charlottesville, VA 22904, USA.
Email: ejs9ea@virginia.edu

subgroup of readers, those who are able to decode words easily but have difficulties comprehending written text, often called hyperlexic readers (for reviews, see Grigorenko et al., 2003; Newman et al., 2007; Saldaña et al., 2009). Specifically, these studies have noted discrepancies between reading comprehension attainment and decoding abilities, when sample averages are considered (Frith and Snowling, 1983; Jones et al., 2009; Nation et al., 2006; Newman et al., 2007; Ricketts et al., 2013); however, the magnitude of the discrepancy varies widely, presumably based on the characteristics of the individuals studied. While these studies provide an important framework describing the reading development of this population, concentration on group means may obfuscate critical differences between categorically heterogeneous subgroups.

A handful of studies have demonstrated a more nuanced picture of reading difficulties in the ASD population by utilizing methodologies that go beyond average performance to identify specific profiles of readers within ASD samples (Davidson and Weismer, 2014; McIntyre et al., 2017b; Nation et al., 2006). These studies have demonstrated that a profile of strong decoding abilities and weak comprehension skills is not universal in samples of individuals with ASD. Instead, reading profiles have emerged suggesting difficulties with decoding (Åsberg and Sandberg, 2012; McIntyre et al., 2017b; Nation et al., 2006; White et al., 2006) and reading fluency, the rate and accuracy with which individuals read connected text (Solari et al., 2017), coupled with comprehension difficulties. Importantly, while there is converging evidence that the majority of samples with ASD struggle with some aspect of reading, studies investigating reading component skills have also found that there is a subgroup of readers that demonstrate reading skills in the average range (Lucas and Norbury, 2014; McIntyre et al., 2017b; Nation et al., 2006); however, the proportion of average readers is much lower than what is seen in typically developing populations.

All existing data demonstrate that reading comprehension difficulties, on average, are more prevalent in individuals with ASD when compared to their typically developing peers; previous studies show 38%–73% of ASD samples with comprehension difficulties (Henderson et al., 2014; Jones et al., 2009; McIntyre et al., 2017a; Nation et al., 2006; Norbury and Nation, 2011) compared to approximately 30% of the typically developing population who develop a reading-related difficulty between second and eighth grades (e.g. Catts et al., 2012). As with their typically developing peers, oral language abilities play an important role in reading comprehension development; investigations into the Simple View have highlighted the importance of oral language abilities in the reading ability of students with ASD (Lindgren et al., 2009; Lucas and Norbury, 2014; McIntyre et al., 2017a; Ricketts et al., 2013).

In addition, one core challenge of individuals diagnosed with ASD, social and communication impairments, could be related to reading comprehension difficulties. Recent studies have investigated this possibility by examining the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2012) relation to reading comprehension outcomes and have shown significant correlations and predictive patterns between ADOS and reading comprehension difficulties (Jones et al., 2009; McIntyre et al., 2017b; Ricketts et al., 2013). There is evidence to support the idea that the social-cognitive difficulties associated with ASD may impact the ability of affected children to comprehend written text (Estes et al., 2011; McIntyre et al., 2017b; Nation et al., 2006; O'Connor and Klein, 2004; Randi et al., 2010). Difficulties in mentalizing (as measured on theory of mind tasks) appear to be a specific correlate of understanding narrative and reading comprehension in children with ASD (Capps et al., 2000; McIntyre et al., 2017a, 2018; Ricketts et al., 2013). This inability to "mentalize" may make it difficult for children with ASD to infer an author's intentions, draw conclusions about a text, and detect causal elements of a story (McIntyre et al., 2018; Randi et al., 2010; Zevenbergen and Whitehurst, 2003). Brown et al. (2013) corroborated these findings when they noted that individuals with ASD are significantly worse at comprehending highly social texts in comparison with nonsocial texts. Another key finding to support this notion is demonstrated through a handful of studies that show that as ASD symptom severity increases in affected individuals, more difficulties with reading comprehension are seen (Åsberg and Sandberg, 2012; McIntyre et al., 2017a; Ricketts et al., 2013). Typically, greater ASD symptom severity indicates higher levels of social-cognitive difficulties (Hughes, 2001; Joseph and Tager-Flusberg, 2004; Travis et al., 2001). Therefore, it is possible that symptom severity is affecting reading comprehension through its relation to social cognitive abilities, such as theory of mind (McIntyre et al., 2017b; Mason et al., 2008; Randi et al., 2010).

Purpose of this study

To date, only two studies (Davidson and Weismer, 2014; McIntyre et al., 2017b) have utilized latent profile analysis (LPA) to investigate language and literacy skills in children with ASD; the Davidson study addressed emergent literacy, while the McIntyre study investigated reading development in school-aged children. An advantage of the LPA approach is that it does not utilize arbitrary cutoff points, but instead derives empirically based profiles of readers. This study seeks to increase the field's understanding of the heterogeneous nature of reading profiles of school-aged children in individuals with ASD, with normative range IQ, and the stability of these distinct profiles over time. This study builds upon the findings of the

McIntyre study to investigate the reader profile assignment of 8- to 18-year-old students with ASD without intellectual disability, across two time points, utilizing a latent transition analysis (LTA) to determine whether the previously identified reading subgroups were stable over time. Work with typically developing students in this age range has shown generally stable reader profiles (Catts et al., 2012); therefore, it was hypothesized that this would be similar for individuals diagnosed with ASD. The stability of reading profiles has important implications for reading instruction; therefore, this study sought to explore the stability of reader profiles longitudinally in students diagnosed with ASD. Previous research with similar populations has shown that some reader profiles exhibit differences in ASD symptom severity (McIntyre et al., 2017b); this study investigated whether ASD symptom severity had longitudinal effects on reading profiles.

Method

Participants

The sample was enrolled in a longitudinal study approved by the university Institutional Review Board, with a community diagnosis of ASD. Subjects were recruited through a university research subject tracking system and from the community through local school districts and word-of-mouth; parent consent was obtained from all participants as well as child assent before testing began. Individuals were included in the sample if they had a full-scale IQ (FIQ) estimate $\geqslant 75$ and their community ASD diagnosis was confirmed by trained researchers using the ADOS-2 (Lord et al., 2012). Exclusion criteria included an identified syndrome other than ASD (e.g. Fragile X, emotional and behavioral disorders (EBD), attention deficit hyperactivity disorder (ADHD)), as ASD was the target sample, significant sensory or motor impairment (e.g. hearing or visual impairments), psychotic symptoms, a neurological disorder (e.g. cerebral palsy), or any major medical disorder that would lead to extended absences from school; children who were not English proficient per parent report were also excluded from the study. Three children in the sample were reported to be bilingual, but English proficient. Exclusionary criteria and language proficiency were determined through parent report of health, medication, developmental and educational history and intervention. At time point 1, 93 individuals between the ages of 8 and 16 years were recruited and met criteria on the ADOS-2, but 12 were excluded due to FIQ <75, and one outlier was removed, resulting in a sample of 80. The outlier was removed because this child had a Rapid Naming score that was outside the range of possible scores. This was likely due to either an error during that assessment or when the data were entered. Scores for all other children and variables were checked and there were no other errors. At time point 2, attrition ($n=16$) resulted in a final sample of 64. An age range of 8–16 years was chosen to compare the development of elementary age (8–11 years) and secondary age students (12–16 years). Sixteen was chosen at the top end to ensure that students could be followed up for 30 months. Eight was chosen to ensure the youngest students could provide valid responses on all measures including self-report measures such as anxiety scales. The majority of the sample (77%) spent much or all of their day in the general education classroom and 91% had an Individualized Education Programs (IEP) or 504 plan. Males made up 81.3% of the sample. In terms of race/ethnicity, 10% of the sample was Latino/a, 5% were Asian, 1.3% were African-American, 66.3% were Caucasian, 16.3% were Other or Mixed, and 1.3% declined to state. Descriptive statistics for the sample, including diagnostic measures, cognition, and reading and language assessments at both time points can be found in Table 1.

Measures and procedures

Trained members of a research group collected all data during assessment sessions in a university-based child assessment laboratory over two 2.5-h sessions at time point 1 and 30 months later, over two 2.5-h sessions at time point 2. Testers were trained through explicit instruction on each assessment, followed by a training procedure comprising observation of experienced testers, then the requirement that the novice tester administers assessments with an experienced tester scaffolding and double scoring until reliable.

Diagnostic measures and sample description. The ADOS-2 (Lord et al., 2012) is a semi-structured diagnostic assessment; scores were utilized to confirm ASD diagnosis at time point 1. The overall total score was used in all analyses. The Wechsler Abbreviated Scales of Intelligence-II (WASI-II; Wechsler, 2011) was administered at time point 1 to assess verbal and nonverbal cognitive abilities.

Language and reading skills. The Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999) was used to assess rapid automatized naming (RAN), phonological memory (nonword repetition), and phonological awareness (Elision) and yielded age-normed scaled scores ($M=10$; standard deviation (SD)$=3$). RAN measured the speed at which participants were able to connect symbols (letters and numbers) to phonological representations (naming the symbol). Elision assessed the ability to repeat a word presented orally, and then say the part of the word that remained after removing a specified sound. Nonword repetition measured the extent to which a participant could repeat increasingly longer and more complex nonwords (3–15 sounds) presented on a CD player. Internal consistency Cronbach's alpha coefficients from our study

Table 1. Descriptive statistics at time points 1 and 2.

	Time 1				Time 2			
	M	*SD*	Min.	Max.	*M*	*SD*	Min.	Max.
Cognitive and diagnostic								
FIQ	100.21	14.89	–	–	–	–	–	–
ADOS-2	10.76	3.48	–	–	–	–	–	–
Age	11.26	2.15	7.92	15.95	13.78	2.12	10.41	18.34
PP and RAN								
RAN	88.97	14.22	58	145	91.48	17.68	55	154
NWR	7.49	2.16	1	13	7.62	1.95	3	12
Elision	10.05	2.93	2	15	9.70	2.61	3	13
Word recognition								
PDE	94.87	14.91	58	127	96.14	12.45	67	122
SWE	93.66	14.47	57	136	94.22	12.94	34	132
GORT accuracy	8.10	2.66	2	16	8.75	3.05	1	14
Linguistic comprehension								
CELF RS	7.44	3.08	1	13	8.02	3.71	1	18
WIAT RV	103.30	17.14	58	140	107.95	16.90	65	146
TAPS AR	6.08	2.76	1	11	5.44	2.67	1	12
Story memory	8.03	3.23	1	15	9.08	3.14	1	15
Reading comprehension								
GORT comprehension	7.45	2.52	1	13	7.81	2.94	1	15

SD: standard deviation; FIQ: full-scale IQ; ADOS-2: Autism Diagnostic Observation Schedule; PP: phonological processing; RAN: rapid automatized naming; NWR: nonword repetition; PDE: phonemic decoding efficiency; SWE: sight word efficiency; GORT: Gray Oral Reading Test; CELF RS: Clinical Evaluation of Language Fundamentals–Recalling Sentences; WIAT RV: Wechsler Individual Achievement Test–Receptive Vocabulary; TAPS AR: Test of Auditory Processing–Auditory Reasoning.

for Elision (0.93) and for nonword repetition (0.76) were generally consistent with publisher reported alphas (0.81–0.91; Wagner et al., 1999). The Test of Word Reading Efficiency, Second Edition (TOWRE-2; Torgesen et al., 2012) provided an age-normed standard score ($M=100$; $SD=15$) measuring accuracy and fluency of phonemic decoding (PDE) and sight word reading (SWE). Participants were asked to read as many decodable nonwords (PDE) or real words (SWE) as they could from a list in 45 s per subtest. Internal consistency Cronbach's alpha coefficients from our study for SWE (0.97) and for PDE (0.96) were consistent with publisher reported alphas for both subtests (>0.90; Torgesen et al., 2012).

Passage-level oral reading fluency was only assessed at time point 2 and was assessed with two passages, consistent with their grade level, from the Dynamic Indicators of Basic Early Literacy Skills–Next (DIBELS; Good et al., 2011) or Florida Oral Reading Fluency (FORF; Florida Department of Education, 2009), and is reported as words read correctly in 1 min. The correlation between the number of words correctly read between passages was strong for the ASD ($r=0.78$, $p<0.005$) sample.

Sentence-level semantic and syntactic language skills were measured with the Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF-4; Semel et al., 2003); age-normed scaled scores were used ($M=10$; $SD=3$). Participants were asked to repeat increasingly longer and more syntactically complex sentences presented by the experimenter. Internal consistency Cronbach's alpha coefficients from our study for Recalling Sentences (0.94) were generally consistent with publisher reported alphas (0.97) (Semel et al., 2003).

Vocabulary was measured using the Receptive Vocabulary (RV) subtest from the Wechsler Individual Achievement Test, Third Edition (WIAT-III; Breaux, 2009), which yielded age-normed standard scores ($M=100$; $SD=15$). Participants were required to point to one of four pictures that best depicted increasingly more semantically sophisticated words. The internal consistency Cronbach's alpha coefficient from our study was 0.80 in the ASD sample.

The Auditory Reasoning subtest from the Test of Auditory Processing, Third Edition (TAPS-3; Martin and Brownwell, 2005) provided an age-normed scaled score ($M=10$; $SD=3$) that measured linguistic comprehension and inference ability. In this subtest, participants were read short (one to three sentences) vignettes and then asked to respond to one question per vignette that required inferential skills or understanding of implied meanings and idioms to answer correctly. The vignettes were a mixture of those which were social in nature and those which were not. The internal consistency Cronbach's alpha coefficient from our study for Auditory Reasoning (0.88) was generally consistent with publisher reported alphas (0.91–0.96; Martin and Brownwell, 2005).

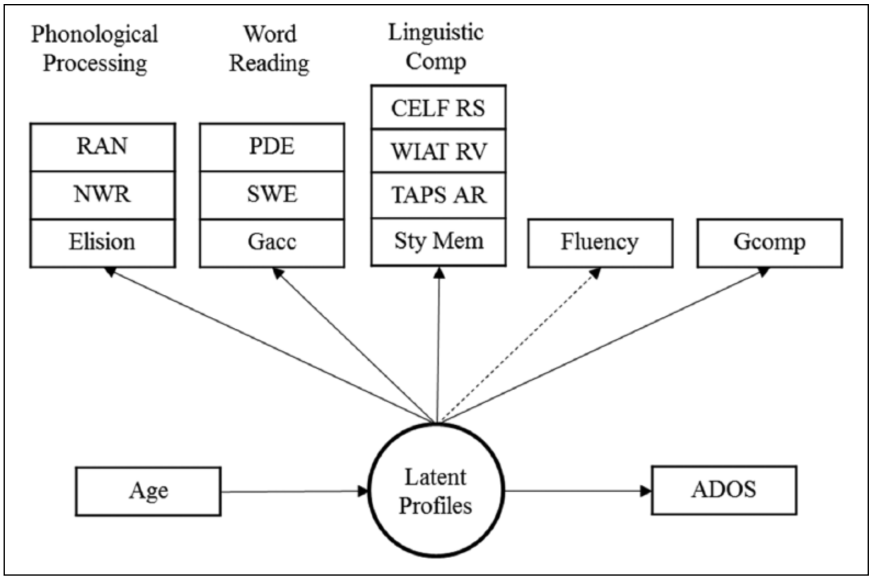

Figure 1. Conceptual diagram of the two LPAs. LPA models were similar across time points 1 and 2, except only time point 2 included fluency (represented by the dashed line). The circle represents the latent profile variable. Boxes above the circle represent observed variables used to measure the latent profile variable. The observed variables are group by reading subdomain. The box to the left of the circle indicates that age was treated as an observed covariate/predictor of the latent profile variable. The box to the right of the circle indicates that ADOS-2 was treated as a distal outcome of the latent profile variable.
RAN: rapid automatized naming; NWR: nonword repetition; PDE: phonemic decoding efficiency; SWE: sight word efficiency; Gacc: Gray Oral Reading Test-5 Accuracy; CELF RS: Clinical Evaluation of Language Fundamentals–Recalling Sentences; WIAT RV: Wechsler Individual Achievement Test-III–Receptive Vocabulary; TAPS AR: Test of Auditory Processing–Auditory Reasoning; Sty Mem: WRAML-2 Story Memory; Gcomp: Gray Oral Reading Test-5 Comprehension; ADOS-2: Autism Diagnostic Observation Schedule-2.

Narrative retelling skills were measured with the Story Recall subtest from the Wide Range Assessment of Memory and Learning, Second Edition (WRAML-2; Sheslow and Adams, 2003), which yielded an age-normed scaled score ($M=10$; $SD=3$). Experimenters read a total of two short stories aloud to each participant and after each story, participants were asked to retell all the parts of the story they could remember. The internal consistency Cronbach's alpha coefficient from our study for Story Recall (0.95) was generally consistent with publisher reported alphas (0.91–0.92; Sheslow and Adams, 2003).

Passage-level reading accuracy and comprehension were assessed with the Gray Oral Reading Tests, Fifth Edition (GORT-5; Wiederholt and Bryant, 2012) which yielded age-normed scaled scores ($M=10$; $SD=3$). This assessment comprises 16 progressively longer and more complex reading passages read aloud by the participant until they reach the ceiling at which they can no longer read with adequate fluency. Accuracy of oral reading is recorded as are the responses to five open-ended reading comprehension questions per passage that were asked by the examiner with the passage removed from view. Cronbach's alpha coefficients from our study for comprehension (0.90) and accuracy (0.92) were generally consistent with publisher reported alphas (ASD = 0.93–0.97) (Wiederholt and Bryant, 2012)

Analytic plan

LPA is a multivariate technique used to empirically identify categorically distinct groups of participants based on patterns of responses across multiple variables; an advantage is that it does not rely on researcher-defined or relatively arbitrary cutoff scores to classify participants. Moreover, it is a probabilistic technique that allows for participants to be classified into a given profile while also quantifying the degree to which profile membership is non-perfect. That is, LPA not only assigns participants to a profile but also provides probabilities that the given participant is also assigned to other profiles. Two LPAs were conducted in this study, each one from data collected at two time points, with 30 months between the initial data collection and the second round of data collection. The second LPA included a measure of reading fluency in addition to all the measures included in the first LPA. A conceptual diagram of the LPAs, which includes each LPA, is presented in Figure 1. LTA is a longitudinal extension of LPA. LTA was used to examine the proportions of participants who did or did not transition between profiles over time. This provided information regarding the stability, and potential malleability, of the reader profiles in the sample.

Scaled scores were converted to z-scores, using the population norms of each test, to facilitate interpretation of the item-profile plots since they were on different scales.

The first step was to conduct two separate LPAs, one from the first data collection time point and the second from the follow-up data collection 30 months later. Sets of models consisting of one to five profiles were conducted with each set using a different variance-covariance specification as recommended by Masyn (2013). The set of models that Masyn calls *class-invariant, diagonal* was chosen as the final set because convergence problems occurred with sets of models utilizing more complex variance-covariance structures. These problems were due to the limited sample size relative to the increasing number of estimable parameters in the more complex models. The class-invariant, diagonal structure specifies that indicator covariances are constrained to zero within profile and indicator variances are constrained to equality across profiles. This is the default specification in M*plus* 8.

All models were conducted using full information maximum likelihood estimation. This estimator uses all available data meaning an individual can be included in the analyses as long as that individual contains data on at least one variable. In this study, all 80 participants contributed data to the first LPA. Individuals who were missing at the second time point did not contribute to the second LPA nor to the LTA. That is, all 80 individuals were used to estimate the latent profiles at the first time point, but only 64 individuals were used to estimate the latent transitions from the first LPA to the second LPA.

Age was treated as a covariate of each latent profile variable to examine whether older participants were significantly more likely to be classified into a given profile relative to a reference profile. Age was centered in both LPAs. ASD symptomatology was treated as a distal outcome—even though it is not truly distal—for estimation purposes to obtain profile-specific means of ASD symptomatology and compare them for significant differences. This decision was made to facilitate a clearer understanding of how ASD symptomatology related to the emergent profiles. That is, interpreting the comparisons of profile-specific means of ASD symptomatology was preferred over comparing log odds of membership in a given profile in relation to a reference profile. These analyses were accomplished using the BCH approach (Asparouhov and Muthén, 2014) to incorporate both analyses into one model. The covariate effects were estimated using a multinomial logistic regression, which presents results on the logit metric and requires a reference profile to compare to all other profiles. Logits can be converted into odds ratios, which may be thought of as an effect size. Odds ratios of 1.0 indicate a 1:1 ratio of being classified into either the reference or the target profile, while values that deviate from 1.0 indicate greater chances of being placed into one of the profiles.

Multiple fit statistics were used for the enumeration process. When fit statistics were ambiguous, substantive checks were employed to ensure the theoretical viability of a particular model, which has been shown to be a key consideration when choosing the preferred model (Muthén, 2003). Information criteria included the Bayesian Information Criterion (BIC; Schwarz, 1978), which has been shown to most consistently identify the best model (Nylund et al., 2007), and the Adjusted Bayesian Information Criterion (ABIC). Minimum values for these fit statistics indicate the preferred model. Two likelihood-based ratio tests—the adjusted Lo–Mendell–Rubin (LMR) and bootstrapped likelihood ratio test (BLRT)—were also used. These indexes provide p values that compare a model with k profiles to a model with $k-1$ profiles. When there is a non-significant p value for the model with k profiles and a significant p value for the model with $k-1$ profiles, the model with k profiles is preferred. Simply put, a non-significant p value indicates the addition of another profile did not significantly improve the model and the more parsimonious version should be preferred. Finally, this study also utilized the Bayes Factor (BF) and correct model probability (cmP), which are information-heuristic statistics derived from the BIC (Masyn, 2013). The BF compares two competing models by providing a ratio of probabilities that a model with $k-1$ profiles is correct compared to a model with k profiles. Values for the $k-1$ model between 1 and 3 are weak evidence, 3 and 10 moderate evidence, and >10 strong evidence. The cmP is interpreted as the probability of a particular model being preferred out of a set of models under consideration. Although not a fit statistic, we report the entropy values for the chosen models as a general measure of strength of classification.

Finally, this study connected the chosen LPAs at both time points using LTA. It was necessary to conduct the measurement models for each individual separately before linking them because it has been shown that estimating structural parameters simultaneously with measurement parameters can cause undesirable shifts in the latent profiles (Nylund, 2007; Nylund-Gibson et al., 2014; Nylund-Gibson and Masyn, 2016; Vermunt, 2010). When estimating the LTA, the primary concern was not with fit statistics, but ensuring the individual LPAs did not influence each other. That is, when two or more LPAs are connected using LTA, there is the potential for profile membership at time 2 to affect profile membership at time 1. This was considered chronologically unviable. To ensure this did not occur, this study employed the three-step method (Nylund-Gibson et al., 2014; Vermunt, 2010). This approach ensured participants were assigned to their latent profiles from the individual LPAs. This method also accounted for classification error resulting from non-perfect assignment to latent profiles. Thus, this study accounted for participants' assignment to profiles at time points 1 and 2, classification error, and the potential for each time point to influence each other unrealistically.

Table 2. Fit statistics for LPAs at time points 1 and 2.

No. of profiles	LL	BIC	ABIC	LMR *p* value	BLRT *p value*	BF	cmP
Time 1							
1	−1184.89	2466.19	2396.81	N/A	N/A	<0.001	<0.001
2	−1089.51	2328.01	2220.79	0.003	<0.001	<0.001	0.261
3	−1064.36	2330.30	2185.24	0.695	<0.001	0.135	0.083
4	−1036.17	2326.29	2143.40	0.311	<0.001	15.959	0.617
5	−1012.55	2331.83	2111.10	0.635	<0.001	N/A	0.039
Time 2							
1	−1060.09	2219.99	2144.46	N/A	N/A	<0.001	<0.001
2	−973.90	2101.68	1985.23	0.211	<0.001	0.020	0.018
3	−942.97	2093.89	1936.52	0.228	<0.001	11.302	0.902
4	−918.37	2098.74	1900.46	0.619	<0.001	N/A	0.080
5	Convergence problems						

LL: log-likelihood; BIC: Bayesian Information Criterion; ABIC: Adjusted Bayesian Information Criterion; LMR: Lo–Mendell–Rubin likelihood ratio test; BLRT: bootstrapped likelihood ratio test; BF: Bayes Factor; cmP: correct model probability.

Results

Latent profile analyses

Fit statistics for the first LPA are presented in the top panel of Table 2. The minimum BIC value occurred with the four-profile model, though values for models consisting of two to five profiles were all similar, with a range of only five points. This relative ambiguity necessitated examination of not only additional fit statistics but also all item-profile plots for theoretical consideration. The ABIC never reached a minimum value. The LMR supported the two-profile model, while the BLRT never became non-significant, so it was considered uninformative. Both the BF and cmP supported the four-profile model. The entropy for this model was 0.90.

Thus, the four-profile model had the most statistical support, but item-profile plots with two to five profiles were examined to be comprehensive in choosing the preferred model. The two-profile model showed rank-ordered profiles across all variables, which was not consistent with the literature citing specific reading difficulties. The five-profile model yielded profiles that overlapped on multiple variables, which suggested overfitting the data. Thus, the two- and five-profile models were removed from further consideration. The three-profile model also yielded ordered profiles, while the four-profile model included a profile that was consistent with prior research, namely, participants with average word reading skills, but below-average reading comprehension skills. In addition, all profiles consisted of sizable proportions of the sample suggesting the emergent profiles were substantively meaningful; this solution was consistent with McIntyre et al. (2017b). The item-profile plot is depicted in the top panel of Figure 2. The profile demarcated by a solid line with circular markers scored near average, and higher than the other profiles, on most variables. This profile was labeled *Average* and consisted of 18.0% of the sample. The profile depicted by a dotted line with square markers exhibited some lower scores in language variables, but was characterized by a particular difficulty in reading comprehension. This profile was labeled *Comprehension Disturbance* and made up 24.2% of the sample. The profile with a dashed line and X markers scored below average on all but one variable and was labeled *Below Average/Intact RV*. This profile contained 23.6% of the sample. The profile at the bottom with a solid line and triangular markers scored the lowest on all variables. It was labeled *Global Disturbance* and consisted of 34.3% of the sample.

Fit statistics for the second LPA are presented in the bottom panel of Table 2. The five-profile model estimated more parameters than the sample size, which led to convergence problems, so it was excluded. The BIC values for the two-, three-, and four-profile models again had similar values, though the minimum value occurred with the three-profile model. The ABIC never reached a minimum value, the LMR never became significant, and the BLRT never became non-significant. These three fit statistics were uninformative. The BF supported the three-profile model, but could not be computed for the four-profile model, due to the exclusion of the five-profile model. Thus, it was difficult to draw a conclusion from the BF. The cmP supported the three-profile model, but this was not surprising as it is derived from the BIC. Therefore, given ambiguity in the fit statistics (particularly the similarity among BIC values for the models with two, three, and four profiles), substantive interpretation of the item-profile plots with two to four profiles was critical. As with the LPA at the first time point, the two- and three-profile plots consisted of rank-ordered profiles inconsistent with previous research. The four-profile plot showed more nuanced profiles that aligned with extant literature and was chosen as the preferred model. The entropy for this model is 0.91.

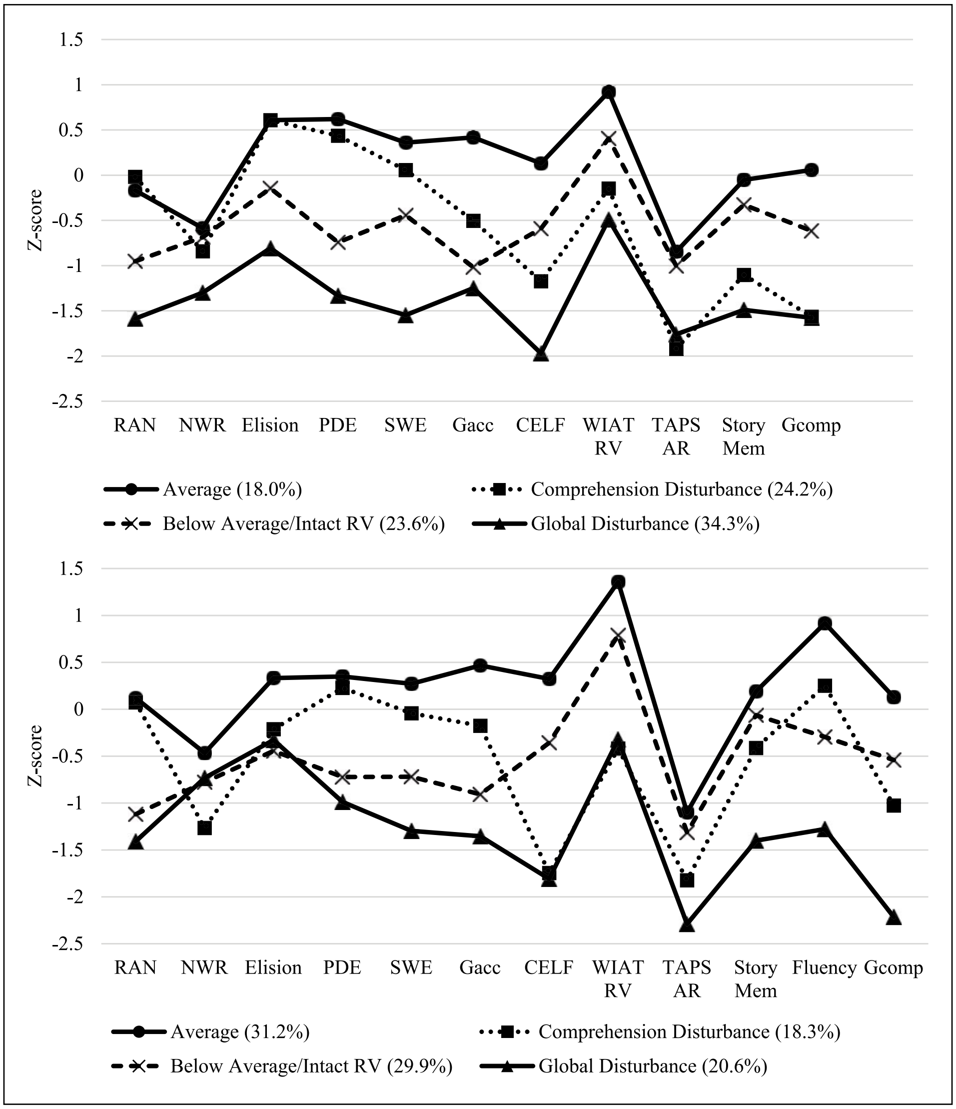

Figure 2. Item-profile plots for each LPA. The first LPA is on top.

The item-profile plot is depicted in the bottom panel of Figure 2. The profiles were fairly similar to the profiles from time point 1. The *Average* profile consisted of 31.2% of the sample. The *Comprehension Disturbance* profile exhibited particular difficulties in CELF Recalling Sentences, TAPS Auditory Reasoning, and GORT Reading Comprehension and consisted of 18.3% of the sample. The third emergent profile is demarcated by dashed lines with an X marker in the bottom panel of Figure 2. This profile was characterized by below average performance across most variables except the WIAT RV subtest. Similar to time point 1, this profile was labeled *Below Average/Intact RV* and consisted of 29.9% of the sample. The final profile represented by a solid line with triangle markers scored one or more standard deviations below average on all but three variables. This profile was termed *Global Disturbance* and consisted of 20.6% of the sample.

Profile separation for each LPA

One method of validating the interpretation of the latent profiles is measuring the amount of profile separation between each profile on each indicator. Profile separation is measured not only by the distance between the profile-specific means of each indicator but also by the variances. This can be quantified by adapting the formula for Cohen's *d* (Masyn, 2013), which yields standardized interclass differences. Coefficients with an absolute value less than 0.85 correspond to a low degree of separation and absolute values greater than 2.0 correspond to a high degree of separation (Masyn, 2013).

Results for all pairwise profile comparisons for each indicator are presented in Table 3. For the first LPA, the greatest separation occurred between the *Average* and *Global Disturbance* profiles as all coefficients were greater than 0.85. There were two comparisons in which the

Table 3. Standardized differences between all pairwise comparisons of profiles for all indicators for each LPA.

Variable	*Comprehension Disturbance* versus *Average*	*Comprehension Disturbance* versus *Global Disturbance*	*Comprehension Disturbance* versus *Below Average/Intact RV*	*Average* versus *Global Disturbance*	*Average* versus *Below Average/Intact RV*	*Global Disturbance* versus *Below Average/Intact RV*
Time 1						
RAN	0.21	2.16	1.29	1.96	1.08	−0.87
NWR	−0.39	0.69	−0.24	1.08	0.15	−0.93
Elision	0.00	1.79	0.95	1.79	0.95	−0.84
PDE	−0.31	2.97	1.98	3.28	2.29	−0.99
SWE	−0.46	2.45	0.76	2.91	1.23	−1.69
Gacc	−1.52	1.23	0.85	2.75	2.37	−0.38
CELF	−1.90	1.16	−0.85	3.06	1.05	−2.01
WIAT RV	−1.06	0.34	−0.55	1.40	0.51	−0.89
TAPS AR	−1.34	−0.20	−1.14	1.14	0.20	−0.95
Sty Mem	−1.16	0.43	−0.85	1.58	0.31	−1.28
Gcomp	−3.16	0.02	−1.85	3.19	1.31	−1.87
Time 2						
RAN	−0.05	−1.57	−1.26	−1.62	−1.31	0.31
NWR	−1.26	0.84	0.77	−0.42	−0.49	−0.07
Elision	−0.65	−0.14	−0.28	−0.79	−0.93	−0.14
PDE	−0.21	−2.07	−1.62	−2.27	−1.83	0.45
SWE	−0.50	−2.01	−1.09	−2.51	−1.59	0.92
Gacc	−0.89	−1.63	−1.02	−2.53	−1.91	0.62
CELF	−2.39	−0.07	1.61	−2.47	−0.79	1.68
WIAT RV	−2.12	0.12	1.44	−2.00	−0.68	1.32
TAPS AR	−0.94	−0.61	0.66	−1.54	−0.28	1.27
Sty Mem	−0.70	−1.14	0.40	−1.83	−0.29	1.54
Fluency	−1.11	−2.54	−0.91	−3.64	−2.01	1.63
Gcomp	−2.11	−2.17	0.89	−4.29	−1.23	3.06

RAN: rapid automatized naming; NWR: nonword repetition; PDE: phonemic decoding efficiency; SWE: sight word efficiency; Gacc: Gray Oral Reading Test accuracy; WIAT RV: Wechsler Individual Achievement Test–Receptive Vocabulary; TAPS AR: Test of Auditory Processing–Auditory Reasoning; Sty Mem: story memory; Gcomp: Gray Oral Reading Test Comprehension; CELF: Clinical Evaluation of Language Fundamentals; LPA: latent profile analysis.
Profile names in parentheses refer to the names of the profiles at time point 2 that differed from time point 2.

absolute values of the coefficients for 6 of the 11 indicators were greater than 0.85 and the remaining were less than 0.85. First, *Comprehension Disturbance* and *Average* were not well separated by the phonological and word reading variables, but were well separated in the linguistic and reading comprehension variables. Second, *Comprehension Disturbance* and *Global Disturbance* were generally well separated by the phonological and reading variables, but not the linguistic and reading comprehension variables.

For the second LPA, the greatest amount of separation was between the *Average* and *Global Disturbance* profiles; absolute values of the coefficients for 10 of the 12 indicators were greater than 0.85. The rest of the pairwise comparisons yielded either seven or eight coefficients whose absolute values were greater than 0.85. These differences largely supported the labels assigned to each latent profile. For example, *Comprehension Disturbance* and *Global Disturbance* were well separated by a mix of word reading and linguistic and reading comprehension variables (as *Global Disturbance* scored lowest across all types of variables). Furthermore, *Comprehension Disturbance* and *Average* were primarily separated by linguistic and reading comprehension variables, but not word reading variables.

Relating age and ASD symptomatology to the LPAs

Results for the age covariate at both time points are presented in Table 4. The reference profile at both time points was *Average*. For both time points, age did not significantly predict profile membership. Participants at time point 1 between ages 8 and 16 years had statistically equal probabilities (logits ranged from −0.09 to 0.01, *p*s ranged from 0.60 to 0.94) of being classified into any profile. For time point 2, logits ranged from −0.17 to 0.01, *p*s ranged from 0.34 to 0.74. This was also supported substantively by the odds ratios, which showed nearly 1:1 chances (0.91:1–1.01:1 at time point 1; .84:1–1.06:1 at time point 2) of being classified into any profile.

Table 4. Age effects for each LPA model.

Profile	Logit	SE	Logit/SE	*p* value	OR
Time 1					
Comprehension Disturbance	−0.09	0.17	−0.53	0.60	0.91
Below Average/Intact RV	−0.03	0.16	−0.16	0.88	0.97
Global Disturbance	0.01	0.17	0.08	0.94	1.01
Time 2					
Comprehension Disturbance	0.06	0.18	0.34	0.74	1.06
Below Average/Intact RV	−0.17	0.18	−0.96	0.34	0.84
Global Disturbance	0.01	0.18	0.34	0.74	1.01

LPA: latent profile analysis; SE: standard error; RV: Receptive Vocabulary; OR: odds ratio.
Reference profile is *Average*.

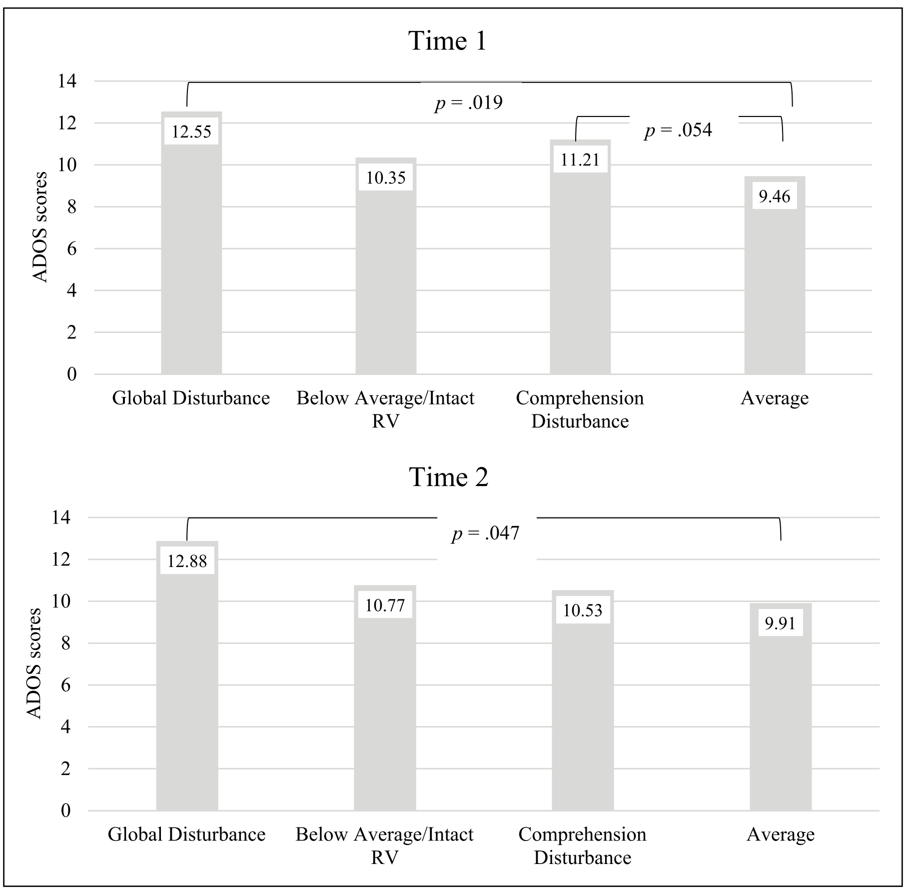

Figure 3. Profile-specific means of ADOS at each time point.

Figure 3 shows significant differences in ADOS scores for all pairwise comparisons between profiles at time point 1 (top panel) and time point 2 (bottom panel). At the first time point, *Global Disturbance* had a significantly higher mean ADOS score than *Average* ($p=0.019$). The ADOS score difference between *Comprehension Disturbance* and *Average* was marginally significant ($p=0.054$). At the second time point, the only significant difference in ADOS scores was between *Global Disturbance* and *Average* with the former scoring higher than the latter ($p=0.047$).

Transitions between profiles over time

Transitions between profiles are presented in Table 5. Values on the diagonal represent proportions of participants who remained in the same profile. Values on the off-diagonals represent proportions of participants who

Table 5. Probabilities of transitioning between profiles.

Time point 1	Time point 2			
	Average	*Comprehension Disturbance*	*Below Average/Intact RV*	*Global Disturbance*
Average	0.89	0.06	0.05	0.00
Comprehension Disturbance	0.28	0.57	0.06	0.09
Below Average/Intact RV	0.21	0.09	0.69	0.01
Global Disturbance	0.00	0.14	0.10	0.76

RV: Receptive Vocabulary.

transitioned to different profiles between time points 1 and 2. The *Comprehension Disturbance* profile had the smallest proportion of participants who remained in the profile at time point 2, indicating it was the least stable, and possibly the most malleable. Of those that were originally in the *Comprehension Disturbance* profile, 28% transitioned to the *Average* profile. Much smaller proportions transitioned to the *Below Average/Intact RV* (6%) and *Global Disturbance* (9%) profiles, which is an optimistic finding.

Of those that were in the *Below Average/Intact RV* profile at time point 1, 21% transitioned into the *Average* profile. The majority were stable (69%), while 9% transitioned to *Comprehension Disturbance* and 1% transitioned to *Global Disturbance*. The *Global Disturbance* profile was largely stable, with 76% remaining in the profile, 10% transitioning to *Below Average/Intact RV*, 14% transitioning to *Comprehension Disturbance*, but no participants transitioning to *Average*.

Discussion

The main goal of this study was to examine heterogeneous profiles of readers with ASD and determine whether these subgroups of readers are stable longitudinally. Latent profile analyses suggested similar reading profiles at time points 1 and 2 in a sample of students with ASD. It is notable that across both time points, all four subgroups had relative difficulty with CTOPP nonword repetition and TAPS Auditory Reasoning. These findings are consistent with prior research. Extant data have demonstrated that nonword repetition tasks are promising markers of language impairment, showing high rates of sensitivity and specificity (Conti-Ramsden et al., 2001). Many children with ASD exhibit developmental language impairments and previous studies have found that many children and adolescents with ASD have difficulty with nonword repetition tasks, but potentially less severely and for different underlying reasons than those with Specific language Impairment (SLI; Riches et al., 2011; Whitehouse et al., 2008; Williams et al., 2013). The Auditory Reasoning subtest required one to apply inferential skills to answer questions about short vignettes. Findings from this study align with previous research into the nature and high prevalence of inference impairments in individuals with ASD (e.g. Ricketts et al., 2013; Saldaña and Frith, 2007; Tirado and Saldaña, 2016).

This study built upon the findings of McIntyre et al. (2017b) and there are differences that merit attention. This study found profiles that were similar to McIntyre et al. (2017b) in terms of shape, but differed in level. For example, the *Global Disturbance* profile in this study was similar in shape to McIntyre et al.'s *Severe Global Disturbance* profile, but *Global Disturbance* in this study scored at a higher level across variables. The *Average* and *Comprehension Disturbance* profiles—at both time points in this study—resembled the analogous profiles found in McIntyre et al.

There was a notable difference between the observed variables McIntyre et al. (2017b) used to measure the latent profiles and those used in this study. McIntyre et al. utilized a measure of expressive vocabulary, while this study used a measure of RV. This did not appear to be a particularly meaningful measure as students' scores within each profile did not seem to align with other measures thought to tap similar skills. This can easily be seen in Figure 2 by the peaks caused by this measure relative to the scores on adjacent measures, for all profiles. With respect to practitioners, RV measures may not provide usable information to guide targeted interventions.

While analyses suggested similar reading profiles at both time points in this study, profile membership was not as stable longitudinally as has been seen in previous literature with students who were not diagnosed with ASD in the same age range (Catts et al., 2012). Subsamples of students in the *Comprehension Disturbance* and *Below Average/Intact RV* profiles improved their reading skills between time points 1 and 2; however, the prevalence of reading difficulties in the overall sample remained much higher than typically developing samples, with 68.8% of the sample experiencing difficulties with comprehension and 50.5% with below-average word reading abilities at time point 2. While *Comprehension Disturbance* has been well documented in the literature, the prevalence of individuals with ASD that also have word level reading difficulties have been investigated to a lesser extent.

The *Comprehension Disturbance* subgroup represents the subgroup of ASD readers that have been most reported in the extant literature, those with intact decoding abilities and comprehension difficulties. While this profile existed across the two time points, it was also the profile with the greatest proportion of participants who transitioned between the two time points. Specifically, between time points 1 and 2, only 57% of the sample remained in the *Comprehension Disturbance* subgroup. Of the remaining 43%, most students transitioned to the *Average* profile (28% of the total students in *Comprehension Disturbance* at time point 1); additionally, the *Comprehension Disturbance* profiles represented the smallest proportion of the sample by time point 2. This finding was somewhat surprising given the extent to which this profile of reader has been reported in the previous studies of reading in ASD samples. Since a sizable proportion of the *Comprehension Disturbance* profile transitioned to the *Average* profile, it would be worthwhile to explore what contributed to these students' transition, though an empirical examination was beyond the scope of this study. One possible explanation may be that social cognition plays a role as it has been suggested that social cognition is related to reading comprehension in students with ASD (Ricketts et al., 2013). As these students mature, they receive more opportunities to develop social cognition or may receive direct intervention in these skills. Thus, students who are successful may have increased chances of improved reading comprehension skills and transition into the *Average* profile. An alternative explanation may be related to the reading instruction the students received. If teachers identified these students' specific comprehension difficulties, they may have provided the students with strategies targeting comprehension. Interestingly, the second largest transition probability was 0.21 (among students who switched profiles), which was the proportion of students in the *Below Average/Intact RV* profile who transitioned into the *Average* profile. Thus, the *Comprehension Disturbance* and *Below Average/Intact RV* profiles contained students who appeared to improve over time in their overall reading development, though this study could not pinpoint the motivating factors for these transitions, whether social cognition, instruction, or other shared characteristics or factors. The field would benefit from future studies that examine the characteristics of children who demonstrate improvement over time in reading skills; the sample size in this study is not large enough to look at within-profile specific characteristics.

The addition of reading fluency at time point 2 allowed for a more nuanced examination of the reading profiles. Previous literature has concentrated on single word decoding abilities as a strength in the ASD population, but has largely ignored the ability to read connected text, or reading fluency; recent literature has shown that reading fluency may play an important role in the reading comprehension abilities of individuals with ASD (Solari et al., 2017). Reading fluency exhibited wider variation than the other variables, which suggests it can be valuable in differentiating readers with ASD. Moreover, the achievement patterns for fluency and the word reading variables were identical across profiles. With this age range, fluency may be a more efficient measure of reading skills since findings were identical to multiple measures of single word reading, which is important for educational practitioners. Notably, fluency and reading comprehension seemed to efficiently summarize the overall profiles. *Average* scored highest on both measures and *Global Disturbance* scored lowest on both measures. *Comprehension Disturbance* scored higher than *Below Average/Intact RV* on fluency, but lower on reading comprehension. The reading fluency measure was not available at time point 1, so this study cannot assess whether these results would have been stable longitudinally. These findings, however, provide initial evidence that fluency may be important for readers with ASD and should be examined in future studies.

Efficient measures to differentiate profiles

Profile separation across indicators provided additional support for the conceptual validity of the profiles and these results may aid practitioners in tailoring reading interventions for students with ASD as well as allocating resources efficiently. Results for profile separation were generally consistent across time. With respect to differentiating the *Comprehension Disturbance* and *Average* profiles, there were generally only small effects in the phonological processing and word reading variables. As would be expected, there were moderate to large effects among the linguistic and reading comprehension variables. However, fluency also had a moderate effect. Therefore, students who demonstrate a specific comprehension difficulty may also require instruction in strategies designed to improve passage-level fluency. The *Comprehension Disturbance* and *Below Average/Intact RV* profiles tended to be separated mostly by word reading variables, which makes intuitive sense given that both profiles would show depressed skills in comprehension, but *Comprehension Disturbance* would be expected to demonstrate adequate word reading. This was also true when comparing the *Comprehension Disturbance* and *Global Disturbance* profiles, but they were further differentiated by fluency and reading comprehension. The *Below Average/Intact RV* and *Global Disturbance* profiles were mostly separated by linguistic comprehension, fluency, and reading comprehension variables.

In terms of informing intervention, it is clear that the majority of students require intervention in linguistic comprehension skills (i.e. *Comprehension Disturbance, Below Average/Intact RV*, and *Global Disturbance* profiles). This may include vocabulary, narrative recall, and inferencing skills. A recent study by Grimm et al. (2018) found that the linguistic comprehension skills of students diagnosed with ASD developed at a similar rate, but at a lower level, compared to same-aged typically developing peers. This study builds upon these findings by providing a more nuanced picture such that linguistic comprehension skills can vary by latent profile. The *Average* profile scored average or above—compared to national norms—on three of the four linguistic comprehension measures (i.e. CELF-4, WIAT-III, and WRAML-2), while the other profiles scored below average on all linguistic comprehension measures, except *Below Average/Intact RV* scored average on the WIAT-III RV. This was true at both time points. Thus, while the majority of students diagnosed with ASD appear to require intervention in linguistic comprehension, it is possible the development of linguistic comprehension skills of students in the *Average* profile may follow a trajectory similar to typically developing students. Screening for this would allow educational practitioners to allocate resources more efficiently.

This study also suggests that word reading and passage fluency may be critical skills for longer periods of time than is typical of other student populations. Both skills differentiated the *Comprehension Disturbance, Below Average/Intact RV*, and *Global Disturbance* profiles. Studies of typically developing students have shown word reading skills are important primarily in the early grades (e.g. Adlof et al., 2006; Hoover and Gough, 1990). This study, however, included a large age range of students diagnosed with ASD, but age did not predict profile membership. This suggests word reading skills remain an important predictor of reading comprehension if students belong to one of the three lower performing profiles regardless of age.

The data demonstrate that both the *Below Average/Intact RV* and *Global Disturbance* profiles require intervention in word reading and linguistic comprehension skills. In typically developing populations, word reading is often viewed as a lower order skill compared to linguistic comprehension. However, this may not be the case for students diagnosed with ASD. Based on this study, students in the *Below Average/Intact RV* and *Global Disturbance* profiles were differentiated by linguistic comprehension, so intervention might best be focused on linguistic comprehension if a practitioner is serving a student in the *Global Disturbance* profile and the goal is to raise the student to the *Below Average/Intact RV* profile. Yet, it should be emphasized that both skill areas would be in need of remediation, but practitioners should be aware of the need to differentiate resource allocation dependent on individual students' strengths and weaknesses.

Relation between reading profiles and ASD symptom severity

Similar to previous findings (McIntyre et al., 2017b; Ricketts et al., 2013), reading impairment was related to ASD symptom severity; in this study, we show that these relations change over time. When measured concurrently at the first time point, ASD symptom severity was significantly different between the *Average* and the *Global Comprehension* readers, while the difference between *Average* and *Comprehension Disturbance* was marginally significant. This latter difference did not occur at the second time point, which reflects the transitions between profiles, specifically, those who transitioned from *Comprehension Disturbance* to *Below Average/Intact RV* and *Average*. However, we also note caution in this interpretation as the difference between *Global Disturbance* and *Average* went from significant at the first time point to marginally significant at the second time point. One possible explanation is that individual participants' ADOS scores did not change over time because ADOS was collected only at the first time point. Rather, the profile-specific average ADOS scores changed. It is possible that the difference between *Global Disturbance* and *Average* may have remained equally significant if ADOS scores were collected at the second time point and included as a covariate in the second LPA. If students' symptomatology evolved over the 30-month period, there may have been a different relationship between ADOS and reading skills at the second time point. Although not statistically significant, the fact that the average ADOS scores of the *Below Average/Intact RV, Comprehension Disturbance*, and *Average* profiles converged at the second time point suggests those with lower ADOS scores (i.e. lower symptom severity) were more likely to transition from *Comprehension Disturbance* to *Average*, while those with higher ADOS scores were more likely to transition from *Comprehension Disturbance* to *Below Average/Intact RV*.

While the trends toward improving reading skills over time are positive, compared to typical reading development trajectories, the reading performance of children diagnosed with ASD, with an IQ in the normative range, remains highly impaired when compared to their cognitive abilities. In this higher functioning sample, with average full-scale IQ, at time point 2, only 31.2% of the sample demonstrated average reading abilities. This is significantly lower than typically developing samples, where approximately 80% of individuals score within the average range when they have comparable cognitive functioning. In this sample, the majority of students spent most of their instructional time in general education settings. The persistent reading difficulties of this sample indicates that a large proportion of individuals with ASD are not responding to the typical reading curriculum. This argues for the importance of comprehensive, evidence-based reading instruction to be provided to individuals diagnosed with ASD from the earliest stages of schooling. Furthermore, the discrepancy between cognitive abilities and reading abilities, and the apparent relation between ASD symptomology and reading comprehension performance suggests the need to develop and empirically test targeted reading interventions specific to the needs of individuals diagnosed with ASD.

Limitations

One limitation of this study is the relatively small sample size. However, in LPA, statistical power is not a function of only sample size. Unlike more traditional analytic techniques, power to detect whether parameter coefficients (e.g. profile-specific indicator means) are statistically significant is not often the focus in LPA. Rather, statistical power in this context generally applies to whether the plurality of fit indexes are able to identify the "correct" model. Tein et al. (2013) found that sample size had minimal effects on statistical power compared to other factors such as interclass distance among the indicators as well as the number of indicators. This study included a large number of indicators and the results demonstrated that 47 out of 66 profile-specific indicator mean comparisons in the first LPA and 46 out of 72 in the second LPA showed standardized interclass differences $>|0.85|$. Finally, the substantive interpretations of the LPAs and LTA were theoretically viable, which is a primary concern in mixture modeling (Muthén, 2003). Thus, it is possible that limitations related to the limited sample size were, at least, partially mitigated by the number of indicators, interclass differences, and substantive interpretation.

Although this study identified two significant differences in autism symptom severity across profiles, it may be more differences among profiles exist and this study did not have the power to detect them. Second, it would have been valuable to measure autism symptom severity at both time points to examine whether individual changes could be related to transitions between profiles. It is possible that individuals with better social communication skills are better able to access the general education curriculum, including reading instruction. This study did not include predictors of the latent transitions; therefore, conclusions concerning what motivated the transitions could not be drawn. Examining such predictors, with a larger sample size, is a promising area of future research.

Conclusion

This study provides evidence that a majority of students diagnosed with ASD, with normative range IQ, experience impaired reading abilities, but also empirically identified four categorically distinct profiles based on varying reading abilities. However, these profiles were not entirely stable over time and participants were able to transition between profiles within the 30-month timeframe of this study. Furthermore, the most often reported profile in the literature, *Comprehension Disturbance*, was the most malleable across time suggesting the potential for the design of effective interventions. Autism symptom severity was related to profile membership, such that lower severity was associated with generally greater reading achievement.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the US Department of Education Institute of Education Sciences (Grant/Award Number: "R324 A120168").

References

Adlof S, Frishkoff G, Dandy J and Perfetti C (2016) Effects of induced orthographic and semantic knowledge on subsequent learning: a test of the partial knowledge hypothesis. *Reading and Writing* 29(3): 475–500.

Åsberg J and Sandberg AD (2012.) Dyslexic, delayed, precocious or just normal? Word reading skills of children with autism spectrum disorders. *Journal of Research in Reading* 35(1): 20–31.

Asparouhov T and Muthén B (2014) Auxiliary variables in mixture modeling: three-step approaches using M Plus. *Structural Equation Modeling: A Multidisciplinary Journal* 21(3): 329–341.

Baio J, Wiggins L, Christensen DL, et al. (2018) Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring Network, 11 Sites, United States, 2014. *MMWR Surveillance Summaries* 67(6): 1–23.

Breaux KC (2009) *Wechsler Individual Achievement Test-3rd Edition: Technical manual*. San Antonio, TX: Pearson.

Brown HM, Oram-Cardy J and Johnson A (2013) A meta-analysis of the reading comprehension skills of individuals on the autism spectrum. *Journal of Autism and Developmental Disorders* 43(4): 932–955.

Capps L, Losh M and Thurber C (2000) "The frog ate the bug and made his mouth sad": narrative competence in children with autism. *Journal of Abnormal Child Psychology* 28(2): 193–204.

Catts HW, Compton D, Tomblin JB, et al. (2012) Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology* 104(1): 166–181.

Catts HW, Hogan TP and Fey ME (2003) Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities* 36(2): 151–164.

Conti-Ramsden G, Botting N and Faragher B (2001) Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry* 42(6): 741–748.

Davidson MM and Weismer SE (2014) Characterization and prediction of early reading abilities in children on the autism spectrum. *Journal of Autism and Developmental Disorders* 44(4): 828–845.

Estes A, Rivera V, Bryan M, et al. (2011) Discrepancies between academic achievement and intellectual ability in higher-functioning school-aged children with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 41(8): 1044–1052.

Frith U and Snowling M (1983) Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology* 1(4): 329–342.

Good RH, Kaminski RA, Cummings K et al. (2011) *DIBELS Next*. Eugene, OR: Dynamic Measurement Group.

Gough PB and Tunmer WE (1986) Decoding, reading, and reading disability. *Remedial and Special Education* 7(1): 6–10.

Grigorenko EL, Klin A and Volkmar F (2003) Annotation: hyperlexia: disability or superability? *Journal of Child Psychology and Psychiatry* 44(8): 1079–1091.

Grimm RP, Solari EJ, McIntyre NS, et al. (2018) Comparing growth in linguistic comprehension and reading comprehension in school-aged children with autism versus typically developing children. *Autism Research* 11(4): 624–635.

Henderson LM, Clarke PJ and Snowling MJ (2014) Reading comprehension impairments in autism spectrum disorders. *L'année Psychologique* 114(4): 779–797.

Hoover WA and Gough PB (1990) The simple view of reading. *Reading and Writing: An Interdisciplinary Journal* 2: 127–160.

Huemer SV and Mann V (2010) A comprehensive profile of decoding and comprehension in autism spectrum disorders. *Journal of Autism and Developmental Disorders* 40(4): 485–493.

Hughes C (2001) Executive dysfunction in autism: its nature and implications for the everyday problems experienced by individuals with autism. In: Burack JA, Charman T, Yirmiya N, et al. (eds) *The Development of Autism: Perspectives from Theory and Research*. Mahwah, NJ: Lawrence Erlbaum Associates, pp.255–275.

Jones CR, Happé F, Golden H, et al. (2009) Reading and arithmetic in adolescents with autism spectrum disorders: peaks and dips in attainment. *Neuropsychology* 23(6): 718–728.

Joseph RM and Tager-Flusberg H (2004) The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and Psychopathology* 16(1): 137–155.

Lindgren KA, Folstein SE, Tomblin JB, et al. (2009) Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research* 2(1): 22–38.

Lord C, DiLavore PC and Gotham K (2012) *Autism Diagnostic Observation Schedule*. Torrance, CA: Western Psychological Services.

Lucas R and Norbury C (2014) Levels of text comprehension in children with autism spectrum disorders (ASD): the influence of language phenotype. *Journal of Autism and Developmental Disorders* 44(11): 2756–2768.

McIntyre NS, Oswald TM, Solari EJ, et al. (2018) Social cognition and reading comprehension in children and adolescents with autism spectrum disorders or typical development. *Research in Autism Spectrum Disorders* 54: 9–20.

McIntyre NS, Solari EJ, Gonzales JE, et al. (2017a) The scope and nature of reading comprehension impairments in school-aged children with higher-functioning autism spectrum disorder. *Journal of Autism and Developmental Disorders* 47(9): 2838–2860.

McIntyre NS, Solari EJ, Grimm RP, et al. (2017b) A comprehensive examination of reading heterogeneity in students with high functioning Autism: distinct reading profiles and their relation to autism symptom Severity. *Journal of Autism and Developmental Disorders* 47(4): 1086–1101.

Martin N and Brownell R (2005) Test of Auditory Processing Skills, Third edition. Novato, CA: Academic Therapy Publications.

Masyn K (2013) Latent class analysis and finite mixture modeling. In: Little T (ed.) *The Oxford Handbook of Quantitative Methods in Psychology*, vol. 2. Oxford: Oxford University Press, pp.375–393.

Mayes SD and Calhoun SL (2006) Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learning and Individual Differences* 16(2): 145–157.

Muthén B (2003) Statistical and substantive checking in growth mixture modeling: comment on Bauer and Curran (2003). *Psychological Methods* 8: 369–377.

Nation K, Clarke P, Wright B, et al. (2006) Patterns of reading ability in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 36(7): 911–919.

Newman TM, Macomber D, Naples AJ, et al. (2007) Hyperlexia in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 37(4): 760–774.

Norbury C and Nation K (2011) Understanding variability in reading comprehension in adolescents with autism spectrum disorders: interactions with language status and decoding skill. *Scientific Studies of Reading* 15(3): 191–210.

Nylund KL (2007) *Latent transition analysis: modeling extensions and an application to peer victimization*. Doctoral Dissertation. Available at: www.statmodel.com

Nylund KL, Asparouhov T and Muthén BO (2007) Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal* 14(4): 535–569.

Nylund-Gibson K, Grimm R, Quirk M, et al. (2014) A latent transition mixture model using the three-step specification. *Structural Equation Modeling: A Multidisciplinary Journal* 21(3): 439–454.

Nylund-Gibson KL and Masyn KE (2016) Covariates and mixture modeling: results of a simulation study exploring the impact of misspecified effects on class enumeration. *Structural Equation Modeling: A Multidisciplinary Journal* 23: 782–797.

O'Connor IM and Klein PD (2004) Exploration of strategies for facilitating the reading comprehension of high-functioning students with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 34(2): 115–127.

Randi J, Newman T and Grigorenko EL (2010) Teaching children with autism to read for meaning: challenges and possibilities. *Journal of Autism and Developmental Disorders* 40(7): 890–902.

Riches NG, Loucas T, Baird G, et al. (2011) Non-word repetition in adolescents with specific language impairment and autism plus language impairments: a qualitative analysis. *Journal of Communication Disorders* 44(1): 23–36.

Ricketts J, Jones CR, Happé F, et al. (2013) Reading comprehension in autism spectrum disorders: the role of oral language and social functioning. *Journal of Autism and Developmental Disorders* 43(4): 807–816.

Saldaña D and Frith U (2007) Do readers with autism make bridging inferences from world knowledge? *Journal of Experimental Child Psychology* 96(4): 310–319.

Saldaña D, Carreiras M and Frith U (2009) Orthographic and phonological pathways in hyperlexic readers with autism spectrum disorders. *Developmental Neuropsychology* 34(3): 240–253.

Schwarz G (1978) Estimating the dimension of a model. *The Annals of Statistics* 6: 461–464.

Seltzer MM, Shattuck P, Abbeduto L, et al. (2004) Trajectory of development in adolescents and adults with autism. *Developmental Disabilities Research Reviews* 10(4): 234–247.

Semel E, Wiig EH and Secord WA (2003) Clinical Evaluation of Language Fundamentals (CELF – 4th Edition). San Antonio, TX: Harcourt Assessment.

Sheslow D and Adams W (2003) Wide range assessment of memory and learning—second edition (WRAML2). Wilmington, DW: Wide Range.

Solari EJ, Grimm R, McIntyre NS, et al. (2017) The relation between text reading fluency and reading comprehension for students with autism spectrum disorders. *Research in Autism Spectrum Disorders* 41: 8–19.

Tein JY, Cox S and Cham H (2013) Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling: A Multidisciplinary Journal* 20: 640–657.

Tirado MJ and Saldaña D (2016) Readers with autism can produce inferences, but they cannot answer inferential questions. *Journal of Autism and Developmental Disorders* 46(3): 1025–1037.

Torgesen JK, Wagner RK and Rashotte CA (2012) *TOWRE-2 Examiner's Manual*. Austin, TX: Pro-Ed.

Travis L, Sigman M and Ruskin E (2001) Links between social understanding and social behavior in verbally able children with autism. *Journal of Autism and Developmental Disorders* 31(2): 119–130.

Whitby PJS and Mancil GR (2009) Academic achievement profiles of children with high functioning autism and Asperger syndrome: a review of the literature. *Education and Training in Developmental Disabilities* 44: 551–560.

White S, Frith U, Milne E, et al. (2006) A double dissociation between sensorimotor impairments and reading disability: a comparison of autistic and dyslexic children. *Cognitive Neuropsychology* 23(5): 748–761.

Vermunt JK (2010) Latent class modeling with covariates: Two improved three-step approaches. Political Analysis 18(4): 450–469.

Wagner R, Torgesen J, Rashotte C, et al. (1999) CTOPP-2: Comprehensive Test of Phonological Processing–Second edition. Austin, TX: PRO-ED.

Wechsler D and Hsiao-pin C (2011) Wechsler Abbreviated Scale of Intelligence (WASI-II). 2nd ed. San Antonio, TX: NCS Pearson.

Whitehouse AJ, Barry JG and Bishop DV (2008) Further defining the language impairment of autism: is there a specific language impairment subtype? *Journal of Communication Disorders* 41(4): 319–336.

Williams D, Payne H and Marshall C (2013) Non-word repetition impairment in autism and specific language impairment: evidence for distinct underlying cognitive causes. *Journal of Autism and Developmental Disorders* 43(2): 404–417.

Zevenbergen AA and Whitehurst GJ (2003) Dialogic reading: a shared picture book reading intervention for preschoolers. In: van Kleeck A, Stahl SA and Bauer EB (eds) *Center for Improvement of Early Reading Achievement, CIERA. On Reading Books to Children: Parents and teachers*. Mahwah, NJ: Lawrence Erlbaum Associates, pp.177–200.