frontiers | Research Topics

Language across neurodevelopmental disorders

Edited by
Marisa Filipe, Lénia Carvalhais, Leonard Abbeduto and Sónia Frota

Published in
Frontiers in Psychology
Frontiers in Communication

ISSN 1664-8714
ISBN 978-2-83251-354-5
DOI 10.3389/978-2-83251-354-5

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Language across neurodevelopmental disorders

Topic editors
Marisa Filipe — Faculdade de Letras da Universidade de Lisboa, Portugal
Lénia Carvalhais — Infante D. Henrique Portucalense University, Portugal
Leonard Abbeduto — University of California, Davis, United States
Sónia Frota — University of Lisbon, Portugal

Citation
Filipe, M., Carvalhais, L., Abbeduto, L., Frota, S., eds. (2023). *Language across neurodevelopmental disorders*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83251-354-5

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 04 January 2023
DOI 10.3389/fpsyg.2022.1121997

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Antonio Benítez-Burraco,
University of Seville, Spain

*CORRESPONDENCE
Marisa G. Filipe
✉ marisa.filipe@campus.ul.pt

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 12 December 2022
ACCEPTED 14 December 2022
PUBLISHED 04 January 2023

CITATION
Filipe MG, Carvalhais L, Abbeduto L and Frota S (2023) Editorial: Language across neurodevelopmental disorders. *Front. Psychol.* 13:1121997. doi: 10.3389/fpsyg.2022.1121997

Editorial: Language across neurodevelopmental disorders

Marisa G. Filipe[1]*, Lénia Carvalhais[2], Leonard Abbeduto[3] and Sónia Frota[1]

[1]Center of Linguistics, School of Arts and Humanities, University of Lisbon, Lisbon, Portugal, [2]Infante D. Henrique Portucalense University, Porto, Portugal, [3]MIND Institute & Department of Psychiatry and Behavioral Sciences, University of California, Davis, Sacramento, CA, United States

KEYWORDS

language, development, impairments, neurodevelopment, neurodevelopmental disorders

Editorial on the Research Topic
Language across neurodevelopmental disorders

The development of language is critical to meet the demands and challenges of contemporary societies. Unfortunately, many children do not master language skills at rates or levels consistent with their chronological ages, and language impairments during childhood tend to persist across the development with lifelong implications for academic, social-emotional, and behavioral functioning (Conti-Ramsden et al., 2018). Although many children diagnosed with neurodevelopmental conditions such as autism and intellectual disability display language impairments, the specific profile of impairments may differ across disorders, making research investigating language across these clinical conditions critical. Neurodevelopmental disorders imply disruptions in the typical growth and development of the central nervous system (e.g., Goldstein and Reynolds, 1999), and the symptoms usually emerge early in childhood (Bishop and Rutter, 2008).

The set of papers gathered on the present Research Topic provides evidence on the relationship between prelinguistic communication, language (oral and written), and several neurodevelopmental conditions, namely autism, fragile X syndrome, Down syndrome, and Developmental Language Disorder. The papers include a multidisciplinary list of contributors with different disciplinary backgrounds (e.g., psychologists, speech and language therapists, linguists, and practitioners), each one contributing in a unique way to our knowledge about language in one or more neurodevelopmental disorders. This collection included original research papers (cross-sectional and longitudinal designs), perspective articles, and reviews dealing with both theoretical and practical issues.

In a perspective article, Weismer and Saffran present a fruitful and interesting line of research. The authors argue that individual differences in statistical learning (i.e., the detection of patterns with stable probabilities) among children with autism, together with the prediction deficits (hyperplasticity) characteristic of autism, may be related to the variability in the structural language in children diagnosed with this condition.

Ravi et al. report on a study in which they found that infants with better social communication skills at 12 months present better language scores at 24 months. However, infants who met criteria for autism did not show this developmental coupling until 24 months of age. The authors suggest that social communication outcomes shape downstream language skills and highlight the need to support the development of social communication skills prior to a formal autism diagnosis.

Mankovich et al., through a recurrence quantification analysis (i.e., a technique to understand how units of speech repeat across stretches of transcriptions), found that grammatical productivity and lexical productivity were related to language competence in different ways in a sample of children with autism. These findings indicated that beyond traditional linguistic analysis, recurrence analysis may be helpful to reveal differences in the spoken language of individuals with autism.

Zheng et al. explored how the measurement of autism symptoms might be affected by language and developmental levels. Even for children with minimal verbal abilities, the authors highlight the need for finer distinctions based on spoken language level and/or mental age to optimize the measurement of autism symptoms.

Reetzke et al. emphasize the need for effective community-based implementation strategies for young autistic children from low-resourced households. In particular, these authors found that young children with autism from the lowest-resource households exhibited the poorest language and social communication skills, as well as the poorest non-verbal problem-solving and fine-motor abilities, along with more features of attention-deficit/hyperactivity disorder and atypical auditory processing.

Venker and Johnson explored the relationship between the use of electronic toys and the quantity and lexical diversity of spoken language produced by children with autism and neurotypical children matched on chronological age (2–5 years). They found that children with autism and their neurotypical peers talked significantly less and produced significantly fewer unique words during electronic toy play compared to traditional toy play. The authors suggest that play-based interventions for children with autism may be most effective when they incorporate traditional toys rather than electronic toys.

Moving from spoken language to reading abilities, Vale et al., in a systematic literature review, showed that the majority of children with autism have well preserved word reading abilities. However, word reading strategies in those with autism are far from being completely understood. The authors emphasize that there is much that remains unknown about the specific word reading difficulties and strengths of children with autism.

The set of papers below focused on other neurodevelopmental conditions beyond autism. Thurman and Nunnally found join attention differences between preschool boys with autism or fragile X syndrome when controlling for the influence of age, non-verbal IQ, and autism symptom severity. In addition, differences between the two groups of children were also observed when considering how joint attention performance related to other aspects of the phenotype. Taken together, these findings have implications for understanding phenotypic differences in the development of joint attention, as well as treatments, for these two conditions.

In a systematic review, Hoffmann explored language patterns of weakness and strengths for individuals with fragile X syndrome, highlining the specific role of cognition, autistic symptomatology, and gender. Importantly, this paper presents implications for assessment and intervention practices.

Filipe et al. explored the evidence for early predictors of language outcomes in infants and toddlers with Down Syndrome. Results indicated that child-related factors such as maternal educational level and parents' translation of their children's gestures into words predict language outcomes in Down syndrome. In addition, the level of adaptive functioning, cognitive function, attention skills, communicative intent, early vocalizations, gestures, baby signs, and vocabulary level of the child are also significant predictors of language outcomes in this population.

Prahl and Schuele explored the reliability and validity of several commonly used measures of listening and reading comprehension in terms of their utility for individuals with Down syndrome. Overall, the authors found strong evidence of reliability and construct validity for three of four measurement methods; namely, non-verbal response, cloze procedure, and passage-level with open-ended questions.

Angulo-Chavira et al. presented clues about how people with Down process and extract information from speech and in context. The authors examined whether young adults with Down syndrome anticipate a referent in the same way as their typical development peers matched by mental age and gender. It was found that participants with Down syndrome predicted nouns in closely related verb-noun pairs but not in pairs that were moderately related and in which they needed visual context to generate the prediction. These processing differences may provide insights into therapeutic targets.

Soares et al. aimed to provide new insights into the role that explicit learning mechanisms play in the implicit learning deficits in preschool children with developmental language disorders by collecting behavioral and neuropsychological data. The findings failed to support the compensatory role of explicit learning mechanisms in the implicit learning impairments characteristic of these children.

Loveall et al. conducted a systematic review and found that the representation of individuals with neurodevelopmental disorders in the normative samples of norm-referenced language assessment tools is very low. The authors argue that test developers should (i) include these individuals as part of the iterative test development, (ii) assess a high number of individuals with disabilities in the normative samples, and (iii) include separate norms for individuals with disabilities. The failure to do so limits the usefulness of these tests for both research and clinical purposes.

Overall, the papers in this Research Topic should be of interest to researchers, teachers, educators, clinicians, and students interested in understanding language across neurodevelopmental disorders. Although the nature and extent of the connections between language and neurodevelopmental disorders continue to be investigated, it is now clear that specific language impairment profiles can be identified, and at different levels, across neurodevelopmental disorders. The papers in this Research Topic reflect recent advances in the field, providing some insight into approaches to inform educational and clinical services related to language intervention. We hope these papers will motivate further research and move the field forward in ways that lead to better functional outcomes for affected children and youth.

Author contributions

MF prepared the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the Portuguese Foundation for Science and Technology (2020.01866.CEECIND; PTDC/LLT-LIN/1115/2021) in conjunction with the European Regional Development Fund from the EU, Portugal 2020, and Lisboa 2020 (PTDC/LLT-LIN/29338/2017).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bishop, D., and Rutter, M. (2008). "Neurodevelopmental disorders: conceptual approaches," in *Rutter's Child and Adolescent Psychiatry,* Rutter, M., Bishop, D., Pine, D., Scott, S., Stevenson, J., Taylor, E., et al. (eds.). Oxford, UK: Blackwell. p. 32–41.

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., and Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *Int. J. Lang. Commun. Disord.* 53, 237–255. doi: 10.1111/1460-6984.12338

Goldstein, S., and Reynolds, C. R. (eds.) (1999). *Handbook of Neurodevelopmental and Genetic Disorders in Children.* The Guilford Press.

PERSPECTIVE
published: 09 June 2022
doi: 10.3389/fpsyg.2022.897187

Differences in Prediction May Underlie Language Disorder in Autism

Susan Ellis Weismer[1,2]* and *Jenny R. Saffran*[1,3]

[1]Waisman Center, University of Wisconsin, Madison, WI, United States, [2]Department of Communication Sciences and Disorders, University of Wisconsin, Madison, WI, United States, [3]Department of Psychology, University of Wisconsin, Madison, WI, United States

Language delay is often one of the first concerns of parents of toddlers with autism spectrum disorder (ASD), and early language abilities predict broader outcomes for children on the autism spectrum. Yet, mechanisms underlying language deficits in autistic children remain underspecified. One prominent component of linguistic behavior is the use of predictions or expectations during learning and processing. Several researcher teams have posited prediction deficit accounts of ASD. The basic assumption of the prediction accounts is that information is processed by making predictions and testing violations against expectations (prediction errors). Flexible (neurotypical) brains attribute differential weights to prediction errors to determine when new learning is appropriate, while autistic individuals are thought to assign disproportionate weight to prediction errors. According to some views, these prediction deficits are hypothesized to lead to higher levels of perceived novelty, resulting in "hyperplasticity" of learning based on the most recent input. In this article, we adopt the perspective that it would be useful to investigate whether language deficits in children with ASD can be attributed to atypical domain-general prediction processes.

Keywords: predictive coding, hyperplasticity, autism, prediction deficits, language processing

OPEN ACCESS

Edited by:
Sónia Frota,
University of Lisbon,
Portugal

Reviewed by:
Lynn K. Perry,
University of Miami,
United States

Correspondence:
Susan Ellis Weismer
susan.ellisweismer@wisc.edu

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *15 March 2022*
Accepted: *19 May 2022*
Published: *09 June 2022*

Citation:
Ellis Weismer S and Saffran JR (2022) Differences in Prediction May Underlie Language Disorder in Autism. Front. Psychol. 13:897187. doi: 10.3389/fpsyg.2022.897187

INTRODUCTION

Autism spectrum disorder (ASD) represents a serious and growing public health concern. According to the latest Center for Disease Control and Prevention (CDC) estimates, one in 44 children is currently diagnosed with ASD (Maenner et al., 2021). Although the symptoms and level of severity vary across individuals, ASD has considerable impact on individuals, families, and society (Howlin et al., 2004; Ganz, 2007; Kogan et al., 2008; Lawer et al., 2009; Bradford, 2010). While social communication is universally impaired in this population, there is considerable variation in structural language abilities, i.e., lexicon, syntax (Eigsti et al., 2011; Ellis Weismer and Kover, 2015; Koegel et al., 2020). Early language/communication delays are among the initial red flags for ASD (NICHD, 2021) and are often one of the first concerns noted by parents (De Giacomo and Fombonne, 1998; Eigsti et al., 2011). Importantly, early language abilities predict broader outcomes for children with ASD, including response to treatment and cognitive outcomes (Stone and Yoder, 2001; Szatmari et al., 2003; Eigsti et al., 2011).

Therefore, language is a critical area to study and further research on the mechanisms underlying language challenges in autistic children is warranted.

In this article, we advocate the study of predictive processing by children with ASD as it applies to language learning. One prominent component of linguistic behavior is the generation and use of predictions or expectations. Language is processed incrementally by both adults and young children. Mismatches between expected and actual input lead to prediction errors, which are hypothesized to drive learning (e.g., Elman, 1990). Differential weights are attributed to prediction errors by neurotypical brains to determine when new learning is appropriate; however, prediction-based accounts of ASD (Sinha et al., 2014; Van de Cruys et al., 2014) suggest that autistic individuals assign disproportionate weight to prediction errors. These differences in prediction are hypothesized to lead to higher levels of perceived novelty, resulting in "hyperplasticity" of learning in ASD—overweighting the most recent input rather than the aggregation of instances (Sinha et al., 2014). On this view, individuals with ASD downweight prior experiences in favor of recent experiences more than neurotypical individuals. These group differences should be especially evident in changing or variable environments, where recent experiences diverge from prior experiences.

Successful prediction requires sensitivity to statistical regularities in the environment. To the extent that challenges for individuals with ASD are related to hyperplasticity, rather than deficits in statistical learning *per se*, we would expect to see group differences emerge in situations where the environment is more variable, but not when the environment remains relatively static. In addition, we would expect any such differences to be both predictive of language abilities and not limited to language learning tasks.

Predictive Processes and ASD

Predictive coding theory characterizes perceptual and cognitive representations in terms of probabilistic prediction (Rao and Ballard, 1999; Friston, 2005, 2010; Friston and Kiebel, 2009; Spratling, 2016). Hierarchical predictive coding (also referred to as "predictive processing," Clark, 2018) entails an implicit process involving the bidirectional flow of information in which bottom–up sensory input is compared to top–down predictions. Incorrect predictions produce error signals. The perceived reliability (or precision) of prediction errors determines the weight assigned to each error with respect to updating predictions. High-precision errors will prompt additional processing and learning whereas low-precision errors typically have less influence.

Several research groups have proposed accounts of ASD (Lawson et al., 2014; Sinha et al., 2014; Van de Cruys et al., 2014) that stem from predictive coding theory. Other studies have employed this theoretical framework to examine more specific areas of functioning within ASD (Gonzalez-Gadea et al., 2015; von der Lühe et al., 2016; Lawson et al., 2017), including auditory processing (Gomot and Wicker, 2012). Prediction deficit accounts of ASD have been linked to previous ASD theories including "extreme male brain" theory (Baron-Cohen, 2002) and intense world theory (Markram and Markram, 2010); prediction models have also been proposed for ASD-typical problems in executive function and central coherence. According to this view, the autism hallmark of restricted and repetitive behaviors serves an adaptive function, compensating for deficits in prediction by insistence on sameness and highly predictable routines.

Sinha et al. (2014) advanced the Predictive Impairment in Autism (PIA) hypothesis as a partial account of the ASD phenotype, focused on the primary diagnostic criteria of social communication deficits and repetitive/restricted behaviors. This framework is based on prior research with neurotypical individuals using Markovian models to characterize gaze behavior, dyadic interactions, mother-infant interactions, smile reciprocity, and communicative interactions. One interesting prediction stemming from the PIA hypothesis (Sinha et al., 2014) has specific implications for learning. Dysfunction in prediction renders higher levels of perceived novelty (see Lawson et al., 2017) leading to hyperarousal of the brainstem and basal ganglia (Joshua et al., 2009) which modulate learning (Schultz et al., 1997; Sutton and Barto, 1998), resulting in hyperplasticity. This hypothesized extreme malleability privileges recent input exposure (disproportionately high weights) and jeopardizes aggregation of prior instances. Hyperplasticity would impair accurate estimation of probabilities when the input changes over time—as in many aspects of the natural world, especially the auditory environment (which is temporally fleeting).

Another prediction account of ASD was proposed by Van de Cruys et al. (2014) based on predictive coding theory (Rao and Ballard, 1999; Friston, 2010). They posit that prior cognitive accounts of ASD—executive dysfunction, theory of mind, and weak central coherence—can all be explained by a core deficit in flexibility in processing violations of expectations. The assumption within this framework is that information is processed by making predictions and detecting violations of expectations (prediction errors). This is an iterative process in which prediction errors (bottom–up information) are influenced by top–down predictions that have been shaped by prior prediction errors. Van de Cruys et al. (2014) suggest that disruptions in bottom-up versus top-down flow of information are reflected in two earlier, opposing cognitive theories of ASD—weak central coherence (bottom-up) versus executive dysfunction (top–down). According to Van de Cruys and colleagues, the core deficit in ASD involves "high, inflexible precision of prediction errors." That is, individuals with ASD assign inflexibly high weights to prediction errors. Predictive coding involves two time scales (Friston, 2010; Dayan, 2012). While predictions are used for processing the immediate environment, prediction errors shape plasticity and learning over a longer time scale. Sometimes prediction errors indicate that the predictive model should be updated, but other times prediction errors should be ignored. Flexible brains attribute differential weights to prediction errors to determine when new learning is appropriate. Van de Cruys and colleagues hypothesize that individuals with ASD assign too much weight to prediction errors (similar to the hyperplasticity in learning hypothesis of Sinha et al., 2014). If unwarranted high precision is assumed for each prediction error, this induces learning for each new event. Consequently, predictions are noisy and lack generalizability. For instance, there is evidence for enhanced perceptual processing of complex acoustic signals such as speech and tones by individuals with ASD (Heaton, 2003; Jarvinen-Pasley et al., 2008) and it has been suggested that generation of overly specific categories of sounds are detrimental to establishing higher order linguistic representations (Bonnel et al., 2010; Crespi, 2013). Another example pertains to difficulties that autistic individuals appear to have with category learning and extracting prototypes from exemplars (Soulières et al., 2011; Gastgeb et al., 2012). In order for categorization to be successful, it is necessary to recognize certain similarities across new instances of input while also ignoring other within-category variation. Routine use of high precision of prediction errors will impede the ability to establish categories (e.g., autistic children's category of "dog" may only extend to dogs whose perceptual features are very similar to their own large brown German Shepard and not to dogs of different sizes, colors, or breeds). It is important to note that according to the model proposed by Van de Cruys and colleagues, computation of prediction errors themselves is not impaired in ASD. Instead, learning is impaired or "short-circuited" by a default of indiscriminately high precision which leads to loss of flexible attention allocation based on informativeness. This model proposes that autism deficits in central coherence, executive functions, and theory of mind can be conceptualized in terms of a core deficit in the ability to flexibly process violations of expectations and further that stereotyped behaviors and restricted interests are actually a secondary symptom of this core deficit in (overly precise) predictive processing.

These prediction deficit accounts of ASD have prompted a flurry of research with autistic adults and children. A recent systematic review of the empirical evidence for prediction deficits in ASD assessed findings from 47 studies (Cannon et al., 2021). These investigations spanned infancy through adulthood, tested visual, auditory and audiovisual modalities, and utilized behavioral and neural indices of prediction. Due to the wide range of experimental paradigms and types of data collected by these studies, a formal meta-analysis was not attempted. Instead, Cannon et al. (2021) provide a detailed narrative review of findings and summarize key points of each study in a table. Although some studies failed to find differences in certain predictive skills for individuals with ASD compared to neurotypical individuals (e.g., Manning et al., 2017; Van de Cruys et al., 2018), there is considerable evidence for impairments in learning predictive pairings between an antecedent and consequence particularly in the context of low salience predictive features or variability (Amoruso et al., 2019; Greene et al., 2019; Ganglmayer et al., 2020). Further, results revealed differences in low-level predictive processing, as reflected by habituation and perceptual adaptation (e.g., Turi et al., 2016; Millin et al., 2018; Ruiz-Martínez et al., 2020). It should be noted that none of these studies of predictive skills addressed language processing or language learning in individuals with ASD.

Statistical Learning and ASD

Successful predictions depend upon sensitivity to patterns in the environment (Van de Cruys et al., 2014). This related area of inquiry has received considerable attention. Statistical learning explanations for ASD have been explored using a range of tasks including serial reaction time (Gordon and Stark, 2007; Travers et al., 2010), contextual cueing (Barnes et al., 2008; Kourkoulou et al., 2012; Travers et al., 2013), probabilistic classification learning (Brown et al., 2010), artificial grammar learning (Brown et al., 2010), speech stream segmentation (Mayo and Eigsti, 2012), and observational learning (Roser et al., 2016). In a meta-analysis, Obeid et al. (Obeid et al., 2016) found no evidence for an overall deficit in statistical learning in ASD. However, a study of visual statistical learning in ASD using ERPs revealed heterogeneity in neural indices of visual statistical learning that were associated with nonverbal IQ and adaptive social function (Jeste et al., 2015).

There are mixed results regarding associations between statistical learning and language abilities in ASD (Scott-Van Zeeland et al., 2010; Mayo and Eigsti, 2012; Haebig et al., 2017). In each of these studies, researchers employed a word segmentation task in which the boundaries between words were indicated by low probabilities of co-occurrences between syllables. Mayo and Eigsti (2012) used a behavioral version of this task with school-aged children with high-functioning autism, and found that their performance was not strongly associated with measures of native language attainment; moreover, their performance was indistinguishable from a comparison group. However, a study using a similar task measuring neural activity *via* fMRI suggested that autistic children may be less sensitive to segmentation cues, with some evidence for a correlation between language measures and recruitment of brain regions believed to be relevant for word segmentation (Scott-Van Zeeland et al., 2010). In a more recent behavioral study with school-aged children, Haebig et al. (2017) also observed an association between statistical learning task performance and measures of English language attainment.

It is important to note that to date, studies of statistical learning in ASD have largely included only participants with relatively strong cognitive and language skills (Obeid et al., 2016) due to the task demands inherent in the statistical learning paradigms commonly used with children and adults. Several studies with (presumably) neurotypical infants and children provide evidence for a relationship between statistical learning performance and language skills (Shafto et al., 2012; Kidd and Arciuli, 2016; Lany et al., 2018), consistent with individual differences observed in adults (Daltrozzo et al., 2017; Siegelman et al., 2017). A similar pattern emerges for children and adolescents with developmental language disorder/specific language impairment who evidence weak performance on both linguistic and nonlinguistic statistical learning tasks (Tomblin et al., 2007; Evans et al., 2009; Obeid et al., 2016; Plante et al., 2017). Taken together, the literature suggests that individual differences in statistical learning task performance may be predictive of language outcomes regardless of whether children have ASD or neurotypical development, at least when the patterns to be learned are relatively static and do not elicit hyperplasticity. The few statistical learning studies that have employed changing input distributions tested only neurotypical adults, and revealed primacy effects in learning (Jungé et al., 2007; Gebhart et al., 2009; Karuza et al., 2016). It remains unclear how individuals with ASD learn from materials in which the input is more variable, such that hyperplasticity may come into play.

Language and Prediction

Predictive coding theory has been applied to language in studies with neurotypical adults and infants (Gagnepain et al., 2012; Ylinen et al., 2016; Zarcone et al., 2016). For example, Gagnepain et al. (2012) used a predictive coding framework to explain spoken word recognition in neurotypical adults. They contrasted two computational models – lexical competition and segment prediction – using novel word training, computational simulation, and neuroimaging (magnetoencephalographic responses). Their findings supported a segment prediction account in which prediction error signals represent the difference between predicted and heard speech sounds. Ylinen et al. (2016) examined word recognition and learning in neurotypical infants within a predictive coding framework. ERPs were recorded while infants heard native language syllables in an oddball paradigm. The data revealed brain responses reflecting predictive inference. There was a significant correlation between parent-reported receptive vocabulary and mismatch response amplitudes reflecting the strength of prediction errors at 12 months.

Viewed more broadly in terms of incremental language processing, there is an abundance of research demonstrating that neurotypical adults (Altmann and Kamide, 1999; DeLong et al., 2005; Levy, 2008; Kleinman et al., 2015) and infants and toddlers (Lew-Williams and Fernald, 2007; Borovsky et al., 2012; Mahr et al., 2015; Reuter et al., 2021) anticipate upcoming speech (at both lexical and sublexical levels). Children's ability to predict upcoming nouns from verb semantics relates strongly to their vocabulary knowledge (Borovsky et al., 2012; Mani and Huettig, 2012). Moreover, nonlinguistic measures of prediction in infancy are related to vocabulary size. Reuter et al. (2018), using a task based on Romberg and Saffran (2013), found that infants' ability to adjust their predictive saccades in a visual task was associated with receptive vocabulary. In sum, the literature to date supports the hypothesis that both linguistic and nonlinguistic predictive processes are related to language processing in neurotypical infants and young children. Currently, we do not have evidence indicating that this link involves a causal relationship or to suggest the direction of influence. Further research involving direct manipulations/longitudinal analyses are needed to ascertain if prediction impacts language gains, language skills influence broader predictive processing, or whether there are, in fact, bidirectional influences between predictive processing and language.

Studies have investigated the ability of children and adolescents with ASD to use semantically informative verbs to predict upcoming nouns (Brock et al., 2008; Bavin et al., 2016; Zhou et al., 2019). Overall, these findings suggest that autistic children can employ predictive language processing, which (like for neurotypical children) is positively associated with their language abilities (Brock et al., 2008; Venker et al., 2019). However, it is important to note that most of this research has only included autistic children with language and nonverbal cognitive abilities in the average range, thereby excluding a significant portion of individuals with ASD and limiting generalizability of the findings. One exception is a recent study by our research team which included a broader sample of young children with ASD (Prescott et al., 2022). Our findings indicated that both the ASD and younger, language-matched neurotypical group made use of semantically informative verbs to predict upcoming nouns as evidenced by anticipatory eye movements. That is, the autistic children (3–4 years of age) performed similarly to language-matched controls (who were on average 18 months younger) in terms of efficiency of predictive language processing. However, regression analyses, when controlling for age, revealed that the ASD group displayed a weaker condition effect (informative vs. neutral verbs) than the neurotypical group, similar to prior research (Zhou et al., 2019). In order to actually evaluate prediction-based accounts of ASD, we need to go beyond these types of incremental language processing tasks. We need to use different paradigms (such as violation of expectation paradigms) that allow us to examine claims about hyperplasticity of learning by manipulating the variability of input and to compare predictive processing for both linguistic and nonlinguistic input. Such tasks will allow us to limit the effects of prior language knowledge (e.g., verb semantics) to isolate predictive behavior based on the statistics of the input, comparing static and variable input distributions.

DISCUSSION

Despite the clear relevance of predictive processes to both language development and to theories of autism, there have been scant attempts to integrate all three areas of study (prediction, language, and ASD). We propose that research addressing this important gap in the literature is warranted. While there are a range of prediction-based theories related to ASD, as noted above, the initial, ongoing attempts by our research team to explore this issue are not designed to adjudicate between them. We follow Kutas et al. (2014, p. 649) who suggest that prediction encompasses any form of cognitive processing involving "the activation of or information about likely upcoming stimuli, prior to their receipt that plays a causal role in stimulus processing." Our theoretical framework draws on various aspects of the prediction deficit accounts of ASD discussed above. We assume predictive coding is a domain-general process, but we are especially interested in its role in acquiring and using spoken language. Prior research lays the foundation for prediction deficits in ASD at the level of neural signaling to auditory stimuli (e.g., Gomot and Wicker, 2012; Font-Alaminos et al., 2020), but has not investigated actual language processing or learning in ASD. We hypothesize that children with ASD will demonstrate hyperplasticity of learning (Sinha et al., 2014) which is assumed to stem from difficulty with precision weighting of prediction errors (van de Cruys et al., 2014). We concur with the claims that prediction deficit accounts accommodate many prior cognitive models of autism (Sinha et al., 2014; van de Cruys et al., 2014), but are cautious about assuming that this framework can explain the totality of the ASD phenotype.

It could be argued that there is an apparent paradox in our hypothesis that prediction deficits both underlie autism and are related to language deficits, given that only some children with ASD exhibit structural language impairments. We assume that prediction deficits in the face of social impairments result in the social communication deficits (pragmatic language) that are central to the ASD phenotype. With respect to structural language (vocabulary/grammar), not only are individual differences in prediction assumed to influence these language abilities, but we speculate that individual differences in statistical learning among children with ASD (detecting patterns with stable probabilities), paired with prediction deficits (hyperplasticity) characteristic of ASD, may be related to the observed variability in structural language in autistic children. We contend that a prediction deficit account of language deficits in ASD may be a fruitful line of investigation and encourage other autism researchers to join us in assessing this claim.

AUTHOR'S NOTE

We alternate between person-first and identify-first language in recognition of the terminology debates among different stakeholders.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SEW and JRS each wrote sections of this paper and wrote portions of the Introduction and Discussion, added substantive edits to the sections drafted by the other person. SEW drafted paragraphs focused on autism spectrum disorders (ASD), predictive processing and ASD, and predictive coding applied to language. JRS drafted paragraphs focused on statistical learning and predictive language processing in neurotypical children. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Institutes of Health grants NIDCD R01 DC17974 (SEW and JRS, MPIs) and NICHD U54 HD090256 (Waisman Center core grant) contributed to the writing of this paper.

ACKNOWLEDGMENTS

We would like to acknowledge the Waisman Center, University of Wisconsin-Madison, and the National Institutes of Health for their support.

REFERENCES

Altmann, G., and Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73, 247–264. doi: 10.1016/S0010-0277(99)00059-1

Amoruso, L., Narzisi, A., Pinzino, M., Finisguerra, A., Billeci, L., Calderoni, S., et al. (2019). Contextual priors do not modulate action prediction in children with autism. *J. Royal Soc. B: Bio. Sci.* 286:20191319. doi: 10.1098/rspb.2019.1319

Barnes, K., Howard, J., Howard, D., Gilotty, L., Kenworthy, L., Gaillard, W., et al. (2008). Intact Iimplicit learning of spatial context and temporal sequences in childhood autism spectrum disorder. *Neuropsychologia* 22, 563–570. doi: 10.1037/0894-4105.22.5.563

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends Cogn. Sci.* 6, 248–254. doi: 10.1016/S1364-6613(02)01904-6

Bavin, E. L., Kidd, E., Prendergast, L. A., and Baker, E. K. (2016). Young children with ASD use lexical and referential information during on-line sentence processing. *Front. Psychol.* 7:171. doi: 10.3389/fpsyg.2016.00171

Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., et al. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia* 48, 2465–2475. doi: 10.1016/j.neuropsychologia.2010.04.020

Borovsky, A., Elman, J., and Fernald, A. (2012). Corrigendum to "Knowing a lot for one's age: vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults". *J. Exp. Child Psychol.* 114, 371–373. doi: 10.1016/j.jecp.2012.09.010

Bradford, K. (2010). Supporting families dealing with autism and Asperger's disorders. *J. Fam. Psychother.* 21, 149–156. doi: 10.1080/08975353.2010.483660

Brock, J., Norbury, C., Einav, S., and Nation, K. (2008). Do individuals with autism process words in context? Evidence from language mediated eye-movements. *Cognition* 108, 896–904. doi: 10.1016/j.cognition.2008.06.007

Brown, J., Aczel, B., Jiménez, L., Kaufman, S., and Grant, K. (2010). Intact implicit learning in autism spectrum conditions. *Q. J. Exp. Psychol.* 63, 1789–1812. doi: 10.1080/17470210903536910

Cannon, J., O'Brien, A. M., Bungert, L., and Sinha, P. (2021). Prediction in autism spectrum disorder: a systematic review of empirical evidence. *Autism Res.* 14, 604–630. doi: 10.1002/aur.2482

Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenom. Cogn. Sci.* 17, 521–534. doi: 10.1007/s11097-017-9525-z

Crespi, B. (2013). Developmental heterochrony and the evolution of autistic perception cognition and behavior. *BMC Med.* 11:119. doi: 10.1186/1741-7015-11-119

Daltrozzo, J., Emerson, S., Deocampo, J., Singh, S., Freggens, M., Branum-Martin, L., et al. (2017). Visual statistical learning is related to natural language ability in adults. *Brain Lang.* 166, 40–51. doi: 10.1016/j.bandl.2016.12.005

Dayan, P. (2012). Twenty-five lessons from computational neuromodulation. *Neuron* 76, 240–256. doi: 10.1016/j.neuron.2012.09.027

De Giacomo, A., and Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *Eur. Child Adolesc. Psychiatry* 7, 131–136. doi: 10.1007/s007870050058

DeLong, K., Urbach, T., and Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat. Neurosci.* 8, 1117–1121. doi: 10.1038/nn1504

Eigsti, I.-M., De Marchena, A., Schuh, J., and Kelley, E. (2011). Language acquisition in autism spectrum disorders: a developmental review. *Res. Autism Spectr. Disord.* 5, 681–691. doi: 10.1016/j.rasd.2010.09.001

Ellis Weismer, S., and Kover, S. (2015). Preschool language variation, growth, and prediction in children on the autism spectrum. *J. Child Psychol. Psychiatry* 56, 1327–1337. doi: 10.1111/jcpp.12406

Elman, J. L. (1990). Finding structure in time. *Cognit. Sci.* 14, 179–211. doi: 10.1207/s15516709cog1402_1

Evans, J., Saffran, J., and Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *J. Speech Lang. Hear. Res.* 52, 321–335. doi: 10.1044/1092-4388(2009/07-0189)

Font-Alaminos, M., Cornella, M., Costa-Faidella, J., Hervás, A., Leung, S., Rueda, I., et al. (2020). Increased subcortical neural responses to repeating auditory stimulation in children with autism spectrum disorder. *Biol. Psychol.* 149:107807. doi: 10.1016/j.biopsycho.2019.107807

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R Soc. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philos. Trans. R Soc. B Biol. Sci.* 364, 1211–1221. doi: 10.1098/rstb.2008.0300

Gagnepain, P., Henson, R., and Davis, M. (2012). Temporal predictive codes for spoken words in auditory cortex. *Curr. Biol.* 22, 615–621. doi: 10.1016/j.cub.2012.02.015

Ganglmayer, K., Schuwerk, T., Sodian, B., and Paulus, M. (2020). Do children and adults with autism spectrum condition anticipate others' actions as goal-directed? A predictive coding perspective. *J. Autism Dev. Disord.* 50, 2077–2089. doi: 10.1007/s10803-019-03964-8

Ganz, M. (2007). The lifetime distribution of the incremental societal costs of autism participants: hypothetical incident autism cohort born in 2000 and diagnosed in 2003. *Arch. Pediatr. Adolesc. Med.* 161, 343–349. doi: 10.1001/archpedi.161.4.343

Gastgeb, H., Dundas, E. M., Minshew, N. J., and Strauss, M. S. (2012). Category formation in autism: can individuals with autism form categories and prototypes of dot patterns? *J. Autism Dev. Disord.* 42, 1694–1704. doi: 10.1007/s10803-011-1411-x

Gebhart, A., Aslin, R., and Newport, E. (2009). Changing structures in midstream: learning along the statistical garden path. *Cognit. Sci.* 33, 1087–1116. doi: 10.1111/j.1551-6709.2009.01041.x

Gomot, M., and Wicker, B. (2012). A challenging, unpredictable world for people with autism Spectrum disorder. *Int. J. Psychophysiol.* 83, 240–247. doi: 10.1016/j.ijpsycho.2011.09.017

Gonzalez-Gadea, M., Chennu, S., Bekinschtein, T., Rattazi, A., Beraudi, A., Tripicchio, P., et al. (2015). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. *J. Neurophysiol.* 114, 2625–2636. doi: 10.1152/jn.00543.2015

Gordon, B., and Stark, S. (2007). Procedural learning of a visual sequence in individuals with autism. *Focus Autism Other Dev. Dis.* 22, 14–22. doi: 10.1177/10883576070220010201

Greene, R. K., Zheng, S., Kinard, J. L., Mosner, M. G., Wiesen, C. A., Kennedy, D. P., et al. (2019). Social and nonsocial visual prediction errors in autism spectrum disorder. *Autism Res.* 12, 878–883. doi: 10.1002/aur.2090

Haebig, E., Saffran, J. R., and Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *J. Child Psychol. Psychiatry* 58, 1251–1263. doi: 10.1111/jcpp.12734

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *J. Child Psychol. Psychiatry* 44, 543–551. doi: 10.1111/1469-7610.00143

Howlin, P., Goode, S., Hutton, J., and Rutter, M. (2004). Adult outcome for children with autism. *J. Child Psychol. Psychiatry* 45, 212–229. doi: 10.1111/j.1469-7610.2004.00215.x

Jarvinen-Pasley, A., Peppe, S., King-Smith, G., and Heaton, P. (2008). The relationship between form and function level receptive prosodic abilities in autism. *J. Autism Dev. Disord.* 38, 1328–1340. doi: 10.1007/s10803-007-0520-z

Jeste, S., Kirkham, N., Senturk, D., Hasenstab, K., Sugar, C., Kupelian, C., et al. (2015). Electrophysiological evidence of heterogeneity in visual statistical learning in young children with ASD. *Dev. Sci.* 18, 90–105. doi: 10.1111/desc.12188

Joshua, M., Adler, A., and Bergman, H. (2009). Novelty encoding by the output neurons of the basal ganglia. *Front. Syst. Neurosci.* 3, 20–28. doi: 10.3389/neuro.06.020.2009

Jungé, J., Scholl, B., and Chun, M. (2007). How is spatial context learning integrated over signal versus noise? A primacy effect in contextual cueing. *Vis. Cogn.* 15, 1–11. doi: 10.1080/13506280600859706

Karuza, E., Li, P., Weiss, D., Bulgarelli, F., Zinszer, B., and Aslin, R. (2016). Sampling over nonuniform distributions: a neural efficiency account of the primacy effect in statistical learning. *J. Cogn. Neurosci.* 28, 1484–1500. doi: 10.1162/jocn_a_00990

Kidd, E., and Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Dev.* 87, 184–193. doi: 10.1111/cdev.12461

Kleinman, D., Runnqvist, E., and Ferreira, V. (2015). Single-word predictions of upcoming language during comprehension: evidence from the cumulative semantic interference task. *Cogn. Psychol.* 79, 68–101. doi: 10.1016/j.cogpsych.2015.04.001

Koegel, L. K., Bryan, K. M., Su, P. L., Vaidya, M., and Camarata, S. (2020). Definitions of nonverbal and minimally verbal in research for autism: a systematic review of the literature. *J. Autism Dev. Disord.* 50, 2957–2972. doi: 10.1007/s10803-020-04402-w

Kogan, M., Strickland, B., Blumberg, S., Singh, G., Perrin, J., and van Dyck, P. (2008). A national profile of the health care experiences and family impact of autism spectrum disorder among children in the United States, 2005-2006. *Pediatrics* 122, e1149–e1158. doi: 10.1542/peds.2008-1057

Kourkoulou, A., Leekam, S., and Findlay, J. (2012). Implicit learning of local context in autism Spectrum disorder. *J. Autism Dev. Disord.* 42, 244–256. doi: 10.1007/s10803-011-1237-6

Kutas, M., Federmeier, K. D., and Urbach, T. P. (2014). "The "negatives" and "positives" of prediction in language," in *The Cognitive Neurosciences*. eds. M. S. Gazzaniga and G. R. Mangun. *5th Edn*. (Cambridge, Mass: Mangun (MIT Press)), 649–656.

Lany, J., Shoaib, A., Thompson, A., and Estes, K. G. (2018). Infant statistical-learning ability is related to real-time language processing. *J. Child Lang.* 45, 368–391. doi: 10.1017/S0305000917000253

Lawer, L., Brusilovskiy, E., Salzer, M., and Mandell, D. (2009). Use of vocational rehabilitative services among adults with autism. *J. Autism Dev. Disord.* 39, 487–494. doi: 10.1007/s10803-008-0649-4

Lawson, R. P., Mathys, C., and Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nat. Neurosci.* 20, 1293–1299. doi: 10.1038/nn.4615

Lawson, R. P., Rees, G., and Friston, K. (2014). An aberrant precision account of autism. *Front. Hum. Neurosci.* 8:302. doi: 10.3389/fnhum.2014.00302

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106, 1126–1177. doi: 10.1016/j.cognition.2007.05.006

Lew-Williams, C., and Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychol. Sci.* 18, 193–198. doi: 10.1111/j.1467-9280.2007.01871.x

Maenner, M., Shaw, K., Bakian, A., Bilder, D., Durkin, M., Esler, A., et al. (2021). Prevalance and characteristic of autism spectrum disorder among children aged 8 years – autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveill. Summ.* 70, 1–6. doi: 10.15585/mmwr.ss7011a1

Mahr, T., McMillan, B., Saffran, J., Ellis Weismer, S., and Edwards, J. (2015). Anticipatory coarticulation facilitates word recognition in toddlers. *Cognition* 142, 345–350. doi: 10.1016/j.cognition.2015.05.009

Mani, N., and Huettig, F. (2012). Prediction during language processing is a piece of cake – but only for skilled producers. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 843–847. doi: 10.1037/a0029284

Manning, C., Kilner, J., Neil, L., Karaminis, T., and Pellicano, E. (2017). Children on the autism spectrum update their behaviour in response to a volatile environment. *Dev. Sci.* 20:e12435. doi: 10.1111/desc.12435

Markram, K., and Markram, H. (2010). The intense world theory – a unifying theory of the neurobiology of autism. *Front. Hum. Neurosci.* 4, 1–29. doi: 10.3389/fnhum.2010.00224

Mayo, J., and Eigsti, I.-M. (2012). Brief report: a comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. *J. Autism Dev. Disord.* 42, 2476–2485. doi: 10.1007/s10803-012-1493-0

Millin, R., Kolodny, T., Flevaris, A. V., Kale, A. M., Schallmo, M.-P., Gerdts, J., et al. (2018). Reduced auditory cortical adaptation in autism spectrum disorder. *eLife* 7:7. doi: 10.7554/eLife.36493

NICHD (2021). Autism spectrum disorder: what are the symptoms of autism? Eunice Kennedy Shriver National Institute of Child Health and Human Development. https://www.nichd.nih.gov/health/topics/autism/conditioninfo/symptoms (Accessed February 20, 2022).

Obeid, R., Brooks, P., Powers, K., Gillespie-Lynch, K., and Lum, J. (2016). Statistical learning in specific language impairment and autism spectrum disorder: a meta-analysis. *Front. Psychol.* 7:1245. doi: 10.3389/fpsyg.2016.01245

Plante, E., Patterson, D., Sandoval, M., Vance, C., and Asbjørnsen, A. (2017). An fMRI study of implicit language learning in developmental language impairment. *Neuroimage* 14, 277–285. doi: 10.1016/j.nicl.2017.01.027

Prescott, K. E., Mathée-Scott, J., Reuter, T., Edwards, J., Saffran, J. R., and Ellis Weismer, S. (2022). Predictive language processing in young autistic children. *Autism Res.* 15, 892–903. doi: 10.1002/aur.2684

Rao, R., and Ballard, D. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580

Reuter, T., Dalawella, K., and Lew-Williams, C. (2021). Adults and children predict in complex and variable referential contexts. *Lang, Cogn. Neurosci.* 36, 474–490. doi: 10.1080/23273798.2020.1839665

Reuter, T., Emberson, L., Romberg, A., and Lew-Williams, C. (2018). Individual differences in nonverbal prediction and vocabulary size in infancy. *Cognition* 176, 215–219. doi: 10.1016/j.cognition.2018.03.006

Romberg, A., and Saffran, J. R. (2013). Expectancy learning from probabilistic input by infants. *Front. Psychol.* 3, 1–16. doi: 10.3389/fpsyg.2012.00610

Roser, M., Aslin, R., Mckenzie, R., Zahara, D., and Fiser, J. (2016). Enhanced visual statistical learning in adults with autism. *Neuropsychologia* 29, 163–172. doi: 10.1037/neu0000137

Ruiz-Martínez, F. J., Rodríguez-Martínez, E. I., Wilson, C. E., Yau, S., Saldaña, D., and Gómez, C. M. (2020). Impaired P1habituation and mismatch negativity in children with autism spectrum disorder. *J. Autism Dev. Disord.* 50, 603–616. doi: 10.1007/s10803-019-04299-0

Schultz, W., Dayan, P., and Montague, P. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Scott-Van Zeeland, A., McNealy, K., Wang, A., Sigman, M., Bookheimer, S., and Dapretto, M. (2010). No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biol. Psychiatry* 68, 345–351. doi: 10.1016/j.biopsych.2010.01.011

Shafto, C., Conway, C., Field, S., and Houston, D. (2012). Visual sequence learning in infancy: domain-general and domain-specific associations with language. *Inf. Dent.* 17, 247–271. doi: 10.1111/j.1532-7078.2011.00085.x

Siegelman, N., Bogaerts, L., Christiansen, M., and Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philos. Trans. R Soc. B Biol Sci.* 372, 1, 20160059–20160010. doi: 10.1098/rstb.2016.0059

Sinha, P., Kjelgaard, M., Gandhi, T., and Held, R. (2014). Autism as a disorder of prediction. *Proc. Natl. Acad. Sci.* 111, 15220–15225. doi: 10.1073/pnas.1416797111

Soulières, I., Mottron, L., Giguère, G., and Larochelle, S. (2011). Category induction in autism: slower, perhaps different, but certainly possible. *Q. J. Exp. Psychol.* 64, 311–327. doi: 10.1080/17470218.2010.492994

Spratling, M. (2016). Predictive coding as a model of cognition. *Cogn. Process.* 17, 279–305. doi: 10.1007/s10339-016-0765-6

Stone, W., and Yoder, P. (2001). Predicting spoken language level in children with autism spectrum disorders. *Autism* 5, 341–361. doi: 10.1177/1362361301005004002

Sutton, R., and Barto, A. (1998). *Reinforcement Learning: An Introduction. 1st Edn*. Cambridge, MA: MIT Press.

Szatmari, P., Bryson, S., Boyle, M., Streiner, D., and Duku, E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome. *J. Child Psychol. Psychiatry* 44, 520–528. doi: 10.1111/1469-7610.00141

Tomblin, J. B., Mainela-Arnold, E., and Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Lang Learn. Dev.* 3, 269–293. doi: 10.1080/15475440701377477

Travers, B., Klinger, M., Mussey, J., and Klinger, L. (2010). Motor-linked implicit learning in persons with autism spectrum disorders. *Autism Res.* 3:77. doi: 10.1002/aur.123

Travers, B., Powell, P., Mussey, J., Klinger, L., Crisler, M., and Klinger, M. (2013). Spatial and identity cues differentially affect implicit contextual cueing in adolescents and adults with autism spectrum disorder. *J. Autism Dev. Disord.* 43, 2393–2404. doi: 10.1007/s10803-013-1787-x

Turi, M., Karaminis, T., Pellicano, E., and Burr, D. (2016). No rapid audiovisual recalibration in adults on the autism spectrum. *Sci. Rep.* 6:21756. doi: 10.1038/srep21756

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de Wit, L., et al. (2014). Precise minds in uncertain worlds: predictive coding in autism. *Psychol. Rev.* 121, 649–675. doi: 10.1037/a0037665

Van de Cruys, S., Vanmarcke, S., Van de Put, I., and Wagemans, J. (2018). The use of prior knowledge for perceptual inference is preserved in ASD. *Clin. Psychol. Sci.* 6, 382–393. doi: 10.1177/2167702617740955

Venker, C. E., Edwards, J., Saffran, J. R., and Ellis Weismer, S. (2019). Thinking ahead: incremental language processing is associated with receptive language abilities in preschoolers with autism spectrum disorder. *J. Autism Dev. Disord.* 49, 1011–1023. doi: 10.1007/s10803-018-3778-4

von der Lühe, T., Manera, V., Barisic, I., Becchio, C., Vogeley, K., and Schilbach, L. (2016). Interpersonal predictive coding, not action perception, is impaired in autism. *Philos. Trans. R Soc. B Biol. Sci.* 371:20150373. doi: 10.1098/rstb.2015.0373

Ylinen, S., Bosseler, A., Junttila, K., and Huotilainen, M. (2016). Predictive coding accelerates word recognition and learning in the early stages of language development. *Dev. Sci.* 20, e12472–e12413. doi: 10.1111/desc.12472

Zarcone, A., van Schijndel, M., Vogels, J., and Demberg, V. (2016). Salience and attention in surprisal-based accounts of language processing. *Front. Psychol.* 7, 1–17. doi: 10.3389/fpsyg.2016.00844

Zhou, P., Zhan, L., and Ma, H. (2019). Predictive language processing in preschool children with autism spectrum disorder: an eye-tracking study. *J. Psycholinguist. Res.* 48, 431–452. doi: 10.1007/s10936-018-9612-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 30 June 2022
doi: 10.3389/fpsyg.2022.905762

Check for updates

Explicit Instructions Do Not Enhance Auditory Statistical Learning in Children With Developmental Language Disorder: Evidence From Event-Related Potentials

Ana Paula Soares[1], Francisco-Javier Gutiérrez-Domínguez[1], Helena M. Oliveira[1], Alexandrina Lages[1], Natália Guerra[1], Ana Rita Pereira[2], David Tomé[3,4] and Marisa Lousada[5]*

[1]Human Cognition Lab, CIPsi, School of Psychology, University of Minho, Braga, Portugal, [2]Psychological Neuroscience Lab, CIPsi, School of Psychology, University of Minho, Braga, Portugal, [3]Department of Audiology, School of Health, Polytechnic Institute of Porto, Porto, Portugal, [4]Neurocognition Group, Laboratory of Psychosocial Rehabilitation, CiR, Porto, Portugal, [5]Center for Health Technology and Services Research (CINTESIS@RISE), School of Health Sciences, University of Aveiro, Aveiro, Portugal

OPEN ACCESS

Edited by:
Marisa Filipe,
Universidade de Lisboa, Portugal

Reviewed by:
Susan Ellis Weismer,
University of Wisconsin-Madison, United States
Andreia Salé Veloso,
University of Porto, Portugal

***Correspondence:**
Ana Paula Soares
asoares@psi.uminho.pt

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: 27 March 2022
Accepted: 26 May 2022
Published: 30 June 2022

Citation:
Soares AP, Gutiérrez-Domínguez F-J, Oliveira HM, Lages A, Guerra N, Pereira AR, Tomé D and Lousada M (2022) Explicit Instructions Do Not Enhance Auditory Statistical Learning in Children With Developmental Language Disorder: Evidence From Event-Related Potentials.
Front. Psychol. 13:905762.
doi: 10.3389/fpsyg.2022.905762

A current issue in psycholinguistic research is whether the language difficulties exhibited by children with developmental language disorder [DLD, previously labeled specific language impairment (SLI)] are due to deficits in their abilities to pick up patterns in the sensory environment, an ability known as statistical learning (SL), and the extent to which explicit learning mechanisms can be used to compensate for those deficits. Studies designed to test the compensatory role of explicit learning mechanisms in children with DLD are, however, scarce, and the few conducted so far have led to inconsistent results. This work aimed to provide new insights into the role that explicit learning mechanisms might play on implicit learning deficits in children with DLD by resorting to a new approach. This approach involved not only the collection of event-related potentials (ERPs), while preschool children with DLD [relative to typical language developmental (TLD) controls] were exposed to a continuous auditory stream made of the repetition of three-syllable nonsense words but, importantly, the collection of ERPs when the same children performed analogous versions of the same auditory SL task first under incidental (implicit) and afterward under intentional (explicit) conditions. In each of these tasks, the level of predictability of the three-syllable nonsense words embedded in the speech streams was also manipulated (high vs. low) to mimic natural languages closely. At the end of both tasks' exposure phase, children performed a two-alternative forced-choice (2-AFC) task from which behavioral evidence of SL was obtained. Results from the 2-AFC tasks failed to show reliable signs of SL in both groups of children. The ERPs data showed, however, significant modulations in the N100 and N400 components, taken as neural signatures of word segmentation in the brain, even though a detailed analysis of the neural responses revealed that only children from the TLD group seem to have taken advantage of the previous knowledge to enhance SL functioning. These results suggest that children with

DLD showed deficits both in implicit and explicit learning mechanisms, casting doubts on the efficiency of the interventions relying on explicit instructions to help children with DLD to overcome their language difficulties.

Keywords: developmental language disorder, statistical learning, implicit learning, explicit learning, SL deficit hypothesis, procedural deficit hypothesis, word predictability, ERP word segmentation correlates

INTRODUCTION

Learning to talk is one of the most astonishing abilities children achieve during infancy. Indeed, within a few years, they go from cooing and babbling to an extraordinary complex use of the sounds of the language spoken around them to communicate their needs, feelings, and thoughts. Although most children acquire this remarkable ability quickly, effortlessly, and with no need for any explicit instructions, a nonnegligible portion (~7%–10%) shows significant problems in using speech and language to communicate (e.g., Tomblin et al., 1997; Norbury et al., 2016), presenting a development language disorder (DLD).

The term DLD was introduced by Bishop et al. (2017) to refer to children who showed significant language difficulties at expressive and/or receptive levels, impacting their daily lives and/or their educational outcomes not only during infancy but also typically throughout their entire lives (see Conti-Ramsden et al., 2018). DLD occurs in the absence of other neurodevelopmental disorders, such as autism spectrum disorder (ASD) or intellectual disability (ID), brain injury, hearing loss, or a known biomedical condition (e.g., genetic conditions). The term was introduced to replace the specific language impairment (SLI; Leonard, 1981, 2014) label, widely used in research since the mid-1980s (see, however, Bishop, 2014 for a review of other terms in clinical and educational contexts), because the strict use of the term SLI excludes from diagnosis a significant number of children who struggle with relevant language difficulties, which might pose a greater challenge for them to have access to specialized health services that could mitigate the detrimental effects this condition brings to these children, their families, and the society as a whole (see Bishop et al., 2016, 2017 and Soares et al., 2021b for details). The definition of DLD is thus broader than SLI since it also includes children whose language problems may co-occur with other motor, cognitive, emotional, and/or behavioral disorders, such as attention-deficit hyperactivity disorder (ADHD), developmental coordination disorder (DCD), or developmental dyslexia (DD). It also includes children scoring one SD below the mean in nonverbal intelligence quotient (NVIQ) standardized scales (i.e., NVIQ < 85) that were excluded from the SLI diagnosis, even though those scoring within the range that qualifies them for intellectual disability (i.e., two SDs below the mean or scores of NVIQ < 70) is still excluded from the DLD diagnosis. This change in the diagnostic criteria responds to compelling evidence, showing that children with selective language impairments (i.e., those who meet the SLI criteria) are relatively rare, and there is no evidence that they respond differently to intervention, or that they present a different psycholinguistic profile than children with language difficulties that do not completely meet SLI criteria (e.g., Dyck et al., 2011; Reilly et al., 2014; Norbury et al., 2016; Lancaster and Camarata, 2019; McGregor et al., 2020).

The etiology of DLD is complex and hotly debated in the current research, with approaches claiming that the language difficulties observed by these children arise from impairments that are specific to grammar (e.g., late parameterization, missing grammatical features, and representational deficits for dependent relationships, e.g., Rice and Wexler, 1996; Clahsen and Hansen, 1997; Leonard et al., 1997) to accounts arguing that the language impairments arise from deficits in the cognitive processes that subserve language but that are not specific to language (e.g., working memory, rapid temporal processing, and attention, e.g., Tallal et al., 1985; Archibald and Gathercole, 2006; Spaulding et al., 2008; Kidd, 2012). Here, we claim that the language difficulties observed in children with DLD might stem from deficits in their ability to extract patterns from the sensory environment without reinforcement or feedback, a cognitive ability known as statistical learning (SL)—see Perruchet and Pacton (2006) and Christiansen (2019) for other terms—and that is assumed to play a critical role in the acquisition of rule-governed aspects of language across phonology, morphology, and grammar (see Romberg and Saffran, 2010; Erickson and Thiessen, 2015; Saffran and Kirkham, 2018; Saffran, 2020). Indeed, in order to use language efficiently, children need to realize that despite the tremendous variability it presents at a surface level, language is a system governed by plenty of rules that define how speech sounds (phonemes) can be combined in the language to which they were exposed to generate words, how parts of words (morphemes) may be (re)arranged to create new words and to adjust them to the syntactic context in which they were used, and ultimately how words should be combined with each other to convey meaning (syntax).

Evidence for the involvement of SL mechanisms in language acquisition comes firstly from the seminal work of Saffran et al. (1996), showing that 8-month-old babies were able to compute the probability of a given syllable to be followed by another syllable in a continuous speech stream made of the repetition of four three-syllable nonsense words (e.g., "*gikobatokibutipolugopilatokibu*"), and to use that statistics, known as Transitional Probability (TP), to extract word-like units from the continuous speech (e.g., "*tokibu*," "*gikoba*," "*gopila*," and "*tipolu*"). Note that, in that artificial language, TPs between syllables were higher within word boundaries (TP = 1.0) than across word boundaries (TP = 0.33), thus making the extraction of TPs a reliable cue for word segmentation. Since then, many other works have provided support for the involvement of SL mechanisms in other levels of language acquisition, such as word-referent associations (e.g., Saffran and Estes, 2006; Estes et al., 2007; Hay et al., 2011; Breen et al., 2019), grammatical categorization (e.g., Mintz, 2003), the establishment of long-distance dependencies in different grammatical structures (e.g., Gómez, 2002; Newport and Aslin, 2004; Gómez and Maye, 2005; Thompson and Newport, 2007; Kidd, 2012; Hsu et al., 2014), and literacy skills (e.g., Arciuli and Simpson, 2012; Spencer et al., 2015; Sawi and Rueckl, 2019; Lages et al., 2022). Statistical learning is, thus, assumed as a powerful mechanism that enables children to detect the regularities embedded in the spoken (and written) language even without awareness or intention to do so, and to use that "knowledge" to make predictions about "what comes next," which not only facilitates language processing but also creates the conditions for children to scale up to the extraction of other (higher) levels of regularities that mastering a language requires.

Because extracting the patterns embedded in a language is assumed to be critical for language acquisition (see Romberg and Saffran, 2010; Erickson and Thiessen, 2015; Saffran and Kirkham, 2018; Saffran, 2020; Siegelman, 2020), and also because SL abilities vary considerably across individuals (e.g., Arciuli and Simpson, 2011; Misyak and Christiansen, 2012; Siegelman and Frost, 2015; Kidd and Arciuli, 2016; Johnson et al., 2020), it is not surprising that deficits in that ability had been put forward as a potential explanation for the difficulties exhibited by children with DLD (e.g., Evans et al., 2009, 2022; Lum et al., 2010, 2012, 2014; Hsu and Bishop, 2014; Arciuli and Conway, 2018; Plante and Gómez, 2018; Saffran, 2018; Soares et al., 2018; Ahufinger et al., 2021; Bogaerts et al., 2021). For instance, in a meta-analysis of studies using the serial reaction time task to test implicit learning in language-impaired participants, Lum et al. (2014) revealed that children with DLD performed poorly than typical language development (TLD) controls, even though the serial reaction time task contains an important motor learning component that seems also to be impaired in children with DLD (see Desmottes et al., 2016), which might have confounded the results. Nevertheless, in another meta-analysis targeting studies using a wide range of SL tasks in the visual and auditory domains (e.g., serial reaction time task and artificial grammar learning task), Obeid et al. (2016) found that children with DLD kept performing significantly below TLD controls, with task modality (visual vs. auditory) not moderating the effects, supporting the claim that SL is a general domain learning mechanism (see, however, Frost et al., 2015 for a discussion). Finally, Lammertink et al. (2017), in another meta-analysis focused on studies using word segmentation tasks (such as the triplet embedded task introduced by Saffran et al., 1996) and the artificial grammar learning task in the auditory domain using verbal materials, showed that children with DLD revealed significant impairments when compared with TLD controls in both tasks, leading the authors to conclude that the level of linguistic processing (word vs. grammar) did not modulate the results. Even though there are also studies showing that children with DLD performed just as well as TLD controls (see Lum and Bleses, 2012; Gabriel et al., 2014; Mayor-Dubois et al., 2016; West et al., 2018; Lammertink et al., 2020), the bulk of the studies conducted so far suggests that, on average, the extraction of the regularities embedded in the input seems to be impaired or, at least, not as effective in children with DLD as in peers controls (e.g., Tomblin et al., 1997; Evans et al., 2009, 2022), hence supporting the view of the existence of an SL deficit in DLD children.

The SL deficit hypothesis is also consistent with the procedural deficit hypothesis (Ullman and Pierpont, 2005; Ullman and Pullman, 2015; Ullman et al., 2020), stating that language difficulties observed in children with language impairments (e.g., DLD and DD) arise from a dysfunction in the procedural memory (PM) system. The PM system is a brain network connecting the cortex with the striatum in the basal ganglia (corticostriatal circuits), thought to be involved in implicit learning and to play a crucial role in the acquisition of the rule-governed aspects of language (e.g., morphosyntax and phonology). Additionally, the PDH also states that language difficulties observed in children with language impairments can be amended by the declarative memory (DM), a neural network located in the medial temporal lobe, thought to be largely spared (or even strengthened) in children with DLD and to assume a crucial role in the acquisition of the lexical-semantic aspects of language. Specifically, within that framework, it is claimed that children with DLD may store complex linguistic structures that normally are processed automatically in the PM system, such as decomposing morphological complex words into their constituents (e.g., "walked" → "walk" + "-ed"), in the DM system by the use of explicit rules (e.g., add "-ed" to a verb if the action has already occurred) or chunking, i.e., by storing these words as a whole (e.g., "walked") in the mental lexicon. Evidence for the compensatory role of DM in children with DLD is, however, contentious. While some studies found intact or even enhanced performance in DM tasks in children with DLD relative to TLD controls, particularly in studies using DM tasks involving nonverbal materials (e.g., Riccio et al., 2007; Lum et al., 2010; Lum and Conti-Ramsden, 2013; Lukács et al., 2017; Earle and Ullman, 2021; see, however, Bishop and Hsu, 2015; Kuppuraj et al., 2016; Lee, 2018), others reported DM impairments, especially those using DM tasks involving verbal materials (e.g., Lum et al., 2010; Lukács et al., 2017; McGregor et al., 2017; Haebig et al., 2019; see, however, Baird et al., 2010; Evans et al., 2022), even though differences tend to vanish when working memory measures were taken into account (e.g., Alt and Plante, 2006; Lum et al., 2012, 2015; Arthur et al., 2021). Thus, it remains largely unknown whether children with DLD show or not deficits in the DM system and even if showing spared or enhanced DM performance, as some studies suggest, the extent to which these abilities can be effectively mobilized by DLD children to compensate for their PM deficits. Note that the studies conducted so far examining DM-PM functioning in children with DLD have relied on the use of different tasks and materials to test each of these functions (typically the serial reaction time task or the artificial grammar learning task to test PM functioning and face, object recognition, or word-lists tasks to test DM functioning) from a wide range of participants (children, adolescents, and adults), which could explain the disparity of the results. It is also important to emphasize that since these studies have relied on the collection of behavioral data (i.e., reaction time and accuracy measures), which can be strongly affected by attention and/or motivational factors, particularly those conducted with young participants, studies using other tasks and techniques, such as the Event-Related Potentials (ERPs) technique, are required to get a deeper understanding of the DM-PM dynamics in children with DLD with important theoretical and clinical implications.

CURRENT STUDY

The work presented here aimed to get new insights into the role that explicit learning mechanisms might play in implicit learning deficits in preschool children with DLD. For this, we resorted to a new approach that involved the collection of ERPs, while children with DLD (relative to TLD controls) were exposed to a continuous auditory stream made of the repetition of three-syllable nonsense words under two different conditions. First, they were exposed to a speech stream without any information regarding the task or the stimuli (i.e., under incidental conditions), and, subsequently, with previous knowledge about the regularities (word-like units) embedded in the input stream (i.e., under intentional conditions). The ERP technique is particularly well suited to study the compensatory role that explicit learning mechanisms might play on DLD since it allows to study the underpinnings of speech processing in the brain in young children with high time (millisecond) precision even in the absence of any overt response. These characteristics make ERPs an exceptional tool to overcome some of the problems that the exclusive use of behavioral measures (reaction times and/or accuracy) with young participants can bring to research (for details, see Daltrozzo and Conway, 2014; Royle and Courteau, 2014).

Following previous works (e.g., Soares et al., 2020, 2021a,b) in each auditory SL task (aSL), the level of predictability of the three-syllable nonsense words embedded in the speech streams was also manipulated (high vs. low) to mimic natural languages closely (see Soares et al., 2020 for details). At the end of the exposure phase of each aSL task, children performed a two-alternative forced-choice (2-AFC) task from which behavioral evidence of SL was obtained, as in most SL studies (for reviews, see Siegelman et al., 2017; Soares et al., 2022a). The collection of both neural and behavioral responses while preschool children, with and without DLD, performed analogous versions of the same task under implicit and explicit conditions. That allowed us to not only control for differences in the results that might have arisen in previous studies, from the use of different tasks and stimuli to test DM and PM functioning but also, importantly, to directly examine the changes that performing analogous versions of the same task, presented under different learning conditions, produced in the SL functioning. Note that although we recognize that the terms "declarative vs. procedural," "explicit vs. implicit," and "intentional vs. incidental" are not exactly the same (the first referring mostly to the brain areas associated with conscious vs. unconscious access, the second to the processes involved in the encoding and storage of information based on a single event vs. extended practice, and the third to participants' passive vs. active orientation toward the encoding and retrieval of the information presented in the task, respectively—see Sawi and Rueckl, 2019 for details), there is a substantial overlap between them. Thus, the terms "incidental-implicit-procedural" and "intentional-explicit-declarative" have been used here, as well as in current SL research, interchangeably (see Conway and Christiansen, 2006; Thiessen, 2017; Soares et al., 2020), even if not necessarily assuming a one-to-one correspondence between them (i.e., a task presented under intentional conditions does not immediately qualify the processes involved in the encoding and storage of the information as explicit nor it would imply the recruitment of brain areas and mechanisms restricted to conscious processing in the medial-temporal lobe).

Moreover, it is also important to point out that we have resorted to the use of an aSL task modeled from Saffran et al. (1996) instead of another implicit learning task (e.g., artificial grammar learning) for several reasons. Firstly, the aSL task allows testing SL skills at a simpler language level of processing (words level), which seems to be particularly appropriate when studying children with language impairments (see also Soares et al., 2018, 2021c; Jiménez et al., 2020 for other arguments justifying why artificial grammar learning tasks were not used). Secondly, recent neuroimaging studies using functional MRI (fMRI) showed that responses to the statistical regularities (TPs) embedded in the input recruit brain areas associated both with procedural and declarative systems, although the reliance on one or another seems also to depend on the type of instructions (implicit vs. explicit) provided to the participants to perform the task (e.g., Karuza et al., 2013; for a review, see Batterink et al., 2019). These features make the aSL an ideal task to test whether children with DLD indeed mobilize the processes and mechanisms associated with declarative learning to compensate for potential procedural deficits, as the PDH claims. Finally, the aSL task has been successfully applied in electrophysiological (ERP) paradigms both with adults (e.g., Abla et al., 2008; François et al., 2014; Batterink et al., 2015a,b; Soares et al., 2020, 2021a, 2022b; Gutiérrez-Domínguez et al., 2022) and young participants (e.g., Bosseler et al., 2016; Mandikal-Vasuki et al., 2017; Choi et al., 2020; Pierce et al., 2021; Soares et al., 2022b,c). This is of special interest because it allows us not only to study the neural underpinnings of speech processing in the brain as exposure to the input stream unfolds with a high time precision, as mentioned above, but also because it allows us to overcome much of the limitations that the use of the 2-AFC post-learning task to test SL entails. Indeed, in a standard aSL experiment, participants are typically tested on their abilities to extract the regularities embedded in the input (TPs) after the exposure phase has occurred by asking them to identify which element of a pair of stimuli (e.g., a three-syllable nonsense word presented during exposure vs. a foil made of the same syllables but never presented together before) resembles most the stream presented before. If performance exceeds the chance level, SL is assumed to have occurred as only the track of the TPs embedded in the input allows a correct "word" discrimination. However, as an increasing number of authors have been pointed out, it is important to consider that a correct "word" discrimination in that task also depends on other cognitive processes (e.g., such as memory and decision making) that might be not fully developed in children of young ages, hence requiring other tasks and techniques to assess SL in a valid and reliable way (see Lammertink et al., 2019; Arnon, 2020; Lukács et al., 2021; Lukics and Lukács, 2021; see also Siegelman et al., 2017 and Soares et al., 2022a for an extended discussion on the limits of the 2-AFC task even with adult participants).

However, despite all these advantages, studies examining the compensatory role of explicit learning mechanisms in the implicit learning deficits in children with DLD by collecting both behavioral and neural data from the same task, presented to the same participants under incidental and intentional conditions, are, to the best of our knowledge, inexistent. We are only aware of a recent brain study conducted by Evans et al. (2022) that have used an aSL task to test procedural learning in adolescents with and without a history of DLD from which behavioral (2-AFC) data were collected, along with behavioral data from the true/false section of the competing language processing (CLPT) task (a task asking participants to judge the veracity of sentences presented in groups of two, three, four, five, and six sentences) to tap declarative knowledge. Participants also performed a semantic congruency task and an auditory lexical decision task as additional indexes of declarative and procedural functioning, respectively, from which behavioral and ERP data were collected. Results showed that adolescents with a history of DLD revealed intact declarative memory, but impaired procedural memory as assessed by CLP and aSL tasks, respectively. Intact lexical-semantic knowledge was observed from the behavioral results of the semantic congruency task, and a less effective lexical-phonological processing was observed from the behavioral results of the auditory lexical decision tasks (as indexed by lower accuracy in words/nonword responses, but an equal sensitivity to high vs. low-frequency words), in adolescents with a history of DLD vs. controls. The neural data revealed that although adolescents with and without a history of DLD showed similar neural responses (i.e., a larger N400 amplitude to incongruent vs. congruent semantic conditions), differences were observed in the location and the time course of the effect. Furthermore, in the auditory lexical decision task, the neural data showed that while adolescents without a history of DLD showed a larger N400 amplitude for low- vs. high-frequency words, as expected, adolescents with a history of DLD did not show any neural signs of such effect. Instead, their neural responses in that ERP component seem to have been modulated by a word imageability and not a word frequency effect. These results were taken by the authors as evidence for the use of a declarative compensatory strategy by adolescents with a history of DLD once they seem to have based their word/nonword responses on their conceptual knowledge rather than on the computation of the phonological patterns of the words used in the auditory lexical decision task. Even though interesting, this interpretation should be taken with caution, as the lexical decision task manipulating the frequency of occurrence of the words might not be the best proxy for the processing of lexical–phonological information (see Quémart and Maillart, 2016 for a study using an auditory lexical decision task but where the phonotactic probability of the non-words was manipulated instead). Moreover, it is also worth noting that Evans et al. (2022) still tested DM and PM functioning by relying on the use of different tasks, and not on the use of the same task manipulating instructions, as we propose in the current work.

Batterink et al. (2015a, see also Batterink, et al., 2015b), in one of the first studies examining the role of implicit and explicit instructions in the context of a typical aSL task to examine the neural underpinnings of the processes recruited to assist SL, collected behavioral (RTs/accuracy) and ERP data while language unimpaired adults performed a speeded target detection task and a 2-AFC task, combined with a remember/know procedure after the exposure phase. Participants were distributed into two learning conditions: in the incidental condition, participants performed the aSL task without any information regarding the task or the stimuli, whereas in the intentional condition learners received explicit training on the six nonsense words embedded in the speech stream previous to the exposure phase. Results from the target detection task showed intentional learners to be faster and to show larger P300 amplitudes to syllables occurring in more predictable than less predictable positions of the triplet, attributable to SL and the greater involvement of controlled and effortful processes. On the 2-AFC task, intentional learners performed more accurately than incidental learners, and their responses were also associated with subjective feelings of stronger recollection, suggesting that the previous knowledge of the nonsense words strengthened participants' explicit memory and boosted SL function, as expected. Although providing interesting insights, the fact that the authors have collected data only after the exposure phase, along with having adopted a between-subject design in the manipulation of the instructions, raises concerns since recent studies showed a lot of variability in the way individuals respond to SL tasks, particularly when using linguistic materials (see Siegelman et al., 2018; Soares et al., 2022b).

To overcome such flaws, Soares et al. (2020) used a within-subject design in which participants were firstly presented with the implicit version of the aSL task with three-syllable nonsense words drawn from one syllabary, and, subsequently, with an explicit version of an analogous aSL task using three-syllable nonsense words generated from another syllabary to avoid confounds. Note that due to the nature of the task, the order of the tasks was not counterbalanced across participants since once the task has been performed explicitly it cannot be performed implicitly anymore. Moreover, it is also important to point out that the fact that participants have performed first the implicit SL task and subsequently the explicit version of an analogous SL task, might have also contributed to making the second task really explicit, as intended. This issue is particularly important as previous studies showed the effect of explicit instructions on the SL function to be only observed when instructions are specific enough to allow participants to use them while dealing with task requirements (see Arciuli et al., 2014 for a discussion). Moreover, in that work, Soares et al. (2020) have also used complex speech streams entailing not only a higher number of three-syllable nonsense words than in previous works (eight) but, importantly, "words" presenting different levels of predictability (four high-TP "words" and four low-TP "words") to mimic natural language closely and to further analyze the limits of SL under more uncertain conditions (see Soares et al., 2020 for details). Although results from the 2-AFC tasks showed that performance was neither affected by the conditions under which the tasks were performed (implicit vs. explicit) nor by the predictability of the nonsense words (high vs. low-TP 'words'), the neural data showed, however, modulations in the N100 and, particularly, in the N400 components, taken as the neural signatures of words' segmentation in the brain (see Sanders et al., 2002; Cunillera et al., 2006; De Diego Balaguer et al., 2007; Abla et al., 2008; Soares et al., 2020, 2021a, 2022b). The auditory N100 ERP component has been associated with the processing of the sensory features of the stimulus and predictive mechanisms involved in the processing of speech streams (e.g., Heinks-Maldonado et al., 2005). In addition, modulations in the N400 have been proposed to reflect processes related to successful online segmentation of the speech stream into its perceptual units and to the emergence of a pre-lexical trace of words in the brain (see Sanders et al., 2002; Cunillera et al., 2006; De Diego Balaguer et al., 2007; Soares et al., 2020).

Of especial relevance for the purposes of this paper, are the results from a follow-up study (Soares et al., 2022b) in which the authors compared the behavioral and the neural correlates of SL in a group of 5-year-old language unimpaired children to the behavioral and the neural correlates of SL in a group of language unimpaired adults to get new insights into the changes SL might undergo throughout development. Although behavioral (2-AFC) signs of SL were only observed for adult participants, evidence of SL was observed in the N100 and N400 ERP components in both groups, even though a detailed analysis of the neural data revealed some differences between adults and children. For instance, although similar modulations were found in the N100 component in both groups, showing a larger amplitude in the last part relative to the first part of the aSL tasks, differences were observed in the N400 component. In this time window, adults revealed a larger N400 amplitude for the high-TP vs. low-TP "words" regardless of the task, replicating Soares et al. (2020) results, while children showed a more intricate pattern that changed as a function of the predictability of the "words," especially in the task presented under explicit conditions (a larger N400 amplitude for the low-TP "words" in the first part of the explicit aSL task and for the high-TP "words" in the last part of the explicit aSL task). These findings led the authors to claim that children and adults rely on different mechanisms to assist the extraction of regularities (TPs) embedded in complex speech streams and that SL with auditory linguistic materials is not age-invariant as some authors state (e.g., Reber, 2013; for a review, see Zwart et al., 2017). Anyway, the important point to stress here is that although preschool language unimpaired children failed to reveal behavioral signs of SL, a result that was also observed in other studies with children below 6 years of age (e.g., Raviv and Arnon, 2018; Shufaniya and Arnon, 2018; van Witteloostuijn et al., 2019), the neural results observed in the N400 ERP component showed critically that preschool language unimpaired children were able to take advantage of the previous knowledge of the "word-like" units embedded in the speech streams to boost SL functioning. Thus, the question at stake in the present study is to analyze whether children of the same age with DLD would show a similar pattern of results. Although this is, to the best of our knowledge, the first study conducted in this regard, we hypothesized that if explicit (declarative) learning mechanisms play indeed a compensatory role in implicit (procedural) learning deficits in children with DLD, as the PDH claims, children with DLD should not only present enhanced modulations in the N400 component when the aSL is performed under intentional (explicit) vs. incidental (implicit) conditions, similarly to TLD controls, but also importantly reveal greater differences between the processing of the speech streams under implicit vs. explicit conditions when compared to children from the TLD group. Moreover, differences across the type of "words," which rely precisely on the computation of syllable TPs, were expected to be lessened in the DLD group due to a strong reliance on explicit (declarative) learning mechanisms to process the speech streams they were exposed to. Although previous studies have failed to show reliable behavioral signs of SL through the use of the offline 2-AFC post-learning task in children below 6 years of age, in this paper we nevertheless opted to collect 2-AFC data from children with and without DLD to further ascertain whether the improvement in SL performance, which children with DLD might reveal when performing the aSL task under explicit conditions, could also be noticed at a behavioral level of analysis.

MATERIALS AND METHODS

Participants

Forty preschool children participated in the study. All were native European Portuguese speakers with normal hearing, as assessed with pure-tone audiometry according to BIAP 02/1 classification (Bureau International Audiofonologie, 2005), and with no neurological or intellectual disabilities. 20 of them, recruited from Speech-Therapist Clinics, presented DLD, while the other 20, recruited from kindergarten institutions, presented typical language development (TLD). Parental informed consent was obtained from all the participants. The study was carried out in accordance with the guidelines of the Declaration of Helsinki and approved by the ethics committee of the local Ethics Committee (University of Minho, SECSH 028/2018).

Children from the DLD and TLD groups were matched on sex, $\chi^2(1) = 2.56$, $p = 0.110$, age, $t(38) = 1.36$, $p = 0.183$, and in non-verbal IQ, $t(38) = 1.84$, $p = 0.073$ as assessed by the Raven's Colored Progressive Matrices—Parallel form (CPM-P; Raven et al., 2009). They were also matched in rapid naming both when the time (in seconds), $t(37) = 1.05$, $p = 0.302$, and the number of errors committed producing the name of colors, $t(37) = 0.86$, $p = 0.396$, were taken into account, and in their visuospatial and kinesthetics short-term memory, as assessed by the Corsi block-tapping test, $t(38) = 0.72$, $p = 0.636$, from the Coimbra Neuropsychological Assessment Battery (CNAB; Simões et al., 2016). Children from both groups were also screened on their language abilities by the use of the Preschool Language Assessment (PLA; Mendes et al., 2014), an European Portuguese instrument measuring preschool receptive [listening comprehension (LC)] and expressive [oral verbal expression (OVE)] language skills and metalinguistic awareness (Metalanguage) in the areas of semantics, morphosyntax, and phonology (see Mendes et al., 2014 for details) and through the use of the nonword repetition task of the Language Skills Screening Test (LSST; Viana, 2004) another European Portuguese instrument targeting preschool children. **Table 1** presents the demographic, cognitive, and linguistic characteristics of the children included in the DLD and TLD groups in the study.

As expected, children from both groups differ on their receptive (LC), $t(38) = 3.22$, $p = 0.003$, expressive (OVE), $t(38) = 3.97$, $p < 0.001$, and metalanguage skills, $t(38) = 3.65$, $p < 0.001$, regardless of the language area; as well as in each of the language areas regardless of being receptive or expressive skills, Total semantics, $t(38) = 2.72$, $p = 0.010$, and Total Morphosyntax, $t(38) = 4.80$, $p < 0.001$. Differences were also observed in the nonword repetition task, $t(38) = 3.67$, $p < 0.001$. Taken together, the results obtained from these measures attested the diagnosis of the children in each group and showed that across groups children were also controlled in important demographic and cognitive measures that could impact the results.

Stimuli

The three-syllable nonsense words used in the implicit and explicit versions of the aSL tasks were drawn from Soares et al. (2020). They were made from 32 unique European Portuguese syllables produced and recorded by a native speaker of European Portuguese with duration of 300 ms each. These syllables were distributed into two different syllabaries (A and B) with 16 syllables each to be used in the implicit and the explicit versions of the aSL tasks (counterbalanced across participants). The syllables were concatenated with the Audacity® software (1999–2019) to ensure the absence of any co-articulation cues to affect word segmentation. In each aSL task, four of the nonsense words present TPs between syllables within a "word"of 1.00 (high-TP "words"), whereas the remaining four present TPs within a "word" of 0.33 (low-TP "words"), as in previous works of Soares et al. (2020, 2021a, 2022b). For instance, the nonsense word "*tucida*" presented in **Figure 1**, which represents a graphic depiction of an auditory stream presented to participants, corresponds to a high-TP "word" as the syllables they entail only appear in that "word" and in that specific syllable positions, while the nonsense word "*migedo*" corresponds to a low-TP "word" as the syllables it entails appear in three different "words" embedded in the stream at different (initial, medial, and final) syllable positions as in the case of the first syllable "*mi*" in the nonsense words "*gemiti*" and "*tidomi*" also presented in **Figure 1** (see Soares et al., 2020 for details).

The streams in each of the aSL tasks were edited to contain 60 repetitions of the same nonsense word distributed over six blocks of 10 repetitions each, lasting 1.4 min per block (8.4 min in total). In each block, "words" were presented in pseudo-randomized order to assure that the same "word" or the same syllable will never appear consecutively in a row. In each of the aSL tasks, the stream was also edited to include a randomly superimposed chirp sound (a 0.1 s sawtooth wave sound from 450 to 1,450 Hz) to provide participants with a cover task (i.e., a click detection task), to ensure appropriate attention to the stream as in previous works (see Arciuli and Simpson, 2012; Soares et al., 2020). The target sound was programmed to appear in the interval between syllables in 30% of each "word" type, counterbalanced across syllables to prevent confounds. Correct detections were 130.6 (±8.6) in the implicit aSL task (90.7% of all responses, including false alarms) and 126.8 (±12.7) in the explicit aSL task (88.1%) in children from the DLD group, whereas they were 136.8 (±4.7) in the implicit aSL task (94.9%) and 132.6 (±5.3) in the explicit aSL task (92.2%) in the children from the TLD group. These findings suggest that children from both groups paid appropriate attention to the speech streams in each of the aSL tasks.

The foils used in the 2-AFC tasks were also drawn from the work of Soares et al. (2020). They were made of the same syllables used in the high- and low-TP "words" presented during the exposure phase of each of the aSL tasks, presented with the same frequency and syllable positions to avoid confounds. However, contrary to the syllables in the high- (TP = 1.00) and low-TP "words" (TP = 0.33), the syllables in the foils were never presented together during exposure (TPs = 0). Four lists of materials were created to counterbalance syllables across positions and the type of "words" in each syllabary. Participants in each group were randomly assigned to one list

TABLE 1 | Descriptive (Frequencies, Means, and SDs—in brackets) of the characteristics of the children in the developmental language disorder (DLD) and typical language developmental (TLD) groups.

		DLD	TLD
Sex (masculine; feminine)		11; 9	6; 14
Age (years; months)		5;8 (0.43)	5;7 (0.34)
CPM-P scores (percentiles)		62.3 (22.2)	73.9 (17.5)
CNAB scores	Rapid naming (time in s.)	81.7 (28.9)	73.9 (16.5)
	Rapid naming (#errors)	0.68 (1.6)	0.35 (0.8)
	Corsi block-tapping test (%accuracy)	31.9 (11.6)	33.4 (11.9)
PLA scores (percentiles)	Listening comprehension (LC)	56.3(20.9)	74.6 (14.2)
	Oral verbal expression (OVE)	44.3 (22.9)	68.9 (16.9)
	Metalanguage	67.8 (25.0)	90.0 (7.9)
	Total Semantics (LC + OVE)	58.6 (20.2)	73.25 (20.0)
	Total Morphosyntax (LC + OVE)	35.3 (29.5)	73.5 (19.3)
LSST	Nonword repetition (%accuracy)	10.0 (12.6)	32.5 (15.1)

DLD, developmental language disorder group; TLD, typical language development group; CPM-P, Raven's colored progressive matrices-parallel form; CNAB, coimbra neuropsychological assessment; PLA, preschool language assessment; and LSST, language skills screening test.

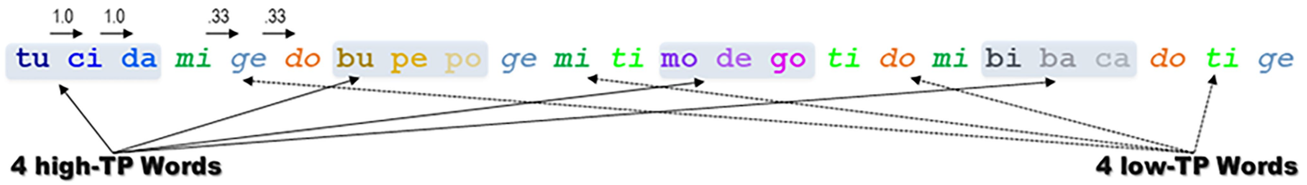

FIGURE 1 | Visual depiction of the high- and low-TP "Words" used in the auditory streams. High-TP, high-transitional probability "words"; Low-TP, low-transitional probability "words."

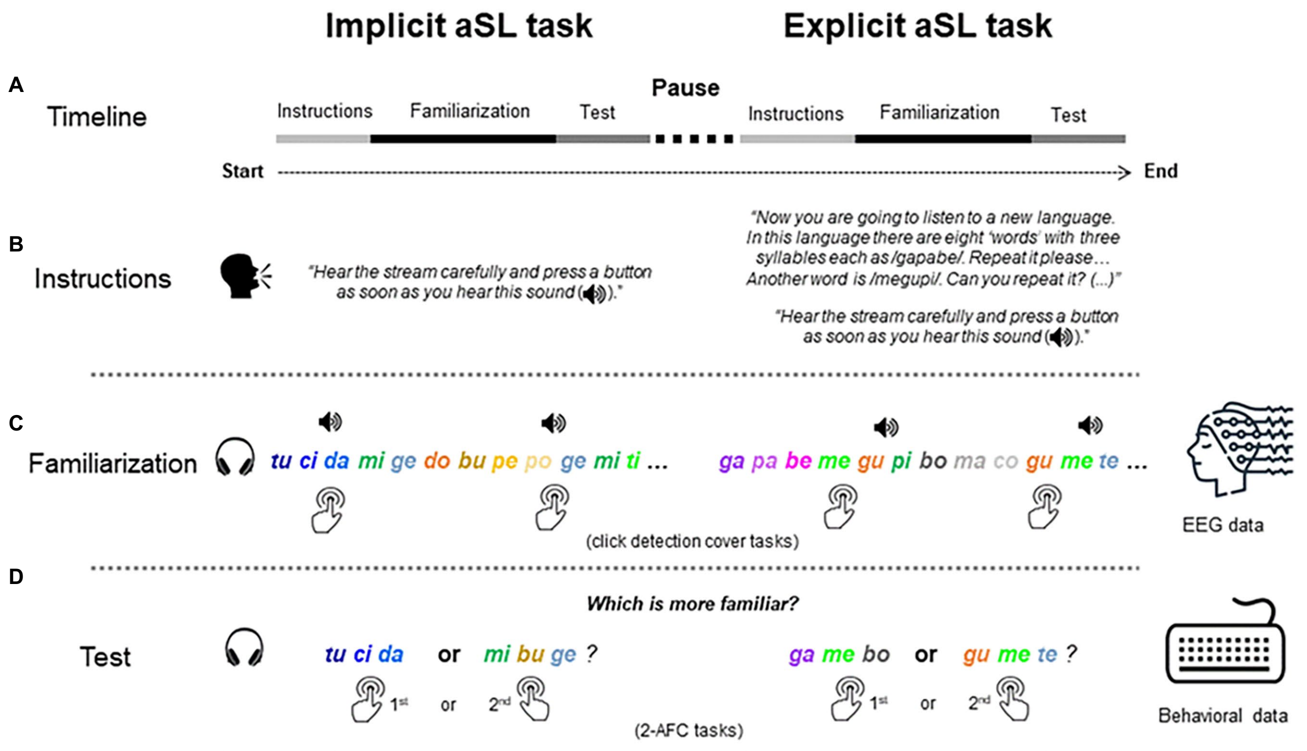

FIGURE 2 | Illustration of the experimental procedure. **(A)** Illustrates the timeline of the experimental procedure in which the implicit and, subsequently, the explicit aSL task were administered. Each aSL task comprised three parts: instructions **(B)**, familiarization **(C)**, and test **(D)** phases. Each task was initiated with specific instructions that determined the conditions under which each of the aSL tasks was performed either without (Implicit aSL) or with the previous knowledge of the task and the structure of the stream used in the experiment (Explicit aSL). In the familiarization or exposure phase of both tasks during which EEG data were collected, participants were presented with a continuous auditory stream of four high-TP and four low-TP "words," with chirp sounds (depicted as a speaker icon on the figure) superimposed over specific syllables. The chirp sounds could emerge at any of the three syllable positions of the "words," which precluded its use as a cue for "word" segmentation. During this phase, participants had to perform a chirp detection task. Then, the test phase in each of the aSL tasks consisted of a two-alternative forced-choice (2-AFC) task asking participants to indicate which of the three-syllable sequences (a "word" and a foil) sounded more familiar based on the stream presented during the previous familiarization phase.

of the Syllabary A and one list of the Syllabary B to perform the aSL tasks either under implicit or explicit conditions. The entire lists of materials are available at https://osf.io/8nx35/?view_only=264c374fa0584584aac85e4b6b39a0b1.

Procedure

EEG data collection was performed in an electric shielded, sound-attenuated room at the facilities of the Psychological Neuroscience Lab, School of Psychology, University of Minho. Participants were seated in a comfortable chair, 1 m away from a computer screen. EEG data were recorded during the exposure phases of each of the aSL tasks with a 64-channel BioSemi Active-Two system (BioSemi, Amsterdam, The Netherlands) according to the international 10–20 system and digitized at a sampling rate of 512 Hz. Electrode impedances were kept below 30 kΩ. EEG was re-referenced offline to the algebraic average of mastoids. Participants were first presented with the implicit version of the aSL task and, subsequently, with the explicit version of an analogous aSL task (see **Figure 2**). In the implicit version of the task, participants were instructed to pay attention to the auditory stream presented at 60 dB SPL *via* binaural headphones, because occasionally a deviant sound (i.e., a click) would appear, and their task would be to detect it as soon and accurately as possible by pressing the spacebar from the computer keyboard. As mentioned, this functioned as a cover task to assure children paid appropriate attention to the speech streams. Following familiarization, participants were asked to perform the 2-AFC (i.e., to decide which of two auditory stimuli, a "word" and a foil, "sounded more like" the stimuli presented before). The 2-AFC comprised 16 trials (two repetitions of the same nonsense words and foils). Each "word" was paired with two different foils. We opted for this solution instead of presenting each "word" paired exhaustively with each foil (64 trials) as in the work of Soares et al. (2020), because Soares et al. (2022a) recently demonstrated that increasing the number of 2-AFC trials by repeating the same stimuli ("words" and foils) several times throughout the 2-AFC task, worsens SL measurement, as it increases the chances of foils being learned as perceptual units and to interfere with correct "word" discrimination (see Soares et al., 2022a for details).

In the 2-AFC tasks, each trial began with the presentation of a fixation point (cross) for 1,000 ms, after which the first stimulus ("word"/foil) was presented, followed by the second stimulus. A 500-ms inter-stimulus interval separated the presentation of both stimuli. The next trial began as soon as participants made a response or 10 s had elapsed. The 16 trials were presented in two blocks of eight trials each. In each block, the order (first or second) by which the stimuli were presented was controlled for, so that in half of the trials half of the high-TP and half of the low-TP "words" were presented firstly and in the other half the other way around. The trials in each block, as well as the blocks across the task, were randomly presented to the participants. After a brief interval, participants underwent the explicit version of the aSL task. This version followed basically the same procedure adopted in the implicit aSL task, except that previously to the exposure phase participants were presented with additional information about the stimuli that they would listen to during another exposure phase with another set of materials. Specifically, during this training phase, participants were presented auditorily with each of eight new "words" (one by one) and asked to repeat each of them correctly before another "word" was presented. As in the implicit version of the task, during the exposure phase, participants were asked to press a button of the keyword whenever they heard the click sound. After familiarization, participants performed another 2-AFC task similar to the one used previously. The procedure took about 90 min to be completed per participant. **Figure 2** presents a visual depiction of the procedure.

RESULTS

Behavioral (2-AFC) and ERP data analyses were performed using the IBM-SPSS® software (Version 27.0). For behavioral data, the proportion of correct responses was computed for each of the 2-AFC tasks and separately for the high-TP and low-TP "words" in each group of participants (coded as 1 for a correct and 0 for an incorrect response). Grand averages waveforms were calculated in each group for each aSL task and type of "word" separately attending to the length of exposure to the stream (first half vs. second half of each task), to get insights into the temporal dynamics of SL as in previous works. Six participants from the DLD group and four from the TLD group were excluded from the EEG (and also from the behavioral) analyses due to artifact rejection. Data were filtered with a bandpass filter of 0.1–30 Hz (zero phase shift Butterworth). ERP epochs were time-locked to the nonsense words' onset, from −300 to 1,200 ms (baseline correction from −300 to 0 ms). Independent component analyses (ICA) were performed to remove stereotyped noise (mainly ocular movements and blinks) by subtracting the corresponding components. After that, epochs containing artifacts (i.e., with amplitudes exceeding +/−100 μV) were removed. EEG data processing was conducted with Brain Vision Analyzer, version 2.1.1. (Brain Products, Munich, Germany).

Based on previous literature, mean amplitudes were measured for the following time windows: 80–120 ms (N100 component) and 350–450 ms (N400 component). To account for the topographical distribution of the abovementioned EEG deflections, mean amplitudes' values were obtained for the topographical regions where amplitudes were maximal: the frontocentral region of interest (ROI; F1, Fz, F2, FC1, FCz, FC2, C1, Cz, and C2) for N100, and the central ROI (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, and CP2) for the N400. Both for behavioral and ERP data, main or interaction effects that reached statistical or marginal significance levels in comparisons of interest are reported. The Greenhouse–Geisser correction for nonsphericity was used when appropriate. *Post hoc* tests for multiple comparisons were reported after the Bonferroni correction. Measures of effect size (partial Eta squared, η_p^2) and observed power (*pw*) for a single effect are reported in combination with the main effects of condition.

Behavioral Data

The mean proportions of correct responses obtained by each group in the 2-AFC tasks performed under implicit and explicit conditions per type of "word" are presented in **Table 2**.

As can be seen in **Table 2**, the results are quite similar across groups and conditions, particularly in the task performed under explicit conditions. In the task performed under implicit conditions, participants from both groups were less accurate at recognizing low- than high-TP "words," as expected. Nevertheless, the results from the one-sample *t*-tests against chance level failed to reach statistical significance in all the conditions (all $ps > 0.144$), indicating that children from each group as a whole failed to reveal reliable behavioral signs of SL. Nonetheless, the analysis of the individual 2-AFC performance of the children in each group and aSL task showed substantial variability with approximately one-third of children in each group showing a 2-AFC performance above the mean group performance, as depicted in **Figure 3**.

ERP Data

The up panel of **Figure 4** depicts the grand-averaged waveforms (central ROI) obtained from children with and without DLD in the aSL tasks performed under implicit (light lines) and explicit (dark lines) conditions and, in each of them, for the high- (solid lines) and low- (dotted lines) TP "words" in the first half (Half 1) and the second half (Half 2) of each of the aSL tasks. The bottom panel displays the topographic maps obtained in these same conditions.

The results of the ANOVAs conducted for each of the time windows of interest revealed a significant main effect of the length of exposure, maximal at the frontocentral ROI, $F(1, 28) = 5.80$, $p = 0.023$, $\eta_p^2 = 0.17$, $pw = 0.64$, in the N100 component. This effect indicates that children from both groups showed a larger amplitude in the second half than in the first half of the aSL tasks, regardless of the conditions under which they were performed (implicit vs. explicit) and the type of "words" (high-TP vs. low-TP).

In the N400 component, a significant main effect of task, maximal at the central ROI, was observed, $F(1,28) = 7.69$, $p = 0.010$, $\eta_p^2 = 0.35$, $pw = 0.76$. This effect revealed that children from both groups as a whole showed a larger amplitude in this component when the task was performed under explicit than implicit conditions, regardless of the type of "words" (high-TP vs. low-TP) and length of exposure to the streams (first half vs. second half). Importantly, the four-way interaction also reached statistical significance in this time window, $F(1,28) = 5.02$, $p = 0.033$, $\eta_p^2 = 0.15$, $pw = 0.58$. Pairwise comparisons revealed that children from the TLD group showed higher amplitudes than children from the DLD group in the first half of the explicit aSL task for the low-TP "words" ($p = 0.020$), and in the second half of the explicit aSL task for the high-TP "words" ($p = 0.018$). Moreover, a marginal significant group effect was also observed in the first half of the implicit aSL task for the high-TP "words" ($p = 0.054$) indicating a tendency for children from the TLD group to show a larger N400 amplitude than children from the DLD group for the high-TP "words." **Figure 5** depicts a graphical representation of the group effect (DLD group = red lines; TLD group = blue lines) in the aSL tasks performed under implicit and explicit conditions per type of "word" (high-TP = solid lines; low-TP "words" = dotted lines) and length of exposure (first vs. second half).

Moreover, the results also revealed that the above-mentioned main effect of task was restricted to children from the TLD group. Indeed, only children from the language unimpaired group showed a larger N400 amplitude in the explicit vs. implicit aSL tasks, even though for the low-TP "words" ($p = 0.002$)

TABLE 2 | Mean (SD) of the number (Proportion) of correct responses for the High- and Low-TP "Words" in the implicit and explicit auditory SL task (aSL) tasks in the DLD and TLD groups.

Type of "Word" Group	aSL task			
	Implicit		Explicit	
	High-TP	Low-TP	High-TP	Low-TP
DLD	0.51 (0.15)	0.48 (0.15)	0.53 (0.16)	0.53 (0.22)
TLD	0.52 (0.14)	0.44 (0.21)	0.55 (0.17)	0.54 (0.19)

DLD, developmental language disorder group; TLD, typical language development group.

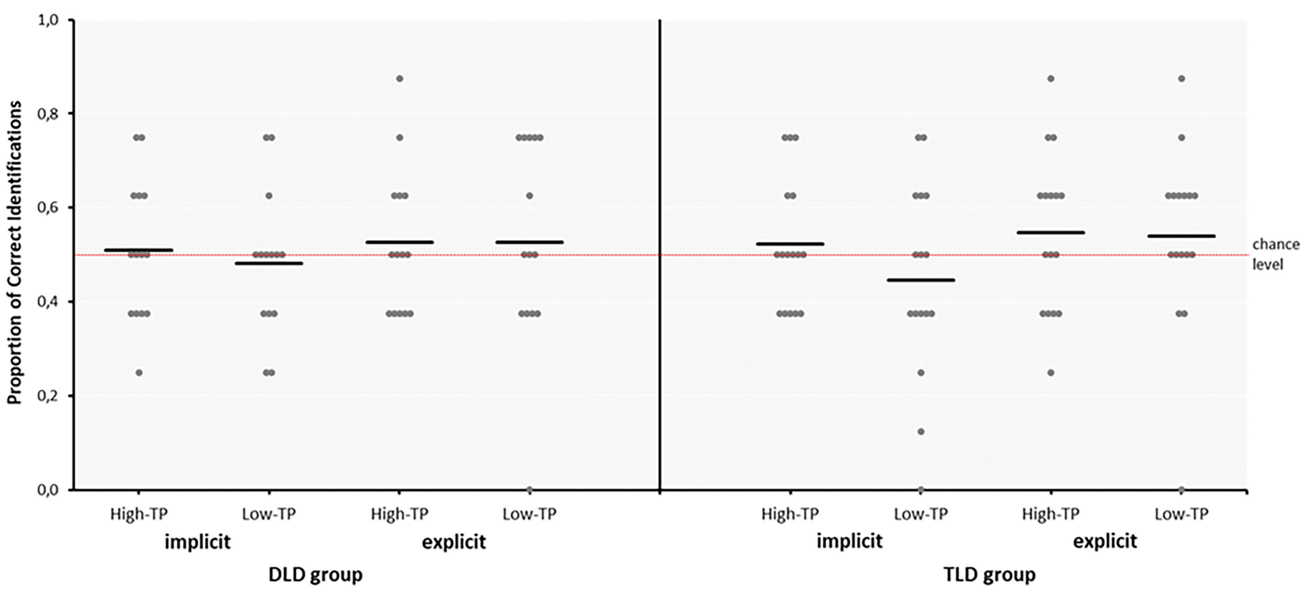

FIGURE 3 | Accuracy Rates (Proportion of Correct Identifications) in the 2-AFC Tasks Performed under Implicit and Explicit Conditions for the high- and low-TP "Words" in the DLD and TLD Groups. DLD, developmental language disorder group; TLD, typical language development group. The dots represent the scores obtained by each participant in each of the conditions (aSL task and type of "word") per group (DLD and TLD) while the horizontal black solid lines in each of these cases represent the mean of the group in each of these conditions.

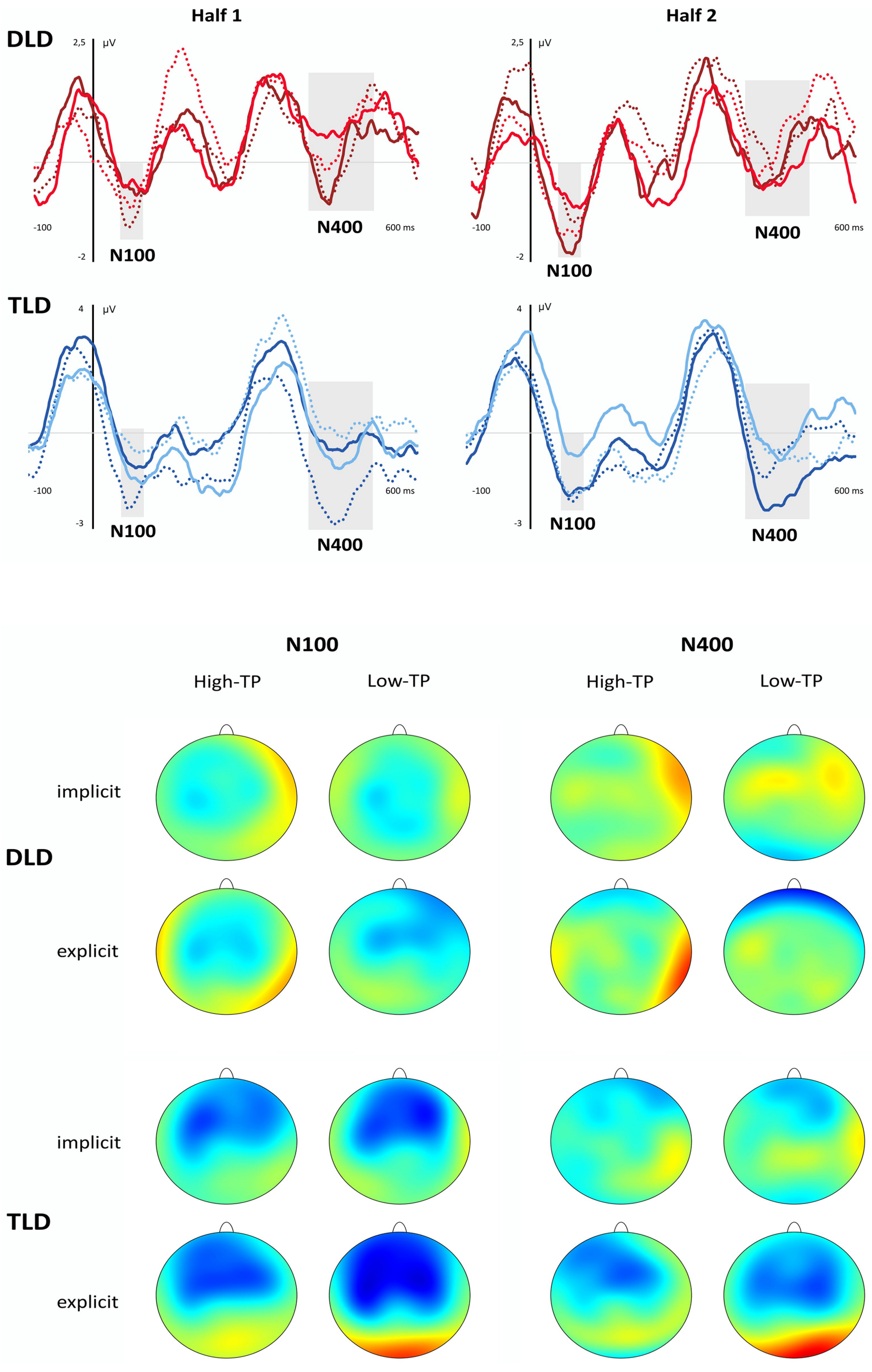

FIGURE 4 | Grand-Averaged Waveforms (Central ROI) and Topographic Maps in the DLD and TLD Groups. In the up panel, the gray shadowed rectangles indicate the analyzed time windows (N100 and N400). DLD, developmental language disorder group: Light red solid line = implicit high-TP condition; Light red dotted line = implicit low-TP condition; Dark red solid line = explicit high-TP condition; Dark red dotted line = explicit low-TP condition. TLD: typical language development group: Light blue solid line = implicit high-TP condition; Light blue dotted line = implicit low-TP condition; Dark blue solid line = explicit high-TP condition; and Dark blue dotted line = explicit low-TP condition. In the bottom panel, values of the topographical images range from −3 to 3 μV in each group and condition.

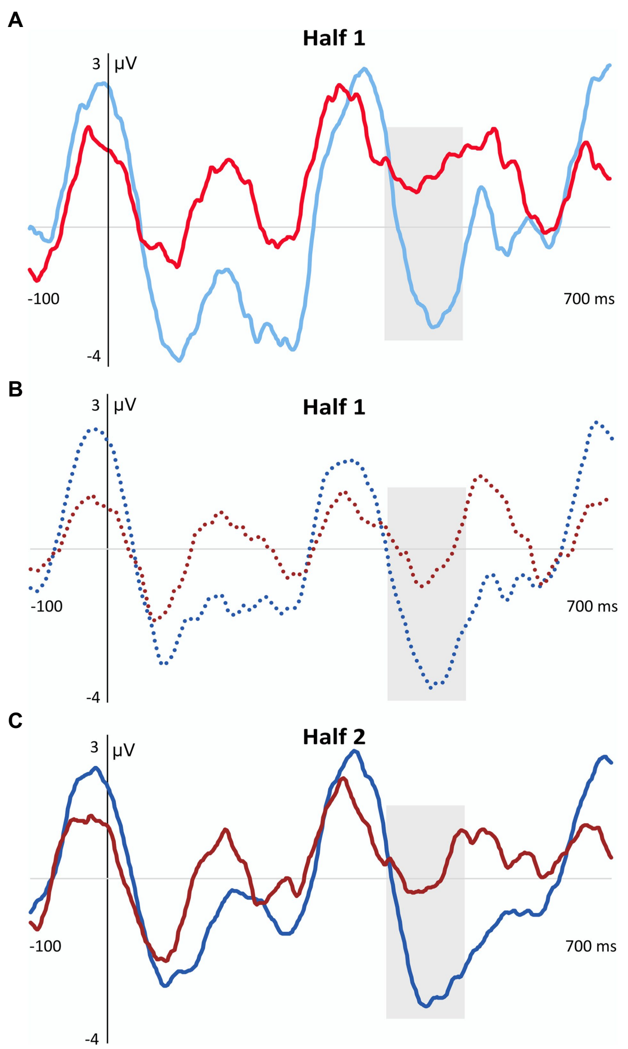

FIGURE 5 | Graphical representation of the Group Effect Observed in the N400 Component for the Implicit and Explicit aSL Tasks as a Function of Type of "Word" and Length of Exposure. Gray shadowed rectangles indicate the N400 time window. **(A)** Group effect in the first half of the implicit aSL task in the high-TP condition. Light red solid line = Development language disorder group; Light blue solid line = Typical language development group. **(B)** Group effect in the first half of the explicit aSL task in the low-TP condition. Dark red dotted line = Development language disorder group; Dark blue dotted line = Typical language development group. **(C)** Group effect in the second half of the aSL explicit task in the high-TP condition. Dark red solid line = Development language disorder group; Dark blue solid line = Typical language development group.

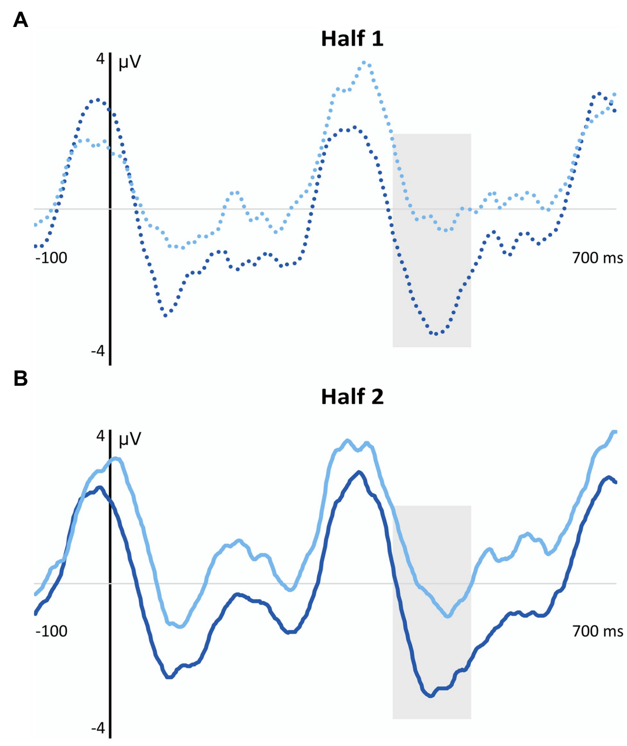

FIGURE 6 | Graphical representation of the aSL Task Effect observed in the N400 Component in the TLD Group as a Function of Type of "Word" and Length of Exposure. Gray shadowed rectangles indicate the N400 time window. **(A)** Task effect in the first half of the explicit aSL task in the TLD group for the low-TP condition. Light blue dotted line = implicit aSL task; Dark blue dotted line = explicit aSL task. **(B)** Task effect in the second half of the explicit aSL in the TLD group for the high-TP condition. Light blue solid line = implicit aSL task; Dark blue solid line = explicit aSL task.

in the first half of the explicit aSL task and for the high-TP words ($p = 0.007$) in the second half of the explicit aSL task. **Figure 6** depicts the task effects observed in the TLD group (aSL implicit = light blue lines; aSL explicit = dark blue lines) per type of "word" (high-TP = solid lines; low-TP "words" = dotted lines) and length of exposure (first vs. second half).

In addition, the four-way interaction revealed that children from the TLD group showed a "word" effect in the second half of the explicit aSL task, as reflected in a larger N400 amplitude for the high-vs. low-TP "words" ($p = 0.044$), and also an exposure length effect indicating a larger N400 amplitude in the first half than in the second half of the explicit aSL task for the low-TP "words" ($p = 0.029$). **Figure 7** depicts these effects.

DISCUSSION

The present work aimed to get new insights on the compensatory role that explicit (declarative) learning might play on implicit (procedural) learning deficits in children with DLD, as the PDH claims. For that purpose, we resorted to a new approach that involved the collection of neural (ERP) data, while preschool children, with and without DLD, were exposed to speech streams made of the repetition of four high- and four low-TP three-syllable nonsense words, first under implicit and afterward under explicit conditions—as in Soares and colleagues previous works (Soares et al., 2020, 2021a, 2022b). At the end of the exposure phase of each aSL task, behavioral data were also collected through the use of a standard 2-AFC task. The

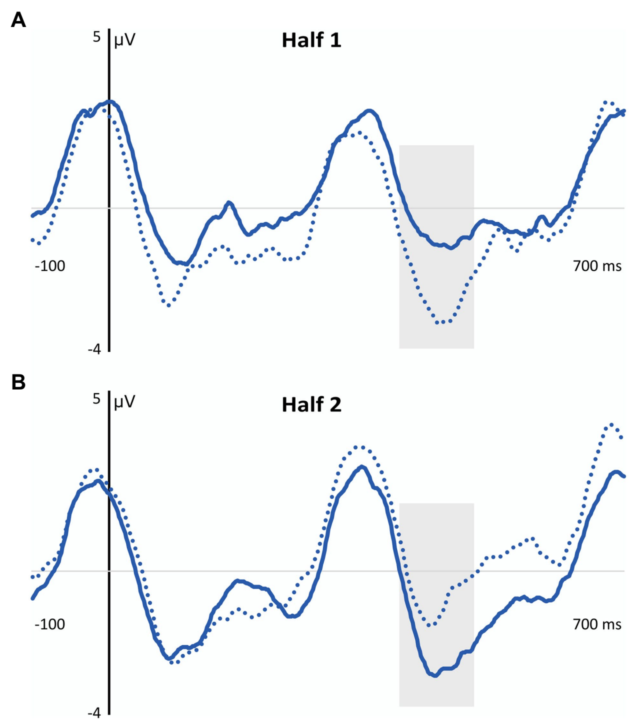

FIGURE 7 | Graphical representation of the Length of Exposure (Panel **A**) and Type of "Word" (Panel **B**) Effects Observed in the N400 Component in the Explicit aSL in the TLD Group. Gray shadowed rectangles indicate the N400 time window. **(A)** Length of exposure effect in the first half of the aSL explicit task in the TLD group. **(B)** "Word" effect in the second half of the explicit aSL task in the DLD group. Dark blue solid line = high-TP condition; Dark blue dotted line = low-TP condition.

combination of behavioral and neural measures in a within-subject design allowed us to overcome some of the limitations of previous works specifically those arising from the use of different tasks and stimuli to test PM vs.DM functioning and the exclusive collection of behavioral (RT/accuracy) responses that are strongly affected by attentional and motivational factors, particularly in studies involving language-impaired children from young ages. This is, to the best of our knowledge, the first study using this approach to further analyze the dynamics of DM-PM learning mechanisms in children with DLD (relative to TLD controls), which might provide compelling evidence to ascertain the extent to which explicit learning mechanisms can be effectively mobilized by these children to compensate for implicit learning deficits, as claimed by the PDH with important theoretical and clinical implications.

The results obtained from the 2-AFC tasks showed that children from each group as a whole failed to show reliable signs of SL even though the analyses of their individual performance showed substantial variability in both groups across aSL tasks and type of "words." The absence of reliable signs of SL even after the "words" have been explicitly taught is not new. Actually, they replicate previous results obtained by our research team with language unimpaired children of the same ages (e.g., Soares et al., 2022b; see also Soares et al., 2021c for similar findings with the artificial grammar learning paradigm), and they also agree with other works using language unimpaired children below 6 years of age (see, for instance, Raviv and Arnon, 2018; Shufaniya and Arnon, 2018 or van Witteloostuijn et al., 2019). Even though it is possible that the absence of behavioral signs of SL in our results may also stem from the complexity of the streams used, which entailed not only a higher number of "words" but "words" more diverse in their internal composition, it is nevertheless important to stress that all those works converge on the view that the 2-AFC task is not well-suited to test SL and that these null results should not be taken as a reflection of "non-learning" but, rather, as the inability of the 2-AFC task to capture SL in children without DLD (Arnon, 2020; Lukics and Lukács, 2021; Soares et al., 2022b).

The ERP data revealed, however, modulations in the N100 and N400 components, taken as the neural signatures of SL in the brain (e.g., Sanders et al., 2002; De Diego Balaguer et al., 2007; Abla et al., 2008; Soares et al., 2020), highlighting, once again, the usefulness of the ERP technique to cope with the limitations of the 2-AFC post-learning tasks to test SL. In particular, the neural results showed enhanced N100 amplitude as exposure to the speech streams unfolded in both groups of participants, regardless of the aSL task and type of "word." These findings are in line with previous studies and suggest that this component indexes transient effects that change as learning/exposure to the speech streams progresses and the regularities embedded in them are extracted (e.g., Sanders et al., 2002; Cunillera et al., 2006; De Diego Balaguer et al., 2007; Abla et al., 2008; Soares et al., 2022b). They also suggest that the task worked appropriately both for children with and without DLD. Although the absence of reliable behavioral signs of SL might raise some concerns about this interpretation, it is important to note that previous studies conducted with adult participants showed modulations in this component to be associated with the 2-AFC performance. For example, Abla et al. (2008) found that participants who have shown the higher performance in the 2-AFC task showed an increased N100 in the first part of the exposure phase of an aSL task with tones stimuli, while learners with an intermediate 2-AFC performance only showed that N100 enhancement in the last part of the aSL task. In the same vein, Soares et al. (2022b) found evidence for an increased N100 when language unimpaired adults were provided with explicit instructions to perform the aSL task, which also agreed with better 2-AFC performance under explicit conditions. Together, these findings seem to support the view that the increased N100 observed in our data for both groups of participants reflects the recruitment of predictive processes associated with the extraction of regularities embedded in the speech input (Heinks-Maldonado et al., 2005), even if behavioral signs of SL were not observed. Critically, they also showed this brain component to be observed not only in 5-year-old children without language impairments, as previously found by Soares et al. (2022b), but, also in children with DLD, suggesting this neural index of SL to be an early-maturing skill supporting language acquisition, as some authors claim (Saffran et al., 1996; Romberg and Saffran, 2010; Saffran, 2018,

2020) even if less efficiently in children with DLD than TLD controls as the results observed in the N400 component seem to suggest.

Indeed, even though the results observed in that time window, assumed to index processes related to a successful segmentation of the speech stream into perceptual units (word-like) in the brain (e.g., Sanders et al., 2002; Cunillera et al., 2006; De Diego Balaguer et al., 2007; Abla et al., 2008; Soares et al., 2020), indicated that children from both groups showed an enhancement in the N400 component when the task was performed under explicit rather than under implicit conditions, a result also observed in previous works conducted with language unimpaired participants (e.g., Daltrozzo and Conway, 2014; Batterink et al., 2015a,b; Soares et al., 2020, 2021a, 2022b); the four-way interaction observed in this ERP component revealed, however, that only children from the TLD group seem to have taken advantage of the previous knowledge to enhance SL functioning. Note that, within our framework, evidence for a compensatory role of explicit (declarative) learning on implicit (procedural) learning deficits would be indexed not only by enhanced modulations in this ERP component when the "word-like" units embedded in the speech streams were explicitly taught (vs. when they were not), but, importantly, that differences between the processing of the speech streams under implicit vs. explicit conditions would be greater for children from the DLD than for children from the TLD group. However, the results showed the reverse. Indeed, not only the group differences reveal that children from the TLD group showed larger N400 amplitudes than children from the DLD group both in the implicit (even though this effect, observed for the high-TP "words" in the first part of the task, was only marginally significant) and explicit aSL task (for low-TP "words" in the first part of the task and for high-TP "words" in the second half of the task) but, notably, that the differences across tasks only reached a statistically significant level for children from the language unimpaired group. These results agree with other works showing DM deficits in children with DLD (e.g., Lum et al., 2010; Bishop and Hsu, 2015; Kuppuraj et al., 2016; Lukács et al., 2017; McGregor et al., 2017; Lee, 2018; Haebig et al., 2019), thus failing to provide support for the compensatory role of DM in DLD, as the PDH claims (Ullman and Pierpont, 2005; Ullman et al., 2020). They also agree with a recent neuroimaging study using the diffusion tensor imaging (DTI) technique (Lee et al., 2020) showing dysfunctions in the white matter of the brain structures supporting both procedural and declarative functioning in adolescents and young adults with DLD relative to TLD controls.

Nonetheless, before strong conclusions can be drawn, it is also important to consider these results to have arisen from the type of stimuli used in our aSL tasks, once evidence showing DM impairments in children with DLD tends precisely to come from studies using verbal materials, as in our case (e.g., Lum et al., 2010; Bishop and Hsu, 2015; Lukács et al., 2017; McGregor et al., 2017; Haebig et al., 2019). Thus, it is possible to argue these results have stemmed from the difficulties that children from the DLD group present in the encoding and storing of the phonological information of the new "words" rather than from difficulties in using explicit knowledge/explicit learning mechanisms to assist SL *per se* (see Alt and Plante, 2006; Lum et al., 2012, 2015). This possibility should be considered, as children from the DLD group present, indeed, lower phonological working memory skills than children from the TLD group, as assessed by the nonword repetition task from the LSST (see **Table 1**), and these skills were proven to be strongly related to declarative memory functioning (e.g., Alt and Plante, 2006; Coady and Evans, 2008; Lum et al., 2012, 2015; Arthur et al., 2021). To explore the role that this variable might have played in the results, we conducted yet another analysis based on the same factorial design reported in the Results section but taking the scores obtained in the nonword repetition task into account (i.e., as a covariable in the ANOVAs). Even though the four-way interaction failed to reach statistical significance, due possibly to the lack of statistical power, further exploration of the results revealed nevertheless that the *post hoc* contrasts where the effects tended to reach statistical significance were exactly the same, thus ruling out the phonological working memory skills as the main driving force behind the results. Moreover, it is also important to consider that presenting such complex speech streams during 8.4 min might not suffice to allow children from the DLD group to use the cues embedded in the speech streams and/or the previous knowledge of the "word-like" units in a more efficient manner. For example, Tomblin et al. (1997), in one of the first studies examining PM deficits in adolescents with and without DLD using a serial reaction time task, found that despite adolescents with DLD showed slower learning rates than controls, at the end of the training, performance did not differ between groups. Also, Evans et al. (2009) using an aSL task similar to the one used here but with a lower number of "words" (six) in children with DLD relative to TLD controls, showed that although after 21 min of exposure children from the DLD group performed at chance in the post-learning 2-AFC task, when the time of exposure was doubled performance was significantly greater than chance. Future research should thus test whether extending the time of exposure to the speech streams would make children from the DLD group show a pattern of neural responses similar to the children from the TLD group, which might have important clinical implications. Note that if the same pattern of results emerges, even with extended exposure, this might suggest that using explicit instructions, a strategy that characterizes most of the language interventions in children with DLD (see Ebbels, 2014), might not be well-suited to help DLD children to overcome their language difficulties once they capitalize on skills that might also be impaired in this group of children. Clinical experiments that contrast the effectiveness of language interventions in children with DLD using implicit vs. explicit methods should also be conducted to address this important issue.

Finally, it is also worth mentioning that the results observed here in children from the TLD group replicate Soares et al. (2022b) findings and suggest that, conversely to children from the DLD group, children from the TLD group seem to have taken advantage of the knowledge generated from the previous presentation of the "word-like" units embedded in the speech streams to boost SL functioning. Moreover, they also showed the advantage of the explicit instructions to have affected first the low-TP "words" and only at a later stage the high-TP "words." This "word" type effect, already observed by Soares et al. (2022b), was accounted by the authors based on two possible explanations: (i) children used the prior knowledge generated from the previous presentation of the "word-like" units to assist the extraction of the most difficult "words"—note that the low-TP "words" are made up of syllables that were also found in other "words" embedded in the stream, which might make these "words" harder to extract and to produce less robust/stable perceptual representations (see Smalle et al., 2016 for evidence of the interference effect generated by item-overlap in a Hebb repetition task); (ii) children relied on syllable frequency instead of syllable TPs to assist word segmentation—note that despite high- and low-TP "words" were presented exactly the same number of times ($N = 60$) during the exposure phase to control for 'word' frequency effects (see Soares et al., 2015, 2019), the fact that low-TP "words" involved the encoding of a smaller number of syllables than high-TP "words" (4 vs. 12, respectively) and syllables that occurred three times more frequently than the syllables of the high-TP "words," might have made children to rely on a simpler strategy to predict the upcoming segment, hence relying on the syllable frequency instead of syllable TPs to create perceptual units beyond the syllable level.

Even though the current work was not designed to disentangle these two proposals, it is nevertheless important to stress that the effect observed in the first part of the implicit aSL task (even if marginal) seems to rule out the second proposal. Indeed, when the aSL task was performed under incidental conditions, children from the TLD group tended to show a larger N400 amplitude for the high-TP "words" in the first part of the implicit aSL task than children from the DLD group, whereas when the aSL task was performed under intentional conditions, children from the TLD group showed a larger N400 amplitude for the low-TP "words" in the first part of the explicit aSL task relative to children from the DLD group.

The result observed in the first part of the implicit aSL task for children from the TLD group suggests that when children performed the task without any information about the task or the stimuli, syllable TPs rather than syllable frequency seems to automatically drive word segmentation. This interesting result suggests that the previous presentation of the "word-like" units embedded in the speech streams might have interfered with the way children usually processed the speech streams to which they were exposed by disrupting a type of processing (based on the extraction of syllable TPs) that might indeed be automatically projected to segment the continuous speech input into word-like units to support language acquisition (Saffran et al., 1996; Romberg and Saffran, 2010; Saffran, 2018, 2020). It is also possible to anticipate that the prior presentation of the word-like units embedded in the speech streams has taxed processing more strongly making children rely on simpler statistics (syllable frequency) to identify the "word-like" units previously presented during exposure. Note that, unlike the implicit aSL task, in the explicit aSL task, children had to simultaneously attend to the "words" previously presented, to the clicks appearing occasionally in the stream, and to the auditory stimuli itself, which was certainly much demanding, justifying the shift in the statistics that children seem to have relied on when the task was performed under implicit vs. explicit conditions, at least when complex speech streams were used. In the same vein, it is possible to anticipate that the capacity limits for information processing that preschool children with DLD typically present in working memory, inhibition, and shifting abilities (see Vissers et al., 2015 for a review), have also hampered the ability of children with DLD to have taken advantage of the previous knowledge of the "word-like" units to boost SL functioning, even if using a simpler strategy as children without language impairments seem to have done. Future research should thus be conducted to analyze if presenting less complex speech streams to children with DLD (made of a lower and/or a less diverse type of "words") and/or with extended exposure to the speech streams would produce similar results. If future research confirms these results, this would also recommend amendments in the PDH, namely regarding two important assumptions: children with DLD have a spared or even an enhanced DM functioning, and these strengthened DM skills can be used to compensate for their PM deficits in language acquisition. Future research should also test whether similar results would be obtained when using other tasks and paradigms, namely those allowing for the counterbalance of the order of the tasks presented to the children once the nature of the SL task used here made the implicit followed by the explicit presentation of the SL task the only viable solution in this type of design.

CONCLUSION

The present study sheds light on the dynamics between implicit–explicit learning mechanisms in children with DLD using a new approach that combined the collection of neural and behavioral data from the same participants (children with DLD and TLD as controls) during the exposure phase of analogous versions of the same aSL task presented under implicit and explicit learning conditions. This new approach allowed us not only to control for differences in the results that might have arisen in previous studies from the use of different tasks and stimuli to test DM and PM functioning but, importantly, to directly examine the changes that performing analogous versions of the same task presented under different learning conditions produced in the SL functioning. This is, to the best of our knowledge, the first study adopting this approach to further examine the compensatory role that explicit learning mechanisms might play on implicit learning deficits in children with DLD, as the PDH claims. Although future studies are required, our findings failed to support the compensatory role of explicit learning mechanisms in the implicit learning deficits in children with DLD, which might have important theoretical and clinical implications.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/8nx35/?view_only=264c374fa0584584aac85e4b6b39a0b1.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Minho, SECSH 028/2018. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AS conceptualized the study and wrote the first draft of the manuscript. F-JG-D and HO implemented the experiment. AL, NG, and AP collected the data. AS, F-JG-D, HO, DT, and ML analyzed and interpreted the data. F-JG-D, HO, DT, and ML critically revised it. All authors contributed to the article and approved the submitted version.

FUNDING

This study was conducted at the Psychology Research Center (PSI/01662), University of Minho, and supported by the Grant POCI-01-0145-FEDER-028212 from the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Science, Technology and Higher Education through national funds, and co-financed by FEDER through COMPETE2020 under the PT2020 Partnership Agreement and within CINTESIS, R&D Unit (references UIDB/4255/2020 and UIDP/4255/2020).

ACKNOWLEDGMENTS

The research team would like to thank all the children, their families, and all the speech therapists, psychologists, and professors who generously shared their time for the purposes of this project.

REFERENCES

Abla, D., Katahira, K., and Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *J. Cogn. Neurosci.* 20, 952–964. doi: 10.1162/jocn.2008.20058

Ahufinger, N., Ferinu, L., Sanz-Torrent, M., Andreu, L., and Evans, J. L. (2021). Statistical word learning in Catalan-Spanish and English-speaking children with and without developmental language disorder. *Int. J. Lang. Commun.* 57, 42–62. doi: 10.1111/1460-6984.12673

Alt, M., and Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. *J. Speech Lang. Hear.* 49, 941–954. doi: 10.1044/1092-4388(2006/068)

Archibald, L. M. D., and Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *Int. J. Lang. Commun.* 41, 675–693. doi: 10.1080/13682820500442602

Arciuli, J., and Conway, C. (2018). The promise—and challenge—of statistical learning for elucidating atypical language development. *Curr. Dir. Psychol. Sci.* 27, 492–500. doi: 10.1177/0963721418779977

Arciuli, J., and Simpson, I. C. (2011). Statistical learning in typically developing children: the role of age and speed of stimulus presentation. *Dev. Sci.* 14, 464–473. doi: 10.1111/j.1467-7687.2009.00937.x

Arciuli, J., and Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cogn. Sci.* 36, 286–304. doi: 10.1111/j.1551-6709.2011.01200.x

Arciuli, J., von Koss Torkildsen, J., Stevens, D. J., and Simpson, I. C. (2014). Statistical learning under incidental versus intentional conditions. *Front. Psychol.* 5:747. doi: 10.3389/fpsyg.2014.00747

Arnon, I. (2020). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behav. Res. Methods* 52, 68–81. doi: 10.3758/s13428-019-01205-5

Arthur, D. T., Ullman, M. T., and Earle, F. S. (2021). Declarative memory predicts phonological processing abilities in adulthood. *Front. Psychol.* 12:658402. doi: 10.3389/fpsyg.2021.658402

Baird, G., Dworzynski, K., Slonims, V., and Simonoff, E. (2010). Memory impairment in children with language impairment. *Dev. Med. Child Neurol.* 52, 535–540. doi: 10.1111/j.1469-8749.2009.03494.x

Batterink, L., Paller, K. A., and Reber, P. J. (2019). Understanding the neural bases of statistical and implicit learning. *Top. Cogn. Sci.* 11, 482–503. doi: 10.1111/tops.12420

Batterink, L., Reber, P. J., Neville, H., and Paller, K. A. (2015a). Implicit and explicit contributions to statistical learning. *J. Mem. Lang.* 83, 62–78. doi: 10.1016/j.jml.2015.04.004

Batterink, L. J., Reber, P. J., and Paller, K. A. (2015b). Functional differences between statistical learning with and without explicit training. *Learn. Mem.* 22, 544–556. doi: 10.1101/lm.037986.114

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *Int. J. Lang. Commun.* 49, 381–415. doi: 10.1111/1460-6984.12101

Bishop, D. V. M., and Hsu, H. J. (2015). The declarative system in children with specific language impairment: A comparison of meaningful and meaningless auditory-visual paired associative learning. *BMC Psychol.* 3:3. doi: 10.1186/s40359-015-0062-7

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., and Greenhalgh, T. The CATALISE Consortium (2016). CATALISE: a multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS One* 11:e0158753. doi: 10.1371/journal.pone.0158753

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., and Greenhalgh, T. The CATALISE Consortium (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: terminology. *J. Child Psychol. Psychiatry Allied Discip.* 11, 1068–1080. doi: 10.1371/journal.pone.0158753

Bogaerts, L., Siegelman, N., and Frost, R. (2021). Statistical learning and language impairments: Toward more precise theoretical accounts. *Perspect. Psychol. Sci.* 16, 319–337. doi: 10.1177/1745691620953082

Bosseler, A. N., Teinonen, T., Tervaniemi, M., and Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS One* 11:e0162177. doi: 10.1371/journal.pone.0162177

Breen, E., Pomper, R., and Saffran, J. (2019). Phonological learning influences label-object mapping in toddlers. *J. Speech Lang. Hear. Res.* 62, 1923–1932. doi: 10.1044/2019_JSLHR-L-18-0131

Bureau International Audiofonologie (2005). Classification audiométrique des deficiencies auditives. BIAP: Recommendation BIAP n° 02/1 bis. Available at: http://www.biap.org/recom02-1.htm

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., and Werker, J. F. (2020). Preverbal infants discover statistical word patterns at similar rates as adults: evidence from neural entrainment. *Psychol. Sci.* 31, 1161–1173. doi: 10.1177/0956797620933237

Christiansen, M. H. (2019). Implicit-statistical learning: a tale of two literatures. *Top. Cogn. Sci.* 11, 468–481. doi: 10.1111/tops.12332

Clahsen, H., and Hansen, D. (1997). "The grammatical agreement deficit in specific language impairment: evidence from therapy experiments," in *The Biological Basis of Language*. ed. M. Gopnik (Oxford, England: Oxford University Press), 148–168.

Coady, J. A., and Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *Int. J. Lang. Commun. Disord.* 43, 1–40. doi: 10.1080/13682820601116485

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., and Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *Int. J. Lang. Commun.* 53, 237–255. doi: 10.1111/460-6984.12338

Conway, C. M., and Christiansen, M. H. (2006). Statistical learning within and between modalities. *Psychol. Sci.* 17, 905–912. doi: 10.1111/j.1467-9280.2006

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., and Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Res.* 1123, 168–178. doi: 10.1016/j.brainres.2006.09.046

Daltrozzo, J., and Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Front. Hum. Neurosci.* 8:437. doi: 10.3389/fnhum.2014.00437

De Diego Balaguer, R., Toro, J. M., Rodriguez-Fornells, A., and Bachoud-Lévi, A. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One* 2:e1175. doi: 10.1371/journal.pone.0001175

Desmottes, L., Meulemans, T., and Maillart, C. (2016). Implicit spoken words and motor sequences learning are impaired in children with specific language impairment. *J. Int. Neuropsychol. Soc.* 22, 520–529. doi: 10.1017/S135561771600028X

Dyck, M. J., Piek, J. P., and Patrick, J. (2011). The validity of psychiatric diagnoses: The case of 'specific' developmental disorders. *Res. Dev. Disabil.* 32, 2704–2713. doi: 10.1016/j.ridd.2011.06.001

Earle, F. S., and Ullman, M. T. (2021). Deficits of learning in procedural memory and consolidation in declarative memory in adults with developmental language disorder. *J. Speech Lang. Hear. Res.* 64, 531–541. doi: 10.1044/2020_JSLHR-20-00292

Ebbels, S. (2014). Effectiveness of intervention for grammar in school-aged children with primary language impairments: a review of the evidence. *Child Lang. Teach. Ther.* 30, 7–40. doi: 10.1177/0265659013512321

Erickson, L. C., and Thiessen, E. D. (2015). Statistical learning of language: theory, validity, and predictions of a statistical learning account of language acquisition. *Dev. Rev.* 37, 66–108. doi: 10.1016/j.dr.2015.05.002

Estes, K. G., Evans, J. L., Alibali, M. W., and Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol. Sci.* 18, 254–260. doi: 10.1111/j.1467-9280.2007.01885.x

Evans, J. L., Maguire, M. J., and Sizemore, M. L. (2022). Neural patterns elicited by lexical processing in adolescents with specific language impairment: support for the procedural deficit hypothesis? *J. Neurodev. Disord.* 14:20. doi: 10.1186/s11689-022-09419-z

Evans, J. L., Saffran, J., and Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *J. Speech Lang. Hear. Res.* 52, 321–335. doi: 10.1044/1092-4388(2009/07-0189)

François, C., Jaillet, F., Takerkart, S., and Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS One* 9:e101340. doi: 10.1371/journal.pone.0101340

Frost, R., Armstrong, B. C., Siegelman, N., and Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cogn. Sci.* 19, 117–125. doi: 10.1016/j.tics.2014.12.010

Gabriel, A., Meulemans, T., Parisse, C., and Maillart, C. (2014). Procedural learning across modalities in French speaking children with specific language impairment. *Appl. Psycholinguist.* 36, 747–769. doi: 10.1017/S0142716413000490

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychol. Sci.* 13, 431–436. doi: 10.1111/1467-9280.00476

Gómez, R., and Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy* 7, 183–206. doi: 10.1207/s15327078in0702_4

Gutiérrez-Domínguez, F. J., Lages, A., Oliveira, H. M., and Soares, A. P. (2022). "Neural signature of statistical learning: proposed signs of typical/atypical language functioning from EEG time-frequency analysis," in *International Handbook of Clinical Psychology*. ed. G. Buela-Casal (New York: Thomson Reuters Editorial).

Haebig, E., Leonard, L. B., Deevy, P., Karpicke, J., Christ, S. L., Usler, E., et al. (2019). Retrieval-based word learning in young typically developing children and children with development language disorder II: A comparison of retrieval schedules. *J. Speech Lang. Hear. Res.* 62, 944–964. doi: 10.1044/2018_JSLHR-L-18-0071

Hay, J. F., Pelucchi, B., Estes, K. G., and Saffran, J. R. (2011). Linking sounds to meanings: infant statistical learning in a natural language. *Cogn. Psychol.* 63, 93–106. doi: 10.1016/j.cogpsych.2011.06.002

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., and Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology* 42, 180–190. doi: 10.1111/j.1469-8986.2005.00272.x

Hsu, H., and Bishop, D. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Dev. Sci.* 17, 352–365. doi: 10.1111/desc.12125

Hsu, H., Tomblin, J., and Christiansen, M. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Front. Psychol.* 5:175. doi: 10.3389/fpsyg.2014.00175

Jiménez, L., Oliveira, H. M., and Soares, A. P. (2020). Surface features can deeply affect artificial grammar learning. *Conscious. Cogn.* 80:102919. doi: 10.1016/j.concog.2020.102919

Johnson, T., Siegelman, N., and Arnon, I. (2020). Individual differences in learning abilities impact structure addition: better learners create more structured languages. *Cogn. Sci.* 44:e12877. doi: 10.1111/cogs.12877

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., and Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* 127, 46–54. doi: 10.1016/j.bandl.2012.11.007

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Dev. Psychol.* 48, 171–184. doi: 10.1037/a0025405

Kidd, E., and Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Dev.* 87, 184–193. doi: 10.1111/cdev.12461

Kuppuraj, S., Rao, P., and Bishop, D. (2016). Declarative capacity does not trade-off with procedural capacity in children with specific language impairment. *Autism Dev. Lang. Impair.* 1, 239694151667441–239694151667517. doi: 10.1177/2396941516674416

Lages, A., Oliveira, H. M., Arantes, J., Gutiérrez-Domínguez, F., and Soares, A. P. (2022). "Drawing the links between statistical learning and children's spoken and written language skills," in *International Handbook Of Clinical Psychology*. ed. G. Buela-Casal (New York: Thomson Reuters Editorial).

Lammertink, I., Boersma, P., Wijnen, F., and Rispens, J. (2017). Statistical learning in specific language impairment: a meta-analysis. *J. Speech Lang. Hear. Res.* 60, 3474–3486. doi: 10.1044/2017_JSLHR-L-16-0439

Lammertink, I., Boersma, P., Wijnen, F., and Rispens, J. (2019). Children with developmental language disorder have an auditory verbal statistical learning deficit: evidence from an online measure. *Lang. Learn.* 70, 137–178. doi: 10.1111/lang.12373

Lammertink, I., Boersma, P., Wijnen, F., and Rispens, J. (2020). Statistical learning in the visuomotor domain and its relation to grammatical proficiency in children with and without developmental language disorder: a conceptual replication and meta-analysis. *Lang. Learn. Dev.* 16, 426–450. doi: 10.1080/15475441.2020.1820340

Lancaster, H. S., and Camarata, S. (2019). Reconceptualizing developmental language disorder as a spectrum disorder: issues and evidence. *Int. J. Lang. Commun. Disord.* 54, 79–94. doi: 10.1111/1460-6984.12433

Lee, J. C. (2018). Episodic memory retrieval in adolescents with and without developmental language disorders (DLD). *Int. J. Lang. Commun.* 53, 271–281. doi: 10.1111/1460-6984.12340

Lee, J. C., Nopoulos, P. C., and Tomblin, J. B. (2020). Procedural and declarative memory brain systems in developmental language disorder (DLD). *Brain Lang.* 205:104789. doi: 10.1016/j.bandl.2020.104789

Leonard, L. B. (1981). Facilitating linguistic skills in children with specific language impairment. *Appl. Psycholinguist.* 2, 89–118. doi: 10.1017/S0142716400000886

Leonard, L. B. (2014). *Children With Specific Language Impairment*. *2nd Edn*. Cambridge: MIT Press.

Leonard, L. B., Eyer, J., Bedore, L., and Grela, B. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *J. Speech Lang. Hear. Res.* 40, 741–753. doi: 10.1044/jslhr.4004.741

Lukács, A., Kemény, F., Lum, J. A. G., and Ullman, M. T. (2017). Learning and overnight retention in declarative memory in specific language impairment. *PLoS One* 12:e0169474. doi: 10.1371/journal.pone.0169474

Lukács, Á., Lukics, K. S., and Dobó, D. (2021). Online statistical learning in developmental language disorder. *Front. Hum. Neurosci.* 15:715818. doi: 10.3389/fnhum.2021.715818

Lukics, K. S., and Lukács, Á. (2021). Tracking statistical learning online: word segmentation in a target detection task. *Acta Psychol.* 215:e103271. doi: 10.1016/j.actpsy.2021.103271

Lum, J. A. G., and Bleses, D. (2012). Declarative and procedural memory in Danish speaking children with specific language impairment. *J. Commun. Disord.* 45, 46–58. doi: 10.1016/j.jcomdis.2011.09.001

Lum, J. A. G., and Conti-Ramsden, G. (2013). Long-term memory: a review and meta-analysis of studies of declarative and procedural memory in specific language impairment. *Top. Lang. Disord.* 33, 282–297. doi: 10.1097/01.tld.0000437939.01237.6a

Lum, J. A. G., Conti-Ramsden, G., Morgan, A., and Ullman, M. (2014). Procedural learning deficits in specific language impairment (SLI): a meta-analysis of serial reaction time task performance. *Cortex* 51, 1–10. doi: 10.1016/j.cortex.2013.10.011

Lum, J. A. G., Conti-Ramsden, G., Page, D., and Ullman, M. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex* 48, 1138–1154. doi: 10.1016/j.cortex.2011.06.001

Lum, J. A. G., Gelgec, C., and Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *Int. J. Lang. Commun.* 45, 96–107. doi: 10.3109/13682820902752285

Lum, J. A. G., Ullman, M. T., and Conti-Ramsden, G. (2015). Verbal declarative memory impairments in specific language impairment are related to working memory deficits. *Brain Lang.* 142, 76–85. doi: 10.1016/j.bandl.2015.01.008

Mandikal-Vasuki, P. R., Sharma, M., Ibrahim, R., and Arciuli, J. (2017). Statistical learning and auditory processing in children with music training: an ERP study. *Clin. Neurophysiol.* 128, 1270–1281. doi: 10.1016/j.clinph.2017.04.0

Mayor-Dubois, C., Zesiger, P., van der Linden, M., and Roulet-Perez, E. (2016). Nondeclarative learning in children with specific language impairment: predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuropsychol.* 20, 14–22. doi: 10.1080/09297049.2012.734293

McGregor, K., Arbisi-Kelm, T., and Eden, N. (2017). The encoding of word forms into memory may be challenging for college students with developmental language impairment. *Int. J. Speech Lang. Pathol.* 19, 43–57. doi: 10.3109/17549507.2016.1159337

McGregor, K., Goffman, L., Owen Van Horne, A., Hogan, T., and Finestack, L. (2020). Developmental language disorder: applications for advocacy, research, and clinical service. *Perspect. Asha Spec. Interest Group* 5, 38–46. doi: 10.1044/2019_PERSP-19-00083

Mendes, A., Lousada, M., Valente, A. R., and Hall, A. (2014). Validity and reliability of the European-Portuguese preschool language assessment ALPE. *Folia Phoniatr. Logop.* 66, 89–94. doi: 10.1159/000365354

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90, 91–117. doi: 10.1016/S0010-0277(03)00140-9

Misyak, J., and Christiansen, M. (2012). Statistical learning and language: an individual differences study. *Lang. Learn.* 62, 302–331. doi: 10.1111/j.1467-9922.2010.00626.x

Newport, E. L., and Aslin, R. N. (2004). Learning at a distance I. statistical learning of non-adjacent dependencies. *Cogn. Psychol.* 48, 127–162. doi: 10.1016/S0010-0285(03)00128-2

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., et al. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *J. Child Psychol. Psychiatry Allied Discip.* 57, 1247–1257. doi: 10.1111/jcpp.12573

Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., and Lum, J. A. (2016). Statistical learning in specific language impairment and autism spectrum disorder: a meta-analysis. *Front. Psychol.* 7:1245. doi: 10.3389/fpsyg.2016.01245

Perruchet, P., and Pacton, P. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn. Sci.* 10, 233–238. doi: 10.1016/j.tics.2006.03.006

Pierce, L. J., Tague, E. C., and Nelson, C. A. III. (2021). Maternal stress predicts neural responses during auditory statistical learning in 26-month-old children: an event-related potential study. *Cognition* 213:104600. doi: 10.1016/j.cognition.2021.104600

Plante, E., and Gómez, R. (2018). Learning without trying: The clinical relevance of statistical learning. *Lang. Speech Hear. Serv. Sch.* 49, 710–722. doi: 10.1044/2018_LSHSS-STLT1-17-0131

Quémart, P., and Maillart, C. (2016). The sensitivity of children with SLI to phonotactic probabilities during lexical access. *J. Commun. Disord.* 61, 48–59. doi: 10.1016/j.jcomdis.2016.03.005

Raven, J., Raven, J., and Court, H. (2009). *CPM-P, Matrizes Progressivas Coloridas (forma paralela) [Coloured Progressive Matrices (parallel version)]*. Edipsico - Edições e investigação em Psicologia, Lda.

Raviv, L., and Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Dev. Sci.* 21:e12593. doi: 10.1111/desc.12593

Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia* 51, 2026–2042. doi: 10.1016/j.neuropsychologia.2013.06.019

Reilly, S., Tomblin, B., Law, J., Mckean, C., Mensah, F., Morgan, A., et al. (2014). SLI: a convenient label for whom? *Int. J. Lang. Commun.* 49, 416–451. doi: 10.1111/1460-6984.12102

Riccio, C. A., Cash, D. L., and Cohen, M. J. (2007). Learning and memory performance of children with specific language impairment (SLI). *Appl. Neuropsychol.* 14, 255–261. doi: 10.1080/09084280701719203

Rice, M., and Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *J. Speech Lang. Hear. Res.* 39, 1239–1257. doi: 10.1044/jshr.3906.1239

Romberg, A. R., and Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 906–914. doi: 10.1002/wcs.78

Royle, P., and Courteau, E. (2014). "Language processing in children with specific language impairment: a review of event-related potential studies," in *Language Processing: New Research*. eds. L. T. Klein and V. Amato (Hauppauge, NY: Nova Science Publishers), 33–64.

Saffran, J. R. (2018). Statistical learning as a window into developmental disabilities. *J. Neurodev. Disord.* 10:35. doi: 10.1186/s11689-018-9252-y

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Dev. Perspect.* 14, 49–54. doi: 10.1111/cdep.1235

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928. doi: 10.1126/science.274.5294.1926

Saffran, J. R., and Estes, K. G. (2006). Mapping sound to meaning: connections between learning about sounds and learning about words. *Adv. Child Dev. Behav.* 34, 1–38. doi: 10.1016/S0065-2407(06)80003-0

Saffran, J. R., and Kirkham, N. Z. (2018). Infant statistical learning. *Annu. Rev. Psychol.* 69, 181–203. doi: 10.1146/annurevpsych-122216-011805

Sanders, L. D., Newport, E. L., and Neville, H. J. (2002). Segmentation nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci.* 5, 700–703. doi: 10.1038/nn873

Sawi, O. M., and Rueckl, J. (2019). Reading and the neurocognitive bases of statistical learning. *Sci. Stud. Read.* 23, 8–23. doi: 10.1080/10888438.2018.1457681

Shufaniya, A., and Arnon, I. (2018). Statistical learning is not age-invariant during childhood: performance improves with age across modality. *Cogn. Sci.* 42, 3100–3115. doi: 10.1111/cogs.12692

Siegelman, N. (2020). Statistical learning abilities and their relation to language. *Lang Ling Compass* 14:e12365. doi: 10.1111/lnc3.12365

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., and Frost, R. (2018). Linguistic entrenchment: prior knowledge impacts statistical learning performance. *Cognition* 177, 198–213. doi: 10.1016/j.cognition.2018.04.011

Siegelman, N., Bogaerts, L., and Frost, R. (2017). Measuring individual differences in statistical learning: current pitfalls and possible solutions. *Behav. Res. Methods* 49, 418–432. doi: 10.3758/s13428-016-0719-z

Siegelman, N., and Frost, R. (2015). Statistical learning as an individual ability: theoretical perspectives and empirical evidence. *J. Mem. Lang.* 81, 105–120. doi: 10.1016/j.jml.2015.02.001

Simões, M. R., Albuquerque, C. P., Pinho, M. S., Vilar, M., Pereira, M., Lopes, A. F., et al. (2016). *Bateria de Avaliação Neuropsicológica de Coimbra (BANC) [Coimbra Neuropsychological Assessment Battery]*. Lisboa: CEGOC-TEA.

Smalle, E. H., Bogaerts, L., Simonis, M., Duyck, W., Page, M., Edwards, M. G., et al. (2016). Can chunk size differences explain developmental changes in lexical learning? *Front. Psychol.* 6:1925. doi: 10.3389/fpsyg.2015.01925

Soares, A. P., França, T., Gutiérrez-Domínguez, F., and Oliveira, H. M. (2022a). As Trials Go By: Effects of 2-AFC Item Repetition on SL Performance for High- and Low-TP 'Words' Under Implicit and Explicit Conditions. School of Psychology, University of Minho.

Soares, A. P., Gutiérrez-Domínguez, F., Lages, A., Oliveira, H. M., Vasconcelos, M., and Jiménez, L. (2022b). Learning words while listening to syllables: electrophysiological correlates of statistical learning in children and adults. *Front. Hum. Neurosci.* 16:805723. doi: 10.3389/fnhum.2022.805723

Soares, A. P., Gutiérrez-Domínguez, F. J., Vasconcelos, M., Oliveira, H. M., Tomé, D., and Jiménez, L. (2020). Not all words are equally acquired: transitional probabilities and instructions affect the electrophysiological correlates of statistical learning. *Front. Hum. Neurosci.* 14:577991. doi: 10.3389/fnhum.2020.577991

Soares, A. P., Lages, A., Oliveira, H. M., and Gutiérrez-Domínguez, F. J. (2021a). "Extracting word-like Units when two concurrent regularities collide: electrophysiological evidence." in *Proceedings of 12th International Conference of Experimental Linguistics: ExLing Society*. ed. A. Botinis; October, 11–13 (Italy: Scuola Universitaria Superiore IUSS di Pavia), 215–218.

Soares, A. P., Lages, A., Oliveira, H. M., and Gutiérrez-Domínguez, F. J. (2022c). "Can explicit instructions enhance auditory statistical learning in children with developmental language disorder?" in *International Handbook of Clinical Psychology*. ed. G. Buela-Casal (New York: Thomson Reuters Editorial).

Soares, A. P., Lages, A., Silva, A., Comesaña, M., Sousa, I., Pinheiro, A. P., et al. (2019). Psycholinguistics variables in the visual-word recognition and pronunciation of European Portuguese words: a megastudy approach. *Lang. Cogn. Neurosci.* 34, 689–719. doi: 10.1080/23273798.2019.1578395

Soares, A. P., Lousada, M., and Ramalho, M. (2021b). ""Perturbação do desenvolvimento da linguagem: Terminologia, caracterização e implicações para os processos de alfabetização" [Language development disorder: Terminology, characterization and implications for literacy]," in *Alfabetização Baseada na Ciência [Science-Based Literacy]*. eds. R. A. Alves and I. Leite (Brasília: Ministério da Educação (MEC), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes)), 441–471.

Soares, A. P., Machado, J., Costa, A., Iriarte, A., Simões, A., Almeida, J. J., et al. (2015). On the advantages of frequency measures extracted from subtitles: The case of Portuguese. *Q. J. Exp. Psychol.* 68, 680–696. doi: 10.1080/17470218.2014.964271

Soares, A. P., Nunes, A., Martins, P., and Lousada, M. (2018). Do children with specific language impairment (SLI) present implicit learning (IL) deficits? Evidence from an artificial grammar learning (AGL) paradigm. *BMC Health Serv. Res.* 18, 1–183. doi: 10.1186/s12913-018-3444-8

Soares, A. P., Silva, R., Faria, F., Oliveira, H. M., and Jiménez, L. (2021c). Literacy effects on artificial grammar learning (AGL) with letters and colors: evidence from preschool and primary school children. *Lang. Cogn.* 13, 534–561. doi: 10.1017/langcog.2021.12

Spaulding, T. J., Plante, E., and Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: evidence for separate attentional capacities. *J. Speech Lang. Hear. Res.* 51, 16–34. doi: 10.1044/1092-4388(2008/002)

Spencer, M., Kaschak, M. P., Jones, J. L., and Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Read. Writ.* 28, 467–490. doi: 10.1007/s11145-014-9533-0

Tallal, P., Stark, R., and Mellits, D. (1985). The relationship between auditory temporal analysis and receptive language development: evidence from studies of developmental language disorder. *Neuropsychologia* 23, 527–534. doi: 10.1016/0028-3932(85)90006-5

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philos. Trans. R. Soc. B: Biol. Sci.* 372:20160056. doi: 10.1098/rstb.2016.0056

Thompson, S., and Newport, E. (2007). Statistical learning of syntax: The role of transitional probability. *Lang. Learn. Dev.* 3, 1–42. doi: 10.1207/s15473341lld0301_1

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., and O'brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *J. Speech Lang. Hear. Res.* 40, 1245–1260. doi: 10.1044/jslhr.4006.1245

Ullman, M., Earle, F., Walenski, M., and Janacsek, K. (2020). The neurocognition of developmental disorders of language. *Annu. Rev. Psychol.* 71, 389–417. doi: 10.1146/annurev-psych-122216-011555

Ullman, M. T., and Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex* 41, 399–433. doi: 10.1016/S0010-9452(08)70276-4

Ullman, M. T., and Pullman, M. Y. (2015). A compensatory role of declarative memory in neurodevelopmental disorders. *Neurosci. Biobehav. Rev.* 51, 205–222. doi: 10.1016/j.neubiorev.2015.01.008

van Witteloostuijn, M., Lammertink, I., Boersma, P., Wijnen, F., and Rispens, J. (2019). Assessing visual statistical learning in early-school-aged children: The usefulness of an online reaction time measure. *Front. Psychol.* 10:2051. doi: 10.3389/fpsyg.2019.02051

Viana, F. (2004). Teste de Identificação de Competências Linguísticas: Caderno de administração [Language Skills Identification Test: Administration Booklet]. Vila Nova de Gaia: Edipsico—Edições e investigação em Psicologia, Lda.

Vissers, C., Koolen, S., Hermans, D., Scheper, A., and Knoors, H. (2015). Executive functioning in preschoolers with specific language impairment. *Front. Psychol.* 6:1574. doi: 10.3389/fpsyg.2015.01574

West, G., Vadillo, M. A., Shanks, D. R., and Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Dev. Sci.* 21:e12552. doi: 10.1111/desc.12552

Zwart, F. S., Vissers, C. T. W., Kessels, R. P., and Maes, J. H. (2017). Procedural learning across the lifespan: a systematic review with implications for atypical development. *J. Neuropsychol.* 13, 149–182. doi: 10.1111/jnp.12139

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AV declared a shared affiliation with the author ML to the handling editor at the time of review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 01 July 2022
doi: 10.3389/fpsyg.2022.929589

Electronic Toys Decrease the Quantity and Lexical Diversity of Spoken Language Produced by Children With Autism Spectrum Disorder and Age-Matched Children With Typical Development

Courtney E. Venker and Jennifer R. Johnson*

Lingo Lab, Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, MI, United States

OPEN ACCESS

Edited by:
Marisa Filipe,
Faculdade de Letras da Universidade de Lisboa, Portugal

Reviewed by:
Asimenia Papoulidi,
University of West Attica, Greece
Bahia Guellai,
Université Paris Nanterre, France
Jorge Oliveira,
Lusophone University, Portugal

***Correspondence:**
Courtney E. Venker
cvenker@msu.edu

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *27 April 2022*
Accepted: *15 June 2022*
Published: *01 July 2022*

Citation:
Venker CE and Johnson JR (2022) Electronic Toys Decrease the Quantity and Lexical Diversity of Spoken Language Produced by Children With Autism Spectrum Disorder and Age-Matched Children With Typical Development. Front. Psychol. 13:929589. doi: 10.3389/fpsyg.2022.929589

Many young children with autism spectrum disorder (ASD) have language delays. Play-based interactions present a rich, naturalistic context for supporting language and communication development, but electronic toys may compromise the quality of play interactions. This study examined how electronic toys impact the quantity and lexical diversity of spoken language produced by children with ASD and age-matched children with typical development (TD), compared to traditional toys without electronic features. Twenty-eight parent-child dyads (14 per group) played with both electronic and traditional toy sets in a counter-balanced order. We transcribed child speech during both play sessions and derived the number of utterances and number of different word (NDW) roots per minute that children produced. Children with ASD and children with TD talked significantly less and produced significantly fewer unique words during electronic toy play than traditional toy play. In this way, children appear to take a "backseat" to electronic toys, decreasing their communicative contributions to play-based social interactions with their parents. These findings highlight the importance of understanding how toy type can affect parent-child play interactions and the subsequent learning opportunities that may be created. Play-based interventions for children with ASD may be most effective when they incorporate traditional toys, rather than electronic toys.

Keywords: autism (ASD), play, toys and games industry, language, intervention

INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social communication impairments, repetitive behaviors, and restricted interests that currently affects 1 in 44 children in the United States (American Psychiatric Association, 2013; Maenner et al., 2021). Impairments in structural language skills (i.e., vocabulary and grammar) are not required for a diagnosis of ASD (American Psychiatric Association, 2013). Nonetheless, many young children with ASD demonstrate severe delays in language development—lagging far behind their peers with typical development (TD) in both receptive and expressive (i.e., spoken) language skills (Charman et al., 2003; Tager-Flusberg et al., 2005; Ellis Weismer and Kover, 2015).

Many intervention approaches have been developed to address early language and communication delays in children with ASD. A growing number of autism interventions promote the creation of naturalistic, play-based interactions to facilitate children's language and communication development (Dawson et al., 2010; Sussman, 2012; Schreibman et al., 2015; Binns and Oram Cardy, 2019; Bruinsma et al., 2020; Rogers et al., 2020). Within these interventions, parents and caregivers are commonly taught to create play interactions that limit distractions and prioritize social communicative exchanges. In line with social interactionist and transactional theories of language development, the rationale is that play provides a developmentally appropriate social context for language learning that maximizes children's engagement and motivation and increases the likelihood that new skills will generalize to everyday settings (Sameroff and Fiese, 2000; Camarata and Yoder, 2002; Schreibman et al., 2015; Bruinsma et al., 2020; Bottema-Beutel and Kim, 2021).

Although play interactions have the potential to serve as an effective learning context for children with ASD, different types of toys may affect the quality of parent-child play interactions and the learning opportunities they provide (O'Brien and Nagle, 1987; Levin and Rosenquest, 2001; Miller et al., 2017). In recent years, electronic toys—toys that talk, sing, play music, and/or have flashing lights—have become increasingly common relative to traditional toys without technological features (Levin and Rosenquest, 2001; Radesky and Christakis, 2016). Contrary to marketing claims that electronic toys offer educational and developmental benefits (Levin and Rosenquest, 2001; Healey et al., 2019; Hassinger-Das et al., 2021; Zero to Three, 2021), numerous studies have shown that electronic toys decrease parent spoken language and responsiveness, compared to traditional toys (Wooldridge and Shapka, 2012; Zosh et al., 2015; Sosa, 2016; Miller et al., 2017; but see Sung, 2018).

Wooldridge and Shapka (2012) conducted an in-home study of parents playing with their typically developing young children (16–24 months old). Relative to traditional toys, electronic toys decreased the quality of parent behaviors associated with responsiveness and teaching. In a study of 10- to 16-month-old infants with TD, Sosa (2016) found that electronic toys were associated with fewer parent words, parent responses, and conversational turns, compared to traditional toys. Similarly, Zosh et al. (2015) found that parents of 24-month-old children with TD who played with electronic toys produced a significantly lower proportion of unique words than parents who played with traditional toys. Overall, these findings suggest that "parents tend to let the toys do the talking for them" (Sosa, 2016, p. 136) when playing with electronic toys, which may have detrimental effects on children's language development (also see Wooldridge and Shapka, 2012; Miller et al., 2017).

Though most research on electronic toys has focused on parents of children with TD, we recently conducted the first published study (Sturman et al., 2022) investigating how electronic toys affect play interactions between children with ASD (2–4 years old) and their parents, compared to traditional toys. We also included a group of children with TD of the same chronological age. Consistent with findings in TD, parents in both groups talked significantly less and produced a significantly fewer unique vocabulary words when playing with electronic toys than traditional toys. Electronic toys also elicited significantly more pause time than traditional toys. Overall, these findings closely align with prior research in suggesting that electronic toys reduce the quality and quantity of parent language input provided to young children.

Understanding the impact of electronic toys on parent spoken language is important, given robust evidence that child language outcomes are closely linked with the quality and quantity of parent language input they receive (Hart and Risley, 1995; Hoff and Naigles, 2002; Huttenlocher et al., 2010; Rowe, 2012; Adamson et al., 2020). However, our recent findings (Sturman et al., 2022) raise an important question: how do electronic toys affect *children's* spoken language, relative to traditional toys? Are differences in parent spoken language paralleled by differences in the spoken language produced by children with ASD or children with TD? Prior studies of children with TD have not investigated the impact of electronic toys on children's spoken language—likely because of the young age of their participants. However, there is evidence that infants with TD produce fewer directed vocalizations and gestures when playing with electronic toys than traditional toys (Miller et al., 2017; also see Sosa, 2016).

The goal of the current study was to determine how toy type (traditional vs. electronic) affects the quantity and lexical diversity of spoken language produced by children with ASD and age-matched children with TD (2–5 years old). Based on prior studies (Wooldridge and Shapka, 2012; Zosh et al., 2015; Sosa, 2016; Miller et al., 2017) and on our findings regarding parent spoken language, we hypothesized that the quantity and lexical diversity of spoken language would be significantly lower during electronic than traditional toy play in both groups.

MATERIALS AND METHODS

General Procedure

The study was approved by the Institutional Review Board at Michigan State University as part of a larger research project focused on language and visual attention in children with ASD (R21 DC 016102; Venker, PI). All parents provided written informed consent before participating. Parent-child dyads visited the lab on two separate days. They completed several activities related to language development, including standardized assessments and parent-child play sessions (described below).

Participants

Twenty-eight parent-child dyads participated ($n = 14$ with ASD, $n = 14$ with TD). Families were recruited through a university email listserv for parents and caregivers, flyers posted in the community, and word of mouth. All children in the ASD group had previously been diagnosed with ASD, per parent report. The Autism Diagnostic Observation Schedule, Second Edition

(ADOS-2; Lord et al., 2012a,b) was administered by a research-reliable examiner to confirm children's existing ASD diagnoses. Module selection was based on age and language level, as described in the ADOS-2 manual. Two children received the Toddler Module (for children 12–30 months old), five children received Module 1: Few to no words, three children received Module 1: Some words, two children received Module 2: Younger than 5, and two children received Module 2: 5 or older. The ADOS-2 also provided calibrated severity scores, which indicate overall autism severity.

Parents reported no developmental concerns for children in the TD group. All families of children with TD completed the Lifetime Form of the Social Communication Questionnaire (Rutter et al., 2003) and scored at or below the cutoff score of 15, which indicated no need for further ASD evaluation.

There were 14 mothers in the ASD group, and 11 mothers and 3 fathers in the TD group. Children in the ASD group (11 male, 3 female; 93% Caucasian, 7% Black or African American; 100% non-Hispanic) and children in the TD group (5 male, 9 female; 93% Caucasian, 7% more than one race; 14% Hispanic, 86% non-Hispanic) were 2–5 years old. A Wilcoxon signed-rank test revealed no significant difference in the mean age of the children with ASD and the children with TD ($p = 0.529$). A Fisher's Exact Test revealed that the proportion of males vs. females in the ASD and TD groups did not significantly differ ($p = 0.054$).

To assess receptive and expressive language abilities, we administered the Auditory Comprehension and Expressive Communication Scales of Pre-school-Language Scales, 5th edition (PLS-5; Zimmerman et al., 2011) to all participants. The PLS-5 provides an in-depth characterization of receptive and expressive language abilities, including vocabulary, grammar, literacy, and narrative skills. We assessed visual organization, memory, sequencing, and spatial awareness using the Visual Reception scale from the Mullen Scales of Early Learning (Mullen, 1995). The children with ASD had significantly lower scores than the children with TD on both the PLS-5 and the Mullen, indicating weaker language and cognitive skills (a topic we return to in the section "Discussion"; see **Table 1**). The number of children with ASD who scored 1.5 *SD* or more below the mean on the PLS-5 was 9/14 for the Expressive Communication Scale and 10/14 for the Auditory Comprehension Scale. Similarly, 10/14 children with ASD scored 1.5 *SD* or more below the mean on the Mullen Visual Reception Scale. In contrast, no child with TD scored more than 1 *SD* below the mean for either measure, indicating language and cognitive skills within the average range.

TABLE 1 | Child demographic information.

	ASD group	TD group
	Mean (SD) **range**	***Mean (SD)*** **range**
Chronological age (months)	43.5 (12.86) 26–71	46 (14.45) 25–67
PLS-5 AC standard score	63.71 (16.37) 50–98	116.71 (8.54) 106–130
PLS-5 AC percentile	6.79 (12.56) 1–45	83.5 (11.97) 66–98
PLS-5 AC age equivalent (in months)	23.57 (14.97) 13–60	58.21 (20.99) 31–95
PLS-5 EC standard score	70.43 (13.02) 50–93	118.79 (16.15) 96–148
PLS-5 EC percentile	6.85 (9.50) 1–32	80.21 (20.95) 39–99
PLS-5 EC age equivalent (in months)	24.43 (13.09) 9–59	58.07 (19.20) 33–95
Mullen VR T-score	27.71 (9.13) 20–46	61.86 (10.63) 42–80
Mullen VR age equivalent (in months)	29.00 (15.42) 14–66	53.29 (12.54) 27–69
ADOS-2 severity score	8.71 (1.33) 6–10	–

ASD, autism spectrum disorder; TD, typical development.
Groups differed significantly at p < 0.001 on all variables except chronological age (p = 0.633). PLS-5, pre-school language scale, 5th Edition; AC, auditory comprehension scale; EC, expressive communication scale; PLS-5, standard scores have a mean of 100 and SD of 15. Mullen VR = mullen scales of early learning, visual reception scale. Mullen t-scores have a mean of 50 and SD of 10. ADOS-2 severity scores = autism diagnostic observation schedule, 2nd Edition, calibrated severity score. Calibrated severity scores range from 1 to 10, with 1–2 indicating minimal-to-no evidence of autism spectrum-related symptoms, 3–4 indicating low evidence, 5–7 indicating moderate evidence, and 8–10 indicating a high level of autism spectrum-related symptoms.

Parent-Child Play Sessions

Parent-child dyads engaged in two, one-on-one play sessions in the lab for 10-min periods, with each session occurring on a different day. Play sessions took place in the laboratory setting, in a quiet room equipped with a table and chairs and a set of toys placed on the floor. Each dyad had the room to themselves. Parent were asked to play with their child as they normally would at home with the set of toys provided. Sessions were recorded using cameras placed around the room and an overhead microphone.

Each parent-child dyad played with the traditional toy set on 1 day and the electronic toy set on the other. The items in the toy sets were closely matched. Each set included a barn with animals, a shape sorter, spiky sensory balls, three vehicles, a puzzle, and a pull toy dog (see **Table 2**). The toys in the electronic toy set talked, sang, made music, and/or flashed lights. Each toy in the electronic toy set except the puzzle had flashing lights. Each toy in the electronic toy set except the sensory balls made sounds, talked, sang, and/or played music. The toys in the traditional toy set had no electronic features or technological enhancements.

We included a variety of toys in each set to increase the likelihood that each dyad would find one or more toys that interested them. We chose these toys because they are developmentally appropriate for children with a broad range of language and cognitive levels. In addition, the toys are representative of the types of toys commonly available for consumers to purchase online and in stores and have been used in previous studies (Wooldridge and Shapka, 2012; Zosh et al., 2015; Sosa, 2016; Miller et al., 2017). The toy sets were presented in counter-balanced order across participants. In the ASD group, the traditional toy set was provided first to eight parent-child

TABLE 2 | The toy sets.

	Electronic	Traditional
Barn with animals		
Shape sorter		
Animal puzzle		
Vehicles		
Pull toy dog		
Spiky sensory balls		

dyads and the electronic toy set was provided first to six dyads. In the TD group, 7 dyads played with the traditional toys first, and 7 dyads played with the electronic toys first.

In the ASD group, the average sample length was 10.26 min (SD = 0.40; range = 10.00–11.32) for traditional toy play and 10.50 min (SD = 0.96; range = 10.00–13.37) for electronic toy play. In the TD group, the average sample length for traditional toy play was 10.46 min (SD = 0.50; range = 10.0–11.52) and the average sample length for electronic toy play was 10.59 (SD = 0.57; range = 10.00–11.67). Dependent variables were represented as a rate (average count per minute) to account for slight variations in sample length.

Transcription

Research assistants used Systematic Analysis of Language Transcripts software (SALT; Miller and Iglesias, 2020) to transcribe the play sessions. Each transcriber completed a comprehensive online training program prior to coding independently. The transcription process involved a first pass by a primary transcriber, review and feedback from a secondary transcriber, and final discussion and consensus transcription by the pair. Coders were aware of the toy condition because the toys were visible (and audible, in the case of electronic toys; also see Sosa, 2016). Following standard SALT procedures, utterances were segmented based on communication units (each independent clause and its modifiers). We derived two variables from SALT that represented the quantity and lexical diversity of child spoken language. Variables were represented as a rate (average count per minute) to account for small variations in sample length. Quantity was measured by the number of child utterances per minute. Lexical diversity was measured by the number of different word (NDW) roots per minute that children produced. Only complete and intelligible child utterances were included in these calculations.

To evaluate inter-transcriber agreement, a separate primary and secondary transcriber independently re-transcribed the videos from 16 randomly selected play sessions (four Traditional and four Electronic videos for the TD and ASD groups). We then compared the number of child utterances and the NDW roots derived from each independent transcription. On average, the transcripts differed by two child utterances in the ASD group and by three child utterances in the TD group. On average, the transcripts differed by two different word roots in the ASD group and four word roots in the TD group. Thus, inter-transcriber outcomes for both of the key dependent variables were closely aligned, differing by no more than an average of 2–3 utterances and 3–4 different word roots.

Analysis Plan

This study involved a within-subject manipulation, wherein each parent-child dyad played with both electronic and traditional toys. Given this within-subject design, as well as the significant differences in language and cognitive skills between the ASD and TD groups, we conducted separate analyses for each group to determine whether the quantity and lexical diversity of children's spoken language differed by toy type. Some children with ASD produced very little (to no) spoken language. For this reason, we were more interested in which toy type elicited the most child spoken language than in the magnitude of these effects (which would be tested by parametric tests). Given this goal, as well as the relatively modest sample sizes, we analyzed the difference between toy types using Wilcoxon rank-sum tests, the non-parametric analog of a paired-samples *t*-test. We set alpha at 0.05. Because there was a clear prediction and expected direction of effect (i.e., that quantity and lexical diversity of child spoken language would be significantly higher during traditional than electronic toy play), we used 1-tailed tests.

RESULTS

The goal of this study was to test the impact of toy type (traditional vs. electronic) on the quantity and lexical diversity of spoken language produced by children with ASD and age-matched children with TD. To examine quantity, we compared the average number of utterances children in each group produced per minute during traditional and electronic toy play (see **Figure 1**). Children with ASD produced, on average, 3.05 utterances per minute during traditional toy play (median = 1.10, $SD = 3.55$, range = 0–9.61) and 2.21 utterances per minute during electronic toy play (median = 0.90, $SD = 2.88$, range = 0–7.90). Children with TD produced, on average, 7.74 utterances per minute during traditional toy play (median = 7.92, $SD = 2.37$, range = 3.84–12.66) and 5.29 utterances per minute during electronic toy play (median = 5.06, $SD = 2.34$, range = 1.30–8.29). Wilcoxon rank sum tests revealed that the mean number of child utterances per minute was significantly lower during electronic toy play than traditional toy play for both the children with ASD (1-tailed $p = 0.025$) and the children with TD (1-tailed $p = 0.004$).

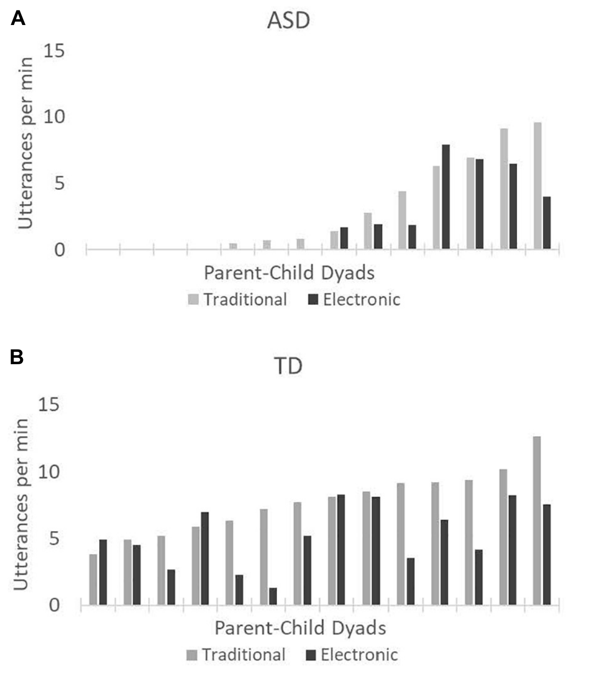

FIGURE 1 | Mean number of utterances per minute produced by the children with autism spectrum disorder (ASD) **(A)** and the children with typical development (TD) **(B)**. Gray bars represent traditional toy play, and black bars represent electronic toy play. Each pair of bars on the *x*-axis represents a single parent-child dyad. Data are presented left to right in order of increasing values for traditional toy play.

To examine lexical diversity, we compared the average number of different (i.e., unique) word roots children produced per minute during traditional and electronic toy play (see **Figure 2**). Children with ASD produced, on average, 2.90 unique words per minute during traditional toy play (median = 1.00, $SD = 3.61$, range = 0–9.75) and 2.06 unique words per minute during electronic toy play (median = 0.56, $SD = 2.85$, range = 0–7.36). Children with TD produced, on average, 9.66 unique words per minute during traditional toy play (median = 9.46, $SD = 3.14$, range = 3.40–15.45) and 7.27 unique words per minute during electronic toy play (median = 7.10, $SD = 3.21$, range = 2.40–12.51). Wilcoxon rank sum tests revealed that the mean number of unique words per minute was significantly lower during electronic toy play than traditional toy play for both the children with ASD (1-tailed $p = 0.021$) and the children with TD (1-tailed $p = 0.005$).

DISCUSSION

To our knowledge, this study provides the first evidence that electronic toys decrease the quantity and lexical diversity of children's spoken language, relative to traditional toys. Children with ASD and age-matched children with TD talked significantly less and produced significantly fewer unique words when playing with electronic toys than with traditional toys. Observations of the electronic play sessions indicated that the talking, singing, music, and animal sounds produced by the toys often left little room for children to contribute. The talking, singing, sounds, and flashing lights of the electronic toys dominated the interaction, interrupting children's utterances and decreasing the space available for parent-child communication. Although electronic toys are often advertised as educational (Levin and Rosenquest, 2001; Healey et al., 2019; Hassinger-Das et al., 2021; Zero to Three, 2021), the current findings add to growing evidence that electronic toys decrease the quality of play interactions between children and their parents (Zosh et al., 2015; Sosa, 2016; Sturman et al., 2022).

In combination with parallel findings in parents (Sturman et al., 2022), these results indicate that electronic toys limit the reciprocal linguistic exchanges between children and parents. This is unfortunate because electronic toys are not a substitute for meaningful communicative exchanges between parents and children. The value of play lies in shared engagement between play partners. When children are engaged in play that encourages linguistic interaction, they have more opportunities to initiate verbal interactions, experiment with grammatical forms, and participate in reciprocal conversational turns. In addition, a reduction in child spoken language limits a parent's ability to respond to and build on children's verbal communication, which is an important avenue for language development (Sameroff and Fiese, 2000; Camarata and Yoder, 2002). Though children with TD may be relatively unaffected by these changes, children

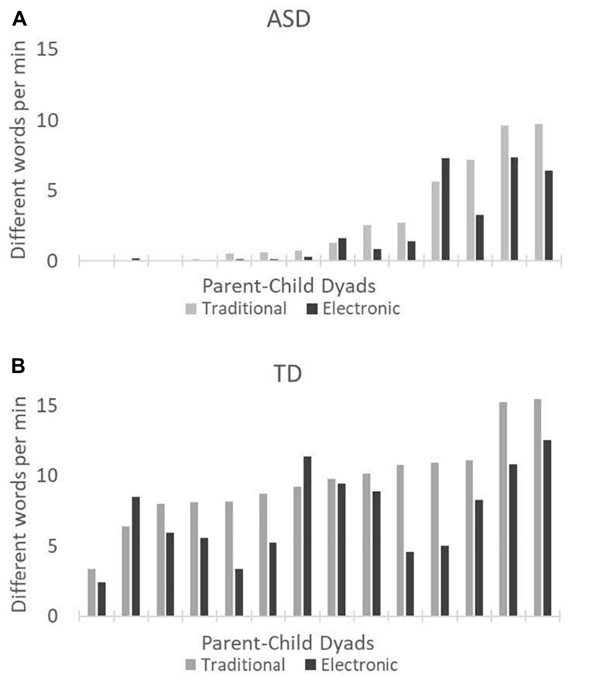

FIGURE 2 | Mean number of different (i.e., unique) words per minute produced by the children with autism spectrum disorder (ASD) **(A)** and the children with typical development (TD) **(B)**. Gray bars represent traditional toy play, and black bars represent the electronic toy play. Each pair of bars on the *x*-axis represents a single parent-child dyad. Data are presented left to right in order of increasing values for traditional toy play.

with ASD are likely to be vulnerable even to seemingly subtle disruptions in parent-child interactions. Electronic toys may compromise the potentially fragile play-based interactions that parents and caregivers create.

Examining the patterns of individual children offers additional insights into the decrease in children's spoken language during electronic toy play. For example, one child with ASD produced an average of four utterances per minute during electronic toy play and 10 utterances per minute during traditional toy play—a more than twofold increase that yielded approximately 40 utterances (electronic) vs. approximately 100 utterances (traditional) over the full 10-min play samples. In addition, three children with ASD produced only a single utterance during electronic toy play, but produced 5, 7, and 8 utterances, respectively, during traditional toy play. For children in the earliest stages of spoken language development, there is a clinically significant difference between producing 1 utterance in a 10-min play session vs. 5, 7, or 8 utterances. It is important to recognize that the play sessions in the current study lasted 10 min; differences between electronic and traditional toys may be even more dramatic as they accumulate over longer periods of time (Zosh et al., 2015).

Though our primary focus was on children with ASD, it is interesting to note that the age-matched children with TD also produced significantly fewer utterances and used significantly fewer unique words when playing with electronic toys than traditional toys. This was the case even though the children with TD had significantly stronger language skills than the children with ASD, suggesting that electronic toys decrease the quantity and lexical diversity of child spoken language regardless of developmental stage. In other words, the current findings suggest that the quantity and lexical diversity of a child's spoken language is likely lower during electronic than traditional toy play whether a child produces single words or 5-word utterances. These findings provide developmental continuity with prior findings that infants with TD produce fewer vocalizations and gestures during electronic toy play than traditional toy play (Sosa, 2016; Miller et al., 2017).

Additional research is needed to determine whether electronic toys disrupt play-based language learning opportunities in other ways. The background noise introduced by electronic toys may make it more difficult for children to understand spoken language, especially when it incorporates speech or other rhythmic sounds (Baker and Holding, 1993; Kirkorian et al., 2009; McMillan and Saffran, 2016; Erickson and Newman, 2017; Godwin et al., 2018; McAuley et al., 2020, 2021). In addition, the salient visual features of electronic toys, such as flashing lights, may compete with other relevant aspects of the child's visual environment (Radesky and Christakis, 2016). Visual salience exerts a strong influence on attention allocation in children with ASD (Sacrey et al., 2014; Venker et al., 2018, 2020). Salient auditory and visual features of electronic toys may decrease the likelihood that children with ASD will engage in joint attention and may cause them to miss important linguistic and social cues (Miller et al., 2017; Healey et al., 2019). These types of effects are important to investigate not only in lab settings, but also in naturalistic contexts, such as homes or classrooms.

An important next step in this line of work is to characterize the beliefs and attitudes of parents of children with ASD regarding toy selection. Parents of young children (without ASD) commonly consider electronic toys an essential teaching tool, with many parents viewing these toys as offering more educational value than themselves (Shah et al., 2018; Healey et al., 2019). Family members seeking to support language development in children with ASD may be particularly susceptible to the claims that electronic toys offer developmental benefits. Clinical practitioners have a responsibility to help parents become informed consumers—for example, by stressing to parents that they, not the toys, are the most important part of play interactions with their child (Wooldridge and Shapka, 2012; Hassinger-Das et al., 2021).

Though the current findings suggest that traditional toy play should be encouraged, it is not necessary (or realistic) to recommend a complete avoidance of electronic toys. Many children enjoy and are highly motivated by electronic toys, and they may be useful when encouraging children (particularly those with ASD) to request or comment on preferred items (Wooldridge and Shapka, 2012). Electronic toys may also facilitate social engagement and shared enjoyment by serving as a source of humor (Bergen et al., 2009). It may be beneficial for parents of children with ASD to make electronic toys available on a limited and purposeful basis, guided by advice from clinical professionals (Healey et al., 2019).

Limitations and Strengths

One limitation of the current study was the relatively small sample size ($n = 28$ parent-child dyads; $n = 14$ per group). Though small sample sizes limit statistical power, we consider the likelihood of replicating the current results to be high based on the robustness of the findings and their consistency with previous studies (also see Zosh et al., 2015; Sosa, 2016). In addition, the racial and ethnic diversity of the participant sample was limited, which may reduce the generalizability of the results. Another potential limitation is that the ASD and TD groups were matched on chronological age, rather than language or cognitive skills. It may be advantageous for future studies to include language-matched comparison groups. Future work focused on naturalistic contexts is also needed to complement lab-based studies like this one. Future studies may also examine more fine-grained patterns of interaction that unfold over the course of a play session. The current study also had several strengths. Its controlled, within-participants design allowed each parent-child dyad to serve as their own control, thereby removing potential confounds introduced by the unique interaction styles of individual dyads (Sosa, 2016). The toy sets were closely matched and included a variety of developmentally appropriate toys. Another strength was the rigorous manual transcription process, which involved a primary and secondary transcriber and consensus coding process.

CONCLUSION

The current findings indicate that electronic toys reduce the quantity and lexical diversity of spoken language produced by children with ASD and age-matched children with TD, thereby undermining play-based language learning opportunities. These findings add to growing empirical evidence that expensive, technologically enhanced toys are not necessary for young children's learning—and, in fact, may be detrimental. Play-based interventions for children with ASD may be most effective when they incorporate traditional toys, rather than electronic toys. These findings also make it possible for clinical practitioners to provide evidence-based recommendations about toy selection to families of children with ASD. Parents should be assured that no toy can take the place of a sensitive, engaged, responsive play partner. Well-targeted, sensitive recommendations will take individual parent and caregiver beliefs into account to ensure practitioners demonstrate respect for parents' efforts to help their children.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Michigan State University IRB. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

CV conceptualized the study, conducted the statistical analyses, and obtained funding for data collection. Both authors formulated the research question and helped write the manuscript.

FUNDING

This work was funded by the National Institute on Deafness and Other Communication Disorders (R21 DC016102; Venker, PI) and by the Center for Research in Autism, Intellectual, and Neurodevelopmental Disabilities (C-RAIND) at Michigan State University. Neither funder was involved in any aspect of the research itself beyond providing financial resources.

ACKNOWLEDGMENTS

We thank the families who took part in this study. We also thank the members of the Lingo Lab who helped with data collection and transcription.

REFERENCES

Adamson, L. B., Kaiser, A. P., Tamis-LaMonda, C. S., Owen, M. T., and Dimitrova, N. (2020). The developmental landscape of early parent-focused language intervention. *Early Child. Res. Q.* 50, 59–67. doi: 10.1016/j.ecresq.2018.11.005

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Washington, DC: American Psychiatric Association.

Baker, M. A., and Holding, D. H. (1993). The effects of noise and speech on cognitive task performance. *J. Gen. Psychol.* 120, 339–355. doi: 10.1080/00221309.1993.9711152

Bergen, D., Hutchinson, K., Nolan, J. T., and Weber, D. (2009). Effects of infant-parent play with a technology-enhanced toy: affordance-related actions and communicative interactions. *J. Res. Child. Educ.* 24, 1–17. doi: 10.1080/02568540903439342

Binns, A. V., and Oram Cardy, J. (2019). Developmental social pragmatic interventions for preschoolers with autism spectrum disorder: a systematic review. *Autism Dev. Lang. Impair.* 4:239694151882449. doi: 10.1177/2396941518824497

Bottema-Beutel, K., and Kim, S. Y. (2021). A systematic literature review of autism research on caregiver talk. *Autism Res.* 14, 432–449. doi: 10.1002/aur.2461

Bruinsma, Y., Minjarez, M. B., Schreibman, L., and Stahmer, A. C. (2020). *Naturalistic Developmental Behavioral Interventions for Autism Spectrum Disorder*. Baltimore, MD: Paul H. Brookes.

Camarata, S., and Yoder, P. (2002). Language transactions during development and intervention: theoretical implications for developmental neuroscience. *Int. J. Dev. Neurosci.* 20, 459–465. doi: 10.1016/S0736-5748(02)00044-8

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A., and Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *Int. J. Lang. Commun. Disord.* 38, 265–285. doi: 10.1080/136820310000104830

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., et al. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics* 125, e17–e23. doi: 10.1542/peds.2009-0958

Ellis Weismer, S., and Kover, S. T. (2015). Preschool language variation, growth, and predictors in children on the autism spectrum. *J. Child Psychol. Psychiatry* 56, 1327–1337. doi: 10.1111/jcpp.12406

Erickson, L. C., and Newman, R. S. (2017). Influences of background noise on infants and children. *Curr. Dir. Psychol. Sci.* 26, 451–457. doi: 10.1177/0963721417709087

Godwin, K. E., Erickson, L. C., and Newman, R. S. (2018). Insights from crossing research silos on visual and auditory attention. *Curr. Dir. Psychol. Sci.* 28, 47–52. doi: 10.1177/0963721418807725

Hart, B., and Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, MD: Paul H. Brookes.

Hassinger-Das, B., Quinones, A., DiFlorio, C., Schwartz, R., Talla Takoukam, N. C., Salerno, M., et al. (2021). Looking deeper into the toy box: understanding caregiver toy selection decisions. *Infant Behav. Dev.* 62:101529. doi: 10.1016/j.infbeh.2021.101529

Healey, A., Mendelsohn, A., and Council on Early Childhood (2019). Selecting appropriate toys for young children in the digital era. *Pediatrics* 143:e20183348. doi: 10.1542/peds.2018-3348

Hoff, E., and Naigles, L. (2002). How children use input to acquire a lexicon. *Child Dev.* 73, 418–433.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., and Hedges, L. V. (2010). Sources of variability in children's language growth. *Cogn. Psychol.* 61, 343–365. doi: 10.1016/j.cogpsych.2010.08.002

Kirkorian, H. L., Pempek, T. A., Murphy, L. A., Schmidt, M. E., and Anderson, D. R. (2009). The impact of background television on parent-child interaction. *Child Dev.* 80, 1350–1359. doi: 10.1111/j.1467-8624.2009.01337.x

Levin, D. E., and Rosenquest, B. (2001). The increasing role of electronic toys in the lives of infants and toddlers: should we be concerned? *Contemp. Issues Early Child.* 2, 242–247. doi: 10.2304/ciec.2001.2.2.9

Lord, C., Luyster, R. J., Gotham, K., and Guthrie, W. (2012a). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 2): Toddler Module*. Torrence, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., and Bishop, S. (2012b). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 1): Modules 1 – 4*. Torrence, CA: Western Psychological Services.

Maenner, M. J., Shaw, K. A., Bakian, A. V., Bilder, D. A., Durkin, M. S., Esler, A., et al. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveill. Summ.* 70, 1–16. doi: 10.15585/MMWR.SS7011A1

McAuley, J. D., Shen, Y., Dec, S., and Kidd, G. R. (2020). Altering the rhythm of target and background talkers differentially affects speech understanding. *Attent. Percept. Psychophys.* 82, 3222–3233. doi: 10.3758/s13414-020-02064-5

McAuley, J. D., Shen, Y., Smith, T., and Kidd, G. R. (2021). Effects of speech-rhythm disruption on selective listening with a single background talker. *Attent. Percept. Psychophys.* 83, 2229–2240. doi: 10.3758/s13414-021-02298-x

McMillan, B. T. M., and Saffran, J. R. (2016). Learning in complex environments: the effects of background speech on early word learning. *Child Dev.* 87, 1841–1855. doi: 10.1111/cdev.12559

Miller, J., and Iglesias, A. (2020). *Systematic Analysis of Language Transcripts (SALT)*. Madison, WI: SALT Software, LLC.

Miller, J. L., Lossia, A., Suarez-Rivera, C., and Gros-Louis, J. (2017). Toys that squeak: toy type impacts quality and quantity of parent–child interactions. *First Lang.* 37, 630–647. doi: 10.1177/0142723717714947

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Minneapolis, MN: AGS edition ed.

O'Brien, M., and Nagle, K. J. (1987). Parents' speech to toddlers: the effect of play context. *J. Child Lang.* 14, 269–279. doi: 10.1017/S0305000900012927

Radesky, J. S., and Christakis, D. A. (2016). Keeping children's attention the problem with bells and whistles. *JAMA Pediatr.* 170, 112–113. doi: 10.1001/jamapediatrics.2015.3877

Rogers, S. J., Yoder, P., Estes, A., Warren, Z., McEachin, J., Munson, J., et al. (2020). A multisite randomized controlled trial comparing the effects of intervention intensity and intervention style on outcomes for young children with autism. *J. Am. Acad. Child Adolesc. Psychiatry* 60, 710–722.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Dev.* 83, 1762–1774. doi: 10.1111/j.1467-8624.2012.01805.x

Rutter, M., Bailey, A., and Lord, C. (2003). *The Social Communication Quesionnaire*. Los Angeles, CA: Western Psychological Services.

Sacrey, L.-A. R., Armstrong, V. L., Bryson, S. E., and Zwaigenbaum, L. (2014). Impairments to visual disengagement in autism spectrum disorder: a review of experimental studies from infancy to adulthood. *Neurosci. Biobehav. Rev.* 47, 559–577. doi: 10.1016/j.neubiorev.2014.10.011

Sameroff, A. J., and Fiese, B. H. (2000). "Transactional regulation: the developmental ecology of early intervention," in *Handbook of Early Childhood Intervention*, eds J. P. Shonkoff and S. J. Meisels (Cambridge: Cambridge University Press), 734.

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., et al. (2015). Naturalistic developmental behavioral interventions: empirically validated treatments for autism spectrum disorder. *J. Autism Dev. Disord.* 45, 2411–2428. doi: 10.1007/s10803-015-2407-8

Shah, R., Gustafson, E., and Atkins, M. (2018). Parental attitudes and beliefs surrounding play among predominantly low-income urban families: a qualitative study. *J. Dev. Behav. Pediatr.* 176, 139–148. doi: 10.1097/DBP.0000000000000708.Parental

Sosa, A. V. (2016). Association of the type of toy used during play with the quantity and quality of parent-infant communication. *JAMA Pediatr.* 170, 132–137. doi: 10.1001/jamapediatrics.2015.3753

Sturman, M., Peffers, K., Johnson, J. R., and Venker, C. E. (2022). Association between toy type and parent language input provided to children with autism spectrum disorder and age-matched children with typical development. *JAMA Pediatr.* e221063. doi: 10.1001/jamapediatrics.2022.1063

Sung, J. (2018). How young children and their mothers experience two different types of toys: a traditional stuffed toy versus an animated digital toy. *Child Youth Care Forum* 47, 233–257. doi: 10.1007/s10566-017-9428-8

Sussman, F. (2012). *More Than Words: A Parent's Guide to Building Interaction and Language Skills for Children with Autism Spectrum Disorder or Social Communication Difficulties*. Toronto, Ont: The Hanen Centre.

Tager-Flusberg, H., Paul, R., and Lord, C. (2005). "Language and communication in autism," in *Handbook of Autism and Pervasive Developmental Disorders*, 3rd Edn. eds F. R. Volkmar, R. Paul, A. Klin, and D. Cohen (Hoboken, NJ: Wiley and Sons), 335–364.

Venker, C. E., Bean, A., and Kover, S. T. (2018). Auditory-visual misalignment: a theoretical perspective on vocabulary delays in children with ASD. *Autism Res.* 11, 1621–1628. doi: 10.1002/aur.2038

Venker, C. E., Mathée, J., Neumann, D., Edwards, J., Saffran, J., and Ellis Weismer, S. (2020). Competing perceptual salience in a visual word recognition task differentially affects children with and without autism spectrum disorder. *Autism Res.* 14, 1147–1162. doi: 10.1002/aur.2457

Wooldridge, M. B., and Shapka, J. (2012). Playing with technology: mother–toddler interaction scores lower during play with electronic toys. *J. Appl. Dev. Psychol.* 33, 211–218. doi: 10.1016/J.APPDEV.2012.05.005

Zero to Three (2021). *Tips for Choosing Toys for Toddlers*. Available online at: https://www.zerotothree.org/resources/1076-tips-for-choosing-toys-for-toddlers (accessed July 14, 2021).

Zimmerman, I. L., Steiner, V. G., and Pond, R. E. (2011). *Preschool Language Scales*, 5th Edn. San Antonio, TX: The Psychological Corporation.

Zosh, J. M., Verdine, B. N., Filipowicz, A., Golinkoff, R. M., Hirsh-Pasek, K., and Newcombe, N. S. (2015). Talking shape: parental language with electronic versus traditional shape sorters. *Mind Brain Educ.* 9, 136–144. doi: 10.1111/mbe.12082

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

TYPE Systematic Review
PUBLISHED 27 July 2022
DOI 10.3389/fpsyg.2022.930275

OPEN ACCESS

EDITED BY
Lénia Carvalhais,
Infante D. Henrique Portucalense University, Portugal

REVIEWED BY
Neelima Wagley,
Vanderbilt University, United States
Isaías Martín-Ruiz,
Universidad de Málaga, Spain

*CORRESPONDENCE
Ana Paula Vale
pvale@utad.pt

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 27 April 2022
ACCEPTED 06 July 2022
PUBLISHED 27 July 2022

CITATION
Vale AP, Fernandes C and Cardoso S (2022) Word reading skills in autism spectrum disorder: A systematic review. *Front. Psychol.* 13:930275. doi: 10.3389/fpsyg.2022.930275

Word reading skills in autism spectrum disorder: A systematic review

Ana Paula Vale[1]*, Carina Fernandes[2] and Susana Cardoso[1]

[1]Dyslexia Unit, Department of Education and Psychology, School of Social and Human Sciences, University of Trás-os-Montes e Alto Douro, Vila Real, Portugal, [2]Laboratory of Neuropsychophysiology, Faculty of Psychology and Education Sciences, University of Porto, Porto, Portugal

A growing body of research suggests that children with autism spectrum disorder (ASD) are at risk of reading and learning difficulties. However, there is mixed evidence on their weaknesses in different reading components, and little is known about how reading skills characterize in ASD. Thereby, the current study aimed to systematically review the research investigating this function in children with ASD. To this purpose, we reviewed 24 studies that compared (1) children with ASD and children with typical development (TD) in word and nonword reading performance, (2) children with ASD and normative data of word and nonword reading tests, and (3) the results obtained by children with ASD in word and nonword reading tests. Most of the comparisons (62%) contrasting the reading performance of children with ASD and children with TD did not find significant differences between groups in both word and nonword reading. However, all the comparisons that reported standardized results showed that children with ASD had scores that fell within population norms. Regarding the third comparison of interest, about 54% of the studies presented data for both word and nonword reading, but only one study tested the difference between them and showed that children with ASD had higher levels of word than of nonword reading. Despite these results, the heterogeneous and small samples do not allow to draw sound conclusions regarding the strategies that children with ASD use to read words. As consequence, the nature of reading difficulties presented by children with ASD are still unknown, requiring future research conducted with larger and well-characterized samples of ASD and TD, using homogeneous specific tasks designed to assess word reading strategies.

KEYWORDS

autism, word reading strategies, decoding, word recognition, methodological features

Introduction

A majority of students with autism spectrum disorder (ASD) has reading difficulties (Ricketts et al., 2013; McIntyre et al., 2017b; Solari et al., 2019). Indeed, there is evidence that even when children and adolescents with ASD perform at the normative range of general cognitive ability and are verbally able, only a small percentage (31.2%) has average reading scores. This data contrasts with what is generally shown in typically developing (TD) peers of similar general cognitive ability, in which at least 80% achieve average scores (Solari et al., 2019). This indicates that a large number of ASD students are not responding well to reading instruction and/or reading instruction may not be well-designed to enable their reading success. One way or another, this argues for the need to understand how children with ASD deal with reading.

Autism spectrum disorder is a neurodevelopmental disorder with early-onset described by some degree of impairment in social interaction and communication. Other characteristics include restricted and repetitive behaviors and interests. Crucially, these features have a detrimental impact on daily life (American Psychiatric Association, 2013; World Health Organization, 2019). Recent ASD prevalence estimates of nationwide and across countries accounts range from 0.8 to 1.5% of school-age children (Lyall et al., 2017; ASDEU, 2018; Baio et al., 2018; Ofner et al., 2018) reported by the Centers for Disease Control and Prevention surveillance network for autism in the USA, being that two-thirds of 8-years old children with ASD did not present intellectual disability (Maenner et al., 2021). Thus, it is conceivable that most children with ASD may be enrolled in core curriculum educational programs to learn how to read (Fleury et al., 2014). Reading has a critical role in adapting to current and future academic, cognitive, and social needs and challenges (Lyon, 2001; Maughan et al., 2020).

Many studies have examined reading skills in ASD (Davidson and Weismer, 2014; Dynia et al., 2014; Bednarz et al., 2017; Nally et al., 2018; Micai et al., 2021). Among studies of reading skills and development in ASD, the majority paid particular attention to children's difficulties in reading comprehension, that is, their struggles to obtain meaning from written passages or texts (Nation et al., 2006; Brown et al., 2013; Ricketts et al., 2013; McIntyre et al., 2017a; Solari et al., 2019). It is often assumed that reading comprehension difficulties of children with ASD derive from oral language limitations (e.g., Nation and Norbury, 2005; Huemer and Mann, 2010; Ricketts, 2011; El Zein et al., 2014; Singh et al., 2021). This has guiding many to assume that readers with ASD have a hyperlexic profile (Fernandes et al., 2015; Ostrolenk et al., 2017; Duncan et al., 2021; Macdonald et al., 2021). That is, they usually present low levels of reading comprehension along with good abilities of word reading. However, it may not be the case (Henderson et al., 2014; Solari et al., 2019; Macdonald et al., 2021). To fully understand the obstacles that children with ASD face to achieve efficient comprehension of written information there is a need to better examine word reading processes in ASD, as much as reading comprehension can be conceived as essentially the same as language comprehension in written format (Hoover and Tunmer, 2020).

According to the Simple View of Reading (SVR, Gough and Tunmer, 1986; Hoover and Tunmer, 2020), a framework of the cognitive capacities needed for reading with wide empirical support (Fernandes et al., 2017a,b; Lonigan et al., 2018; Nation, 2019; Verhoeven et al., 2019; Kim et al., 2021), reading comprehension is the product of two sets of skills: word reading and oral language skills. This means that neither of these two components is sufficient *per se* for achieving reading comprehension and if one of them is somehow failing, reading comprehension will fail too. Thus, SVR establishes that reading difficulties may be dependent on word reading problems, language comprehension problems, or both (Hoover and Gough, 1990).

A recent meta-analysis (Duncan et al., 2021) aiming to clarify the role of these two skills in reading comprehension in ASD computed data of 26 studies that included both a measure of word reading and reading comprehension. Their analyses showed that each of the SVR components made a similar size contribution to the statistical model, demonstrating that word reading is as critical as the oral language to achieve reading comprehension for children with ASD. Other studies had previously shown that word reading has an important role in reading comprehension in ASD (Brown et al., 2013; Ricketts et al., 2013). For example, Brown et al.'s (2013) meta-analysis reported that, although children with ASD had word reading standard scores within the average range, which were better than their weak oral language scores, word reading explained a comparable amount of unique variance as the oral language (57%) in reading comprehension. Moreover, word reading was strongly associated with reading comprehension ($r = 0.77$, $n = 1{,}469$).

These findings concur with emerging evidence indicating that there is more than one type of reading profile among children and adolescents with ASD. For instance, Solari et al.'s (2019) found four different reading profiles: (a) a group with average word reading and reading comprehension (18%); (b) a group with specific difficulties in reading comprehension (poor comprehenders; 24%); (c) a group with low scores in both word reading and reading comprehension but good receptive vocabulary (23.6%) and (d) a group with a profile of generalized low scores in word reading, reading comprehension and oral language (34.3%). Interestingly, a former study (Henderson et al., 2014) similarly found that only 24.5% of their sample of children and adolescents with ASD could be characterized as poor reading comprehenders, presenting word reading accuracy on the average range, reading comprehension below a standard score of 89 and a discrepancy of at least 1 standard deviation (SD) between the two. In this study, 57% scored more than 1 SD, and 31% more than 2 SD, below the mean on word reading. These and other studies (Nation et al., 2006; White et al., 2006; Jonhels and Sandberg, 2012; McIntyre et al., 2017a) point out that, contrary to the widespread oral language deficit only explanation for reading comprehension problems (Singh et al., 2021), many children and adolescents with ASD have word reading difficulties.

Since the majority of studies on reading in ASD have been designed to address children's reading comprehension difficulties, many enrolled children with average to high levels of word reading (Engel, 2018; Ibrahim, 2020; Macdonald et al., 2022) and did not examine specific effects on single-word reading, such as frequency, orthographic consistency, and length of words. This trend contributed to a scarce knowledge about how children with ASD process written words.

One of the key hypotheses regarding word reading is that its development evolves through the emergence of two broad mechanisms described by the main theories of skilled reading. Essentially, the three most acknowledge computational models of fluent reading—Dual route cascaded model (DRC; Coltheart et al., 2001), the Triangle model (Plaut et al., 1996; Harm and Seidenberg, 2004), and the Connectionist Dual Process model (CDP++; Perry et al., 2010)—agree on a need for two-pathways to read words, independent of the particular orthography to be learnt: a direct, lexical process that merge the words spellings to the meanings usually preferred for familiar words and irregular words (such as *have, come*, and *eye,* that cannot be correctly read using only grapheme - phoneme conversions); and a sub-lexical indirect way *via* the phonological serial conversion of each grapheme into a phoneme, to obtain a pronunciation and then the word meaning, used mostly for new and low frequency words.

Thus, in the same vein, single-word reading is considered to involve at least two main cognitive mechanisms: decoding and recognition (Castles et al., 2018; Miles and Ehri, 2019). Decoding is employed when children use a phonological/alphabetic approach in which a pattern of grapheme—phoneme correspondences are assembled sequentially to sound out a word. This mechanism requires conscious cognitive effort and, consequently, it is also a time-consuming procedure. Recognition, on the other hand, is an almost effortless automatic process of accurately matching a written word with an orthographic pattern stored in long-term memory combined with phonological and semantic information (Miles and Ehri, 2019), called orthographic strategy. As Miles and Ehri (2019) detailed, this strategy does not equate with visual memory processes. Instead, it depends on a tuned representation of the specific string of amalgamated grapho-phonemic structures composing a word that draws on high levels of orthographic knowledge.

At the beginning of reading acquisition TD children rely predominantly on the alphabetic/phonological strategy to decode most of the words they encounter. Gradually, with further instruction and great amounts of exposure they gain sophisticated knowledge about the specificities of the orthographic system functioning and eventually achieve the automaticity that characterizes written word recognition (Castles et al., 2018; Miles and Ehri, 2019). Importantly, decoding is deemed to be a crucial skill to develop written word recognition (Share, 1995). Thus, the strategies children use to read single words adjust to their reading ability (Ehri, 2013).

Data on word reading strategies of children with ASD is very scarce and offers mixed evidence on the relative strengths and weaknesses of word reading skills among these readers.

Some small-scale studies (Frith and Snowling, 1983; Minshew et al., 1995) found that children with ASD who had word reading levels within the expected range for their age were also able of decoding nonwords (a string of letters, such as *slint*, that do not exist in the lexicon and thus is virtually independent of the memory for individual words, requiring decoding skills in order to be read). Later studies have also found good levels of nonword reading in groups of school-age children with ASD (Gabig, 2010; McIntyre et al., 2017b). On the contrary, Nation et al. (2006) noted that many ASD children had difficulties when reading nonwords and White et al. (2006) reported that more than half of their sample of children with ASD presented word decoding difficulties and poor phonological awareness, a skill that enables isolating each phoneme in a word and that is robustly related to word decoding (Melby-Lervåg et al., 2012). More recently, Henderson et al. (2014) found that although word and nonword reading were strongly correlated in a group of children with ASD, nonword reading scores were significantly lower for the ASD group than for a group of their TD peers matched by word reading level. Also Westerveld et al. (2018) found significant floor effects for nonword reading but not for words in first graders with ASD. Thus, because nonwords reading is thought to require the use of the sub-lexical indirect route according to the computational theories of reading mentioned above these results suggests that for many ASD children the indirect/phonological path for reading may present some degree of dysfunctionality and decoding appears to be a challenging task, being unclear how they read unfamiliar words.

Considering the above-mentioned results, we may argue that children with ASD that achieve typical scores on word reading may not be using phonological strategies but instead a direct access procedure based on their visual memory, possibly supported by intact or enhanced associative learning mechanisms (Walenski et al., 2008) and/or enhanced processing of broad visual aspects of written material (Samson et al., 2012; Ostrolenk et al., 2017) along with a detail-focused style of cognitive processing (Happé and Frith, 2006) that may favors word patterns recognition. In line with this, Macdonald et al. (2021) observed preschool children with ASD and hyperlexia exhibiting advanced word reading and letter naming in tandem with low phonological awareness and letter-sound correspondence skills. Other studies (Hooper et al., 2006; White et al., 2006; Gabig, 2010; Jonhels and Sandberg, 2012) have also reported that children with ASD showed poorer phonemic awareness than their age-matched peers, conflicting with Frith and Snowling (1983) findings.

Together, these results suggest that children with ASD may be employing their own strategies to process word reading. However, Cardoso-Martins et al. (2015) reported that a group of Brazilian Portuguese speaking children with ASD, varying considerably in nonverbal intelligence and word reading ability, did not differ from their TD colleagues, matched for word reading accuracy, on nonword reading. In addition, the ASD group presented an equivalent reading accuracy in word and nonword reading. Also, likewise to their TD peers, word reading was strongly correlated with nonword reading for the ASD group. The authors argued that participants with ASD used a similar phonological-based sub-lexical strategy as their peers for reading, which contrasts with the formerly mentioned evidence. In face of this results and in consonance with the Psycholinguistic Grain Size Theory of Reading (Ziegler and Goswami, 2005) mentioned by the authors, we could reason that, since the match between letters and sounds is more fixed in the Portuguese orthography than in the English one, the learning and use of the sub-lexical indirect-phonological route could be easier in Portuguese than in English (Duncan et al., 2013) that often requires reliance on orthographic structures larger than single letters to achieve word reading. Although this was never tested with ASD children, it may not be the unique explanation for the Brazilian results. As a matter of fact, Frith and Snowling (1983) reported a pattern of results similar to those of Cardoso-Martins et al. (2015), showing that ASD English children did not differ from their TD peers on using the sub-lexical procedure better than the lexical one. Thus, it seems that there is a number of discordant findings regarding word reading skills of ASD children challenging the reaching of a coherent description.

A robust predictor of reading automaticity and therefore word recognition (Landerl et al., 2019) is Rapid Naming (RAN). RAN, a task requiring the serial naming of arrays of familiar pictures of objects or colors or letters in a speedy manner, is supposed to involve, like reading, the lexical retrieval of familiar phonological sequences. While there is consistent evidence that children with ASD perform more poorly than their peers presenting longer naming times (White et al., 2006; Gabig, 2010; McIntyre et al., 2017b; Nayar et al., 2021), RAN was shown to be significantly associated with word reading fluency, but not accuracy (Johnels et al., 2021). This suggests that many children with ASD could experience difficulty in building word reading automaticity; that is, difficulty in using the direct/lexical procedure hypothesized by the computational theories of reading (Plaut et al., 1996; Coltheart et al., 2001; Perry et al., 2010).

Although scarcely, other psycholinguistic effects were examined in ASD reading studies. For instance, Welsh et al. (1987) showed a significant frequency effect indicating that ASD children had grew a lexicon for written words and used it successfully. Still, the same children read regular words better than irregular ones suggesting that they were applying grapheme-phoneme conversion rules more effectively than using lexical orthographic knowledge. Earlier work (Frith and Snowling, 1983) had reported the same pattern of results, showing that children with ASD presented an advantage of regular words both in accuracy and time scores. According with the computational theories of reading mentioned above, these frequency and regularity effects point out that some ASD children can use both direct/lexical and indirect/phonological sub-lexical routes to read words. Yet, because the studies have a small number of participants and have large age ranges, this body of findings does not clearly elucidate about the word reading skills of ASD children.

Thus, studies on reading acquisition and development in ASD offer divergent evidence of what might be the word reading strategies of those children. However, knowing how children with ASD read words and, complementarily, identifying the challenges they meet in that endeavor is vital to assist in teaching them to read and in designing solid remediation interventions when they are needed. Furthermore, as word reading is necessary for reading comprehension that, in turn, is determinant to progress in other academic subjects and to increase the knowledge of the world (Hoover and Tunmer, 2020), it is critical to systematize what is the current knowledge concerning the word reading processes used by children with ASD.

Methods

Systematic search strategy

This review was performed according to the actualized Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). Articles published until December 2021 were selected from PubMed, Web of Science, and EBSCOhost (including the Academic Search Complete, Psychology and Behavioral Sciences Collection, CINAHL Plus with Full Text, Fonte Acadêmica, MedicLatina, PsycARTICLES, PsycBOOKS, and PsycINFO databases). The search expression was "(autis* OR ASD) AND (read* OR literacy OR "word decoding" OR "word recognition"). In addition, we screened the reference list of reviews in this field and of all included studies.

Selection criteria

We included experimental and quasi-experimental studies that have assessed word reading abilities in children with autism spectrum disorders (ASD). In this study, we included children aged up to 12 years. Considering that at the age of 6 reading skills are becoming reasonably well-established (Nation et al., 2006; Henderson et al., 2014), an upper limit of 12 years was chosen to avoid ceiling effects.

After being included for reporting research in the topic of the review, articles were excluded according to the following criteria: (a) articles without a group of children with ASD (children with ASD mixed with other diagnosis, children

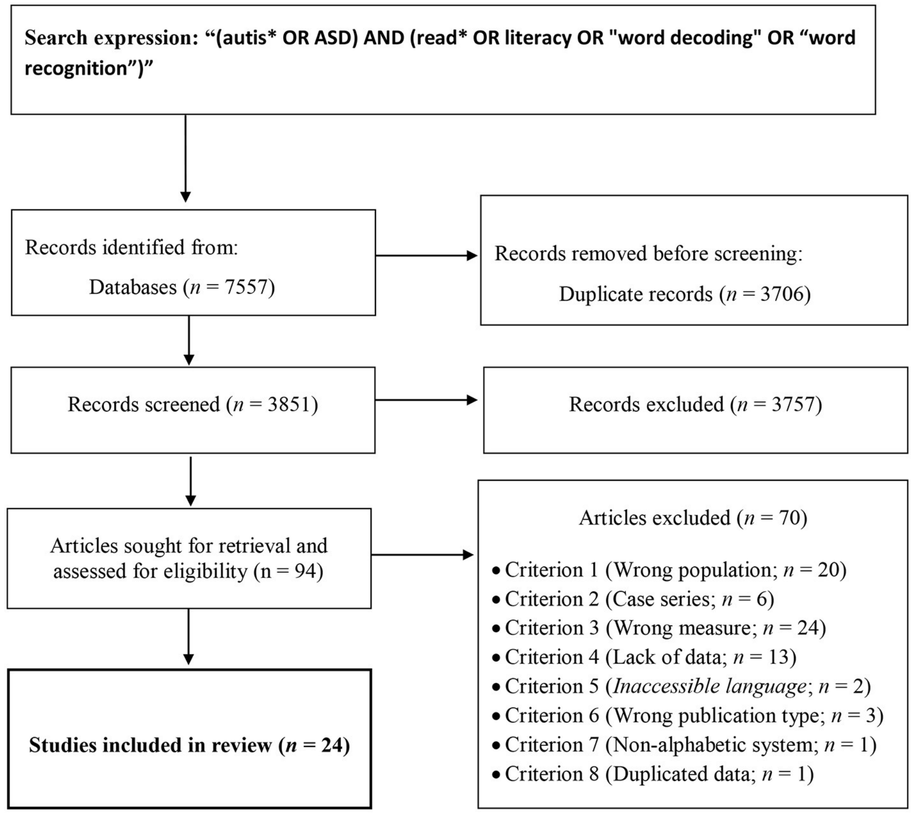

FIGURE 1
Flow diagram illustrating the systematic search, results and the selection of the studies included in this systematic review.

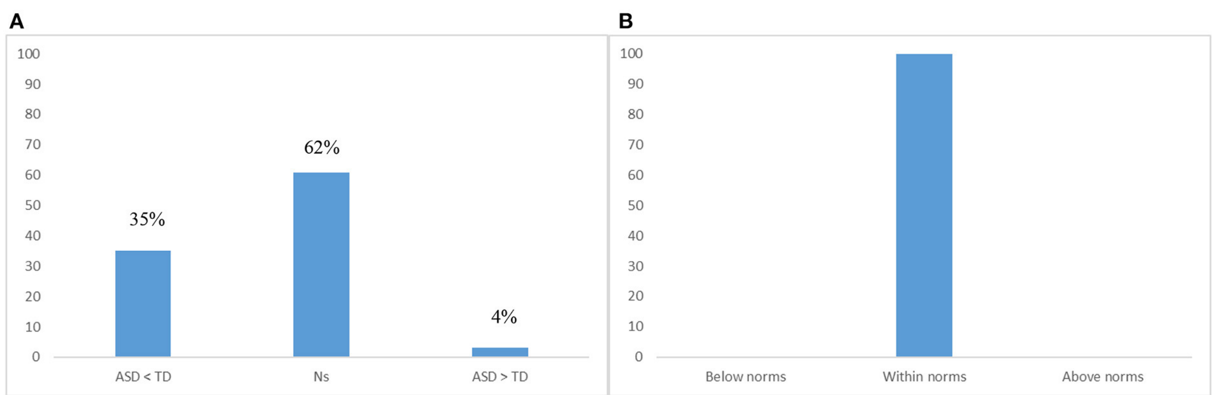

FIGURE 2
(A) Percentage of tests or subtests that compared children with autism spectrum disorders (ASD) and children with typical development (TD); **(B)** Percentage of tests or subtests that compared children with ASD with normative scores or percentiles.

with typical development or participants with diagnosis of ASD with mean age > 12 years old); (criterion 1: wrong population); (b) articles <10 ASD participants (criterion 2: case series); (c) articles that did not assess word reading abilities (criterion 3: wrong measure); (d) inaccessible articles or studies without information about word reading abilities in children with ASD (criterion 4: lack of data); (e) articles published in other languages than English, Portuguese, Spanish, or French (criterion 5: inaccessible language); (f) other orthographic system rather than alphabetic (criterion 6: non-alphabetic system) and (g) abstracts, reviews, commentaries, or methods (criterion 7: wrong publication type). Articles reporting only duplicated data were also excluded, and when articles reported an expansion of previously conducted research, data from the most recent article were selected (criterion 8: duplicated data).

Screening and selection of records

The results of the literature search were compiled on Rayyan QCRI (Ouzzani et al., 2016). On this platform, two researchers blindly screened the titles and abstracts, excluded the articles out of topic, and retained the remaining studies. When this task was completed, the screening was unblinded. The reference lists of the included empirical studies and reviews were also screened, retaining titles in the topic that did not appear in the systematic search. Two authors read all the retained studies and, independently, decided to include or exclude them. Disagreements in both stages were solved by consensus.

Data collection and analysis

The data of each included article were added to an extraction sheet developed for this review and refined when necessary. When available, the following variables were extracted from each article: year of publication; diagnostic/inclusion criteria used by the authors to compose the ASD samples; sample size of each group (children with ASD and typically developed controls, when present) and number of female participants; mean age and standard deviation per group; mean years of education and standard deviation per group; name and description of word reading tools or tasks; results obtained to each dependent variable (means and standard deviations per group); *p*-values and direction of the significant differences between groups or conditions.

Considering the goals of this systematic review, the results will be reported by comparison of interest: (1) comparison between children with ASD and children with typical development, and (2) comparison between children with ASD and normative data. Since one of our goals is also to understand the reading strategies of children with ASD, we will also report (3) the comparisons between the results obtained in word and nonword reading.

Results

The systematic search provided 3,444 titles. The search in the reference lists and other sources provided four additional studies. After excluding duplicates, 7,557 studies were screened based on titles and abstracts. A total of 94 articles were selected for full-text assessment of eligibility, and the remaining articles were excluded for being off-topic. From the full text assessment, 24 articles were included in the review. The entire selection process is represented in the flowchart of Figure 1.

The 24 included studies were published between 2006 and 2021 and provide data from 1,549 children with ASD (about 19% females; $M_{pooledage}$= 7.58, $SD_{pooledage}$ = 1.10) and from 1,187 children with typical development (about 55% females; $M_{pooledage}$ = 9.96, $SD_{pooledage}$ = 1.16). The entire sample of children with ASD was composed of verbal children.

Fifteen studies (62.5%) compared the performance of children with ASD and children with typical development, while nine studies (37.5%) compared the performance of children with ASD with the normative results. In these cases, the raw scores were converted to standardized scores (mean is 100 and standard deviation is 15) or were presented through the percentile. Two study presented longitudinal research, but only baseline data were analyzed (Solari et al., 2019, 2022).

As these studies were composed of several reading tests and tasks, they provided 41 comparisons of interest. Specifically, we found that: 26 tests or subtests were used to compare the performance of children with ASD and children with typical development, while 15 tests or subtests provided standardized data from children with ASD (Figure 2). Of note, we included studies that were designed to assess word reading abilities or word reading strategies in children with ASD (e.g., Nation et al., 2006; Henderson et al., 2014; Cardoso-Martins et al., 2015), and also studies designed with other goals (e.g., Arciuli et al., 2013) but had at least one measure of word or nonword reading in their assessment.

The Test of Word Reading Efficiency (TOWRE; Torgesen et al., 1999) was the most used test, followed by the Woodcock Reading Mastery Tests (WRMT; Woodcock, 2011). Table 1 presents a description of all the tasks identified in the present review, along with their distribution by comparison of interest. The totality of the tests and tasks demanded a verbal response.

Regarding the results of the individual studies, 62% of the comparisons (n = 16) contrasting the reading performance of children with ASD and children with TD did not find significant differences between groups in both word and nonword reading. However, 35% of the comparisons (n = 9) found that children with ASD had a significantly worse performance than children with TD, while 4% of the comparisons (n = 1) found the opposite pattern of results. These results are presented in detail on Table 2. With one exception (Lucas and Norbury, 2014b), the results obtained for word reading were consistent with the results obtained for nonword reading. Only Lucas and Norbury (2014b) found that children with ASD had worse performance than children with TD for word reading, although the groups did not differ significantly in nonword decoding.

TABLE 1 Description of the tests included in this review and the number of comparisons (*N*) in which they were used.

Test	Subtest	Test description	*N* ASD vs. TD	*N* standardized scores
Test of Word Reading Efficiency (TOWRE)	Sight word efficiency (SWE)	Participants read as many real words as they could in 45 s	5	2
	Phonemic decoding efficiency (PDE)	Participants read as many decodable nonwords as they could in 45 s	4	3
H4 test (Franzén, 1997) and LS test (Johansson, 1992)	–	Timed tests of single, out of context, word reading. It assesses word decoding efficiency. The H4 test was used for girls in grades 2–6 (8–12 years), while the LS test was used for the older girls (13–17 years)	1	–
Wordchains test (Jacobson, 1996)	–	Assesses word decoding ability and fluency. The task is to mark with a pencil where divisions should be made in a chain of three words without inter-word blank spaces (e.g., carhousetree). Task duration = 90 s	1	–
Phonological judgment (Auphan et al., 2018)	–	The task is to judge if a word and a pseudoword (a list of pairs) sound equal. Task duration = 2 min	1	–
Woodcock Reading Mastery Tests-Revised (WRMT-R)	Word identification	Assesses the child's ability to recognize sight word vocabulary of increasing difficulty	4	1
	Word Attack	Assesses the ability to phonetically decode pseudowords	4	3
TOWRE and WRMT-R	–	In one study (Lu et al., 2016), word reading was measured with the average of the standard scores of word identification, word attack, sight word efficiency, and phonemic decoding proficiency	1	–
Woodcock Johnson Test of Achievement—IV edition	Letter-Word Identification	Assesses single word reading	1	1
Test of School Performance	Word reading	Comprises 70 words printed in lower-case letters on a card in order of increasing difficulty	1	–
	Nonword reading	Assesses phonological decoding: the child was asked to read 20 pseudowords	1	–
Phonological Awareness Literacy Screening for Kindergarten (PALS-K; Invernizzi et al., 2015)	PALS-K—word Identification	Literacy screening tool that measures kindergarteners' developing literacy skills. The PALS-K—word identification assesses a student's ability to recognize words in text	1	–
Graded Nonword Reading Test (GNWRT; Snowling et al., 1996)	–	Involves reading nonwords presented in isolation	1	1
British Ability Scales (BAS-II; Elliot et al., 1996)	Word reading	Involves reading words presented in isolation that gradually increase in difficulty	1	1
Wide Range Achievement Test-IV	Word reading	Involves reading aloud letters and words	–	1
Illinois test of psycholinguistic abilities (ITPA-3; Hammill et al., 2001)	Sight decoding	Involves reading a list of printed words	–	1

N, frequency of use the respective task; ASD, children with autism spectrum disorder; TD, children with typical development.

TABLE 2 Results of the studies that compared the performance of children with ASD with children with typical development.

References	Participants *N* (female)	Age *M* (*SD*)	Reading test	Dependent variable	Effect direction	Main results
Johnels and Sandberg (2008)	TD: 19 (3) ASD: 37 (4)	8.81 (1.34) 9.74 (1.87)	Wordchains test	Word reading	Ns	Strong association between word decoding fluency and sentence reading comprehension in ASD group even after the effect of age and Verbal IQ was partialled out.
Johnels et al. (2010)	TD: 54 (54) ASD: 20 (20)	12.5 (2.6) 11.8 (2.7)	H4 test and LS test	Word reading	Ns	The TD and ASD girls performed very close to the normative mean on the literacy tests.
Auphan et al. (2018)	TD: 89 (56) ASD: 10 (1)	10.5 10.27	Phonological judgment	Phonological judgment		Analyses were carried out case by case. 3/10 of children with ASD had word reading decoding difficulties.
Davidson (2016)	TD: 21 (7) ASD: 21 (3)	–	Word Identification—WRMT-III Word Attack—WRMT-III	Word reading Nonword decoding	Ns Ns	*
Davidson et al. (2018)	TD: 24 (11) ASD: 19 (4)	10.97 (1.04) 11.21 (1.48)	Word Identification—WRMT-III Word Attack—WRMT-III	Word reading Nonword decoding	ASD < TD ASD < TD	Word decoding was not significantly related to reading comprehension in the TD group. In the ASD group, word decoding significantly correlated with age, reading comprehension, word reading cluster, word recognition and vocabulary.
Gabig (2010)	TD: 10 (3) ASD: 14 (2)	6.8 (0.89) 6.5 (0.72)	Word Identification—WRMT-R-NU Word Attack—WRMT-R-NU	Word reading Nonword decoding	Ns Ns	Children with ASD performed better when decoding words than nonwords: 60% had slow, labored, and inaccurate decoding attempts; 22% attempted to parse the individual graphemes/phoneme relationship and sound out the nonword but could not blend the individual phonemes into a whole; 22% were able to decode the nonwords efficiently.
Henderson et al. (2014)	TD: 49 ASD: 49	–	GNWRT BAS-II	Word reading Nonword decoding	ASD < TD ASD < TD	To examine the discrepancy between word and nonword reading, 25 children with ASD were pair-wise matched to 25 children with TD on raw word reading scores. The ASD group obtained significantly lower nonword reading scores than TD, suggesting that word reading skills are not supported by adequate phonological decoding skills in ASD.
Lu et al. (2016)	TD: 20 ASD: 25	10.3 (3.57) 11.3 (3.48)	TOWRE and WRMT composite score	Word reading	ASD < TD	The reading scores of children with ASD were near the standardized mean of 100, but significantly lower than the scores of the TD group.

(Continued)

TABLE 2 Continued

References	Participants N (female)	Age *M* (*SD*)	Reading test	Dependent variable	Effect direction	Main results
Lucas and Norbury (2014a)	TD: 30 (12) ASD-ALN: 25 (3) ASD-ALI: 12 (4)	10.47 (1.01) 11.21 (1.9) 11.77 (1.38)	sight word efficiency—TOWRE phonemic decoding efficiency—TOWRE	Word recognition Nonword decoding	ALI < (ALN = TD) ALI < (ALN = TD)	*
Lucas and Norbury (2014b)	TD: 21 (9) ASD: 20 (5)	10.46 (0.92) 10.57 (1.37)	Sight word efficiency—TOWRE Phonemic decoding efficiency—TOWRE	Word recognition Nonword decoding	ASD < TD[a] Ns	[a]However, groups did not differ significantly when analyzing the raw score of this subtest.
Macdonald et al. (2021)	TD: 15 (11) ASD: 15 (1)	4.08 (0.67) 4.58 (0.83)	Letter-Word Identification	Word reading	ASD < TD	The ASD group was divided in a subgroup of children with and without hyperlexia. This analysis showed that the group with both ASD and hyperlexia exhibited advanced word reading and letter naming skills that TD and ASD without hyperlexia but did not demonstrate commensurate phonological awareness, letter-sound correspondence, or language skills.
Cardoso-Martins et al. (2015)	TD: 19 (0) ASD: 19 (3)	6.5 (0.38) 11.5 (3.9)	Word reading—TDE Nonword reading	Word reading Nonword decoding	Ns Ns	The ability to read and spell words with accuracy was strongly correlated with the ability to read pseudowords in ASD and TD.
McIntyre et al. (2017a)	TD: 44 (16) ASD: 81 (15)	11.59 (2.25) 11.24 (2.19)	Sight word efficiency—TOWRE Phonemic decoding efficiency—TOWRE	Word recognition Nonword decoding	NS Ns	*
Solari et al. (2022)	TD: 735 ASD: 616	–	PALS-K—word Identification	Word identification	ASD > TD	*
Weissinger (2013)	TD: 37 (18) ASD: 10 (2)	–	Word Identification—WRMT-III Word Attack—WRMT-III sight word efficiency—TOWRE	Word reading Nonword decoding Word recognition	NS NS NS	*

ASD, children with autism spectrum disorder; TD, children with typical development; *Ns*, non-significant differences between groups; WRMT-R-NU, Woodcock Reading Mastery Tests-Revised-Normative Update; TOWRE, Test of Word Reading Efficiency; PALS-K, Phonological Awareness Literacy Screening for Kindergarten; GNWRT, Graded Nonword Reading Test; BAS-II, British Ability Scales; TDE, Teste de Desempenho Escolar [Test of School Performance]; ALN, ASD children with age-appropriate structural language skills; ALI, ASD children with language impairments.
*The article does not provide further qualitative information regarding word reading skills beyond the scores obtained in the reading tests.

Regarding the results of the studies that presented standardized data, the findings were more consistent as 100% of the comparisons showed that children with ASD had scores that fell within population norms. These results are presented in detail on Table 3.

About 54% of the studies reviewed ($n = 13$) presented data for both word and nonword reading. Regarding the results, we found that only 4% ($n = 1$) directly tested the difference between them showing that children with ASD had higher levels of word than of nonword reading. Adding to these results, it was noticeable that both word and nonword reading fell within the normal range values in 11 of those studies, being that nonwords presented slightly smaller standardized values than words in 5 studies, slightly bigger in 3 and virtually the same in another 3. In the other two studies, the raw data presented was converted in percentages which allowed to observe that nonwords had lower values than words in one study and nearly the same in the other. These results are presented in detail on Table 4.

TABLE 3 Results of the studies that compared the performance of children with ASD with normative data.

References	Participants *N* (female)	Age *M* (*SD*)	Reading test	Language of the study	Dependent variable	Effect direction	Main results
Arciuli et al. (2013)	21 (3)	7.8 (1.75)	Word Reading—WRAT-IV	English	Word reading	Within population norms	Significant correlation between word-level accuracy and adaptive communication domain of adaptative behavior as assessed by the parent self-report of children's adaptative behavior
Cronin (2014)	13 (2)	9.7	Word Attack—WRMT-III Phonemic decoding efficiency—TOWRE	English	Nonword decoding Nonword decoding	Within population norms Within population norms	No significant correlation between phonology and decoding or comprehension. Strong correlation between semantics and decoding, as well as decoding and comprehension
Jones (2007)	27	–	Word Identification—WRMT-III Word Attack—WRMT-III	English	Word reading Nonword decoding	Within population norms Within population norms	*
Johnels et al. (2021)	40	12	Sight decoding—ITPA-3	English	Word decoding	Within population norms	*
Knight (2016)	201	–	Word Attack—WRMT-III Letter-Word Identification	English	Word reading Word reading	Within population norms Within population norms	*
McIntyre et al. (2017a)	81 (15)	11.24 (2.19)	Sight word efficiency—TOWRE Phonemic decoding efficiency—TOWRE	English	Word recognition Nonword decoding	Within population norms Within population norms	Four profiles of readers: (1) Comprehension Disturbance; (2) Global Disturbance; (3) Severe Global Disturbance; (4) Average Readers. All but the Severe has normative or near normative word reading scores. None manifested a profile of good comprehension and poor word reading
Nation et al. (2006)	32	–	BAS-II GNWRT	English	Word reading Nonword decoding	Within population norms Within population norms	*
Quan (2014)	29 (2)	–	Letter-Word Identification	English	Word reading	Within population norms	Majority of children (61%) falling within one SD of population norms; 1 student performed above one SD. Six students (21%) had standard scores below one SD of population norms, three students (11%) fell below two SDs, and one student fell below three SDs
Solari et al. (2019)	80 (15)	11.26 (2.15)	Sight word efficiency—TOWRE Phonemic decoding efficiency—TOWRE	English	Word recognition Nonword decoding	Within population norms Within population norms	Similar reading profiles at time points 1 and 2 of assessment

ASD, children with autism spectrum disorder; TD, children with typical development; *Ns*, non-significant differences between groups; WRMT-R-NU, Woodcock Reading Mastery Tests-Revised-Normative Update; TOWRE, Test of Word Reading Efficiency; WRAT-IV, Wide Range Achievement Test; ITPA-3, Illinois test of psycholinguistic abilities; GNWRT, Graded Nonword Reading Test; BAS-II, British Ability Scales.
*The article does not provide further qualitative information regarding word reading skills beyond the scores obtained in the reading tests.

10.3389/fpsyg.2022.930275

TABLE 4 Results of the studies that compared the performance of children with ASD in tests assessing word and nonword reading.

References	Participants N (female)	Age M (SD)	Reading test	Language of the study	Dependent variable	Word reading	Nonword decoding	Comparison word vs. nonword
Davidson (2016)	21 (3)	–	Word Identification—WRMT-III Word Attack—WRMT-III	English	Word reading Nonword decoding	102.90 (12.57)	95.14 (13.91)	Not tested
Davidson et al. (2018)	19 (4)	11.21 (1.48)	Word Identification—WRMT-III Word Attack—WRMT-III	English	Word reading Nonword decoding	98.42 (13.26)	89.47 (11.97)	Not tested
Gabig (2010)	14 (2)	6.5 (0.72)	Word Identification—WRMT-R-NU Word Attack—WRMT-R-NU	English	Word reading Nonword decoding	115 (10.3)	104 (11.2)	ASD: Words > Nonwords TD: NS
Henderson et al. (2014)	49	–	BAS-II GNWRT	English	Word reading Nonword decoding	69.56 (12.58)/90	14.92 (6.92)/25	Not tested
Jones (2007)	27	–	Word Identification—WRMT-III Word Attack—WRMT-III	English	Word reading Nonword decoding	100.17 (15.54)	96.41 (24.08)	Not tested
Lucas and Norbury (2014a)	ALN: 25 (3) ALI: 12 (4)	11.21 (1.9) 11.77 (1.38)	sight word efficiency—TOWRE phonemic decoding efficiency—TOWRE	English	Word recognition Nonword decoding	ALN: 104.69 (12.63) ALI: 91.83 (8.09)	ALN: 109.02 (12.10) ALI: 94.89 (12.03)	Not tested
Lucas and Norbury (2014b)	20 (5)	10.57 (1.37)	sight word efficiency—TOWRE phonemic decoding efficiency—TOWRE	English	Word recognition Nonword decoding	95.48 (13.11)	101.13 (16.94)	Not tested
Cardoso-Martins et al. (2015)	19 (3)	11.5 (3.9)	Word reading—TDE Nonword reading	Portuguese	Word reading Nonword decoding	46.5 (20.43)/70	11.89 (5.71)/20	Not tested; ASD > phonological errors in reading
McIntyre et al. (2017a)	81 (15)	11.24 (2.19)	Sight word efficiency—TOWRE Phonemic decoding efficiency—TOWRE	English	Word recognition Nonword decoding	93.29 (14.75)	94.89 (14.81)	Not tested
Nation et al. (2006)	32	–	BAS-II GNWRT	English	Word reading Nonword decoding	96.56 (23.37)	90.83 (17.87)	Not tested; 64 % of children was 1 SD below norms on nonword reading
Weissinger (2013)	10 (2)	–	Sight word efficiency—TOWRE Word Identification—WRMT-R-NU Word Attack—WRMT-R-NU	English	Word recognition Word reading Nonword decoding	92.00 (11.039) 101.40 (15.63)	111.3 (20.7)	Not tested

(Continued)

TABLE 4 Continued

References	Participants N (female)	Age M (SD)	Reading test	Language of the study	Dependent variable	Word reading	Nonword decoding	Comparison word vs. nonword
Solari et al. (2019)	80 (15)	11.26 (2.15)	Sight word efficiency—TOWRE Phonemic decoding efficiency—TOWRE	English	Word recognition Nonword decoding	94.87 (14.91)	93.66 (14.47)	Not tested

ASD, children with autism spectrum disorder; TD, children with typical development; ALN, ASD children with age-appropriate structural language skills; ALI, ASD children with language impairments; WRMT-III, Woodcock Reading Mastery Tests-Third Edition; WRMT-R-NU, Woodcock Reading Mastery Tests-Revised-Normative Update; TOWRE, Test of Word Reading Efficiency; GNWRT, Graded Nonword Reading Test; BAS-II, British Ability Scales; TDE, Teste de Desempenho Escolar (Test of School Performance).

Discussion

In the present review we analyzed empirical research in order to systematize the current knowledge concerning the word reading processes used by children with ASD. Twenty-four articles published from 2006 to December 2021 were selected and three data comparisons were analyzed: (1) comparisons between children with ASD and children with typical development regarding word and nonword reading, (2) comparisons between children with ASD and normative data, and (3) comparisons between the results obtained in word and nonword reading in the ASD group.

Considering the comparisons with the typical development peers, it was possible to observe that although a majority of them have shown that the children with ASD achieved similar levels of performance, a considerable percentage (35%) showed lower reading levels. However, 10 out of 15 of those comparisons were conducted with small samples (between 10 and 21 participants), which may have affected the statistical power of the results. The effect of small samples may explain the divergence between the actual results and those found by Solari et al. (2019) study, which indicated that 58% of their 80 participants showed word reading difficulties. In addition, there is evidence that even when there are no significant differences between children with ASD and their typically developing peers on reading accuracy and speed, results from psychophysiological measures (such as eye fixations and regressions) have been showing that reading is often a more effortful task for them than for their colleagues (Howard et al., 2017).

In turn, the totally of the studies that reported standardized measures showed that the reading levels of children with ASD were within the normal range. Although the findings of these two type comparisons are not paradoxical, they should be interpreted with caution. First, it indicates that the presence of a control group seems to be critical to better understand the challenges that children with ASD face when reading. Moreover, putting the methodological precision aside, using only population norms to characterize the average reading profiles may be rather imprecise since the individual variability among children with ASD is large as the standard deviations presented indicate (see, for instance, Table 4).

Analyzing the comparisons between word and nonword reading performances among children with ASD aimed to investigate the reading strategy(ies) more successfully used by these children—decoding (using phonological sub-lexical units) and/or recognition (using lexical patterns). It is important to note that reading words may be achieved by recognition or by the sequential process of converting graphemes into phonemes and assembling them to pronouncing the word—the decoding strategy. It depends on the familiarity of the word and the reading level of the reader. Thus, when authors present data on word reading most of the time there is ambiguity about the cognitive processes in course when children underwent the task. Some tests, however, are more evident. For instance, when it is said that children were asked to read sight words or visual vocabulary it is more likely they were using word recognition than decoding. To the contrary, when children are said to be reading nonwords, they will need to decode them using sub-lexical procedures in order to pronounce them since nonwords does not exist in the mental lexicon and cannot profit from memory for words. Indeed, unless the reader is very unskilled, words require less of decoding than nonwords do (Weekes, 1997) and nonwords cannot be recognized.

As it was mentioned, only 13 studies reported both the results that assessed word and nonword reading. From this small pool of studies, it seems that children with ASD can use both lexical and sub-lexical phonological knowledge to read at a similar easy. It is noteworthy though that there was a patent paucity of direct assessment of the processes underlying word reading performances since all but four studies (Nation et al., 2006; Gabig, 2010; Henderson et al., 2014; Cardoso-Martins et al., 2015) aimed to study other aspects of reading, like comprehension or the relationship between ASD characteristics and comprehension processes. Thus, the characteristics of items to be read were not specifically designed for comparing word with nonword reading and this limits the soundness of the outcome of these comparisons.

Nevertheless, some of the studies suggest that children with ASD may be more prone to rely on recognition processes than in decoding ones. This appears so when, in addition to the reading scores, it is pointed out that children with ASD produced more phonological errors when reading (Cardoso-Martins et al., 2015) and presented a higher percentage of nonword reading levels below the norms, although they did not differentiate from their peers with typical reading levels in word and nonword reading (Nation et al., 2006). Indeed, there is good evidence that children with ASD are skillful in visual patterns processing (Ostrolenk et al., 2017), which can aid their lexical orthographic learning. On the other hand, there is also some research indicating that phonological processes and alphabet knowledge seem to be areas of strength for children with ASD (Frith and Snowling, 1983; Lucas and Norbury, 2014b; Dynia et al., 2016).

These apparent discrepancies suggest that word reading strategies in ASD are far from being well-understood and that studying them should consider environmental variables like reading instruction and the orthographic consistency of the language to be read. Indeed, it is important to point out that the teaching methods the children undertook have probably influenced their performances. Although scarce, there is evidence that children with ASD may have limited instructional conditions (Spector and Cavanaugh, 2015). These authors concluded that children with ASD had a reading instructional time of 60–30 min per day, contrary to recommendations of 90–120 min for K-3 students. Also, there were reports of an instructional overemphasis on narrow skills such as sight-word knowledge which, with time passing, has becoming more combined with code instruction. In the studies reviewed there were no references to teaching methods. Regarding the orthographic consistency, the degree to which each letter has one or more phonological correspondences, and considering that orthographies differ greatly in their consistency and therefore in their correspondent easy to learn to read (Duncan et al., 2013), it is remarkable that we have found only one study that was not run in English, the Brazilian Portuguese study of Cardoso-Martins et al. (2015). Having studies of ASD children's word reading skills run in orthographies more consistent than English, the most inconsistent orthography, is necessary because it is not possible to have a clear idea of reading in ASD without contrasting different orthographies. As it is put by the Psycholinguistic Grain Size Theory of Reading (Ziegler and Goswami, 2005), it is possible that in more consistent orthographies ASD children could, like their TD peers, learn how to read faster and easier than in English showing a different and clearer pattern when considering the lexical and sub-lexical pathways.

Related with the lack of solid data on word reading strategies in ASD is the confusing language of many studies when referring to word reading. In analyzing the selected studies, it was witnessed a generalized interchangeable use of the words "decoding" and "recognition", as if they were synonyms, to the point that it was often impossible to understand exactly what the authors were talking about. For instance, the phrase "accuracy in word recognition was measured for both real words and for nonwords" is a unique extreme example of incorrectness, but the profusion of inaccuracy and ambiguity in the usage of "decoding" and "word recognition" is probably contributing to curtail a clear understanding on word reading strengths and challenges in ASD.

Given the goals of the research reviewed it is unsurprising that, almost all the studies selected children with ASD and, at least, some reading knowledge. Nevertheless, since there are estimates of about 30% of children with ASD being nonverbal or minimally verbal (Tager-Flusberg and Kasari, 2013), what is somewhat unexpected is that all the studies analyzed required the children to orally respond to the reading tasks. However, some individuals with ASD can read and write meaningfully despite not using spoken language (Goh et al., 2013) and minimally verbal children with ASD were proven to be able to learn to read words and to discriminate between words and nonwords (Serret et al., 2017).

Thus, in addition to little solid evidence about the word reading strategies of children with ASD, this review highlighted the dearth of knowledge about nonverbal children with ASD reading profiles. All the studies included in the review had participants who were verbal and required a verbal response to the reading assessments, and this may be a major limitation of the previous research in this field. Moreover, this fact may explain the evidence that all the reviewed studies that compared the performance of children with ASD with normative data fell within the population norms. If confirmed by future studies, the evidence of a lack of difference between children with ASD and normative data may be a sample effect instead of the true effect of ASD in word reading. Thereby, our results show that the current research is very limiting for understanding autism functioning and diversity. As challenging as it may be, this calls for further research employing valid assessment reading tools that can fairly be used with a larger extent of ASD heterogeneity like silent reading tests and methods like eye-tracking that do not demand verbal answers which may also be conducted with nonverbal children.

Another issue regarding the weaknesses of the results analyzed is that the female samples are very under-represented, considering the known ratio male-female being 3:1 (Loomes et al., 2017). Future studies may take this ratio into account during participants recruitment to constitute samples that are as representative as possible of the real world. This lack of representativeness, associated with the fact that the samples of ASD children were treated as wholes, instead of being grouped according to relevant characteristics (such as oral language level, attentional difficulties, years of schooling, among others), suggests that different results could be shown if these constraints were controlled.

Therefore, future studies should be conducted with a larger number of gender matched participants, using tasks specifically designed to study word reading. They also should include a typical developing control group and collect psychophysiological data, such as eye-tracking measures, to increase the understanding of behavioral outcomes. In addition, samples of children with ASD should be specifically characterized in order to better realize how the natural variability of ASD concur with word reading performances.

This review suggests that children with ASD may have preserved word reading abilities, being a further step in the direction of a better understanding of autism associated word reading challenges, identifying weaknesses of existing studies and opening new directions for future research. However, given the weaknesses found, it is not possible to identify which strategies children with ASD use better to identify written words, nor can we thoroughly deduce what are exactly the word reading difficulties and strengths of children with ASD.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

AV conducted the literature review, did the systematic search and the inclusion of the articles, analyzed the data, wrote the introduction and the discussion, and revised the manuscript. CF did the systematic search and the inclusion of the articles, analyzed the data, wrote the method, and revised the manuscript. SC analyzed the data, wrote the discussion, and revised the manuscript. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Washington, DC: American Psychiatric Pub.

*Arciuli, J., Stevens, K., Trembath, D., and Simpson, I. C. (2013). The relationship between parent report of adaptive behavior and direct assessment of reading ability in children with autism spectrum disorder. *J. Speech Lang. Hear. Res.* 56, 1837–1844 doi: 10.1044/1092-4388(2013/12-0034)

ASDEU (2018). *Autism Spectrum Disorder in Europe*. Executive Summury. The European Commission Action scheme: Call for tender n° SANCO/2014/C2/035. Available online at: http://asdeu.eu (accessed January 20, 2020).

*Auphan, P., Potocki, A., Ecalle, J., and Magnan, A. (2018). Evaluation informatisée de la compréhension en lecture d'enfants autistes sans déficience intellectuelle. *Rev. Psychoéduc.* 47, 333–356. doi: 10.7202/1054064ar

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., 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 Surveill. Summar.* 67, 1–28. doi: 10.15585/mmwr.ss6706a1

Bednarz, H. M., Maximo, J. O., Murdaugh, D. L., O'Kelley, S., and Kana, R. K. (2017). "Decoding versus comprehension": brain responses underlying reading comprehension in children with autism. *Brain Lang.* 169, 39–47. doi: 10.1016/j.bandl.2017.01.002

Brown, H. M., Oram-Cardy, J., and Johnson, A. (2013). A meta-analysis of the reading comprehension skills of individuals on the autism spectrum. *J. Autism Dev. Disord.* 43, 932–955. doi: 10.1007/s10803-012-1638-1

*Cardoso-Martins, C., Gonçalves, D. T., de Magalhães, C. G., and da Silva, J. R. (2015). Word reading and spelling ability in school-age children and adolescents with autism spectrum disorders: evidence from Brazilian Portuguese. *Psychol. Neurosci.* 8, 479. doi: 10.1037/pne0000029

Castles, A., Rastle, K., and Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychol. Sci. Public Interest* 19, 5–51. doi: 10.1177/1529100618772271

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256. doi: 10.1037/0033-295X.108.1.204

*Cronin, K. A. (2014). The relationship among oral language, decoding skills, and reading comprehension in children with autism. *Exceptionality* 22, 141–157. doi: 10.1080/09362835.2013.865531

*Davidson, M. M. (2016). *Reading for Meaning: Reading Comprehension Skills in ASD and the Role of Oral Language, Central Coherence, and Executive Function* (dissertation thesis). The University of Wisconsin-Madison, Madison, WI, United States.

*Davidson, M. M., Kaushanskaya, M., and Ellis Weismer, S. (2018). Reading comprehension in children with and without ASD: the role of word reading, oral language, and working memory. *J. Autism Dev. Disord.* 48, 3524–3541. doi: 10.1007/s10803-018-3617-7

Davidson, M. M., and Weismer, S. E. (2014). Characterization and prediction of early reading abilities in children with autism spectrum. *J. Autism Dev. Disord.* 44, 828–845. doi: 10.1007/s10803-013-1936-2

Duncan, L. G., Casto, S. L., Defior, S., Seymour, P. H. K., Baillie, S., and Serrano, F. (2013). Phonological development in relation to language and literacy: Variations on a theme in six alphabetic orthographies. *Cognition* 127, 398–419. doi: 10.1016/j.cognition.2013.02.009

Duncan, T. M., Karkada, M.,Deacon, S., and Smith, I. (2021). Building meaning: meta-analysis of component skills supporting reading comprehension in children with autism spectrum disorder. *Autism Res.* 14, 840–858. doi: 10.1002/aur.2483

Dynia, J. M., Brock, M. E., Logan, J. A., Justice, L. M., and Kaderavek, J. N. (2016). Comparing children with ASD and their peers' growth in print knowledge. *J. Autism Dev. Disord.* 46, 2490–2500. doi: 10.1007/s10803-016-2790-9

Dynia, J. M., Lawton, K., Logan, J. A. R., and Justice, L. M. (2014). Comparing emergent-literacy skills and home-literacy environment of children with autism and their peers. *Topics Early Child. Spec. Educ.* 34, 142–153. doi: 10.1177/0271121414536784

Ehri, L. (2013) "Aquisição da habilidade de leitura de palavras e sua influência na pronúncia e na aprendizagem do vocabulário [Word reading acquisition and its influence on vocabulary pronunciation and learning]," in *Alfabetização no s?culo XXI: Como se aprende a ler e a escrever*, eds M. R. Maluf and C. Cardoso-Martins (Porto Alegres - RS: Penso), 49–81.

El Zein, F., Solis, M., Vaughn, S., and McCulley, L. (2014). Reading comprehension interventions for students with autism spectrum disorders: a synthesis of research. *J. Autism Dev. Disord.* 44, 1303–1322. doi: 10.1007/s10803-013-1989-2

Elliot, C. D., Smith, P., and McCulloch, K. (1996). *British Ability Scales. 2nd Edn.* Windsor: NFER.

Engel, K. S. (2018). *Reading Comprehension Instruction for Young Students With High Functioning Autism: Forming Contextual Connections* (ProQuest Dissertations and Theses Global Database) (Proquest ID No. 2155988501). The City University of New York, New York, NY, United States.

Fernandes, F. D. M., Amato, C. A. D. L. H., Cardoso, C., Navas, A. L. G. P., & Molini-Avejonas, D. R. (2015). Reading in autism spectrum disorders: a literature review. *Folia Phoniatrica et Logopaedica*, 67, 169?-177. doi: 10.1159/000442086

Fernandes, S., Querido, L., Verhaeghe, A., Marques, C., and Araújo, L. (2017b). Reading development in European Portuguese: Relationships between oral reading fluency, vocabulary and reading comprehension. *Read. Writing Interdiscipl. J.* 30, 1987–2007. doi: 10.1007/s11145-017-9763-z

Fernandes, T., Araújo, S., Sucena, A., Reis, A., and Castro, S. L. (2017a). The 1-min screening test for reading problems in college students: psychometric properties of the 1-min TIL. *Dyslexia* 23, 66–87. doi: 10.1002/dys.1548

Fleury, V. P., Hedges, S., Hume, K., Browder, D. M., Thompson, J. L., Fallin, K., et al. (2014). Addressing the academic needs of adolescents with autism spectrum disorder in secondary education. *Remed. Special Educ.* 35, 68–79. doi: 10.1177/0741932513518823

Franzén, L. (1997). *Diagnostisering av elever med läs-och skrivsvårigheter: skolår 4-6: diagnostisering, tolkning och åtgärder*. Ekelund.

Frith, U., and Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *Br. J. Dev. Psychol.* 1, 329–342. doi: 10.1111/j.2044-835X.1983.tb00906.x

Gabig, C. (2010). Phonological awareness and word recognition in reading by children with autism. *Commun. Disord. Q.* 31, 67–85. doi: 10.1177/1525740108328410

Goh, S., Whitaker, A., Feldman, J., Cull, M. B., Hoyte, K., Algermissen, M., et al. (2013). Teaching non-verbal children with autistic disorder to read and write: a pilot study. *Int. J. Dev. Disabil.* 59, 95–107. doi: 10.1179/2047387713Y.0000000014

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remed. Special Educ.* 7, 6–10. doi: 10.1177/074193258600700104

Hammill, D. D., Mather, N., and Roberts, R. (2001). *ITPA-3: Illinois Test of Psycholinguistic Abilities.* Austin, TX: *Pro-ed*.

Happé, F., and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0

Harm, M. W., and Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychol. Rev.* 111, 662–720. doi: 10.1037/0033-295X.111.3.662

*Henderson, L. M., Clarke, P. J., and Snowling, M. J. (2014). Reading comprehension impairments in autism spectrum disorders. *L'Annee Psychol.* 114, 779–797. doi: 10.3917/anpsy.144.0779

Hooper, S. R., Poon, K. K., Marcus, L., and Fine, C. (2006). Neuropsychological characteristics of school-age children with high-functioning autism: performance on the NEPSY. *Child Neuropsychol.* 12, 299–305. doi: 10.1080/09297040600737984

Hoover, W., and Tunmer, W. (2020). *The Cognitive Foundations of Reading and Its Acquisition. A Framework with Applications Connecting Teaching and Learning.* Cham: Spinger.

Hoover, W. A., and Gough, P. B. (1990). The simple view of reading. *Read. Writ.* 2, 127–160. doi: 10.1007/BF00401799

Howard, P. L., Liversedge, S. P., and Benson, V. (2017). Benchmark eye movement effects during natural reading in autism spectrum disorder. *J. Exp. Psychol. Learn. Mem. Cogn.* 43, 109–127. doi: 10.1037/xlm0000289

Huemer, S. V., and Mann, V. (2010). A comprehensive profile of decoding and comprehension in autism spectrum disorders. *J. Autism Dev. Disord.* 40, 485–493. doi: 10.1007/s10803-009-0892-3

Ibrahim, I. (2020). Specific learning disorder in children with autism spectrum disorder: current issues and future implications. *Adv. Neurodev. Disord.* 4, 103–112. doi: 10.1007/s41252-019-00141-x

Invernizzi, M., Swank, L., Juel, C., and Meier, J. (2015). *Phonological Awareness Literacy Screening Kindergarten*. University Printing

Jacobson, C. (1996). *Manual: Ordkedjor (Manual: Wordchains).* Stockholm: Psykologiförlaget.

Johansson, M. G. (1992). *Klassdiagnoser i läsning och skrivning för högstadiet och gymnasiet* (Stockholm: Psykologiförlaget).

Johnels, J., Fernell, E., Kjellmer, L., Gillberg, C., and Norrelgen, F. (2021). Language/cognitive predictors of literacy skills in 12-year-old children on the autism spectrum. *Logopedics Phoniatr. Vocol.* 4, 1–5. doi: 10.1080/14015439.2021.1884897 [Epub ahead of print].

Johnels, J., Kopp, S., Berg-Kelly, K., and Gillberg, C. (2010). Reading comprehension, word decoding and spelling in girls with autism spectrum disorders (ASD) or attention-deficit/hyperactivity disorder (AD/HD): performance and predictors. *Int. J. Lang. Commun. Disord.* 45, 61–71. doi: 10.3109/13682820902745438

Johnels, J., and Sandberg, A. D. (2008). Basic reading skills in high-functioning Swedish children with autism spectrum disorders or attention disorder. *Res. Autism Spectr. Disord.* 2, 95–109. doi: 10.1016/j.rasd.2007.03.006

*Jones, G. W. (2007). *Written and Oral Language in Autism* (dissertation thesis). University of Connecticut. Available online at: https://opencommons.uconn.edu/dissertations/AAI3279276 (accessed March 14, 2022).

Jonhels, J., and Sandberg, A. D. (2012). Dyslexic, delayed, precocious or just normal? Word reading skills of children with autism spectrum disorders. *J. Res. Read.* 35, 20-?31. doi: 10.1111/j.1467-9817.2010.01452.x

Kim, Y. S. G., Petscher, Y., and Vorstius, C. (2021). The relations of online reading processes (eye movements) with working memory, emergent literacy skills, and reading proficiency. *Sci. Stud. Read.* 25, 351–369. doi: 10.1080/10888438.2020.1791129

*Knight, E. M. (2016). *Development of Reading, Language, and Social Skills in Young Children With ASD*. Riverside, CA: University of California.

Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., et al. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Sci. Stud. Read.* 23, 220–234. doi: 10.1080/10888438.2018.1510936

Lonigan, C. J., Burgess, S. R., and Schatschneider, C. (2018). Examining the simple view of reading with elementary school children: still simple after all these years. *Remed. Special Educ.* 39, 260–273. doi: 10.1177/07419325187 64833

Loomes, R., Hull, L., and Mandy, W. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *J. Am. Acad. Child Adolesc. Psychiatry* 56, 466–474. doi: 10.1016/j.jaac.2017.03.013

*Lu, C., Qi, Z., Harris, A., Weil, L. W., Han, M., Halverson, K., et al. (2016). Shared neuroanatomical substrates of impaired phonological working memory across reading disability and autism. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 1, 169–177. doi: 10.1016/j.bpsc.2015.11.001

*Lucas, R., and Norbury, C. F. (2014a). Levels of text comprehension in children with autism spectrum disorders (ASD): The influence of language phenotype. *J. Autism Dev. Disord.* 44, 2756–2768. doi: 10.1007/s10803-014-2133-7

*Lucas, R., and Norbury, C. F. (2014b). Orthography facilitates vocabulary learning for children with autism spectrum disorders (ASD). *Q. J. Exp. Psychol.* 67, 1317–1334. doi: 10.1080/17470218.2013.859714

Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., et al. (2017). The changing epidemiology of autism spectrum disorders. *Annu. Rev. Public Health* 38, 81–102. doi: 10.1146/annurev-publhealth-031816-044318

Lyon, G. R. (2001). *Measuring Success: Using Assessments and Accountability to Raise Student Achievement Statement of Dr. G. Reid Lyon, Chief, Child Development and Behavior Branch, National Institute of Child Health and Human Development, National Institutes of Health.* Subcommittee on Education Reform, Committee on Education and the Workforce, U.S. House of Representatives. Washington, DC. Available online at: http://projectpro.com/ICR/Research/Releases/NICHD_Testimony1.htm (accessed March 16, 2022).

*Macdonald, D., Luk, G., and Quintin, E. M. (2021). Early word reading of preschoolers with ASD, both with and without HYPERLEXIA, compared to typically Developing Preschoolers. *J. Autism Dev. Disord.* 51, 1598–1612. doi: 10.1007/s10803-020-04628-8

Macdonald, D., Luk, G., and Quintin, E. M. (2022). Early reading comprehension intervention for preschoolers with autism spectrum disorder and hyperlexia. *J. Autism Dev. Disord.* 52, 1652–1672. doi: 10.1007/s10803-021-05057-x

Maenner, M., Shaw, K., Bakian, A., Bilder, D., and Durkin, M., Esler, A., et al. (2021). Prevalence and Characteristics Of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR Surveill. Summ.* 70, 1–16. doi: 10.15585/mmwr.ss7011a1

Maughan, B., Rutter, M., and Yule, W. (2020). The Isle of Wight studies: the scope and scale of reading difficulties. *Oxf. Rev. Educ.* 46, 429–438. doi: 10.1080/03054985.2020.1770064

McIntyre, N. S., Solari, E. J., Gonzales, J. E., Solomon, M., Lerro, L. E., Novotny, S., et al. (2017b). The scope and nature of reading comprehension impairments in school-aged children with higher-functioning autism spectrum disorder. *J. Autism Dev. Disord.* 47, 2838–2860. doi: 10.1007/s10803-017-3209-y

*McIntyre, N. S., Solari, E. J., Grimm, R. P., E., Lerro, L., E., et al. (2017a). A comprehensive examination of reading heterogeneity in students with high functioning autism: Distinct reading profiles and their relation to autism symptom severity. *J. Autism Dev. Disord.* 47, 1086–1101. doi: 10.1007/s10803-017-3029-0

Melby-Lervåg, M., Lyster, S. A. H., and Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychol. Bull.* 138, 322–352. doi: 10.1037/a0026744

Micai, M., Vulchanova, M., and Saldaña, D. (2021). Reading goals and executive function in autism: an eye-tracking study. *Autism Res.* 14, 1007–1024. doi: 10.1002/aur.2447

Miles, K., and Ehri, L. (2019). "Orthographic mapping facilitates sight word memory and vocabulary learning," in *Reading Development and Difficulties*, eds D. Kilpatrick, R. Joshi, and R. Wagner (Springer), 63–82.

Minshew, N. J., Goldstein, G., and Siegel, D. J. (1995). Speech and language in high-functioning autistic individuals. *Neuropsychology* 9, 255. doi: 10.1037/0894-4105.9.2.255

Nally, A., Healy, O., Holloway, J., and Lydon, H. (2018). An analysis of reading abilities in children with autism spectrum disorders. *Res. Autism Spectr. Disord.* 47, 14–25. doi: 10.1016/j.rasd.2017.12.002

*Nation, K. (2019). Children's reading difficulties, language, and reflections on the simple view of reading. *Aust. J. Learn. Difficult.* 24, 47–73, doi: 10.1080/19404158.2019.1609272

Nation, K., Clarke, P., Wright, B., and Williams, C. (2006). Patterns of reading ability in children with autism spectrum disorder. *J. Autism Dev. Disord.* 36, 911–919. doi: 10.1007/s10803-006-0130-1

Nation, K., and Norbury, C. F. (2005). Why reading comprehension fails: Insights from developmental disorders. *Top. Lang. Disord.* 25, 21–32. doi: 10.1097/00011363-200501000-00004

Nayar, K., Kang, X., Xing, J., Gordon, P., Wong, P., and Losh, M. (2021). A cross-cultural study showing deficits in gaze-language coordination during rapid automatized naming among individuals with ASD. *Sci. Rep.* 11, 13401. doi: 10.1038/s41598-021-91911-y

Ofner, M., Coles, A., Decou, M. L., Do, M. T., Bienek, A., Snider, J., Ugnat, A. M. (2018). *Autism Spectrum Disorder Among Children and Youth in Canada*. Available online at: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/autism-spectrum-disorder-children-youth-canada-2018.html (accessed April 23, 2022).

Ostrolenk, A., d'Arc, B. F., Jelenic, P., Samson, F., and Mottron, L. (2017). Hyperlexia: systematic review, neurocognitive modelling, and outcome. *Neurosci. Biobehav. Rev.* 79, 134–149. doi: 10.1016/j.neubiorev.2017.04.029

Ouzzani, M., Hammady, H., Fedorowicz, Z., and Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Syst. Rev.* 5, 210. doi: 10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372, n71. doi: 10.1136/bmj.n71

Perry, C., Ziegler, J. C., and Zorzi, M. (2010). Beyond single syllables: large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cogn. Psychol.* 61, 106–151. doi: 10.1016/j.cogpsych.2010.04.001

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., and Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol. Rev.* 103, 56–115. doi: 10.1037/0033-295X.103.1.56

*Quan, J. L. (2014). *Reading Achievement for Students With Autism and Students With Learning Disability: A Comprehensive Examination of Five Key Areas of Reading* (dissertation thesis). Los Angeles, CA: California State University.

Ricketts, J. (2011). Research review: reading comprehension in developmental disorders of language and communication. *J. Child Psychol. Psychiatry Allied Disciplines* 52, 1111–1123. doi: 10.1111/j.1469-7610.2011.02438.x

Ricketts, J., Jones, C. R. G., Happé, F., and Charman, T. (2013). Reading comprehension in autism spectrum disorders: the role of oral language and social functioning. *J. Autism Dev. Disord.* 43, 807–816. doi: 10.1007/s10803-012-1619-4

Samson, F., Mottron, L., Souliéres, I., Zeffiro, T.A. (2012). Enhanced visual functioning in autism: an ALE meta-analysis. *Hum. Brain Mapp.* 33, 1553–1581. doi: 10.1002/hbm.21307

Serret, S., Hun, S., Thümmler, S., Pierron, P., Santos, A., Bourgeois, J., et al. (2017). Teaching literacy skills to French minimally verbal school-aged children with autism spectrum disorders with the serious game SEMA-TIC: an exploratory study. *Front. Psychol.* 8, 1523. doi: 10.3389/fpsyg.2017.01523

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition* 55, 151–226. doi: 10.1016/0010-0277(94)00645-2

Singh, B. D., Moore, D. W., Furlonger, B. E., Anderson, A., Fall, R., and Howorth, S. (2021). Reading comprehension and autism spectrum disorder: a systematic review of interventions involving single-case experimental designs. *Rev. J. Autism Dev. Disord.* 8, 3–21. doi: 10.1007/s40489-020-00200-3

Snowling, M., Stothard, J., and McLean, J. F. (1996). *Graded Nonword Reading Test*. London: Thames Valley Test.

*Solari, E. J., Grimm, R. P., McIntyre, N. S., Zajic, M., and Mundy, P. C. (2019). Longitudinal stability of reading profiles in individuals with higher functioning autism. *Autism* 23, 1911–1926. doi: 10.1177/1362361318812423

*Solari, E. J., Henry, A. R., Grimm, R. P., Zajic, M. C., and McGinty, A. (2022). Code-related literacy profiles of kindergarten students with autism. *Autism* 26, 230–242. doi: 10.1177/13623613211025904

Spector, J. E., and Cavanaugh, B. J. (2015). The conditions of beginning reading instruction for students with autism spectrum disorder. *Remed. Special Educ.* 36, 337–346. doi: 10.1177/0741932515579277

Tager-Flusberg, H., and Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism Res.* 6, 468–478. doi: 10.1002/aur.1329

Torgesen, J., Wagner, R., and Rashotte, C. (1999, 2012). *Test of Word Reading Efficiency. 2nd Edn.* Austin, TX: Pro-Ed Inc.

Verhoeven, L., Voeten, M., and Vermeer, A. (2019). Beyond the simple view of early first and second language reading: the impact of lexical quality. *J. Neurolinguist.* 50, 28–36. doi: 10.1016/j.jneuroling.2018.03.002

Walenski, M., Mostofsky, S. H., Gidley-Larson, J. C., and Ullman, M. T. (2008). Brief report: enhanced picture naming in autism. *J. Autism Dev. Disord.* 38, 1395–1399. doi: 10.1007/s10803-007-0513-y

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Q. J. Exp. Psychol.* 50, 439–456. doi: 10.1080/713755710

*Weissinger, K. (2013). *Factors Associated With Individual Differences in Reading Comprehension for Typically-Developing Students and for a Pilot Sample of Students Diagnosed With Autism Spectrum Disorder*. University of Rhode Island. Available online at: https://digitalcommons.uri.edu/oa_diss/70 (accessed April 08, 2022).

Welsh, C. M., Pennington, B. F., and Rogers, S. (1987). Word recognition and comprehension skills in hyperlexic children. *Brain Lang.* 32, 76–96. doi: 10.1016/0093-934X(87)90118-0

Westerveld, M. F., Paynter, J., O'Leary, K., and Trembath, D. (2018). Preschool predictors of reading ability in the first year of schooling in children with ASD. *Autism Res.* 11, 1332–1344. doi: 10.1002/aur.1999

White, S., Frith, U., Milne, E., Rosen, S., Swettenham, J., and Ramus, F. (2006). A double dissociation between sensorimotor impairments and reading disability: a comparison of autistic and dyslexic children. *Cogn. Neuropsychol.* 23, 748–761. doi: 10.1080/02643290500438607

Woodcock, R. (2011). *Woodcock Reading Mastery Tests, 3rd Edn*. Bloomington, MN: Pearson, Inc.

World Health Organization (2019). *Autism Spectrum Disorders*. Available from: https://www.who.int/newsroom/fact-sheets/detail/autism-spectrum-disorders (accessed March 16, 2021).

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

* Study included in the review.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 29 July 2022
DOI 10.3389/fpsyg.2022.927847

OPEN ACCESS

EDITED BY
Lénia Carvalhais,
Infante D. Henrique Portucalense University, Portugal

REVIEWED BY
Zachary J. Williams,
Vanderbilt University, United States
Courtney Venker,
Michigan State University, United States
Meagan Talbott,
University of California, Davis, United States

*CORRESPONDENCE
Shuting Zheng
shuting.zheng@ucsf.edu

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 25 April 2022
ACCEPTED 11 July 2022
PUBLISHED 29 July 2022

CITATION
Zheng S, Kaat A, Farmer C, Thurm A, Burrows CA, Kanne S, Georgiades S, Esler A, Lord C, Takahashi N, Nowell KP, Will E, Roberts J and Bishop SL (2022) Bias in measurement of autism symptoms by spoken language level and non-verbal mental age in minimally verbal children with neurodevelopmental disorders. *Front. Psychol.* 13:927847.
doi: 10.3389/fpsyg.2022.927847

Bias in measurement of autism symptoms by spoken language level and non-verbal mental age in minimally verbal children with neurodevelopmental disorders

Shuting Zheng[1]*, Aaron Kaat[2], Cristan Farmer[3], Audrey Thurm[3], Catherine A. Burrows[4,5], Stephen Kanne[6], Stelios Georgiades[7], Amy Esler[5], Catherine Lord[8], Nicole Takahashi[9], Kerri P. Nowell[9,10], Elizabeth Will[11], Jane Roberts[11] and Somer L. Bishop[1]

[1]Department of Psychiatry and Behavioral Sciences, Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, United States, [2]Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, [3]Neurodevelopmental and Behavioral Phenotyping Service, National Institute of Mental Health, Bethesda, MD, United States, [4]Department of Pediatrics, University of Minnesota, Minneapolis, MN, United States, [5]Masonic Institute for the Developing Brain, University of Minnesota, Minneapolis, MN, United States, [6]Center for Autism and the Developing Brain, Weill Cornell Medical College, White Plains, NY, United States, [7]Offord Centre for Child Studies, McMaster University, Hamilton, ON, Canada, [8]UCLA Semel Institute for Neuroscience & Human Behavior, Center for Autism Research and Treatment, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [9]Thompson Center for Autism and Neurodevelopmental Disorders, University of Missouri, Columbia, MO, United States, [10]Department of Health Psychology, University of Missouri, Columbia, MO, United States, [11]Department of Psychology, University of South Carolina, Columbia, SC, United States

Increasing numbers of children with known genetic conditions and/or intellectual disability are referred for evaluation of autism spectrum disorder (ASD), highlighting the need to refine autism symptom measures to facilitate differential diagnoses in children with cognitive and language impairments. Previous studies have reported decreased specificity of ASD screening and diagnostic measures in children with intellectual disability. However, little is known about how cognitive and language abilities impact the measurement of specific ASD symptoms in this group. We aggregated a large sample of young children (N = 1196; aged 31–119 months) to examine measurement invariance of ASD symptoms among minimally verbal children within the context of the Autism Diagnostic Observation Schedule (ADOS) Module 1. Using confirmatory factor analysis (CFA) and moderated non-linear factor analysis (MNLFA), we examined how discrete behaviors were differentially associated with the latent symptom domains of social communication impairments (SCI) and restricted and repetitive behaviors (RRB) across spoken language levels and non-verbal mental age groupings. While the two-factor structure of SCI and RRB held consistently across language and cognitive levels, only partial invariance was observed for both ASD symptom domains of SCI and RRB. Specifically, four out of the 15 SCI items and one out of the three RRB items examined showed differential item functioning between children with "Few to No Words" and those with "Some Words"; and one SCI item and one RRB item showed differential item functioning across non-verbal mental age groups. Moreover, even after adjusting for the differential item functioning to reduce measurement bias across groups, there were still differences in ASD symptom domain scores across spoken language levels. These findings further underscore the influence of spoken language level on measurement of ASD symptoms and the importance of measuring ASD symptoms within refined spoken language levels, even among those with minimal verbal abilities.

KEYWORDS

autism symptoms, measurement invariance, language level, non-verbal mental age, ADOS

Introduction

Evidence of social communication impairments (SCI) and restricted and repetitive behaviors (RRB) is required for a diagnosis of autism spectrum disorder (ASD) (American Psychiatric Association., 2013). However, symptoms in these two domains occur commonly in children with a range of other neurodevelopmental disorders (NDDs), greatly complicating differential diagnosis (Grzadzinski et al., 2011; Hepburn and Moody, 2011; Bishop et al., 2019; Lord and Bishop, 2021). Differential diagnosis of ASD is especially challenging in the context of intellectual disability (ID) (Moss and Howlin, 2009; Thurm et al., 2019). By definition, children with ID exhibit delays in social communication relative to same-aged peers (American Psychiatric Association., 2013), and they often present with RRBs (Evans and Gray, 2000; Moss et al., 2009; Burbidge et al., 2010; Wolff et al., 2012; Hoch et al., 2016). Not surprisingly, therefore, children with lower IQ/mental age often receive elevated scores on ASD symptom measures, regardless of whether they ultimately receive a clinical diagnosis of ASD (Havdahl et al., 2016).

Decreased specificity (i.e., higher false positive rate) of commonly used diagnostic instruments such as the Autism Diagnostic Interview-Revised (ADI-R) (Lord et al., 1994; Rutter et al., 2005) and Autism Diagnostic Observation Schedule (ADOS)/ADOS-2 (Lord et al., 2000, 2012) is particularly pronounced among children with very low mental ages and/or non-verbal IQ below 50 (Risi et al., 2006). Thus, the authors have cautioned against interpreting scores in children with non-verbal mental ages below 15–18 months for the ADOS and below 24 months for the ADI-R (Lord et al., 1994, 2012). Nevertheless, these measures are still widely applied in clinical and research samples of children with very low levels of language and cognitive abilities. Especially as DSM-5 now explicitly allows for the diagnosis of ASD with a range of other conditions, a growing number of children with known genetic diagnoses, many of whom have severe to profound intellectual disability (ID), are being referred for assessment of ASD (Hepburn and Moody, 2011; King et al., 2014; Richards et al., 2015; Abbeduto et al., 2019).

Understanding how cognitive and/or language ability affects the measurement of ASD symptoms has implications for clinical practice and research involving children with ASD, other NDDs, and/or genetic conditions (Thurm et al., 2019). Inaccurate diagnosis may lead to delayed or inappropriate clinical services, and in the research context, presents a serious threat to the validity of ASD case vs. control studies. Further, if measures systematically provide higher or lower symptom scores for individuals with certain characteristics (regardless of ASD status), the score differences will fail to represent true differences in abilities/impairments across groups. Numerous studies have established that both language and cognitive ability influence the manifestation of ASD-related symptoms, which in turn may affect accuracy of classifications yielded by ASD symptom measures in certain groups (Risi et al., 2006; Corsello et al., 2007; Gotham et al., 2007; Kim and Lord, 2012; Hus et al., 2013; Havdahl et al., 2016). However, there is much less work on how specific aspects of ASD symptom measurement are affected by developmental and/or language level. This information is needed to increase precision of measurement of ASD symptoms in the context of extreme developmental variability that characterizes NDD clinical and research populations.

Examining measurement invariance (MI)/differential item functioning (DIF) across groups defined based on certain characteristics is one way to advance ASD measurement in this area. MI refers to "the situation in which scales provide the same results across different samples or populations" (Zedeck, 2014, p. 211), which is a critical property of measures that allows factor scores to be compared meaningfully across groups. MI is often tested in a stepwise fashion with increasingly strict standards for equivalence. Specifically, MI is commonly tested across three levels of equivalence: (1) configural invariance of the number of factors and loading pattern, (2) metric invariance of the factor loadings, which reflect the strength of the associations between the items and the factors (i.e., latent constructs), and (3) scalar invariance of the intercepts, which indicates the means of item scores across groups were reflective of means of the latent construct. Adequate MI is established by demonstrating that constraints on each of the parameters described above do not significantly worsen model fit. For more information on MI and differential item functioning, please see Widaman and Reise (1997), Teresi and Fleishman (2007), and Bauer et al. (2020).

In recent years, multiple studies on MI/DIF have been carried out with different ASD symptom measures, including the Childhood Autism Rating Scale (CARS) (Schopler et al., 1988, 2010), ADOS (Lord et al., 2000, 2012), Social Responsiveness Scale, and ADI-R (Constantino, 2005; Constantino and Gruber, 2012). These studies primarily focused on the effects of race/ethnicity, sex/gender and chronological age on scores (ADOS: Harrison et al., 2017; Ronkin et al., 2021; Burrows et al., 2022; Kalb et al., 2022; CARS: Stevanovic et al., 2021; SRS and ADI-R: Frazier and Hardan, 2017), with a few studies also investigating MI across groups with or without ID (Sturm et al., 2017; Dovgan et al., 2019). While these studies provided preliminary evidence that ASD symptom measures should take the impact of cognitive abilities into account, understanding of how cognitive or language abilities influence the measurement of specific ASD symptoms is still limited. Thus, the current study chose to focus on children with developmental delays to clarify the impact of finer divisions of cognitive and language abilities on the measurement of ASD symptom domains within this population. This information is necessary to improve the measurement of ASD symptoms within this special group, wherein differential diagnosis of ASD is especially challenging.

The ADOS is one of the most commonly used measures in the diagnostic assessment of ASD. Module 1 is designed for individuals with chronological age over 31 months who are not yet using flexible phrase speech; thus, children receiving Module 1 present with clinically significant delays in language and/or overall development. However, even among this group, there is substantial variability in age and non-verbal cognitive ability, as well as in expressive language ability (i.e., from no word approximations or words to beginning use of multiple word combinations). Therefore, examining MI of the latent constructs of ASD symptom domains in the context of the ADOS Module 1 provides a unique opportunity to elucidate the impact of mental age and spoken language level on the measurement of ASD symptoms in children with developmental delays.

Materials and methods

Participants

Data for the current analyses were aggregated from multiple sites to obtain a large sample of children who received ADOS Module 1 as part of a comprehensive diagnostic evaluation. Participants were included in the current analysis if they: (1) were between 31 and119 months at the time of ADOS administration; (2) had undergone a comprehensive diagnostic evaluation to determine a best-estimate diagnosis of ASD or another non-ASD NDD; (3) had complete data on the selected items from ADOS Module 1; (4) received a developmental/cognitive assessment at the time of the ADOS Module 1 administration; and (5) had cognitive assessment information available for the calculation of non-verbal age equivalents. This resulted in 1043 children with ASD and 153 without ASD from seven sites (see **Supplementary material** for details about data sources and sample aggregation). **Table 1** shows the demographic characteristics of the study sample.

Measures

The Autism Diagnostic Observation Schedule (Lord et al., 2000, 2012) is a standardized, semi-structured observational assessment designed to elicit social communication and restricted and repetitive behaviors associated with a diagnosis of ASD. It was designed to accommodate the assessment of ASD symptoms across language levels, with developmentally appropriate activities and codes organized into Modules (Lord et al., 2000, 2012). In the current analysis, we only included participants who were administered Module 1, designed for individuals who do not yet use flexible phrase speech. Consistent with scoring conventions, item scores of 0,1, and 2 were included in the analysis as they were, scores of 3 were converted to 2s for analysis, and scores of 8 ("Not applicable") and 9 ("Unknown") were converted to 0s.

As reflected in DSM-5 diagnostic criteria for ASD (American Psychiatric Association., 2013), previous factor analyses of the ADOS have consistently identified two core symptom domains of SCI and RRB (Gotham et al., 2007, 2008; Huerta et al., 2012; Harrison et al., 2017). Therefore, the current analyses focused on a subset of items mapping onto the two latent constructs of interest, SCI and RRB. Items on play (Section C) and other abnormal behaviors (Section E) were excluded. We also excluded the following items, as they were later added in the ADOS-2 and therefore missing for older cases who received ADOS-G: B13a *Amount of Social Overtures/Maintenance to Attention: Examiner*; B13b *Amount of Social Overtures/Maintenance to Attention: Parent/Caregiver*; B14 *Quality of Social Response*; B15 *Level of Engagement*; *B16*

TABLE 1 Sample characteristics.

		Non-ASD (*n* = 153)	ASD (*n* = 1043)
ADOS Module 1 language levels	Few to No word	49(32%)	474 (45.5%)
	Some words	104 (68%)	569 (54.5%)
Non-verbal mental age[a]	Below 24 months	62 (40.5%)	244 (23.4%)
	24 months and above	91 (59.5%)	799 (76.6%)
Sex	Male	104 (68.0%)	847 (81.2%)
	Female	49 (32.0%)	196 (18.8%)
Race	White	109 (71.2%)	688 (66.0%)
	Black	15 (9.8%)	116 (11.1%)
	AAPI	4 (2.6%)	71 (6.8%)
	AIAN	1 (0.7%)	5 (0.5%)
	Other	18 (11.8%)	116 (11.1%)
	Missing	6 (3.9%)	47 (4.5%)
Ethnicity	Hispanic	19 (12.4%)	117 (11.2%)
	Non-Hispanic	117 (76.5%)	868 (83.2%)
	Missing	17 (11.1%)	58 (5.6%)
Primary diagnosis[b]	Down Syndrome	37 (24.2%)	
	Language Disorders	21 (13.7%)	
	ID unknown etiology	15 (9.8%)	
	Fragile X Syndrome	11 (7.2%)	
	Williams Syndrome	7 (4.6%)	
	Global Developmental Delay	5 (3.3%)	
	Others	3 (2.0%)	
	Not Specified	54 (35.3%)	
		N, Mean (SD), Range	**N, Mean (SD), Range**
	Age	153, 46.15 (13.94), 31-115	1043, 62.11 (22.32), 31-119
	Non-verbal mental age	153, 25.89 (8.01), 6-58	1043, 32.61 (14.29), 2-104
	Non-verbal IQ	133, 59.85 (21.44), 13-133	1037, 55.66 (20.52), 2-144
	Verbal IQ	132, 52.70 (20.63), 11.83-110	1025, 38.26 (20.26), 3-103

[a]NVMA $<$ 15months: $N_{(non-ASD)}$ = 8 (5.2%), $N_{(ASD)}$ = 26 (2.5%); 15 months $\leq$ NVMA $<$ 18 months: $N_{(non-ASD)}$ = 12 (7.8%), $N_{(ASD)}$ = 38 (3.6%).
[b]Other primary diagnoses for the non-ASD group include one Cerebral Palsy, one Behavioral Disorder, and one genetic syndrome. Cases from all data sources have clinical best-estimate diagnoses of ASD and non-ASD, but some did not have primary diagnosis information available.

Overall Quality of Rapport. Item A6 *Use of Another's Body* was excluded as, unlike the other SCI items, it reflects the presence of abnormal behavior rather than the absence of developmentally expected behavior. For RRB, we excluded items that were dependent on sufficient spoken language to exhibit the abnormality (A3 *Intonation of Vocalizations*, A4 *Immediate Echolalia*, A5 *Stereotyped/Idiosyncratic Use of Words or Phrases*). We also excluded Item D3 *Self-Injurious Behavior* due to an extremely low rate of endorsement (<9% endorsing 1s or 2s). In total, 15 items assessing SCI and three items assessing RRB were included in the analyses (see **Table 2**).

Spoken Language Level. We derived the language level classification based on Item A1 "*Overall Level of Non-Echoed Spoken Language*" from the ADOS Module 1. Consistent with instructions for use of the revised algorithms (Gotham et al., 2007), participants who received scores of 3 or 4 were assigned to "Few to No words" and participants who received scores of 0, 1, or 2 were assigned to "Some words" group. The validity of these spoken language groups is further supported by previous studies showing differences between "Few to No Words" and "Some Words" on other measures of expressive language and cognitive ability (Bal et al., 2016; Mazurek et al., 2019).

Non-verbal mental age. Participants included in the aggregated dataset were administered at least one measure of cognitive ability based on site-specific protocols and/or clinician judgment about the developmentally appropriate test: the Mullen Scales of Early Learning (MSEL; (Mullen, 1995), the Differential Ability Scales (DAS) (Elliott, 2007), and/or the Merrill-Palmer Scales of Development (Roid and Sampers, 2004). The MSEL was used for 89% of the non-ASD sample and 75% of the ASD sample. For each participant, a non-verbal mental age was derived based on averaging available age equivalents from the non-verbal subtests. For those who received the MSEL, the age equivalents from the Fine Motor and

TABLE 2 ADOS Module 1 Items included in the analyses.

	Item level	Item description
Social communication impairments	A2	*Frequency of Spontaneous Vocalization Directed to Others*
	A7	*Pointing*
	A8	*Gestures*
	B1	*Unusual Eye Contact*
	B2	*Responsive Social Smile*
	B3	*Facial Expressions Directed to Others*
	B4	*Integration of Gaze and Other Behaviors During Social Overtures*
	B5	*Shared Enjoyment in Interaction*
	B6	*Response to Name*
	B7	*Requesting*
	B8	*Giving*
	B9	*Showing*
	B10	*Spontaneous Initiation of Joint Attention*
	B11	*Response to Joint Attention*
	B12	*Quality of Social Overtures*
Repetitive behaviors and restricted interests	D1	*Unusual Sensory Interests in Play Material/Person*
	D2	*Hand and Finger and Other Complex Mannerisms*
	D4	*Unusually Repetitive Interests or Stereotyped Behaviors*

For detailed item description and scoring instruction of each item, please refer to ADOS Module 1 scoring protocol (Lord et al., 2012).

Visual Reception subscales were averaged to represent NVMA (see Bishop et al., 2011; Farmer et al., 2016).

We dichotomized the NVMA variable to NVMA under 24 months vs. NVMA of 24 months and above for both practical and theoretical reasons: (1) given different tests were administered across sites, the binary categories will achieve more reliable grouping by avoiding the point estimates of the NVMA; (2) the cut point at 24 months allows sufficient sample sizes in both groups; (3) moreover, 24 months is an age at which children would be expected to use phrase speech in typical development (Sheldrick et al., 2019); thus, children with a non-verbal mental age above 24 months who receive Module 1 (rather than Module 2 or 3) show evidence of a discrepancy between their non-verbal mental age and their spoken language level. Therefore, we might expect that items developed for children with a very low spoken language level (i.e., language abilities characteristic of children under 24 months) might function differently in those with higher NVMA.

Best Estimate Diagnosis of ASD. All participants underwent multi-disciplinary evaluations by experienced clinicians and/or researchers who had established and maintained research reliability on the ADOS/ADOS-2. Best-estimate clinical diagnoses of ASD or the absence of ASD (i.e., Non-ASD) were determined based on all available information, including parent interviews of developmental history and direct observation of ASD symptoms (including the ADOS), as well as tests of cognitive and adaptive functioning.

Statistical analyses

Confirmatory factor analyses

Separate CFA with two factors (SCI and RRB; see **Table 2** for ADOS Module 1 item mapping onto the two factors) were conducted across two spoken language level groups and two NVMA groups, respectively, to examine configural invariance (i.e., the number of factors and loading pattern) (Widaman and Reise, 1997). Factor analyses were conducted in *Mplus* with WLSMV estimator for ordered categorical variables. The chi-square statistics, comparative fit index (CFI), the root mean square error of approximation (RMSEA) and its 90% confidence interval (CI), and the standardized root mean square residual (SRMR) were examined for CFA model fit, with CFI larger than 0.95, RMSEA and SRMR smaller than 0.08 indicating a good fit (Hu and Bentler, 1999).

Moderated non-linear factor analysis

Once configural invariance was established through the CFA across NVMA and spoken language level groups, we proceeded to examine higher levels of structural validity testing of the two latent constructs across covariate groupings of interests (both of which were analyzed using effects coding): spoken language level (i.e., −1 = Few to No words vs. 1 = Some words) and developmental level (i.e., NVMA: −1 = under 24 months vs. 1 = 24 months and above). Moderated Non-linear Factor Analysis (MNLFA) is similar to both the multiple-group CFA and the multiple-indicator multiple-cause (MIMIC) methods for evaluating measurement invariance, but it extends both to multiple groups, categorical or count data, and the inclusion of multiple grouping variables at the same time. In the MNLFA model, MI/DIF is viewed as a form of parameter moderation; and thus, tested in the model for statistical significance as moderators of factor and item parameters. That is, moderation of the intercepts would indicate uniform DIF, whereas moderation of the factor loadings would indicate non-uniform DIF. We recommend that interested readers refer to Bauer (2017) for more details. Since MNLFA only accommodates unidimensional factor structure, separate analyses were conducted for SCI and RRB. The MNLFA method allows testing of the impact of spoken language levels and NVMA groups at the same time on the mean and variances of latent constructs, as well as their impacts on the intercept and loading of each item on the latent constructs. MNLFA involves an iterative process where each item is tested independently, then the significant ($p < 0.05$) effects are retained and tested simultaneously in one model. Lastly, a final model was estimated using the statistically significant parameters after the Benjamini-Hochberg false discovery rate correction to adjust for multiple comparisons. Moderated item effects were examined and reported to understand the impact of NVMA and spoken language level. The resulting model was then used to estimate the factor scores of the two latent constructs of SCI and RRB.

We employed an updated version of the R package *aMNLFA* Version 1.1.2 (Gottfredson et al., 2019)[1] to streamline the generation of the *MPlus* codes and automate the process of integrating all effects into one model. We carefully reviewed and modified the automated *MPlus* codes to fit our dataset and research questions.

While there are multiple measures of impact of DIF on the overall measurement of the latent constructs (Meade, 2010), no recommended metric is available for the assessment of overall differential test functioning (DTF) in the context of MNLFA with simultaneous testing of multiple grouping variables. Therefore, to evaluate the differences between DIF-adjusted latent construct scores based on the group-specific information and the factor scores of latent constructs assuming full measurement invariance, we chose to adapt the Root Expected Mean Square Difference (REMSD) which was developed to index subpopulation invariance of linking and equating relationships (Dorans and Holland, 2000). Although MNLFA and equating analyses are distinct, the contrast between group-specific (i.e., with DIF) and overall (i.e., invariant) item parameters in the MNLFA context is comparable to the group-specific and overall equating relationship from which the REMSD statistic was originally derived. The adapted REMSD metric was calculated as the square root of the expected value of squared differences between the DIF-adjusted latent construct scores (FS_{mnlfa}) and the factor scores assuming full measurement invariance of the item parameters (FS_{FI}), divided by the standard deviation of the latent factor score (fixed to 1):

$$REMSD = \sqrt{E\left(\left(FS_{mnlfa} - FS_{FI}\right)^2\right)}/\sigma_{FS}$$

Further, effect sizes (Cohen's *d*) were calculated for group comparisons of the latent construct factor scores.

Results

The majority of the aggregated sample was diagnosed with ASD, which is expected given the data were mostly drawn from autism specialty clinics or research projects focused on ASD (see **Table 1**). The descriptive statistics showed that, compared to those without ASD, children in the ASD group were more likely to be male ($\chi^2 = 13.05$, $p = 0.003$), to have "Few to No words" ($\chi^2 = 10.02$, $p = 0.002$), and to have an NVMA of 24 months and over ($\chi^2 = 18.92$, $p < 0.001$).

The two-factor structure with SCI and RRB showed a good fit, supporting configural invariance of the ADOS across the two spoken language levels and the two NVMA groups (see **Table 3**). **Table 4** shows item factor loadings onto the two factors of SCI and RRB, respectively, across the two spoken language levels and two NVMA groups.

For each latent construct, ensuing MNLFAs were conducted separately. For the latent construct of SCI, we observed a significant effect of spoken language level on the measured SCI scores (Estimate = −0.45, SE = 0.034, $p < 0.001$), with individuals with Few to No words showing higher levels of SCI symptoms. Multiple items showed loading and intercept DIF across language levels on the latent construct of SCI, including Unusual Eye Contact, Integration of Gaze and Other Behaviors during Social Overtures, Requesting, and Showing. Only one item, Frequency of Vocalization, showed significant loading and intercept DIF across the NVMA groups on the SCI (see **Table 5** upper panel for parameter estimates and **Figure 1** for the final SCI measurement model). For the latent construct of RRB, the mean level of measured RRB differed across language levels (Estimate = −0.249, SE = 0.046, $p < 0.001$). There were also loading DIFs of Item "Hand/finger and Other Complex Mannerisms" across spoken language levels and "Unusually Repetitive Interests or Stereotyped Behaviors" across NVMA groups (see **Table 5** bottom panel for parameter estimates and **Figure 2** for the final RRB measurement models of the two latent constructs). That is, these items show different levels of associations with the latent constructs of SCI and RRB, as well as varying item difficulties. In sum, metric invariance did not hold for several items on both SCI and RRB latent constructs, with subsets of items functioning differently across groups.

Item-level DIF has moderate to large impact on the score of the two latent constructs: $REMSD_{SCI} = 0.66$ and $REMSD_{RRB} = 0.74$, indicating the need to consider measurement bias in interpreting the measured scores of the two latent constructs. Effect sizes of the DIF-adjusted SCI factor scores indicated that children with "Few to No Words" scored about 1 standard deviation (SD) higher than those with "Some Words" in SCI severity factor scores (Cohen's $d = 1.01$); similarly, for the RRB factor scores, children with "Few to No Words" scored 0.75 SD higher. On the other hand, small ES were observed for the group comparisons across NVMA of both latent constructs (SCI: ES = 0.34, RRB: ES = 0.27).

Discussion

The current study was conducted to provide more explicit guidance about how the measurement of ASD symptoms (as indexed by selected items from Module 1 of the ADOS) might be affected by language and developmental level.

1 https://github.com/cran/aMNLFA

TABLE 3 Fit statistics of two-factor CFA models.

		χ^2 (df = 134)	CFI	SRMR	RMSEA
Non-verbal mental age	Below 24 months	260.02, $p < 0.001$	0.978	0.057	0.055 [0.045, 0.065]
	24 months and above	428.11, $p < 0.001$	0.975	0.048	0.050 [0.044, 0.055]
Language level	Few to No words	349.48, $p < 0.001$	0.956	0.063	0.055 [0.048, 0.063]
	Some words	345.50, $p < 0.001$	0.978	0.047	0.048 [0.042, 0.055]

TABLE 4 Item factor loadings on the two factors from CFA across the groups.

		Non-verbal mental age		Language level	
Factor names		**Below 24 months**	**24 months and above**	**Few to No words**	**Some words**
Social communication impairment scores	A2	0.91	0.79	0.81	0.79
	A7	0.77	0.70	0.65	0.67
	A8	0.68	0.61	0.60	0.64
	B1	0.91	0.94	0.85	0.99
	B2	0.56	0.55	0.48	0.55
	B3	0.79	0.76	0.73	0.77
	B4	0.88	0.77	0.82	0.78
	B5	0.75	0.65	0.63	0.69
	B6	0.51	0.45	0.41	0.49
	B7	0.73	0.70	0.69	0.67
	B8	0.59	0.55	0.54	0.55
	B9	0.87	0.78	0.85	0.74
	B10	0.78	0.76	0.70	0.75
	B11	0.58	0.54	0.48	0.48
	B12	0.89	0.87	0.87	0.85
Repetitive behaviors and restricted interests scores	D1	0.61	0.62	0.56	0.59
	D2	0.55	0.34	0.56	0.25
	D4	0.57	0.54	0.48	0.62

While decades of research indicate that both language and cognitive ability influence the manifestation and measurement of ASD-related symptoms, there is much less work about specific aspects of ASD symptom measurement that may be problematic when comparing children developmental and spoken language levels. Greater understanding of this issue is important given the extreme developmental heterogeneity that characterizes ASD and NDD clinical and research populations.

Consistent with previous studies of ASD symptom structure, which ultimately informed DSM-5 diagnostic criteria (Huerta et al., 2012; Frazier et al., 2014), findings from the current CFA of the ADOS Module 1indicate two core symptom domains (i.e., SCI and RRB). This structure held across spoken language levels and NVMA groupings, supporting configural invariance of the measure. However, when examining the mean levels of latent constructs for both SCI and RRB, children with "Few to No Words" scored systematically higher (i.e., more impairments) than those with "Some Words".

When looking at where the ASD symptom measurements showed biases, stricter levels of measurement invariance did not hold at the item level for some items in the MNLFA models for either SCI or RRB latent construct. For the measurement of SCI, loading and intercept DIF was observed for four items across spoken language levels [*Unusual Eye Contact* (B1), *Integration of Gaze and Other Behaviors during Social Overtures* (B4), *Requesting* (B7), and *Showing* (B9)], and one item [*Frequency of Vocalization*(A2)] across NVMA groups. All four SCI items that showed DIF across spoken language levels involved the use of eye contact with the examiner, highlighting the potential role of spoken language level even when measuring basic non-verbal social communication skills such as eye contact. Even though only a small subset of items (n = 5) showed any measurement bias on the latent construct SCI, the DIFs showed impact on the overall latent construct scores, underscoring the need to carefully consider

TABLE 5 Parameter estimates of the resulting MNLFA model.

Parameter type	Variables	Estimate	SE	*p*-value
Social communication impairments				
Intercept	ETA	0		
Loading	*Frequency of Vocalization* [a]	2.794	0.180	<0.001
	Pointing	1.624	0.100	<0.001
	Gestures	1.229	0.081	<0.001
	Unusual Eye Contact [a]	2.890	0.281	<0.001
	Responsive Social Smile	1.006	0.074	<0.001
	Facial Expressions Directed to Others	1.835	0.115	<0.001
	Integration of Gaze and Other Behaviors During Social Overtures [a]	2.517	0.163	<0.001
	Shared Enjoyment in Interaction	1.463	0.093	<0.001
	Response to Name	0.845	0.069	<0.001
	Requesting [a]	1.869	0.127	<0.001
	Giving	1.063	0.077	<0.001
	Showing [a]	2.075	0.144	<0.001
	Spontaneous Initiation of Joint Attention	1.744	0.110	<0.001
	Response to Joint Attention	1.128	0.083	<0.001
	Quality of Social Overtures	2.621	0.165	<0.001
Mean Factor	ETA on Language Levels	-0.450	0.034	<0.001
Intercept DIF	*Frequency of Vocalization* on NVMA	−0.369	0.102	<0.001
	Unusual Eye Contact on Language Levels	1.595	0.367	<0.001
	Integration of Gaze and Other Behaviors During Social Overtures on Language Levels	0.242	0.087	0.005
	Requesting on Language Levels	0.067	0.081	0.409
	Showing on Language Levels	−0.064	0.106	0.547
Loading DIF	*Frequency of Vocalization* on NVMA	−0.347	0.127	0.006
	Unusual Eye Contact on Language Levels	0.780	0.274	0.004
	Integration of Gaze and Other Behaviors During Social Overtures on Language Levels	−0.498	0.122	<0.001
	Requesting on Language Levels	−0.228	0.102	0.025
	Showing on Language Levels	−0.302	0.123	0.014
Repetitive behaviors and restricted interest				
Intercept	ETA	0		
Loading	*Unusual Sensory Interests in Play Material/Person*	1.697	0.302	<0.001
	Hand and Finger and Other Complex Mannerisms [a]	0.664	0.114	<0.001
	Unusually Repetitive Interests or Stereotyped Behaviors [a]	1.139	0.178	<0.001
Mean Factor	ETA on Language Levels	−0.258	0.047	<0.001
Intercept DIF	*Hand and Finger and Other Complex Mannerisms* on Language Levels	−0.219	0.069	0.001
	Unusually Repetitive Interests or Stereotyped Behaviors on NVMA	0.060	0.080	0.458
Loading DIF	*Hand and Finger and Other Complex Mannerisms* on Language Levels	−0.221	0.106	0.037
	Unusually Repetitive Interests or Stereotyped Behaviors on NVMA	−0.413	0.142	0.004

[a]Unweighted grand mean of loading across groups.
NVMA groups: −1 = under 24 months, 1 = 24 months and above; Language levels: −1 = Few to No Words; 1 = Some Words.

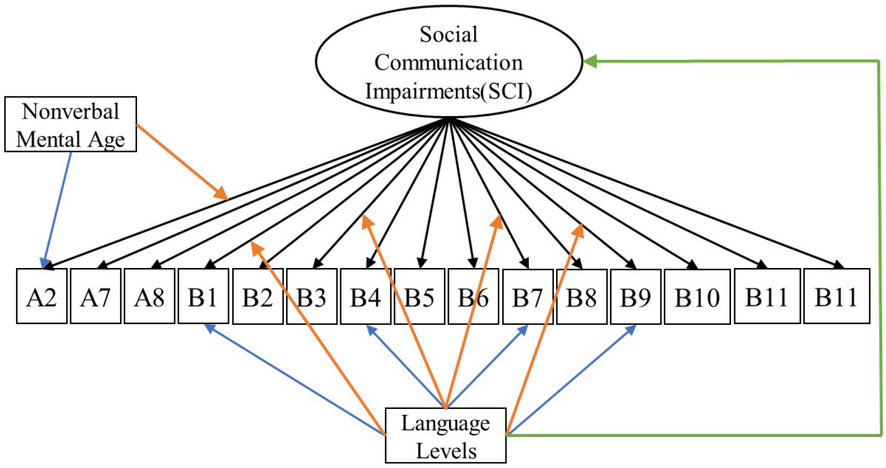

FIGURE 1
Measurement model for social communication impairments (SCI). Black arrows indicate factor loadings of each item examined on the SCI latent construct. Colored Arrows in the figure showing significant impact of the covariate on the factor and item parameters: (1) Green arrow represents the impact of language level on the mean of the latent construct; (2) Orange arrows represent the impact of covariates (NVMA and language level groups) on the relationships between the item and the latent construct (non-uniform DIF); (3) Blue arrows represent the impact of covariates on the levels of items when the overall level of the latent construct is similar across groups (uniform DIF). For specific item names, please refer to **Table 2**.

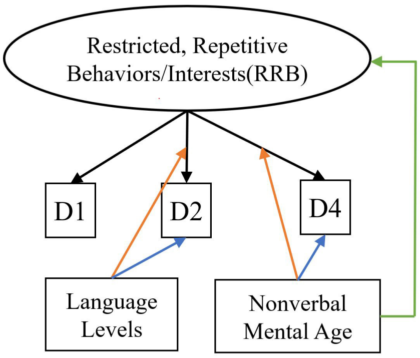

FIGURE 2
Measurement model for restricted, repetitive behaviors/interests. Black arrows indicate factor loadings of each item examined on the RRB latent construct. Colored Arrows in the figure showing significant impact of the covariate on the factor and item parameters: (1) Green arrow represents the impact of language level on the mean of the latent construct; (2) Orange arrows represent the impact of covariates (NVMA and language level groups) on the relationships between the item and the latent construct (non-uniform DIF); (3) Blue arrows represent the impact of covariates on the levels of items when the overall level of the latent construct is similar across groups (uniform DIF). For specific item names, please refer to **Table 2**.

the impact of spoken language levels when making score comparisons between individuals. On the other hand, two out of three RRB items (i.e., *Hand and Finger and Other Complex Mannerisms* and *Unusually Repetitive Interests or Stereotyped Behaviors*) included in the analyses showed bias across either spoken language or NVMA, indicating that the measurement of RRBs with only the three selected items is likely problematic. This is consistent with previous item response theory analyses done with ADOS Modules 3 and 4 (Kuhfeld and Sturm, 2018).

To further understand different levels of autism symptoms across spoken language levels, we compared SCI and RRB factor scores after adjusting for measurement biases identified at the item level, and found that they still differed significantly across spoken language levels, with higher severity scores seen in children with "Few to No Words". These findings suggest that there are likely true differences in the levels of SCI and RRB symptom severity, as measured using Module 1 of the ADOS, between children with "Few to No Words" vs. "Some Words". This provides further evidence for the decision to create separate algorithms based on finer language-level divisions within Module 1 (Gotham et al., 2007, 2008). Given that some items on the ADOS Module 1 function differently for children of different spoken language levels, even among those with minimal verbal abilities, clinicians and researchers should follow the algorithm guidelines to derive scores for the two spoken language levels separately and only interpret scores at the domain and scale levels, but not at the item level.

To our knowledge, this study is the first to examine MI of ASD symptoms within children with developmental delays across cognitive and spoken language levels. A deeper understanding of how ASD symptom measurement is affected by developmental level is critical, particularly given increased interests in behavioral phenotyping of rare genetic conditions, many of which are associated with severe to profound ID (Arvio and Sillanpää, 2003; Richards et al., 2015; Abbeduto et al., 2019; Burdeus-Olavarrieta et al., 2021). We focused on ADOS Module 1 to specifically home in on the effects of mental age and expressive language in children with lower cognitive and language abilities. However, this sample does not represent the full range of minimally verbal individuals who have even more severe delays. Importantly, valid administration of the ADOS requires that a child be able to walk, see, and hear at the time of assessment, meaning that it is not even valid for a significant proportion of children with severe to profound ID. Moreover, given the reduced specificity of the measure, the test developers advised against using the ADOS in children with NVMA below 15 months, resulting in very few such cases available for the current analyses: Non-ASD = 8 (5.2%), ASD = 26 (2.5%). Therefore, the present findings have limited applicability to individuals with severe to profound ID and/or sensory and motor impairments, and do not change the recommendation that ADOS scores may not be valid in this group. Yet, the fact remains that clinicians and researchers are increasingly faced with the challenges of assessing ASD symptoms in individuals for whom current measures were not validated, highlighting the need for empirical evidence to measure ASD symptoms validly and reliably in this population. Further, children develop over time and some gain cognitive and language skills as they grow and receive intervention. Thus, future longitudinal studies should examine intra-individual changes as children shift from "Few to No Words" to "Some Words" and/or from lower NVMA group to higher NVMA levels.

The current study represents a first step in understanding ASD symptom measurement for those who are minimally verbal. Even within Module 1, which is already only applicable to children within a relatively narrow developmental range, our findings highlight the need for finer divisions based on spoken language level (e.g., "Few to No Words" and "Some Words") and/or mental age to optimize measurement of ASD symptoms. Thus, to advance measurement of SCI and RRB in the extremely heterogeneous population of children with neurodevelopmental disorders, the field must work to enhance developmentally appropriate measurement strategies (Bishop et al., 2019). Moreover, it is imperative that clinicians and researchers implement best-practice methods for carefully considering developmental profiles, including cognitive and spoken language levels, in their assessment of ASD-related symptoms and behaviors.

Data availability statement

Publicly available datasets were analyzed in this study. Data from the Simons Simplex Collections is available to qualified researchers. Approved researchers can obtain the SSC population dataset described in this study by applying at https://base.sfari.org. Data from other sources can be requested by contacting the principal investigators and directors at each site.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants or their legal guardian/next of kin.

Author contributions

SZ and SB conceptualized and designed the study with consultation from AK, CF, and AT. SZ aggregated datasets, conducted statistical analyses, interpreted the results, and drafted the manuscript. AK and CB provided consultation on data analysis and result interpretation. SB together with AK, CF, and AT provided feedback on previous versions of the manuscript. AT, CB, SK, CL, AE, NT, KN, EW, JR, and SB contributed to data collection. SB secured funding for the current study. All authors provided feedback on and approved the final version of the manuscript.

Funding

This work was supported by grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD; R01HD093012 to SB), and in part by the Intramural Research Program of the NIMH (1ZICMH002961 to AT).

Acknowledgments

We are grateful to all of the families at the participating Simons Simplex Collection (SSC) sites, as well as the principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, E. Wijsman). We appreciate obtaining access to phenotypic data on the Simons Foundation Autism Research Initiative (SFARI) Base. Approved researchers can obtain the

SSC population dataset described in this study by applying at https://base.sfari.org. We are also grateful to all the clinicians (Robin Rumsey, Desirae Rambeck, Chimei Lee, Rebekah Hudock, and Jane Nofer) from the University of Minnesota Autism and Neurodevelopment Clinic and all other clinics who contributed to the data collection.

Conflict of interest

CL and SB have received royalties from the ADOS-2, and all profits are donated to charity.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.927847/full#supplementary-material

References

Abbeduto, L., Thurman, A. J., McDuffie, A., Klusek, J., Feigles, R. T., Brown, W. T., et al. (2019). ASD comorbidity in Fragile X Syndrome: symptom profile and predictors of symptom severity in adolescent and young adult males. *J. Autism Dev. Disord.* 49, 960–977. doi: 10.1007/s10803-018-3796-2

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Washington, DC: American Psychiatry Publishing.

Arvio, M., and Sillanpää, M. (2003). Prevalence, aetiology and comorbidity of severe and profound intellectual disability in Finland. *J. Intellect. Disabil. Res.* 47, 108–112. doi: 10.1046/j.1365-2788.2003.00447.x

Bal, V. H., Katz, T., Bishop, S. L., and Krasileva, K. (2016). Understanding definitions of minimally verbal across instruments: Evidence for subgroups within minimally verbal children and adolescents with autism spectrum disorder. *J. Child Psychol. Psychiatry* 57, 1424–1433. doi: 10.1111/jcpp.12609

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychol. Methods* 22, 507–526. doi: 10.1037/met0000077

Bauer, D. J., Belzak, W. C. M., and Cole, V. T. (2020). Simplifying the assessment of measurement invariance over multiple background variables: using Regularized Moderated Nonlinear Factor Analysis to detect differential item functioning. *Struct. Equ. Model.* 27, 43–55. doi: 10.1080/10705511.2019.1642754

Bishop, S. L., Guthrie, W., Coffing, M., and Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the Differential Ability Scales in children with autism spectrum disorders. *Am. J. Intellect. Dev. Disabil.* 116, 331–343. doi: 10.1352/1944-7558-116.5.331

Bishop, S., Farmer, C., Kaat, A., Georgiades, S., Kanne, S., and Thurm, A. (2019). The need for a developmentally based measure of social-communication skills. *J. Am. Acad. Child Adolesc. Psychiatry* 58, 555–560. doi: 10.1016/j.jaac.2018.12.010

Burbidge, C., Oliver, C., Moss, J., Arron, K., Berg, K., Furniss, F., et al. (2010). The association between repetitive behaviours, impulsivity and hyperactivity in people with intellectual disability. *J. Intellect. Disabil. Res.* 54, 1078–1092. doi: 10.1111/j.1365-2788.2010.01338.x

Burdeus-Olavarrieta, M., San José-Cáceres, A., García-Alcón, A., González-Peñas, J., Hernández-Jusdado, P., and Parellada-Redondo, M. (2021). Characterisation of the clinical phenotype in Phelan-McDermid syndrome. *J. Neurodev. Disord.* 13:26. doi: 10.1186/s11689-021-09370-5

Burrows, C. A., Grzadzinski, R. L., Donovan, K., Stallworthy, I. C., Rutsohn, J., St. John, T., et al. (2022). A data driven approach in an unbiased sample reveals equivalent sex ratio of autism spectrum disorder associated impairment in early childhood. *Biol. Psychiatry* doi: 10.1016/j.biopsych.2022.05.027

Constantino, J. (2005). *Social Responsiveness Scale*. Torrance, CA: Western Psychological Services.

Constantino, J. N., and Gruber, C. P. (2012). *Social Responsiveness Scale, Second Edition (SRS-2)*. Los Angeles, CA: Western Psychological Services.

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook, E. H., Leventhal, B. L., et al. (2007). Between a ROC and a hard place: Decision making and making decisions about using the SCQ. *J. Child Psychol. Psychiatry* 48, 932–940. doi: 10.1111/j.1469-7610.2007.01762.x

Dorans, N. J., and Holland, P. W. (2000). Population invariance and the equatability of tests: basic theory and the linear case. *J. Educ. Measur.* 37, 281–306.

Dovgan, K., Mazurek, M. O., and Hansen, J. (2019). Measurement invariance of the child behavior checklist in children with autism spectrum disorder with and without intellectual disability: follow-up study. *Res. Autism Spectr. Disord.* 58, 19–29. doi: 10.1016/j.rasd.2018.11.009

Elliott, C. D. (2007). *Differential Ability Scales*, 2nd Edn. San Antonio, TX: Harcourt Assessment.

Evans, D. W., and Gray, F. L. (2000). Compulsive-like behavior in individuals with Down Syndrome: its relation to mental age level, adaptive and maladaptive behavior. *Child Dev.* 71, 288–300. doi: 10.1111/1467-8624.00144

Farmer, C., Golden, C., and Thurm, A. (2016). Concurrent validity of the differential ability scales, second edition with the Mullen Scales of Early Learning in young children with and without neurodevelopmental disorders. *Child Neuropsychol.* 22, 556–569. doi: 10.1080/09297049.2015.1020775

Frazier, T. W., and Hardan, A. Y. (2017). Equivalence of symptom dimensions in females and males with autism. *Autism* 21, 749–759. doi: 10.1177/1362361316660066

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., and Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism* 18, 31–44. doi: 10.1177/1362361313500382

Gotham, K., Risi, S., Dawson, G., Tager-flusberg, H., Joseph, R., Carter, A., et al. (2008). A Replication of the Autism Diagnostic Observation Schedule (ADOS) Revised Algorithms. *J. Am. Acad. Child Adolesc. Psychiatry* 47, 642–651. doi: 10.1097/CHI.0b013e31816bffb7

Gotham, K., Risi, S., Pickles, A., and Lord, C. (2007). The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *J. Autism Dev. Disord.* 37, 613–627. doi: 10.1007/s10803-006-0280-1

Gottfredson, N. C., Cole, V. T., Giordano, M. L., Bauer, D. J., Hussong, A. M., and Ennett, S. T. (2019). Simplifying the implementation of modern scale scoring methods with an automated R package: automated moderated nonlinear factor analysis (aMNLFA). *Addict. Behav.* 94, 65–73. doi: 10.1016/j.addbeh.2018.10.031

Grzadzinski, R., Di Martino, A., Brady, E., Mairena, M. A., O'Neale, M., Petkova, E., et al. (2011). Examining autistic traits in children with ADHD: does the autism spectrum extend to ADHD? *J. Autism Dev. Disord.* 41, 1178–1191. doi: 10.1007/s10803-010-1135-3

Harrison, A. J., Long, K. A., Tommet, D. C., and Jones, R. N. (2017). Examining the role of race, ethnicity, and gender on social and behavioral ratings within the Autism Diagnostic Observation Schedule. *J. Autism Dev. Disord.* 47, 2770–2782. doi: 10.1007/s10803-017-3176-3

Havdahl, K. A., Hus Bal, V., Huerta, M., Pickles, A., Øyen, A.-S., Stoltenberg, C., et al. (2016). Multidimensional influences on autism symptom measures: Implications for use in etiological research. *J. Am. Acad. Child Adolesc. Psychiatry* 55, 1054-1063.e3. doi: 10.1016/j.jaac.2016.09.490

Hepburn, S. L., and Moody, E. J. (2011). Diagnosing autism in individuals with known genetic syndromes: clinical considerations and implications for intervention. *Int. Rev. Res. Dev. Disabil.* 40, 229–259. doi: 10.1016/B978-0-12-374478-4.00009-5

Hoch, J., Spofford, L., Dimian, A., Tervo, R., MacLean, W. E., and Symons, F. J. (2016). A direct comparison of self-injurious and stereotyped motor behavior between preschool-aged children with and without developmental delays. *J. Pediatr. Psychol.* 41, 566–572. doi: 10.1093/jpepsy/jsv102

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Huerta, M., Bishop, S. L., Duncan, A., Hus, V., and 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. *Am. J. Psychiatry* 169, 1056–1064. doi: 10.1176/appi.ajp.2012.12020276

Hus, V., Bishop, S., Gotham, K., Huerta, M., and Lord, C. (2013). Factors influencing scores on the social responsiveness scale. *J. Child Psychol. Psychiatry* 54, 216–224. doi: 10.1111/j.1469-7610.2012.02589.x

Kalb, L. G., Singh, V., Hong, J. S., Holingue, C., Ludwig, N. N., Pfeiffer, D., et al. (2022). Analysis of race and sex bias in the Autism Diagnostic Observation Schedule (ADOS-2). *JAMA Pediatr.* 5:e229498. doi: 10.1001/jamanetworkopen.2022.9498

King, B. H., de Lacy, N., and Siegel, M. (2014). Psychiatric assessment of severe presentations in autism spectrum disorders and intellectual disability. *Child Adolesc. Psychiatr. Clin. North Am.* 23, 1–14. doi: 10.1016/j.chc.2013.07.001

Kim, S. H., and Lord, C. (2012). Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and young preschoolers from 12 to 47 months of age. *J. Child Psychol. Psychiatry* 53, 143–151. doi: 10.1111/j.1469-7610.2011.02458.x

Kuhfeld, M., and Sturm, A. (2018). An examination of the precision of the Autism Diagnostic Observation Schedule using item response theory. *Psychol. Assess.* 30, 656–668. doi: 10.1037/pas0000512

Lord, C., and Bishop, S. L. (2021). Let's Be Clear That "Autism Spectrum Disorder Symptoms". Are Not Always Related to Autism Spectrum Disorder. *Am. J. Psychiatry* 178, 680–682. doi: 10.1176/appi.ajp.2021.21060578

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule—Generic: a Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *J. Autism Dev. Disord.* 30, 205–223. doi: 10.1023/A:1005592401947

Lord, C., Rutter, M., and Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* 24, 659–685. doi: 10.1007/BF02172145

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., and Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part I): Modules 1-4*. Torrance, CA: Western Psychological Services.

Mazurek, M. O., Baker-Ericzén, M., and Kanne, S. M. (2019). Brief Report: Calculation and convergent and divergent validity of a new ADOS-2 expressive language score. *Am. J. Intellect. Dev. Disabil.* 124, 438–449. doi: 10.1352/1944-7558-124.5.438

Meade, A. W. (2010). A taxonomy of effect size measures for the differential functioning of items and scales. *J. Appl. Psychol.* 95, 728–743. doi: 10.1037/a0018966

Moss, J., and Howlin, P. (2009). Autism spectrum disorders in genetic syndromes: implications for diagnosis, intervention and understanding the wider autism spectrum disorder population. *J. Intellect. Disabil. Res.* 53, 852–873. doi: 10.1111/j.1365-2788.2009.01197.x

Moss, J., Oliver, C., Arron, K., Burbidge, C., and Berg, K. (2009). The prevalence and phenomenology of repetitive behavior in genetic syndromes. *J. Autism Dev. Disord.* 39, 572–588. doi: 10.1007/s10803-008-0655-6

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: AGS.

Richards, C., Jones, C., Groves, L., Moss, J., and Oliver, C. (2015). Prevalence of autism spectrum disorder phenomenology in genetic disorders: a systematic review and meta-analysis. *Lancet Psychiatry* 2, 909–916. doi: 10.1016/S2215-0366(15)00376-4

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., et al. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *J. Am. Acad. Child. Adolesc. Psychiatry* 45, 1094–1103. doi: 10.1097/01.chi.0000227880.42780.0e

Roid, G., and Sampers, J. (2004). *Merrill-Palmer Revised Scales of Development*. Wood Dale, IL: Stoelting Co.

Ronkin, E., Tully, E. C., Branum-Martin, L., Cohen, L. L., Hall, C., Dilly, L., et al. (2021). Sex differences in social communication behaviors in toddlers with suspected autism spectrum disorder as assessed by the ADOS-2 toddler module. *Autism* [Epub ahead of print]. doi: 10.1177/13623613211047070

Rutter, M., Le Couteur, A., and Lord, C. (2005). *Autism Diagnostic Interview-Revised*. Los Angeles, CA: Western Psychological Services.

Schopler, E., Reichler, R. J., and Renner, B. R. (1988). *The Childhood Autism Rating Scale (CARS)*. Los Angeles, CA: Western Psychological Services.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., and Love, S. R. (2010). *The Childhood Autism Rating Scale, Second Edition (CARS 2)*. Los Angeles, CA: Western Psychological Services.

Sheldrick, R. C., Schlichting, L. E., Berger, B., Clyne, A., Ni, P., Perrin, E. C., et al. (2019). Establishing new norms for developmental milestones. *Pediatrics* 144:e20190374. doi: 10.1542/peds.2019-0374

Stevanovic, D., Costanzo, F., Fucà, E., Valeri, G., Vicari, S., Robins, D. L., et al. (2021). Measurement invariance of the Childhood Autism Rating Scale (CARS) across six countries. *Autism Res.* 14, 2544–2554. doi: 10.1002/aur.2586

Sturm, A., Kuhfeld, M., Kasari, C., and McCracken, J. T. (2017). Development and validation of an item response theory-based Social Responsiveness Scale short form. *J. Child Psychol. Psychiatry* 58, 1053–1061. doi: 10.1111/jcpp.12731

Teresi, J. A., and Fleishman, J. A. (2007). Differential item functioning and health assessment. *Qual. Life Res.* 16, 33–42. doi: 10.1007/s11136-007-9184-6

Thurm, A., Farmer, C., Salzman, E., Lord, C., and Bishop, S. (2019). State of the Field: differentiating intellectual disability from autism spectrum disorder. *Front. Psychiatry* 10:526. doi: 10.3389/fpsyt.2019.00526

Widaman, K. F., and Reise, S. P. (1997). "Exploring the measurement invariance of psychological instruments: Applications in the substance use domain," in *The Science of Prevention: Methodological Advances from Alcohol and Substance Abuse Research*, eds K. J. Bryant, M. Windle, and S. G. West (Washington, DC: American Psychological Association), 281–324. doi: 10.1037/10222-009

Wolff, J. J., Bodfish, J. W., Hazlett, H. C., Lightbody, A. A., Reiss, A. L., and Piven, J. (2012). Evidence of a distinct behavioral phenotype in young boys with Fragile X Syndrome and Autism. *J. Am. Acad. Child Adolesc. Psychiatry* 51, 1324–1332. doi: 10.1016/j.jaac.2012.09.001

Zedeck, S. (2014). *APA Dictionary of Statistics and Research Methods*. Washington, DC: American Psychological Association, doi: 10.1037/14336-000

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 08 August 2022
DOI 10.3389/fpsyg.2022.918181

Check for updates

OPEN ACCESS

EDITED BY
Marisa Filipe,
Faculdade de Letras da Universidade de Lisboa, Portugal

REVIEWED BY
Shruthi Ravi,
The University of Texas at Dallas, United States
Peter Gerhardstein,
Binghamton University, United States

*CORRESPONDENCE
Angela John Thurman
ajthurman@ucdavis.edu

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 12 April 2022
ACCEPTED 11 July 2022
PUBLISHED 08 August 2022

CITATION
Thurman AJ and Dimachkie Nunnally A (2022) Joint attention performance in preschool-aged boys with autism or fragile X syndrome.
Front. Psychol. 13:918181.
doi: 10.3389/fpsyg.2022.918181

Joint attention performance in preschool-aged boys with autism or fragile X syndrome

Angela John Thurman[1,2]* and Amanda Dimachkie Nunnally[1,2]

[1]Department of Psychiatry and Behavioral Sciences, University of California, Davis, Sacramento, CA, United States, [2]Medical Investigation of Neurodevelopmental Disorders (MIND) Institute, University of California, Davis, Sacramento, CA, United States

Early development marks a period of rapid learning facilitated by children's natural curiosity about the people around them. In children with typical development, these early social attentional preferences set the foundation for learning about and from the surrounding world of people. Much of this learning happens using joint attention, the ability to coordinate attention between people and objects of mutual interest. It is well documented that decreased gaze use is commonly observed in individuals with autism and individuals with fragile X syndrome (FXS). Despite the growing body of research comparing phenotypic similarities between individuals with autism and individuals with FXS, no studies have directly compared joint attention performance between these groups. In the present study, we considered the similarities and differences in joint attention between preschool-aged boys with autism or FXS, and the relation between joint attention, language, and other phenotypic characteristics known to differ between boys with autism and boys with FXS. Although joint attention appeared similar, between-group differences emerged when controlling for the influence of age, non-verbal IQ, and autism symptom severity. Differences were also observed when considering how joint attention performance related to other aspects of the phenotype. For example, strong positive associations were observed between joint attention and language performance in boys with autism but not boys with FXS, even after controlling for non-verbal IQ. In contrast, the negative association between joint attention and anxiety symptom severity was significant and stronger in boys with FXS than in autism. These data offer preliminary insights into the similarities and differences between the autism and FXS phenotypes.

KEYWORDS

autism spectrum disorder, fragile X syndrome, joint attention, language, anxiety

Introduction

Social attention plays a pivotal role in children's early learning. For example, the ability to coordinate attention between a social partner and objects or events of mutual interest, known as joint attention (JA), is crucial for socio-cognitive development (Morales et al., 2000; Adamson et al., 2009; Mundy and Bullen, 2022). The development of JA skills may reflect the maturation of socio-cognitive and attentional capabilities that facilitate language learning (Mundy and Gomes, 1998). By adopting a shared frame of reference, JA skills allow children to participate in social learning situations that facilitate mapping words onto their intended referents (Morales et al., 2000; Adamson et al., 2009). Early challenges in JA may create a developmental cascade altering children's early language development (Mundy and Bullen, 2022). In the present study, we considered JA performance in boys with autism and boys with fragile X syndrome (FXS). Despite the growing body of research comparing the autism and FXS phenotypes, direct comparisons between JA skills across phenotypes are limited. Understanding how JA skills compare between boys with autism and boys with FXS and the associations between JA and other developmental characteristics will aid our understanding of the mechanisms underlying both phenotypes.

Joint attention and autism symptomatology

JA delays are among the earliest emerging behavioral features associated with autism (Werner et al., 2005). Moreover, in individuals with autism, early JA skills have been shown to predict autism symptomatology in adulthood (Gillespie-Lynch et al., 2012). These early delays in JA development are considered a critical indicator of, and contributor to, a modified trajectory of social learning in individuals with autism (Mundy and Bullen, 2022). Indeed, children with autism display fewer JA acts than children with typical development and children with other developmental disabilities (Wetherby and Prutting, 1984; Loveland and Landry, 1986; Mundy et al., 1986; Dawson et al., 1998; Stallworthy et al., 2022). Moreover, JA delays in children with autism are observable by 9 months of age (Chawarska et al., 2013; Gangi et al., 2016; Stallworthy et al., 2022). Because JA difficulties are among the earliest emerging features associated with the autism phenotype, considerations of JA skills are found in diagnostic and screening assessments of autism symptomatology (Constantino and Gruber, 2012; Lord et al., 2012).

Autism symptomatology is also a hallmark feature of the FXS phenotype. FXS is the most common monogenetic cause of autism and results from a trinucleotide (CGG) expansion in the *FMR1* gene, located on the X chromosome (Oostra and Willemsen, 2003). Because females with FXS have a second, unaffected X chromosome, which can continue to serve as a protective factor, the FXS is more pronounced in males (Tassone et al., 1999; Loesch et al., 2004; Stembalska et al., 2016). Indeed, nearly all males with FXS demonstrate symptoms consistent with the autism phenotype, and 60% or more meet the criteria for an autism diagnosis when using gold-standard assessment tools (Clifford et al., 2007; Harris et al., 2008; Kaufmann et al., 2008; Abbeduto and McDuffie, 2010; Budimirovic and Kaufmann, 2011; Klusek et al., 2014). Although there are numerous similarities between the autism and FXS phenotypes, significant differences are also observed. At the group level, autism symptomatology is milder in FXS than in autism, even when only considering children with FXS who meet diagnostic criteria for ASD (Wolff et al., 2012; McDuffie et al., 2015). Differences have also been observed across groups in the developmental features correlated with autism symptomatology (Thurman et al., 2015). These subtle differences may significantly impact the developmental trajectories associated with these phenotypes. Outside of comparisons of diagnostic assessments that include items focused on JA performance on the ADOS-2, there are no direct comparisons of JA performance between boys with autism and boys with FXS. Understanding the similarities and differences in JA performance can provide essential insights into these phenotypes' developmental mechanisms.

Language difficulties associated with autism or fragile X syndrome

Delays in language development are common in young children with autism or FXS (De Giacomo and Fombonne, 1998; McDuffie et al., 2017) and are often observed well beyond early childhood (Hartley et al., 2011; Howlin et al., 2014). For individuals with autism, language delays are often noticed early in development (Tager-Flusberg et al., 2009; Boucher, 2012; Talbott et al., 2015). Approximately 30% of individuals with autism demonstrate limited spoken language skills into the school-age years (Tager-Flusberg and Kasari, 2013). Moreover, approximately half for those who do acquire spoken language demonstrate language delays relative to both chronological age or non-verbal cognitive ability expectations (Kjelgaard and Tager-Flusberg, 2001; Boucher, 2012).

Similarly, early expressive language delays are often observed in individuals with FXS (Roberts et al., 2001; Kover et al., 2015), with up to 30% of individuals still demonstrating limited spoken language skills into adolescence (Levy et al., 2006). Moreover, in individuals with FXS, delays in language performance are often more significant than expected relative to achievements in non-verbal cognition; however, there is some variation as a function of language domain and developmental period considered (Abbeduto et al., 2016).

Association between joint attention and language skills in autism or fragile X syndrome

Because JA delays are among the earliest emerging behavioral features associated with autism (Werner et al., 2005), there is a robust literature considering the association between JA and language performance in children with autism (Bottema-Beutel, 2016). Numerous investigations have demonstrated a significant association between JA and language performance. Some types of JA (e.g., responding to JA) demonstrate lasting and significant associations even after controlling for non-verbal cognitive ability (Charman, 2003; Sigman and McGovern, 2005; Yoder et al., 2015; Bottema-Beutel, 2016). Moreover, interventions targeting JA skills have positively impacted language outcomes, offering additional support for the role of JA in language development (Kasari et al., 2006, 2012).

In FXS, however, few studies have considered JA skills or the association between JA and language performance. Nonetheless, there is reason to posit that children with FXS may demonstrate delays in JA skills. For example, nearly all males with FXS have been found to have limited gaze use (Murphy et al., 2007; Hall et al., 2009), a critical JA skill. Moreover, Hahn et al. (2016) found that joint engagement (i.e., states in which the child is actively engaged with objects and people) during play with a caregiver was indeed positively associated with expressive language performance and negatively associated with autism symptomatology scores.

Finally, data from the limited studies that have directly compared performance between boys with autism and boys with FXS suggest potential between-group differences in JA performance. For example, when considering autism symptomatology between the two groups, Wolff et al. (2012) found that boys with FXS + ASD were less impaired than boys with autism in response to JA, and McDuffie et al. (2015) found that boys with FXS + ASD were less impaired in showing and sharing attention than were boys with autism. In addition, some recent studies have documented strength in language performance (e.g., vocabulary) in boys with FXS relative to boys with autism, particularly when you control for between-group differences in non-verbal cognitive ability (McDuffie et al., 2013; Thurman et al., 2017; Sterling, 2018; Thurman and Hoyos Alvarez, 2020). It is plausible that early between-group differences in JA skills may contribute to some of the strengths in language development observed in boys with FXS. However, there are no direct comparisons between boys with autism and boys with FXS in the associations between JA and language performance.

Other attention-modifying phenotypic considerations

Notably, other behavioral similarities are observed between the autism and FXS phenotypes in domains that may also impact the development and/or the measurement of JA performance, such as inattention/hyperactivity (Mayes et al., 2012; Thurman et al., 2014) and anxiety (Leyfer et al., 2006; de Bruin et al., 2007; Cordeiro et al., 2011). Consistent with findings in other domains, despite the similarities observed, differences are also noted in the presence of symptoms of inattention/hyperactivity and anxiety. For example, Thurman et al. (2014) found, while controlling for various developmental characteristics, that parent ratings of anxiety and manic/hyperactive behaviors were significantly higher for boys with FXS than for boys with autism. In addition, the authors found that the association between social avoidance and general anxiety was significantly higher for boys with FXS than for boys with autism. Indeed, increased rates of attentional or anxiety symptomatology may modify how children engage with others and, in turn, influence the development or measurement of JA performance (Salley and Colombo, 2016). Considering these developmental differences may reveal whether similar or different developmental mechanisms underlie shared symptomatology between boys with autism and boys with FXS (Thurman et al., 2015).

Present study

Despite the growing body of research comparing the autism and FXS phenotypes, direct comparisons of JA skills are limited. These comparisons are needed to clarify the associations between JA, attention-modifying phenotypic characteristics, and language. Moreover, such data can provide insights into these phenotypes' developmental mechanisms. Research comparing JA skills and their associations to language and other attention-modifying phenotypic characteristics can provide insights into the developmental mechanisms underlying these phenotypes. In the present study, we provide a preliminary examination of JA performance in preschool-aged boys with autism or FXS to answer the following questions:

1. Does JA performance differ between preschool-aged boys with autism and boys with FXS?
2. Is JA performance concurrently associated with language ability, specifically vocabulary ability, in boys with autism and boys with FXS after controlling for the influences of non-verbal cognitive ability? Note, because of the age and language delays associated with both phenotypes, the language measures used in the present study focused heavily on vocabulary ability.
3. Is JA performance concurrently associated with other child characteristics, such as autism symptom severity or other attention-modifying behavioral features observed to differ between boys with autism and boys with FXS (i.e., ADHD and anxiety symptomatology)?

Materials and methods

Participants

Fifty-one participants between 3.00 and 5.50 years of age were included in the present study, 30 males with autism and

21 males with FXS. Descriptive statistics for the sample are presented in Table 1. Significant between group differences were observed between participants with autism and participants with FXS differed on Non-verbal IQ scores [$t(1, 49) = 3.35$, $p = 0.002$, Cohen's $d = 0.95$] and on autism symptom severity scores [$t(1, 49) = 2.13$, $p = 0.04$, Cohen's $d = 0.62$]. For the participants with autism, the racial/ethnic composition of the sample was 25% Hispanic, with 6.7% Asian, 6.7% Black/African American, 63.3% Caucasian, 20.0% Multi-racial, and 3.3% preferring not to answer. For participants with FXS, the racial/ethnic composition of the sample was 19% Hispanic, with 4.8% Black/African American, 71.4% Caucasian, 19% Multi-racial, and 4.8% preferring not to answer.

Participants were drawn from a longitudinal study focused on elucidating the mechanisms underlying word learning in boys with autism or FXS. Documentation of an existing diagnosis was provided by families of participants with autism (i.e., existing community diagnosis of autism spectrum disorder) and families of participants with FXS (i.e., documentation of a diagnosis of the FMR1 full mutation, with or without mosaicism). Diagnoses for participants with autism were confirmed through administration of the Autism Diagnostic Observation Schedule-2 (Lord et al., 2012). In addition, all participants enrolled in the present study met the following criteria (based on parent report): (a) English is the primary language of exposure; (b) no sensory or physical impairments that would limit participation in project activities; and (c) no medical conditions (e.g., severe and frequent seizures) that prevented them from meeting the demands of the testing protocol.

Multiple sources were utilized for recruitment, including the MIND Institute's IDDRC Clinical Translational Core Registry, parent listservs, the National Fragile X Foundation, and clinics and preschools specialized in working with children with NDDs. Due to differences in prevalence rates, participants with autism were more likely to reside locally than were those with FXS.

TABLE 1 Descriptive statistics for participant groups.

	Autism (*n* = 30) ***M* (*SD*, range)**	**FXS (*n* = 21)** ***M* (*SD*, range)**
CA	4.37 (0.82, 3.08–5.56)	4.20 (0.83, 3.02–5.55)
Non-verbal cognition (IQ)	78.17 (20.88, 30–113)	59.48 (17.53, 30–91)
Verbal cognition (IQ)	69.30 (21.75, 30–100)	62.62 (20.51, 30–93)
Autism symptom severity	7.07 (1.84, 4–10)	5.79 (2.35, 2–10)
Receptive vocabulary knowledge (growth score)	91.10 (29.74, 12–137)	81.10 (30.11, 12–122)
Expressive vocabulary knowledge (growth score)	101.03 (31.42, 42–150)	85.90 (29.98, 45–129)
Play: # utterances	113.27 (70.76, 0–225)	114.38 (99.24, 0–378)
Play: weighted comm score	164.63 (110.97, 0–367)	162.76 (150.22, 0–565)
ADHD symptomatology (total score)	27.13 (12.20, 7–58)	30.48 (8.64, 15–45)
Anxiety symptomatology (total score)	30.54 (18.30, 9–75)	36.15 (16.04, 2–67)

PPVT-4, Peabody Picture Vocabulary Test-4; EVT-2, Expressive Vocabulary Test-2; DAS-II SNC, Differential Ability Scales-II Special Non-verbal Composite.

Methods

Approval from the Institutional Review Board, as well as written informed consent from the parent/guardian of all participants; verbal assent by the child was not required due to the chronological age and developmental delays displayed by the children in the present study. Visit study measures were administered over the course of two consecutive days. All assessments took place in the research laboratory and were conducted by PhD-level study personnel or trainees under their supervision. Multiple direct assessment and caregiver-report measures were utilized in the present study.

Joint attention

The child's ability to coordinate attention between the examiner and an object was examined during a semi-structured play assessment, using the procedures outlined by Thurman and Mervis (2013). Specifically, two sets (version A and version B) of 30 toys/objects were created, each divided into six blocks of five. Four of the five toys/objects in each block were assigned to an elicitation condition in which the child's gaze behavior was considered following the presentation of a specific gesture made by the examiner (i.e., giving, blocking, teasing, or point/gaze gesture). In each block, elicitation condition was randomized. In addition, the semi-structured play assessment was completed over the course of 2 days in order to minimize testing fatigue and maximize the naturalistic quality of the elicitation conditions. Version was counterbalanced across participants in each group.

The Blocking/Teasing/Giving conditions were based on the goal ambiguity task developed by Phillips et al. (1992), which was designed to assess the child's used of JA in response to gestures made by an adult, which varied with regard to the ambiguity of the adult's intention. The Response to JA trials (Carpenter et al., 1998; Brooks and Meltzoff, 2002) were designed to examine the child's ability to monitor the adult's looking/gazing behavior.

(1) Block: Once the child was manually and visually engaged with the toy, the examiner covered the child hands, blocking the child from further activity.
(2) Tease: The examiner offered the child the toy. Once the child reached for the toy, the examiner quickly withdrew the toy out of the child's reach.
(3) Give: The examiner handed the child the toy and allowed the child to play with it.
(4) Response to JA: Trials were administered during transitions between toys presented to the child. The

examiner made eye contact with the child, and once eye contact was established, turned to the images on either the right or left side of the table while demonstrating the appropriate cue (i.e., cross-midline point + gaze following vs. gaze following only) in conjunction with saying, "Oh, wow!).

During each elicitation, the examiner looked at the child (or looked at the object, during response to JA trials) with neutral affect during a 4-s wait period or until the child initiated/responded with JA (whichever came first). Once the 4-s wait-period elapsed (or the child initiated JA), the examiner resumed play (e.g., giving the toy to child or playing with the toys at the table). The child received a point for each trial in which they demonstrated a JA response during the 4-s wait period; a point was not assigned on trials in which the child did not demonstrate a JA response during the wait period. The total score (across all trials) was used to assess JA in the present study. These tasks were selected as a starting point for considering JA performance, because at the time of data collection, the Phillips et al. (1992) Blocking Tasks had been integrated into the ADOS-2 as method of eliciting JA (Lord et al., 2012) and were used to by Thurman and Mervis (2013) to characterize between group differences between children in the same age range with two other neurodevelopmental disabilities associated with varying social communication phenotypes.

Language measures

Receptive vocabulary knowledge

The Peabody Picture Vocabulary Test–4th Edition (Dunn and Dunn, 2007) was used to assess receptive vocabulary knowledge in the present study. The PPVT-4 is an individually administered assessment designed for children and adults aged 2.5–90 years and older. Two versions of this measures (i.e., Version A and Version B) are available and were alternated across participants in each group; thus, approximately half of the participants in each group received Version A and half of the participants received Version B of this measure. Growth scores were utilized in study analyses.

Expressive vocabulary knowledge

The Expressive Vocabulary Test–2nd Edition (Williams, 2007) was used to assess expressive vocabulary knowledge in the present study. The EVT-2 is an individually administered assessment designed for children and adults aged 2.5–90 years and older. Two versions of this measures (i.e., Version A and Version B) are available and were alternated across participants in each group; thus, approximately half of the participants in each group received Version A and half of the participants received Version B of this measure. Growth scores were utilized in study analyses.

Naturalistic language sample

The Abbreviated- Communication Play Protocol, a 20-min semi-structured play sample with a caregiver, was used as the naturalistic language sample (Adamson et al., 2009). The CPP-A consists of four 5-min activities in which caregivers press for different communicative functions by modifying how they use each toy set. Samples start with a 5-min free play activity, where the parent plays with the child as they usually would. Three additional communication contexts are considered: (1) Social Interaction—parents engage in turn-taking games with the child; (2) Requesting—toys placed out of child reach; and (3) Commenting—parents share a series of objects with the child (Adamson et al., 2009). During this sample, each occurrence of the child's single-word utterances and multiple-word utterances were coded using the Behavioral Observation Research Interactive Software (Friard and Gamba, 2016). Utterances were segmented into C-Units, providing an objective criteria for segmenting utterances (Abbeduto et al., 2020). Specifically, at the upper bound, C-units include an independent clause and any of its modifiers; at the lower bound, any sentence fragment and elliptical response also constitutes a C-unit (Thurman et al., in press). The frequency of each utterance was weighted, such that single-word utterances counted as one point and multiple-word utterances counted as two points. The Weighted frequency total and the Unweighted frequency total (i.e., the number of utterances produced without weighting) were considered in the study analyses.

Other child characteristics

Non-verbal cognition

The Differential Ability Scales-II Upper-Level Early Years (Elliott, 2007) was used to assess non-verbal cognitive ability in the present study. The DAS-II in an individually administered assessment of general intellectual functioning, designed for children aged 2.5–8 years of age. The Special Non-verbal Composite, which reflects non-verbal cognition using both non-verbal reasoning and non-verbal spatial subtests was used in study analyses.

Autism symptomatology

The ADOS-2 (Lord et al., 2012) was used to assess autism symptom severity in the present study. The ADOS-2 is a semi-structured observational assessment, administered by a trained examiner, which is designed to press for the social affective and restricted and repetitive behaviors associated with autism. In the present project, 35 participants received a Module 1 (Autism: $n = 20$, FXS: $n = 15$) and 15 participants received a Module 2 (Autism: $n = 10$, FXS: $n = 5$). Data was missing for two participants due to fatigue. All examiners on the project met research reliability training standards.

Attention/hyperactivity symptomatology

The ADHD Rating Scale-IV Preschool Version (McGoey et al., 2007), an 18-item caregiver report measure designed to assess attention deficit and hyperactivity disorder symptoms outlined in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (American Psychiatric Association, 2000) ADHD. Items are rated on a scale of 0 (not at all) to 3 (very often); total score was used in study analyses.

Anxiety symptomatology

The Revised Preschool Anxiety Scale (Edwards et al., 2010), a 30-item caregiver report measure designed to assess anxiety symptoms associated with social anxiety, separation anxiety, general anxiety, specific fears, and obsessive-compulsive symptoms. Items are rated on a scale from 1 (not at all true) to 4 (very often true); total score was used in study analyses.

Analysis plan

A series of analyses were used to consider whether JA scores differed between boys with autism and boys with FXS. First, a univariate analysis of variance was used to directly compare JA scores across the two samples. Second, a regression analysis was used to compare performance across the groups, after controlling for the effects, of CA, non-verbal IQ, and overall autism symptom severity. We corrected for multiple comparisons by using Benjamini and Hochberg's false discovery rate [FDR; 39] procedures in order to maintain a familywise alpha rate of $p \leq 0.05$. Finally, descriptive statistics (i.e., means and standard errors) regarding the patterns observed across the different JA elicitation contexts are provided.

Parametric correlations were used to evaluate the correlations between total JA scores and performance on the language ability measures, after controlling for the influence of non-verbal cognitive ability. In each of these analyses, the FDR was used to maintain a familywise alpha rate of $p \leq 0.05$. Parametric correlations were also used to evaluate the correlations between total JA scores and other child characteristics (i.e., autism symptom severity, ADHD symptomatology, and anxiety symptomatology).

Results

Between-group comparisons of joint attention

First, we considered whether JA scores differed between boys with autism and boys with FXS, using a series of analyses. Overall JA scores between boys with autism and boys with FXS were directly compared. Results indicated that mean overall JA scores were slightly higher in boys with FXS (M = 14.16, SD = 5.75) than in boys with autism (M = 11.57, SD = 5.10), but the statistical comparison of scores did not reach criterion for a significant between-group effect [F (1, 49) = 2.87, p = 0.10, partial eta squared = 0.06]. That said, the regression model considering overall JA scores, after controlling for the effects of CA, non-verbal IQ, and overall autism symptom severity, were significant [F (4, 48) = 12.76, $p < 0.001$, $R^2_{adj} = 0.54$]. Specifically, diagnostic group uniquely accounted for approximately 16% of the variance in overall JA scores, with overall JA scores approximately four points higher for boys with FXS than for boys with autism (p = 0.006 and remained significant after FDR correction). See **Figure 1** for graphs of JA score comparisons before and after controlling for CA, non-verbal IQ and autism symptom severity.

Finally, to provide preliminary data on the patterns observed across JA contexts, we considered participant performance across the different elicitation conditions. **Figure 2** presents comparisons of mean JA scores as function of task for boys with autism and boys with FXS. Across all comparisons JA scores were slightly higher for boys with FXS than for boys with autism.

Association between joint attention and language performance

Analyses were then conducted to consider the relations between JA and language performance for boys with autism and boys with FXS. The contributions of non-verbal cognitive ability were partialed out of this correlation due to its association with both language ability (ASD: rs: 0.59–0.70, ps < 0.001; FXS: rs: 0.55–0.70, ps: 0.01–0.004) and JA performance (ASD: r = 0.54, p = 0.002; FXS: r = 0.42, p = 0.06). After controlling for the influence of non-verbal cognitive ability, overall JA scores were strongly correlated with all language measures considered in boys with autism, with all associations significant after applying the FDR correction (see **Table 2**). For boys with FXS, JA scores were not observed to be significantly associated with performance on any of the language measures.

Relations between joint attention and other child characteristics

Finally, we considered the relations between JA and other child characteristics known to influence social attention (see **Table 3**). In both boys with autism and boys with FXS, JA performance was negatively associated with autism symptomatology. In boys with autism, but not boys with FXS, JA performance was negatively correlated with ADHD symptomatology. Finally, in boys with FXS, JA performance was negatively associated with anxiety symptomatology in boys with FXS; this association was not significant for boys with

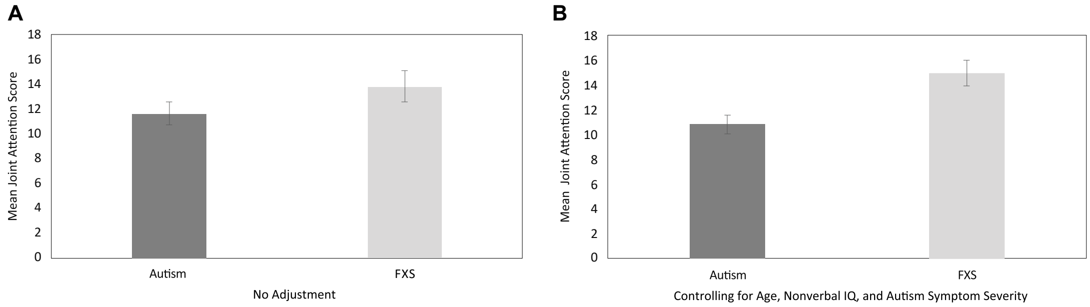

FIGURE 1
Between-group descriptive comparisons of total joint attention performance, with no adjustment **(A)** and after controlling for the effects of age, non-verbal IQ and autism symptom severity **(B)**.

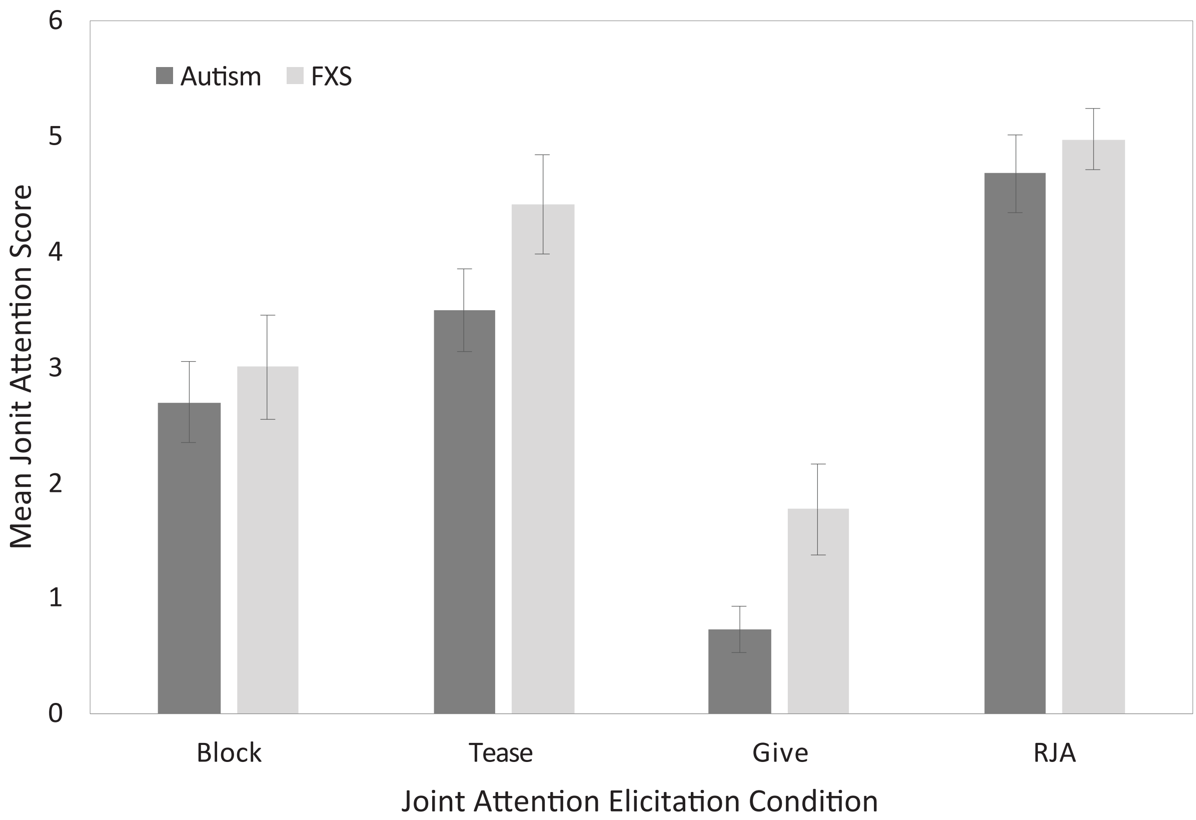

FIGURE 2
Uncorrected between-group descriptive comparisons of joint attention performance as a function of elicitation condition.

autism. All significant correlations remained after applying the FDR correction.

Discussion

The goal of the present study was to provide preliminary insights into the similarities and differences in the use of JA skills between boys with autism or FXS and the phenotypic and behavioral characteristics that were concurrently associated with children's use of JA. Several findings emerged from this study, including group differences in JA performance between boys with autism and those with FXS. In addition, data from the current study suggests that there may be between-group differences in the phenotypic factors associated with JA performance. Altogether, these findings begin to elucidate the different contributors to JA performance in boys with autism and boys with FXS, which has implications for

TABLE 2 Correlations between joint attention performance and language measures after controlling for non-verbal IQ.

	PPVT-4 growth score	EVT-2 growth score	Play: # utterances	Play: weighted comm score
	Correlations controlling for DAS-II SNC			
Autism	0.72***	0.73***	0.67***	0.72***
FXS	0.14	0.20	−0.14	−0.19

***p < 0.001.
PPVT-4, Peabody Picture Vocabulary Test-4; EVT-2, Expressive Vocabulary Test-2; DAS-II SNC, Differential Ability Scales-II Special Non-verbal Composite.

TABLE 3 Correlations between joint attention performance and phenotype control measures.

	Autism	FXS
Autism symptoms	−0.39*	−0.61**
ADHD symptoms	−0.44*	−0.11
Anxiety symptoms	0.27	−0.58**

*p < 0.05, **p < 0.01.
DAS-II SNC, Differential Ability Scales-II Special Non-verbal Composite; ADOS-2 Overall CSS, Autism Diagnostic Observation Schedule-2 Overall Comparison Severity Scores.

understanding phenotypic differences in the development of JA for these populations.

Although the literature indicates that JA delays are common in children with autism (Wetherby and Prutting, 1984; Loveland and Landry, 1986; Mundy et al., 1986; Dawson et al., 1998; Chawarska et al., 2013; Gangi et al., 2016), relatively few studies have explored this construct in children with FXS (Murphy et al., 2007; Hall et al., 2009; Hahn et al., 2016). Moreover, although autism symptoms are commonly observed in males with FXS, no studies directly compare JA performance across the two groups. The current study used an examiner-delivered task-based measure of JA skills in boys with autism and boys with FXS. When compared directly, no between-group differences in JA performance were observed. However, when considering boys of the same age, one must acknowledge that boys with autism typically demonstrate higher IQ scores and more autism symptomatology than boys with FXS. Taking this into consideration, we conducted comparisons controlling for the influences of age, non-verbal IQ, and autism symptom severity, and found that boys with FXS earned significantly higher JA scores than boys with autism. These data add to a growing body of research suggesting that, even though both boys with autism and boys with FXS demonstrate reduced gaze use in social interactions, a between-group difference may emerge between these two groups, particularly when other developmental factors are considered (Wolff et al., 2012; McDuffie et al., 2015).

Next, we considered the association between JA performance and language in boys with autism and boys with FXS. In boys with autism, JA performance was strongly associated with all language measures, even after controlling for non-verbal cognitive abilities. This finding is consistent with the extensive literature documenting the association of language abilities and JA performance in children with autism (Charman, 2003; Sigman and McGovern, 2005; Yoder et al., 2015; Bottema-Beutel, 2016). In contrast, the associations between JA performance and language were not significant in boys with FXS, after controlling for the influence of non-verbal cognitive ability and were weaker (*rs*: −0.19–0.20) than the associations observed for boys with autism (*rs*: 0.67–0.73). There is limited data considering the association between JA and language performance in FXS. Hahn and colleagues found that time spent in joint engagement states with caregivers (i.e., both the child and caregiver engaged with the same object) was positively associated with expressive language abilities (Hahn et al., 2016). Notably, joint engagement does not require the child to initiate JA; rather, the caregiver can support these states by following into the child's attention to an object. These data suggest differences in the association between JA and language between boys with autism and boys with FXS. However, more research is needed to confirm these findings and to elucidate the nature of any differences and their contributions to language delays or any other similarities and differences observed between the autism and FXS phenotypes.

Lastly, we considered the associations between JA performance and other child characteristics such as autism symptom severity and other attention-modifying behavioral characteristics. Indeed, in addition to the link established between JA performance and autism (Mundy and Bullen, 2022), other factors such as inattention/hyperactivity and anxiety may also impact the development or measurement of JA (Leyfer et al., 2006; de Bruin et al., 2007; Cordeiro et al., 2011; Mayes et al., 2012; Thurman et al., 2014). These characteristics must be considered when comparing JA performance across conditions, such as autism and FXS, because the co-occurrence of these features likely differs across groups (McDuffie et al., 2013; Thurman et al., 2014, 2015). JA performance was associated with autism symptom severity for both groups. In addition, for boys, ADHD symptom severity was also associated with JA performance. Finally, for boys with FXS, anxiety symptom severity was significantly negatively associated with JA performance ($r = -0.58$); for boys with autism a non-significant positive correlation was observed between these variables ($r = 0.27$).

Data from the present study adds to the growing body of literature documenting an association between JA and both autism symptomatology and non-verbal cognitive development in autism and adds to the limited research considering JA skills in children with FXS (Sullivan et al., 2007; Constantino and Gruber, 2012; Lord et al., 2012; Brewe et al., 2018; Mundy and Bullen, 2022). Moreover, these findings not only point to potential underlying differences in factors contributing to JA performance in boys with autism and FXS, but also highlight the need for considering behavioral characteristics that may impact the development of JA in cross-syndrome research. Indeed, the co-occurring presence of attentional difficulties or anxiety can modify how children engage with others. Clarifying similarities and differences in the factors influencing JA will provide insight into when similar treatment or measurement approaches can be utilized when working with boys with autism or FXS and when phenotypic differences must be considered. For example, data from the present study suggests that the co-occurrence of anxiety in individuals with FXS may impact JA performance in this population. For boys with FXS, gaze avoidance has often been described to be related to anxiety. This may be attributable to higher rates of anxiety disorders, especially social anxiety, in FXS relative to the general population (Shaffer et al., 1996) and to other NDD groups (Dekker and Koot, 2003). Moreover, Thurman et al. (2014) previously observed maternal ratings of child social avoidance was related to child anxiety in FXS but not ASD. Taken together, it is posited that the co-occurrence of anxiety differentially impacts the measurement of JA in boys with autism and boys with FXS. More research is needed to disentangle the impact of anxiety from the association between JA and language development. For example, it may be that interactions with a more familiar partner, such as a caregiver, can provide additional insights into JA performance, and its role in language learning, in boys with FXS. Finally, ADHD symptomatology, not anxiety, was found to be associated with JA performance in boys with autism. Given the co-occurrence of ADHD and autism, it is necessary to better understand how ADHD symptomatology may impact the development of JA in boys with autism. Additional research elucidating the interrelations between these different domains may also provide insight into the developmental mechanisms underlying the autism phenotype.

Findings from the present study provide some important initial insights into potential similarities and differences in JA performance between boys with autism and boys with FXS. Nonetheless, there is much that remains to be understood. For example, in recent years, newer methods of considering joint attention have been developed and psychometrically validated (e.g., Nowell et al., 2020). It will be important for future studies comparing JA skills between individuals with autism and individuals with FXS to utilize a robust battery of JA measures, across different contexts, to ensure a thorough characterization of the similarities and differences observed across these groups. Moreover, larger longitudinal investigations of the associations between JA, language, and other phenotypic characteristics, are needed to more carefully consider findings suggested by the present study and to provide additional insights into potential cascading impacts of JA challenges on later skills across groups and to consider the intricate ways in which JA skills interact with other phenotypic characteristics (e.g., anxiety). Finally, because FXS is an X-linked disorder, and females with FXS have a second, unaffected X chromosome that can continue to serve as a protective factor, the present study focused on males only (Tassone et al., 1999; Loesch et al., 2004; Stembalska et al., 2016). It is vital that future research also consider JA performance in females with autism and females with FXS to consider whether findings in males also apply to females.

In summary, the present study compared JA performance between boys with autism and boys with FXS, as well as associations between JA, language and other child characteristics. Although overall JA performance was similar across the groups, JA scores were higher for boys with FXS when controlling for the influences of CA, non-verbal IQ, and autism symptom severity. In addition, potential between-group differences may emerge when considering the child characteristics associated with JA performance. Specifically, the positive association between JA performance and language was stronger in boys with autism than boys with FXS, after controlling for the influences of non-verbal IQ. In contrast, the negative association between JA performance and anxiety was stronger in boys with FXS than in boys with autism. These data suggest crucial differences in the contributors to JA performance, or the measurement thereof, and highlight the importance of considering whether similar or different developmental mechanisms underlie shared symptomatology between boys with autism and boys with FXS. Additional research is needed to elucidate the intricate associations between phenotypic features and JA is necessary to clarify the role of JA learning for boys with FXS and the extent to which differences in JA performance, and predictors thereof, contribute to the similarities and differences observed between the autism and FXS phenotypes.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

The studies involving human participants were reviewed and approved by the University of California Davis Clinical Committee B. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AT conceived of the study, participated in its design and coordination, performed the statistical analyses, and drafted the manuscript. AD helped design the analytic plan and draft the manuscript. Both authors read and approved the final manuscript.

Funding

This research was supported by the National Institute of Deafness and Communication Disorders (Grants R03DC014543), the National Institute of Child Health and Human Development (Grants U54HD079125), and the National Institute of Child Health and Human Development (Grants UL1TR000002).

Acknowledgments

We thank the children and their families for their participation in this study. We also thank Cynde Josol, Amy Banasik, Theofilos Mazidzoglou, and Claudine Anglo for assisting with data collection; Karina Gonzalez and Vivian Nguyen for coordinating all study visits; and Danielle Harvey for assisting with data analyses.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abbeduto, L., and McDuffie, A. S. (2010). "Genetic syndromes associated with intellectual disabilities," in *Handbook of Medical Neuropsychology: Applications of Cognitive Neuroscience* eds C. Armstrong, and L. Morrow (Cham: Springer). doi: 10.1007/978-1-4419-1364-7_11

Abbeduto, L., Berry-Kravis, E. M., Sterling, A. M., Sherman, S. L., Edgin, J. O., McDuffie, A. S., et al. (2020). Expressive language sampling as a source of outcome measures for treatment studies in fragile X syndrome: feasibility, practice effects, test-retest reliability, and construct validity. *J. Neurodev. Disord.* 12:10. doi: 10.1186/s11689-020-09314-5

Abbeduto, L., McDuffie, A., Thurman, A. J., and Kover, S. T. (2016). Language development in individuals with intellectual and developmental disabilities: from phenotypes to treatments. *Int. Rev. Res. Dev. Disabil.* 50, 71–118. doi: 10.1016/bs.irrdd.2016.05.006

Adamson, L. B., Bakeman, R., Deckner, D. F., and Romski, M. (2009). Joint engagement and the emergence of language in children with autism and Down Syndrome. *J. Autism Dev. Disord.* 39, 84–96. doi: 10.1007/s10803-008-0601-7

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn. Washington, DC: American Psychiatric Association.

Bottema-Beutel, K. (2016). Associations between joint attention and language in autism spectrum disorder and typical development: a systematic review and meta-regression analysis. *Autism Res.* 9, 1021–1035. doi: 10.1002/aur.1624

Boucher, J. (2012). Research review: structural language in autistic spectrum disorder - characteristics and causes. *J. Child Psychol. Psychiatry* 53, 219–233. doi: 10.1111/j.1469-7610.2011.02508.x

Brewe, A. M., Reisinger, D. L., Adlof, S. M., and Roberts, J. E. (2018). Initiating joint attention use in infants at high-risk for autism spectrum disorder. *J. Intell. Disabil. Res.* 62, 842–853. doi: 10.1111/jir.12539

Brooks, R., and Meltzoff, A. N. (2002). The importance of eyes: how infants interpret adult looking behavior. *Dev. Psychol.* 38:958. doi: 10.1037/0012-1649.38.6.958

Budimirovic, D. B., and Kaufmann, W. E. (2011). What can we learn about autism from studying fragile X syndrome? *Dev. Neurosci.* 33, 379–394. doi: 10.1159/000330213

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., and Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, 1–143. doi: 10.2307/1166214

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 315–324. doi: 10.1098/rstb.2002.1199

Chawarska, K., Macari, S., and Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biol. Psychiatry* 74, 195–203. doi: 10.1016/j.biopsych.2012.11.022

Clifford, S., Dissanayake, C., Bui, Q. M., Huggins, R., Taylor, A. K., and Loesch, D. Z. (2007). Autism spectrum phenotype in males and females with fragile X full mutation and premutation. *J. Autism Dev. Disord.* 37, 738–747. doi: 10.1007/s10803-006-0205-z

Constantino, J. N., and Gruber, C. P. (2012). *The Social Responsiveness Scale*, 2nd Edn. Torrance, CA: Western Psychological Services.

Cordeiro, L., Ballinger, E., Hagerman, R., and Hessl, D. (2011). Clinical assessment of DSM-IV anxiety disorders in fragile X syndrome: prevalence and characterization. *J. Neurodev. Disord.* 3, 57–67. doi: 10.1007/s11689-010-9067-y

Dawson, G., Meltzoff, A. N., Osterling, J., and Rinaldi, J. (1998). Neuropsychological correlates of early symptoms of autism. *Child Dev.* 69, 1276–1285. doi: 10.2307/1132265

de Bruin, E. I., Ferdinand, R. F., Meester, S., de Nijs, P. F., and Verheij, F. (2007). High rates of psychiatric co-morbidity in PDD-NOS. *J. Autism Dev. Disord.* 37, 877–886. doi: 10.1007/s10803-006-0215-x

De Giacomo, A., and Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *Eur. Child Adolesc. Psychiatry* 7, 131–136. doi: 10.1007/s007870050058

Dekker, M. C., and Koot, H. M. (2003). DSM-IV disorders in children with borderline to moderate intellectual disability. I: prevalence and impact. *J. Am. Acad. Child Adolesc. Psychiatry* 42, 915–922. doi: 10.1097/01.CHI.0000046892.27264.1A

Dunn, D. M., and Dunn, L. M. (2007). *Peabody Picture Vocabulary Test*, 4th Edn. Minneapolis, MN: Pearson. doi: 10.1037/t15144-000

Edwards, S. L., Rapee, R. M., Kennedy, S. J., and Spence, S. H. (2010). The assessment of anxiety symptoms in preschool-aged children: the revised preschool anxiety scale. *J. Clin. Child Adolesc. Psychol.* 39, 400–409. doi: 10.1080/15374411003691701

Elliott, C. D. (2007). *Differential Abilities Scale - Second Edition (DAS-II)*. Minneapolis, MN: Psychological Corporation. doi: 10.1037/t15074-000

Friard, O., and Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol. Evol.* 7, 1325–1330. doi: 10.1111/2041-210X.12584

Gangi, D. N., Messinger, D. S., Martin, E. R., and Cuccaro, M. L. (2016). Dopaminergic variants in siblings at high risk for autism: associations with initiating joint attention. *Autism Res.* 9, 1142–1150. doi: 10.1002/aur.1623

Gillespie-Lynch, K., Sepeta, L., Wang, Y., Marshall, S., Gomez, L., Sigman, M., et al. (2012). Early childhood predictors of the social competence of adults with autism. *J. Autism Dev. Disord.* 42, 161–174. doi: 10.1007/s10803-011-1222-0

Hahn, L. J., Brady, N. C., Fleming, K. K., and Warren, S. F. (2016). Joint engagement and early language in young children with fragile X syndrome. *J. Speech Lang. Hear. Res.* 59, 1087–1098. doi: 10.1044/2016_JSLHR-L-15-0005

Hall, S. S., Lightbody, A. A., Huffman, L. C., Lazzeroni, L. C., and Reiss, A. L. (2009). Physiological correlates of social avoidance behavior in children and adolescents with fragile X syndrome. *J. Am. Acad. Child Adolesc. Psychiatry* 48, 320–329. doi: 10.1097/CHI.0b013e318195bd15

Harris, S. W., Hessl, D., Goodlin-Jones, B., Ferranti, J., Bacalman, S., Barbato, I., et al. (2008). Autism profiles of males with fragile X syndrome. *Am. J. Ment. Retard.* 113, 427–438. doi: 10.1352/2008.113:427-438

Hartley, S. L., Seltzer, M. M., Raspa, M., Olmstead, M., Bishop, E., and Bailey, D. B. (2011). Exploring the adult life of men and women with fragile X syndrome: results from a national survey. *Am. J. Intell. Dev. Disabil.* 116, 16–35. doi: 10.1352/1944-7558-116.1.16

Howlin, P., Savage, S., Moss, P., Tempier, A., and Rutter, M. (2014). Cognitive and language skills in adults with autism: a 40-year follow-up. *J. Child Psychol. Psychiatry* 55, 49–58. doi: 10.1111/jcpp.12115

Kasari, C., Freeman, S., and Paparella, T. (2006). Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. *J. Child Psychol. Psychiatry* 47, 611–620. doi: 10.1111/j.1469-7610.2005.01567.x

Kasari, C., Gulsrud, A., Freeman, S., Paparella, T., and Hellemann, G. (2012). Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. *J. Am. Acad. Child Adolesc. Psychiatry* 51, 487–495. doi: 10.1016/j.jaac.2012.02.019

Kaufmann, W. E., Capone, G. T., Clarke, M., Budimirovic, D. B., and Tarsy, D. (2008). "Autism in genetic intellectual disability," in *Autism - Current Theories and Evidence*, ed. A. W. Zimmerman (Totowa, NJ: Humana Press), 81–108. doi: 10.1007/978-1-60327-489-0_4

Kjelgaard, M. M., and Tager-Flusberg, H. B. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Lang. Cogn. Process.* 16, 287–308. doi: 10.1080/01690960042000058

Klusek, J., Martin, G. E., and Losh, M. (2014). Consistency between research and clinical diagnoses of autism among boys and girls with fragile X syndrome. *J. Intell. Disabil. Res.* 58, 940–952. doi: 10.1111/jir.12121

Kover, S. T., McCary, L. M., Ingram, A. M., Hatton, D. D., and Roberts, J. E. (2015). Language development in infants and toddlers with fragile X syndrome: change over time and the role of attention. *Am. J. Intell. Dev. Disabil.* 120, 125–144. doi: 10.1352/1944-7558-120.2.125

Levy, Y., Gottesman, R., Borochowitz, Z., Frydman, M., and Sagi, M. (2006). Language in boys with fragile X syndrome. *J. Child Lang.* 33, 125–144. doi: 10.1017/S030500090500718X

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., et al. (2006). Comorbid psychiatric disorders in children with autism: interview development and rates of disorders. *J. Autism Dev. Disord.* 36, 849–861. doi: 10.1007/s10803-006-0123-0

Loesch, D. Z., Huggins, R. M., and Hagerman, R. J. (2004). Phenotypic variation and FMRP levels in fragile X. *Ment. Retard. Dev. Disabil. Res. Rev.* 10, 31–41. doi: 10.1002/mrdd.20006

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., and Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule-Second Edition (ADOS-2).* Torrance, CA: Western Psychological Services.

Loveland, K. A., and Landry, S. H. (1986). Joint attention and language in autism and developmental language delay. *J. Autism Dev. Disord.* 16, 335–349. doi: 10.1007/BF01531663

Mayes, S. D., Calhoun, S. L., Mayes, R. D., and Molitoris, S. (2012). Autism and ADHD: overlapping and discriminating symptoms. *Res. Autism Spectr. Disord.* 6, 277–285. doi: 10.1016/j.rasd.2011.05.009

McDuffie, A. S., Kover, S. T., Hagerman, R., and Abbeduto, L. (2013). Investigating word learning in fragile X syndrome: a fast-mapping s tudy. *J. Autism Dev. Disord.* 43, 1676–1691. doi: 10.1007/s10803-012-1717-3

McDuffie, A. S., Thurman, A. J., Channell, M. M., and Abbeduto, L. (2017). *Language Disorders in Children with Intellectual Disability of Genetic Origin, Handbook of Child Language Disorders*, 2nd Edn. London: Psychology Press. doi: 10.4324/9781315283531-2

McDuffie, A. S., Thurman, A. J., Hagerman, R. J., and Abbeduto, L. (2015). Symptoms of autism in males with fragile X syndrome: a comparison to nonsyndromic ASD using current ADI-R scores. *J. Autism Dev. Disord.* 45, 1925–1937. doi: 10.1007/s10803-013-2013-6

McGoey, K. E., DuPaul, G. J., Haley, E., and Shelton, T. L. (2007). Parent and teacher ratings of attention-deficit/hyperactivity disorder in preschool: the ADHD rating scale-IV preschool version. *J. Psychopathol. Behav. Assess.* 29, 269–276. doi: 10.1007/s10862-007-9048-y

Morales, M., Mundy, P., Delgado, C. E., Yale, M., Messinger, D., Neal, R., et al. (2000). Responding to joint attention across the 6-through 24-month age period and early language acquisition. *J. Appl. Dev. Psychol.* 21, 283–298. doi: 10.1016/S0193-3973(99)00040-4

Mundy, P., and Bullen, J. (2022). The bidirectional social-cognitive mechanisms of the social-attention symptoms of autism. *Front. Psychiatry* 12:752274. doi: 10.3389/fpsyt.2021.752274

Mundy, P., and Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behav. Dev.* 21, 469–482. doi: 10.1016/S0163-6383(98)90020-0

Mundy, P., Sigman, M., Ungerer, J., and Sherman, T. (1986). Defining the social deficits of autism: the contribution of non-verbal communication measures. *J. Child Psychol. Psychiatry* 27, 657–669. doi: 10.1111/j.1469-7610.1986.tb00190.x

Murphy, M. M., Abbeduto, L., Schroeder, S., and Serlin, R. (2007). Contribution of social and information-processing factors to eye-gaze avoidance in fragile X syndrome. *Am. J. Ment. Retard.* 112, 349–360. doi: 10.1352/0895-8017(2007)112[0349:COSAIF]2.0.CO;2

Oostra, B. A., and Willemsen, R. (2003). A fragile balance: FMR1 expression levels. *Hum. Mol. Genet.* 12(Suppl. 2), R249–R257. doi: 10.1093/hmg/ddg298

Nowell, S. W., Watson, L. R., Crais, E. R., Baranek, G. T., Faldowski, R. A., and Turner-Brown, L. (2020). Joint attention and sensory-regulatory features at 13 and 22 months as predictors of preschool language and social-communication outcomes. *J. Speech Lang. Hear. Res.* 63, 3100–3116.

Phillips, W., Baron-Cohen, S., and Rutter, M. (1992). The role of eye contact in goal detection: evidence from normal infants and children with autism or mental handicap. *Dev. Psychopathol.* 4, 375–383. doi: 10.1017/S0954579400000845

Roberts, J. E., Mirrett, P., and Burchinal, M. (2001). Receptive and expressive communication development of young males with fragile X syndrome. *Am. J. Ment. Retard.* 106:216. doi: 10.1352/0895-8017(2001)106<0216:RAECDO>2.0.CO;2

Salley, B., and Colombo, J. (2016). Conceptualizing social attention in developmental research. *Soc. Dev.* 25, 687–703. doi: 10.1111/sode.12174

Shaffer, D., Fisher, P., Dulcan, M. K., Davies, M., Piacentini, J., Schwab-Stone, M. E., et al. (1996). The NIMH diagnostic interview schedule for children version 2.3 (DISC-2.3): description, acceptability, prevalence rates, and performance in the MECA study. *J. Am. Acad. Child Adolesc. Psychiatry* 35, 865–877. doi: 10.1097/00004583-199607000-00012

Sigman, M., and McGovern, C. W. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *J. Autism Dev. Disord.* 35, 15–23. doi: 10.1007/s10803-004-1027-5

Stallworthy, I. C., Lasch, C., Berry, D., Wolff, J. J., Pruett, J. R., Marrus, N., et al. (2022). Variability in responding to joint attention cues in the first year is associated with autism outcome. *J. Am. Acad. Child Adolesc. Psychiatry* 61, 413–422. doi: 10.1016/j.jaac.2021.03.023

Stembalska, A., Łaczmańska, I., Gil, J., and Pesz, K. A. (2016). Fragile X syndrome in females-A familial case report and review of the literature. *Dev. Period Med.* 20, 99–104.

Sterling, A. (2018). Grammar in boys with idiopathic autism spectrum disorder and boys with fragile X syndrome plus autism spectrum disorder. *J. Speech Lang. Hear. Res.* 61, 857–869. doi: 10.1044/2017_JSLHR-L-17-0248

Sullivan, M., Finelli, J., Marvin, A., Garrett-Mayer, E., Bauman, M., and Landa, R. (2007). Response to joint attention in toddlers at risk for autism spectrum disorder: a prospective study. *J. Autism Dev. Disord.* 37, 37–48. doi: 10.1007/s10803-006-0335-3

Tager-Flusberg, H. B., and Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism Res.* 6, 468–478. doi: 10.1002/aur.1329

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., et al. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *J. Speech Lang. Hear. Res.* 52, 643–652. doi: 10.1044/1092-4388(2009/08-0136)

Talbott, M. R., Nelson, C. A., and Tager-Flusberg, H. (2015). Maternal gesture use and language development in infant siblings of children with autism spectrum disorder. *J. Autism Dev. Disord.* 45, 4–14. doi: 10.1007/s10803-013-1820-0

Tassone, F., Hagerman, R. J., Ikle, D. N., Dyer, P. N., Lampe, M., Willemsen, R., et al. (1999). FMRP expression as a potential prognostic indicator in fragile X syndrome. *Am. J. Med. Genet.* 84, 250–261. doi: 10.1002/(SICI)1096-8628(19990528)84:3<250::AID-AJMG17>3.0.CO;2-4

Thurman, A. J., and Hoyos Alvarez, C. (2020). Language performance in preschool-aged boys with nonsyndromic autism spectrum disorder or fragile X syndrome. *J. Autism Dev. Disord.* 50, 1621–1638. doi: 10.1007/s10803-019-03919-z

Thurman, A. J., Hoyos Alvarez, C., and Nguyen, V. (in press). Using weighted communication scoring procedures in naturalistic play samples: preliminary validation in preschool-aged boys with autism or fragile X syndrome. *Autism Res.*

Thurman, A. J., McDuffie, A. S., Hagerman, R. J., Josol, C. K., and Abbeduto, L. (2017). Language skills of males with fragile X syndrome or nonsyndromic autism spectrum disorder. *J. Autism Dev. Disord.* 47, 728–743. doi: 10.1007/s10803-016-3003-2

Thurman, A. J., McDuffie, A. S., Kover, S. T., Hagerman, R. J., and Abbeduto, L. (2015). Autism symptomatology in boys with fragile X syndrome: a cross-sectional developmental trajectories comparison with nonsyndromic autism spectrum disorder. *J. Autism Dev. Disord.* 45, 2816–2832. doi: 10.1007/s10803-015-2443-4

Thurman, A. J., McDuffie, A., Hagerman, R., and Abbeduto, L. (2014). Psychiatric symptoms in boys with fragile X syndrome: a comparison with nonsyndromic autism spectrum disorder. *Res. Dev. Disabil.* 35, 1072–1086. doi: 10.1016/j.ridd.2014.01.032

Thurman, A. J., and Mervis, C. B. (2013). The regulatory function of social referencing in preschoolers with Down syndrome or Williams syndrome. *J. Neurodevelop. Dis.* 5, 1–20.

Werner, E., Dawson, G., Munson, J., and Osterling, J. (2005). Variation in early developmental course in autism and its relation with behavioral outcome at 3–4 years of age. *J. Autism Dev. Disord.* 35, 337–350. doi: 10.1007/s10803-005-3301-6

Wetherby, A. M., and Prutting, C. A. (1984). Profiles of communicative and cognitive-social abilities in autistic children. *J. Speech Lang. Hear. Res.* 27, 364–377. doi: 10.1044/jshr.2703.364

Williams, K. T. (2007). *Expressive Vocabulary Test*, 2nd Edn. Minneapolis, MN: Pearson. doi: 10.1037/t15094-000

Wolff, J. J., Bodfish, J. W., Hazlett, H. C., Lightbody, A. A., Reiss, A. L., and Piven, J. (2012). Evidence of a distinct behavioral phenotype in young boys with fragile X syndrome and autism. *J. Am. Acad. Child Adolesc. Psychiatry* 51, 1324–1332. doi: 10.1016/j.jaac.2012.09.001

Yoder, P., Watson, L. R., and Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *J. Autism Dev. Disord.* 45, 1254–1270. doi: 10.1007/s10803-014-2286-4

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 August 2022
DOI 10.3389/fpsyg.2022.905273

Check for updates

OPEN ACCESS

EDITED BY
Lénia Carvalhais,
Infante D. Henrique Portucalense University, Portugal

REVIEWED BY
Frances A. Conners,
University of Alabama, United States
Susan Loveall,
University of Nebraska-Lincoln, United States
Marie Moore Channell,
University of Illinois at Urbana-Champaign, United States

*CORRESPONDENCE
Alison Prahl
alison_prahl@baylor.edu

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 26 March 2022
ACCEPTED 25 July 2022
PUBLISHED 12 August 2022

CITATION
Prahl A and Schuele CM (2022) A pilot study assessing listening comprehension and reading comprehension in children with down syndrome: Construct validity from a multi-method perspective.
Front. Psychol. 13:905273.
doi: 10.3389/fpsyg.2022.905273

A pilot study assessing listening comprehension and reading comprehension in children with down syndrome: Construct validity from a multi-method perspective

Alison Prahl[1]* and C. Melanie Schuele[2]

[1]Department of Communication Sciences and Disorders, Baylor University, Waco, TX, United States, [2]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, United States

Obtaining valid assessments of language and literacy skills in children with Down syndrome (DS) presents a challenge as there is a paucity of information about the psychometrics of measures that are commonly used to measure listening and reading comprehension in this population. Evaluating the construct validity of measures that employ different methods is essential to ascertain the optimal method of assessment in individuals with DS and with typical developmental histories (TD). This pilot study aimed to evaluate the construct validity of four parallel measures of listening and reading comprehension. Participants included 19 individuals with DS ($M = 17$ years, 3 months; $SD = 3$ years, 6 months) and 19 word-level reading-matched children with TD ($M = 7$ years, 2 months; $SD = 7$ months). Participants completed norm-referenced assessments for four parallel measures of listening and reading comprehension. The four measurement methods were: (1) non-verbal response, (2) cloze procedure, (3) passage-level with close-ended questions, and (4) passage-level with open-ended questions. Participants completed additional assessments (e.g., cognition, language, and speech) for descriptive purposes. Construct validity was assessed using the Multitrait-Multimethod Matrix, a correlation matrix arranged to facilitate the assessment and interpretation of construct validity of measures across various formats. For both study groups, we observed strong evidence of construct validity for three out of four measurement methods. Results using the multimethod perspective also indicated that the listening and reading comprehension constructs were not separable. The findings from this pilot study represent a first step toward determining optimal methods of listening and reading comprehension assessment for individuals with DS. Additionally, these results can inform outcome measure selection in future language and literacy research with children with DS.

KEYWORDS

Down syndrome, listening comprehension, reading comprehension, construct validity, psychometrics

Introduction

Down syndrome (DS), the most common known genetic cause of intellectual disability, is characterized by a behavioral phenotype consisting of a pattern of strengths and weaknesses across multiple domains (e.g., cognitive, linguistic, speech-motor, and social-emotion; Chapman and Hesketh, 2000; Fidler, 2005; Fidler and Philofsky, 2009). In relation to reading outcomes, one of the hallmark DS phenotypic characteristics is that individuals often present with language and literacy deficits that are disproportionate to their broader cognitive profiles (Byrne et al., 1995). Although there is substantial literature on language development in children with DS, children and young adults with DS are quite underrepresented in reading research (Afacan et al., 2018). In this study, we aimed to gain insight into evidence-based language and literacy assessment approaches to inform educational and clinical services.

Despite perpetuated beliefs that children with DS cannot learn to read and comprehend text, an emerging body of evidence challenges this assumption (Buckley, 2001; Byrne et al., 2002; Lemons et al., 2017). For example, Buckley (2001) found that 60–70% of individuals with DS in Australia and the United Kingdom have attained functional levels of literacy. Byrne and colleagues found that some children with DS (ages 4–12) demonstrate word-level reading developmental trajectories that are not significantly different compared with development in reading-matched children with typical development (TD; ages 4–10). They also found that children with DS presented with word-level reading standard scores that were higher than their intelligence quotient scores (e.g., Byrne et al., 1995, 2002). Based on current evidence, many individuals with DS present with a relative strength in word-level reading as compared with other reading skills; however, they often experience persistent difficulties with reading comprehension. Research with children and young adults with DS, though limited, demonstrates that reading comprehension growth tends to progress slowly and achievement rarely reaches levels commensurate to word-level reading skills or oral language abilities (e.g., Byrne et al., 2002; Groen et al., 2006; Nash and Heath, 2011).

Reading comprehension–the construction of meaning from written text and the ultimate goal of reading (Catts and Kamhi, 1999)–requires the coordination of multiple underlying cognitive and linguistic processes (Kintsch, 1998; Snow, 2002; Elleman and Compton, 2017; Fuchs et al., 2018). Thus, across multiple theoretical models of reading that place reading comprehension as the outcome of interest, reading comprehension is viewed as a multidimensional construct (Gough and Tunmer, 1986). Within these theoretical models, proficient word recognition (i.e., decoding) and listening comprehension are widely recognized competencies that underlie reading comprehension. Decoding involves context-free word recognition measured by production of real or pseudo words and listening comprehension is the process by which lexical information, sentences, and discourse are interpreted (Gough and Tunmer, 1986). Generally for readers with typical development, once a word is accurately decoded a few times, it is likely to be recognized immediately without any conscious effort, leading to efficient word recognition. As such, across typical development, the influence of word recognition on reading comprehension decreases, whereas the contribution of listening comprehension on reading comprehension increases over time (e.g., Gough and Tunmer, 1986; Catts et al., 2006; García and Cain, 2014; Hogan et al., 2014). Although less is known about the relation between these constructs across development in DS, cross-sectional studies provide useful information. For readers with DS, listening comprehension is reported to predict reading comprehension and is more strongly correlated with reading comprehension in children and young adults with DS than TD peers (e.g., Roch and Levorato, 2009; Prahl and Schuele, 2022). As a result and given that individuals with DS often have a relative strength in word recognition as detected on word recognition tasks rather than decoding tasks (Fidler, 2005; Martin et al., 2009), listening comprehension is hypothesized as the main barrier to reading comprehension. Because individuals with DS often engage in the task of learning to read with difficulties in listening comprehension (Cossu et al., 1993; Roch and Levorato, 2009), evaluating listening comprehension using psychometrically sound measures is an important consideration to understand reading outcomes for individuals with DS. However, it is challenging to obtain valid estimates of these language skills in individuals with DS, as there is limited information specific to DS about the psychometrics of common measures of listening comprehension and reading comprehension. Thus, the purpose of this pilot study was to explore the psychometric properties of commonly used measures of listening comprehension and reading comprehension for individuals with DS.

Three challenges to valid assessment of listening comprehension and reading comprehension include (1) challenges with the constructs, (2) challenges with measures commonly used to assess the constructs of interest, and (3) challenges specific to the DS phenotype. First, given that reading comprehension and listening comprehension are multidimensional constructs, the degree to which measures tap various underlying cognitive and linguistic processes differs based on how listening comprehension or reading comprehension is operationalized with a specific measure. To illustrate the challenges that emerge with measuring reading comprehension, in a study of 97 school-age children with TD, Cutting and Scarborough (2006) found that the relative contributions of word reading (R^2s = 6.1–11.9%) and oral language (R^2s = 9–15%) to reading comprehension varied substantially across three reading comprehension measures: Wechsler Individual Achievement Test (WIAT; Wechsler, 1992), Gates-MacGinitie Reading Test (MacGinitie et al., 2000), and Gray Oral Reading Test (GORT; Wiederholt and Bryant, 1992). Additionally, in their sample of 510 school-age twin sibling pairs with TD, Keenan et al. (2008) found only modest intercorrelations (*rs* = 0.31–0.70) among four commonly used reading comprehension measures: GORT (Wiederholt and Bryant, 1992), Qualitative Reading Inventory (QRI; Lauren and Caldwell, 2001), Woodcock-Johnson Test of Achievement Passage Comprehension subtest (Woodcock et al., 2001), and Peabody Individual Achievement Test Reading Comprehension subtest (Dunn and Markwardt, 1970). Thus, based on these findings, various reading comprehension measures do not seem to be converging on the same construct. Rather, these tests differentially measure the multiple aspects of reading comprehension. Researchers have identified several additional reader characteristics that contribute to comprehending written text, some of which may account for the lack of association across reading comprehension measures (Miller et al., 2013). These characteristics include reading fluency, working memory, verbal reasoning, background knowledge, motivation and engagement, and executive functioning (Perfetti et al., 1996; Snow, 2002; Kintsch and Kintsch, 2005; Cutting and Scarborough, 2006).

It is not surprising that there is similarly a lack of consensus among researchers on how to operationalize listening comprehension and whether listening comprehension and oral language are distinct constructs. Some researchers propose that oral language contributes to listening comprehension, or the opposite, that listening comprehension is part of a broader construct of oral language, and yet others suggest that oral language and listening comprehension are separate constructs (Hogan et al., 2014; Kim and Phillips, 2014; Catts et al., 2015; Gray et al., 2017). In a large-scale longitudinal study designed to increase the field's understanding of this topic, the Language and Reading Research Consortium, 2017 evaluated the dimensionality of oral language and listening comprehension based on confirmatory factor analysis of data from a population-based sample of preschool through third grade children (n = 1,869). Evidence of oral language and listening comprehension operating as a single construct was stronger in the preschool and kindergarten data as compared with the first through third grade data. Although the best fitting model at all grade levels included two separate factors (i.e., expressive and receptive vocabulary and grammar) for oral language and listening comprehension, oral language and listening comprehension were highly correlated (r = 0.87–0.91). The LARRC concluded that oral language and listening comprehension were best characterized as a single oral language construct, and thus measures of oral language and listening comprehension were argued to assess the same underlying construct. Based on this conclusion, measures of oral language and measures of listening comprehension can presumably be used interchangeably, as they all would yield an estimate of "listening comprehension." Given the lack of convergence in measures of reading comprehension and listening comprehension, it is essential that constructs such as these are operationally defined to promote clarity.

For the purpose of the current study, **reading comprehension** was operationalized as constructing meaning from written text and **listening comprehension** was operationalized as constructing meaning from read-aloud written text. These definitions align with how Gough and Tunmer (1986) originally defined reading and listening comprehension within the Simple View of Reading. Gough and Tunmer (1986) further argued that parallel definitions, and thus, parallel measures of listening comprehension and reading comprehension are essential to adequately capturing the relation between these two constructs. Given the parallel nature of the operational definitions, it is not surprising that listening comprehension and reading comprehension have been found to be highly correlated in studies of children with TD and children with DS (Sinatra, 1990; Nation and Snowling, 2004; Roch and Levorato, 2009; Laws et al., 2016). Further, listening comprehension operationalized in this manner is distinct from listening comprehension in conversation or as operationalized in some oral language measures. Unlike listening comprehension as operationalized here, listening comprehension in the context of conversation includes a certain level of redundancy, additional non-verbal cues, and the opportunity to repair any lapses in comprehension that are not available in text. Further, listening comprehension operationalized as such is distinct from other oral language measures (e.g., vocabulary, grammar comprehension) that do not necessarily require text-level processing (Catts and Kamhi, 1986) and instead often involve comprehension of language at the single word or phrase level. Whereas, common measures of oral language (e.g., grammar comprehension and vocabulary) do not have parallel formats with typical measures of reading comprehension, the measures included in this study reflect parallel measures of listening comprehension and reading comprehension that align with the operational definitions above.

Second, listening comprehension and reading comprehension assessment is complicated by substantial variation across measurement methods. To illustrate, Francis et al. (2005) reported a stronger association between decoding and reading comprehension when comprehension was assessed with a cloze-procedure measurement method compared with a multiple-choice question method among children with TD. Commonly used measures vary in the (a) text format that is presented at the single word, phrase, sentence, or paragraph/passage level and (b) response format that requires the test taker, for example, to point to a picture to identify the referent or to verbally answer multiple-choice, close-ended, or open-ended questions. Further, many commonly used standardized measures have psychometric weaknesses for test reliability and validity (Paris and Stahl, 2005). Petersen and Stoddard (2018) argued that because emphasis has been placed on test reliability, many reading comprehension measures with weak validity have emerged. In particular, content validity—how well test items adequately represent the entirety of the measured construct—as well as construct validity—the degree to which a test measures what it claims to be measuring—comes into question. Due to weaknesses in content and construct validity, any conclusions about listening comprehension and reading comprehension must be considered in the context of the specific measure used. For any particular measure of comprehension, it is important to evaluate how the construct is operationalized (e.g., recalling facts and constructing inferences), presentation of the test stimuli (e.g., visual or oral), the response format (e.g., oral or written; multiple-choice; or open-ended), and the test format (e.g., timed or untimed; individual or group administration; Fuchs et al., 2018).

Third, listening comprehension and reading comprehension assessment for individuals with DS warrants careful consideration because most measures were not developed with sufficient attention to the myriad characteristics of individuals with disabilities. Given phenotypic characteristics of DS (e.g., cognitive and linguistic deficits), norm-referenced assessments may not yield valid measurement for this population, despite the demonstration of validity for other populations. In previous studies of reading comprehension in DS, authors do not consistently report reliability scores and validity scores. The DS behavioral phenotype consists of patterns of strengths and challenges across not only cognitive and linguistic domains, but also speech-motor and social-emotional domains. Two challenges characteristic of the DS phenotype, but perhaps not of other groups of individuals with intellectual disabilities, may contribute to underestimation of skills. First, the speech of individuals with DS is characterized by persistent, atypical phonological error patterns that have an adverse impact on intelligibility (Stoel-Gammon, 1997). Reduced speech intelligibility may be a confounding factor for reading comprehension measures requiring a verbal response. Second, when faced with cognitive challenges, individuals with DS are more likely than TD peers to engage in positive and negative behaviors to avoid tasks (Wishart, 1996). This behavior reflects overall poor task persistence and higher levels of off-task social behaviors, especially when cognitive processes are strained, for example, in reading comprehension assessment (Wishart, 1996; Fidler, 2006).

Historically, researchers have not considered behavioral phenotypes in selecting or developing assessment measures to address these challenges (Lemons et al., 2017). Thus, the purpose of this pilot study was to evaluate the validity of listening comprehension and reading comprehension measures for individuals with DS. We evaluated the construct validity, the degree to which a test measures what it claims to be measuring, for four parallel measures of listening comprehension and reading comprehension. The Multitrait-Multimethod matrix (MTMM; Campbell and Fiske, 1959) is an approach using a matrix of correlations to facilitate the assessment and interpretation of the construct validity of measures across various methods. Within the MTMM, convergent validity and discriminant validity is assessed. Convergent validity refers to the degree to which there is empirical evidence that a measure correlates with other measures of the same construct which are assumed to relate based on theory. Discriminant validity refers to the degree to which there is empirical evidence that constructs can be meaningfully differentiated (i.e., not highly correlated) from other theoretically distinct constructs (Campbell and Fiske, 1959). Several traits and several methods are measured and evaluated within the MTMM. In this study, we evaluated two traits—listening comprehension and reading comprehension—and four methods (non-verbal response, cloze-procedure, passage-level with close-ended questions, and passage-level with open-ended questions), resulting in an 8×8 matrix.

We addressed two research questions for two groups of participants—individuals with DS and word-level reading-matched children with typical development (TD): (1) For both groups, are measures of the same construct that use different methods (monotrait-heteromethod) more strongly correlated than (a) measures of different constructs that use the same method (heterotrait-monomethod) and (b) measures of different constructs that use different methods (heterotrait-heteromethod)? And (2) Is evidence of construct validity moderated by group?

Methods

The study procedures were approved by the university Institutional Review Board. The data reported here are part of a study on listening comprehension and reading comprehension in DS (e.g., Hessling, 2020; Prahl and Schuele, 2022). In this article, we present data related to the construct validity and reliability of measures of listening and comprehension in DS and their TD peers.

Participants and procedure

The study was conducted with two groups in which participants were matched on word-level reading: the first group consisted of 19 individuals with DS (32% boys) ages 10 to 22 years ($M = 17$ years, 3 months; $SD = 3$ years, 6 months) and the second group was comprised of 19 children with TD (42% boys) ages 6–8 years ($M = 7$ years, 2 months; $SD = 7$ months). Because listening comprehension and reading comprehension were the outcomes of interest in this pilot study, participants were matched on the remaining variable most-often included in theoretical models of reading comprehension—word-level reading. To form the TD control group, each participant with DS was matched to one TD participant (i.e., a TD participant could only be paired with a single DS participant) based on word-level reading and sex when possible. A TD child was considered an eligible match if his or her raw score on the Word Identification subtest of the Woodcock Reading Mastery Tests-III (WRMT-III; Woodcock, 2011) was within three points of the raw score for a participant with DS. See **Table 1** for participant demographic information. Significant between-group differences were observed on all descriptive measures except word level reading, the matching criteria (see **Table 2**).

TABLE 1 Participant demographic information.

	DS group ($n = 19$)	TD group ($n = 19$)
Sex		
Male	8	6
Female	11	13
Race		
American Indian/Alaska Native	0	0
Asian	0	0
Black/African American	1	1
Hispanic	0	2
Native Hawaiian/Other Pacific Islander	0	0
White	17	15
Multiple races	1	1
Not reported	0	0
Ethnicity		
Hispanic or Latino	1	3
Not Hispanic or Latino	17	15
Not reported	1	1
Mother's education level		
Some high school	0	0
High school diploma/GED	1	0
Some college	2	3
Associate's degree	3	0
Bachelor's degree	6	9
Master's degree	5	4
Professional degree	2	3

This content has been adapted from Prahl and Schuele (2022).

Participants with DS were recruited by distributing study flyers at private schools, on electronic mailing lists, and with DS community organizations in the Nashville, TN and Dallas/Fort Worth, TX regions as well as with families whose children had participated in previous research studies in the lab. Participants with TD were recruited solely from the Nashville, TN metropolitan area by distributing flyers on electronic mailing lists, to families whose children had participated in previous research studies in the lab, to community organizations, and families of local elementary school first- and second-grade students who were reading on grade level. Participants were compensated $15 for completing the eligibility session and $40 for completing the assessment session as well as an additional $20 if participants traveled to the university lab to complete the study activities.

Individuals with DS were eligible to participate if they (a) had been diagnosed with DS by a physician per parent report, (b) were monolingual English speakers and used spoken language as a primary form of communication, (c) successfully completed the screening battery (i.e., listened to directions, completed assessments), and (d) had normal or corrected-to-normal vision per parent report. Hearing status inclusionary criteria was not used for the DS group to ensure inclusion of a representative sample of participants with DS, who frequently present with mild-to-moderate hearing loss (Roizen et al., 1993). Children with TD were eligible to participate if they (a) demonstrated oral language skills within normal limits (i.e., standard score = 85) and neurotypical development per parent report; (b) were monolingual English speakers; (c) successfully completed the screening battery (i.e., listened to directions, completed assessments); (d) passed hearing screening in at least one ear, unaided using American Speech-Language-Hearing Association (ASHA) (2022); and (e) had normal or corrected-to-normal vision per parent report. Exclusionary criteria for both groups included correctly reading fewer than 80% of words on the Phonological Awareness Literacy Screening—Kindergarten (PALS-K; Invernizzi et al., 1997) and children with TD were excluded if they scored more than 1.5 standard deviations below the mean on the measure of non-verbal cognition. Seven consented individuals with DS were not eligible to participate; one individual did not successful complete the screening battery,

TABLE 2 Participant characteristics in raw scores, standard deviations, and ranges.

	DS group ($n = 19$)			TD group ($n = 19$)			
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**	***P***
Age (years; months)	17; 3	3; 6	11; 1–22; 9	7; 2	0; 7	6; 6–8; 1	0.00*
KBIT-2	16.21	5.02	10–28	25.42	5.63	14–34	0.00*
ROWPVT-4	77.58	27.85	22–132	101.47	8.71	82–117	0.00*
EOWPVT-4	82.95	19.40	50–117	96.79	14.32	68–122	0.01*
TACL-4 Grammatical Morphemes	35.53	8.73	19–54	48.16	4.71	41–54	0.00*
WRMT-3 Word Identification	21.32	6.79	12–37	20.84	6.90	11–34	0.83
Arizona-4	88.92	7.27	74–100	97.90	3.34	88–100	0.00*

DS, Down syndrome; TD, Typically developing; SD, Standard deviation; KBIT-2, Kaufman Brief Intelligence Test-Second Edition (Kaufman, 2004); ROWPVT-4, Receptive One Word Picture Vocabulary Test-Fourth Edition (Martin and Brownell, 2011b); EOWPVT-4, Expressive One Word Picture Vocabulary Test-Fourth Edition (Martin and Brownell, 2011a); TACL-4, Test of Auditory Comprehension of Language-Fourth Edition (Carrow-Woolfolk, 2014); *TACL-4 Scaled scores not reported for DS Group because the age range of the DS group extended beyond the TACL-4 normative age; WRMT-III, Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); Arizona-4, Arizona Articulation Phonology Scale-Fourth Edition (Fudala and Stegall, 2017). This content has been adapted from Prahl and Schuele (2022).

and six individuals did not meet the word reading criteria. Five consented children with TD were not eligible to participate; four did not meet the word reading criteria, and two were not monolingual English speakers.

Participants completed two individual sessions (eligibility and assessment) at the university lab, school, community location (e.g., public library), or in their home. Parents or guardians provided written consent (or participants/power of attorneys for individuals over the age of 18), and participants provided written assent. Each participant's guardian provided demographic background information by completing an intake questionnaire. Eligibility measures included a hearing screening (for the TD group only), word-level reading screening, and measure of non-verbal cognition. To match participants in the TD and DS groups, a word-level reading measure was also administered during the eligibility session. Additional descriptive measures administered at the eligibility session included measures of oral language (receptive and expressive vocabulary and grammar comprehension) and speech accuracy. All eligibility session measures were administered in the same fixed order. The eligibility session lasted 45–60 min. The first author, a certified speech-language pathologist, collected all study data. See **Tables 2**, **3** for participant raw scores and standard scores, respectively, on the descriptive measures.

Assessment measures included four methods of measuring listening comprehension and four methods of measuring reading comprehension. The selected methods represent a range of text and response formats that may frequently be encountered in academic and vocational settings (see **Table 4**). The specific measures were selected because the initial test items at lower levels of difficulty and complexity and the amount of scaffolding provided (i.e., illustrated items on the WRMT-III Passage Comprehension subtest, non-verbal response required on the KABC Reading/Understanding subtest) were expected to reduce task demands to minimize floor effects. Assessment order was counterbalanced across participants in each group to control for order effects. Participants were given breaks between tasks as needed to maintain attention and on-task behavior. The assessment session for each participant lasted 75–100 min. All eligibility and assessment measures were administered in accordance with the manualized directions unless otherwise noted.

Measures

Descriptive measures

Hearing screening

Pure tone audiometry with a standard hand-raising response was used to screen hearing acuity in both ears at frequencies of 500, 1,000, 2,000, and 4,000 Hz at 30 dB. For the DS group, when a participant failed to respond to a particular frequency at 30 dB, the intensity of the tone was increased until a reliable response was obtained for descriptive purposes. The highest intensity necessary to elicit a passing response (two out of three presentations) was recorded. The participants with DS' responses to the tones ranged from 30 to 70 dB ($M = 35$, $SD = 10$) at 500 Hz, 30–60 dB ($M = 34$, $SD = 9$) at 1,000 Hz,

TABLE 3 Participant characteristics in standard score or scaled score means, standard deviations, and ranges.

	DS group ($n = 19$)			TD group ($n = 19$)		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
KBIT-2	52.37	12.25	40–80	109.47	13.26	82–127
ROWPVT-4	59.63	7.87	55–81	112.89	7.80	96–127
EOWPVT-4	62.67	10.81	55–86	111.32	14.57	85–131
TACL-4 grammatical morphemes*				11.58	2.22	8–15
WRMT-3 word identification	61.68	11.07	55–86	110.21	15.91	75–138
Arizona-4	57.5	15.82	50–96	99.58	1.16	96–100

DS, Down syndrome; TD, Typically developing; SD, Standard deviation; KBIT-2, Kaufman Brief Intelligence Test-Second Edition (Kaufman, 2004); ROWPVT-4, Receptive One Word Picture Vocabulary Test-Fourth Edition (Martin and Brownell, 2011b); EOWPVT-4, Expressive One Word Picture Vocabulary Test-Fourth Edition (Martin and Brownell, 2011a); TACL-4, Test of Auditory Comprehension of Language-Fourth Edition (Carrow-Woolfolk, 2014); *TACL-4 Scaled scores not reported for DS Group because the age range of the DS group extended beyond the TACL-4 normative age; WRMT-III, Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); Arizona-4, Arizona Articulation Phonology Scale-Fourth Edition (Fudala and Stegall, 2017). This content has been adapted from Prahl and Schuele (2022).

TABLE 4 Methods of measuring listening comprehension and reading comprehension.

Method	Text format	Response format	Listening comprehension measure	Reading comprehension measure
Non-verbal response	Phrase and sentence	Non-verbal (pointing, acting out)	WJ IV Test of Oral Language Understanding Directions subtest	KABC Reading/ Understanding subtest
Cloze-procedure	Sentence and paragraph	Verbal, one word	WJ IV Test of Oral Language Oral Comprehension subtest	WRMT-III Passage Comprehension subtest
Passage-level with close-ended questions	Paragraph	One word, verbal or pointed	TILLS Listening Comprehension subtest	TILLS Reading Comprehension subtest
Passage-level with open-ended questions	Paragraph	Verbal	WIAT-III Listening Comprehension subtest	WIAT-III Reading Comprehension subtest

WJ IV, Woodcock-Johnson IV (Schrank et al., 2014); KABC, Kaufman Assessment Battery for Children (Kaufman and Kaufman, 1983); WRMT-III, Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); TILLS, Test of Integrated Language and Literacy (Nelson et al., 2015); WIAT-III, Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009). This content has been adapted from Prahl and Schuele (2022).

30–55 dB ($M = 32$, $SD = 6$) at 2,000 Hz, and 30–70 dB ($M = 36$, $SD = 11$) at 4,000 Hz.

Non-verbal intelligence

The Kaufman Brief Intelligence-Second Edition Matrices subtest (KBIT-2; Kaufman, 2004) was administered as a measure of non-verbal intelligence. Test takers infer a relation or rule in a set of pictures or patterns and point to the picture or pattern that best fits the relation or rule. The KBIT-2 includes simple oral instructions and only requires test takers to answer with a meaningful gesture such as pointing. The K-BIT is normed for individuals ages 4–90 and is ideal for those with limited language ability. The mean internal-consistency reliability by age was 0.88 and the mean test-retest reliability by age was 0.83, as reported in the K-BIT manual.

Oral language

The Receptive and Expressive One Word Picture Vocabulary Tests-Fourth Editions (ROWPVT-4 and EOWPVT-4; Martin and Brownell, 2011a,b) were administered as measures of receptive and expressive semantic knowledge. For the ROWPVT-4, test takers point to the picture (out of a field of four) that corresponds with the word the examiner says aloud. The ROWPVT-4 manual reported median internal consistency reliability coefficient by age of 0.97 and the test-retest reliability coefficient of 0.97. For the EOWPVT-4, test takers name pictures. The EOWPVT-4 manual reported median internal consistency reliability coefficient by age of 0.95 and the test-retest reliability coefficient of 0.98. These measures are normed for individuals ages 2–70. The Test of Auditory Comprehension of Language-Fourth Edition Grammatical Morphemes subtest (TACL-4; Carrow-Woolfolk, 2014) was administered as a measure of grammar comprehension. Test takers point to the picture (out of a field of three) that corresponds to stimuli of increasing grammatical complexity presented orally by the examiner. The TACL-4 is normed for individuals ages 3–12. Due to limited grammar comprehension characteristic of the DS phenotype, participants with DS did not reach ceiling levels on this measure despite that the DS participant age range extended beyond the normative age range. The TACL-4 manual reported Grammatical Morphemes mean internal consistency reliability of 0.95 and test-retest reliability of 0.71. The TACL-4 is a valid measure of oral language based on strong evidence of content-description, criterion-prediction, and construct-identification validity.

Word-level reading

On the PALS-K primer list (eligibility measure), test takers read a list of 20 isolated, real words. Each word read accurately via decoding or automatic recognition is scored as correct; percent correct was calculated. On the WRMT-III Word Identification subtest test takers read isolated, real words. A word is scored correct if read accurately within approximately 5 s, whether it is decoded or automatically recognized. In addition to participant matching, the WRMT-III Word Identification raw scores and standard scores are reported for descriptive purposes. Each DS participant began reading at one of the first three entry points depending on the ease with which they read the PALS-K words and each TD participant began reading at their respective grade level entry point. The manualized instructions were then followed to establish the basal and ceiling. The manual reported mean internal-consistency reliability by school-level socioeconomic status of 0.93 and the mean test-retest reliability by age of 0.92. In addition to participant matching, the WRMT-III Word Identification raw scores and standard scores are reported for descriptive purposes. The WRMT-III is normed for individuals ages 4; 6–79. The manual reported mean split-half reliability coefficient by age of 0.93 and the test-retest reliability coefficient of 0.95 for pre-kindergarten through Grade 2, 0.90 for Grades 3–8, and 0.88 for Grades 9–12.

Speech

The Arizona Articulation and Phonology Scale-Fourth Revision (Arizona-4; Fudala and Stegall, 2017) was administered as a measure of speech accuracy. Test takers label pictures. If the child does not provide the intended label, the label is modeled by the examiner and repeated by the test taker. The examiner notes speech sound production errors. The Word Articulation Total Score was calculated based on the weighted values (a reflection of how frequently the sound occurs in American speech) of the sounds that were produced accurately. The Arizona-4 is normed for individuals ages 18 months to 21 years. Internal consistency coefficients reported in the manual ranged from 0.90 to 0.97 depending on age and test-retest reliability was 0.96. The Arizona-4 has strong evidence of content, response process, construct, and convergent validity.

Dependent variable measures

Listening comprehension

Raw scores were calculated for all four listening comprehension measures to capture incremental differences between participants that would be obscured by using standard scores for individuals with DS (Mervis and Klein-Tasman, 2004). Two subtests from the Woodcock-Johnson IV Test of Oral Language (WJ IV; Schrank et al., 2014), normed for individuals ages 2–90 years, were administered. The Understanding Directions subtest requires a non-verbal response. Test takers follow single-sentence directions, presented orally via an audio recording to point to familiar objects with varying characteristics (e.g., size and location) in a picture scene. This subtest has a median reliability of 0.86 in the 5–19 age range and 0.87 in the adult age range as reported in the manual. The Oral Comprehension subtest uses a cloze procedure. Test takers listen to a short audio-recorded passage and supply the missing word from the final sentence in a one- or two-sentence passage. This subtest has a median reliability of 0.82 in the 5–19 age range and 0.80 in the adult age range. The Test of Integrated Language and Literacy Skills (TILLS; Nelson et al., 2015). Listening Comprehension subtest was administered as a measure that used passage-level text paired with close-ended questions. It is normed for individuals ages 6;0 to 18;11. On this subtest, test takers selected "yes," "no," or "maybe" to answer questions about passage-level text read aloud by the examiner. As an accommodation, a card with the three choices (yes, no, and maybe) was placed on the table in front of the examiner as additional visual support and to provide a non-verbal response option. The mean intraclass reliability coefficient reported in the manual was 0.95 and test-retest reliability was 0.77. The TILLS was found to have strong construct and concurrent validity. The Wechsler Individual Achievement Test-III (WIAT-III; Wechsler, 2009) Listening Comprehension Oral Discourse Comprehension subtest was administered. Test takers listen to audio-recorded passage-level text and then verbally answer open-ended questions read aloud by the examiner. Test takers' answers were scored according to the possible correct answers listed on the Record Form; one point was awarded for each correct answer and zero points for incorrect answers. The mean internal reliability coefficient reported in the manual was 0.83 and test-retest reliability was 0.75. The WIAT-III was found to have strong evidence of validity based on content, response process, and internal structure.

Reading comprehension

The Kaufman Ability Battery for Children (KABC; Kaufman and Kaufman, 1983) Reading/Understanding subtest requires a non-verbal response. Test takers act out written directions. The Reading/Understanding subtest is normed for individuals ages 7–12. The manual reported mean internal consistency coefficient based on the split-half reliability method based on age of 0.90 for preschool children and 0.93 for children ages 5–12 years and the test-retest reliability coefficient of 0.83, 0.88, and 0.92 for ages 2; 6–4, 5–8, and 9–12; 6, respectively. The WRMT-III Passage Comprehension subtest uses cloze procedure. Initial passages are single sentences and passages increase in length across the subtest. Initial passages are accompanied by a picture, but pictures are phased out as passages increase in length. Test takers supply the missing word located anywhere in the passage to complete the meaning of a sentence or paragraph that they read. The manual reported mean internal consistency coefficient based on the split-half reliability method based on age of 0.90 and the test-retest reliability coefficient of 0.86 for Pre-Kindergarten-Grade 2, 0.88 for grades 3–8, and 0.81 for grades 9–12. Raw scores were calculated on the KABC Reading/Understanding and WRMT-III Passage Comprehension subtests. The TILLS Reading Comprehension subtest was administered as a measure that used passage-level text paired with close-ended questions. It is normed for individuals ages 6; 6–18; 11. On this subtest, test takers read passage-level text and questions and then selected "yes," "no," or "maybe" to answer the questions. In accordance with the manualized directions, the TILLS Reading Comprehension subtest was discontinued if test takers made seven or more miscues when reading the first passage. Rather than assigning a raw score of 0, for the purpose of this study, if the discontinue rule was met, the participant was considered to score at the floor level and a score was not included on this measure for the construct validity analyses. The mean intraclass reliability coefficient reported in the manual was 0.99 and test-retest reliability was 0.86. The WIAT-III Reading Comprehension subtest uses passage-level text paired with open-ended questions. Test takers read passage-level text and then verbally answer open-ended questions read aloud by the examiner. Test takers' answers were scored according to the criteria provided on the Record Form; answers could be scored as 2-points, 1-point, or 0-points for some questions and scored as 2-point or 0-points on other questions. Four to eight questions were asked per passage. For participants with DS, the entry point was based on their word-level reading grade equivalent based on the WRMT-III Word Identification subtest and for TD participants, the entry point was based on their current grade level. Because WIAT-III Reading Comprehension scores are based on the particular item set administered and the total raw scores from different item sets are not directly comparable, vertically scaled scores (i.e., weighted scores) were used as outlined in the assessment manual. The WIAT-III is normed for individuals ages 4–50. The mean internal reliability coefficient reported in the manual was 0.86 and test-retest reliability was 0.90.

Design and variables

To establish inter-rater reliability, initially the first author scored all measures. A graduate student reliability coder with formal training in psychoeducational assessment was trained on the scoring procedures for the dependent measures. She then independently scored the participants' assessment sessions from video and audio recordings for a random selection (=25%) of participants; only video recordings with camera angles that allowed for valid assessment scoring were eligible for random selection. Inter-rater reliability was estimated using intraclass correlation coefficients (ICCs). ICCs account for differences in scores between coders as well as the variance among participants on the measures of interest. For the dependent measures, the mean ICC value was 0.99 for the DS group and 0.93 for the TD group (Hessling, 2020) and thus, the primary coder's scoring was used in the analyses. The ICC values were all excellent for the DS group (0.94–1.00) and the values ranged from good to excellent for the TD group (0.80–1.00). For both groups, the lowest ICC values were observed for the WIAT-III measures which is not surprising given that the response format is an open-ended verbal response, and thus the rubric requires decisions by the coder which can lead to lack of agreement across scorers. See the blue cells in Figures 1, 2 for the ICC values for each measure by group. The primary scorer and reliability scorer double scored all measures from the assessment protocols (93% inter-rater agreement) and discrepancies were resolved by consensus before data was double entered for analysis.

To answer research question one, separate MTMM were created for the DS and TD groups and analyzed based on Campbell and Fiske's (1959) guidelines. Within the matrices, four classes of cells are distinguished. *Monotrait-monomethod cells* (reliability diagonal, blue cells) constitute the main diagonal of the matrix and contain the reliability coefficient of each trait in each method, as measured by interclass correlations as an estimate of inter-rater reliability. Because a high consistency of scores is an essential requirement for test validity, the monotrait-monomethod cells are expected to be the highest values in the MTMM. *Monotrait-heteromethod cells* (validity diagonal, yellow cells) reflect the correlation between measures of the same trait measured using different methods (convergent validity). Because the two measures are of the same trait, strong correlations are expected. *Heterotrait-monomethod cells* (purple cells) reflect the correlation among measures that share the

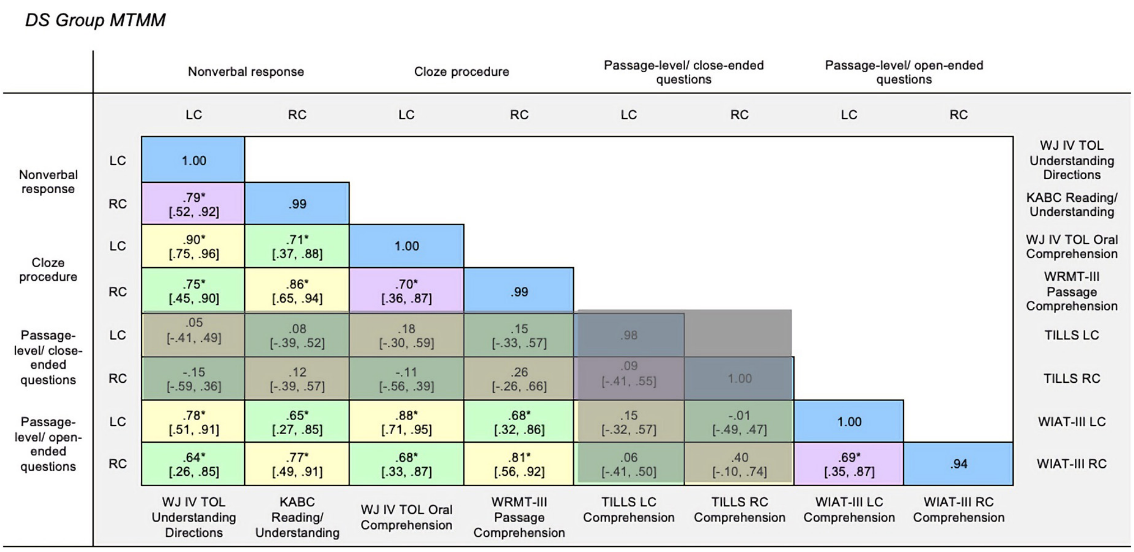

FIGURE 1
DS group MTMM. Multitrait-multimethod matrix for Down syndrome group. Monotrait-monomethod cells (reliability diagonal) marked in blue, monotrait-heteromethod cells (validity diagonal) marked in yellow (heterotrait-monomethod cells marked in purple, and heterotrait-heteromethod cells marked in green. LC, Listening Comprehension; RC, Reading Comprehension; WJ IV TOL, Woodcock-Johnson Test of Oral Language IV (Schrank et al., 2014); KABC, Kaufman Assessment Battery for Children (Kaufman and Kaufman, 1983); WRMT-III, Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); WIAT-III, Wechsler Individual Achievement Test-Third Edition (Wechsler, 2009). *$p > 0.05$.

TD Group MTMM

		Nonverbal response LC	Nonverbal response RC	Cloze procedure LC	Cloze procedure RC	Passage-level/ close-ended questions LC	Passage-level/ close-ended questions RC	Passage-level/ open-ended questions LC	Passage-level/ open-ended questions RC	
Nonverbal response	LC	.99								WJ IV TOL Understanding Directions
	RC	.41 [-.06, .73]	.99							KABC Reading/ Understanding
Cloze procedure	LC	.62* [.22, .84]	.43 [-.03, .74]	.96						WJ IV TOL Oral Comprehension
	RC	.41 [-.05, .73]	.78* [.50, .91]	.58* [.17, .82]	1.00					WRMT-III Passage Comprehension
Passage-level/ close-ended questions	LC	.15 [-.33, .56]	.49* [.04, .77]	.39 [-.08, .72]	.57* [.16, .81]	1.00				TILLS LC
	RC	.66* [.22, .88]	.64* [.18, .87]	.66* [.23, .88]	.63* [.17, .86]	.53* [.02, .82]	1.00			TILLS RC
Passage-level/ open-ended questions	LC	.70* [.36, .88]	.38 [-.09, .71]	.66* [.29, .86]	.44 [-.02, .74]	.17 [-.30, .58]	.43 [-.10, .77]	.80		WIAT-III LC
	RC	.55* [.13, .80]	.88* [.70, .95]	.60* [.21, .83]	.81* [.56, .92]	.59* [.19, .82]	.76* [.41, .92]	.57* [.16, .81]	.84	WIAT-III RC
		WJ IV TOL Understanding Directions	KABC Reading/ Understanding	WJ IV TOL Oral Comprehension	WRMT-III Passage Comprehension	TILLS LC Comprehension	TILLS RC Comprehension	WIAT-III LC Comprehension	WIAT-III RC Comprehension	

FIGURE 2

TD group MTMM. Multitrait-multimethod matrix for Down syndrome group. Monotrait-monomethod cells (reliability diagonal) marked in blue, monotrait-heteromethod cells (validity diagonal) marked in yellow, heterotrait-monomethod cells marked in purple, and heterotrait-heteromethod cells marked in green. LC, Listening Comprehension; RC, Reading Comprehension; WJ IV TOL, Woodcock-Johnson Test of Oral Language IV (Schrank et al., 2014); KABC, Kaufman Assessment Battery for Children (Kaufman and Kaufman, 1983); WRMT-III, Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); WIAT-III, Wechsler Individual Achievement Test-Third Edition (Wechsler, 2009). *$p > 0.05$.

same measurement method, but measure different traits. These values are considered an index of discriminant validity and thus should be weaker than the correlations in the yellow cells. If, however, these correlations are high, it is because measuring different constructs with the same methods results in correlated measures. *Heterotrait-heteromethod cells* (green cells) reflect the correlation among measures that differ in trait and method (discriminant validity). Because these correlations share neither trait nor method, the heterotrait-heteromethod cells are expected to be the lowest values in the MTMM. The degree to which these cells are smaller than the heterotrait-monomethod cells is considered an index of the influence of the methods factor. Summary level statistics are reported for each matrix to ascertain the extent to which the cells overlap or differ from one another according to Campbell and Fiske's (1959) guidelines. To determine whether the correlations were significantly different, we evaluated whether there was overlap in the confidence intervals around the correlation coefficients. In addition, we demonstrated sufficient power (=0.80) to interpret at least moderate correlation coefficients (0.50–0.70) within the MTMM using G*Power 3.1 Software (Faul et al., 2009). Cook's distance was used to monitor for undue influence across analyses relevant to each cell within the MTMM. There was no evidence that any individual data points were leveraging regression lines. Because scores were not reported for participants who met the discontinue rule on the TILLS Reading Comprehension measure, follow-up analyses demonstrated that the study results were robust to listwise deletion. To answer the second research question, a combined MTMM with data from both groups was created. We conducted 56 separate regression analyses for each cell to evaluate whether the evidence of construct validity was moderated by group for each cell in the MTMM. In addition to monitoring the data for outliers, scores for both groups were determined to be normally distributed based on visual analysis of histograms. For each regression, we were interested in evaluating the interaction effect—whether the effect of one measure on another measure changed depending on group membership (DS vs. TD).

Results

Mean listening comprehension and reading comprehension raw scores as well as the number of participants who completed the measures and yielded scores above the floor level (i.e., raw score > 0) are displayed in **Table 5**.

Evaluating construct validity in Down syndrome group

The DS group MTMM is displayed in **Figure 1**. The reliability diagonal marked in blue reflects the

TABLE 5 Participant reading comprehension and listening comprehension raw scores and participants scoring above floor level.

Measure	TD group ($n = 19$)				DS group ($n = 19$)			
	Mean	SD	Range	# (%) of participants above floor level	Mean	SD	Range	# (%) of participants above floor level
Listening comprehension								
WJ-IV TOL Understanding Directions	35.47	6.53	22–50	19 (100)	17.74	9.89	2–37	19 (100)
WJ-IV TOL Oral Comprehension	15.26	2.62	10–20	19 (100)	7.84	4.62	0–17	17 (89)
TILLS Listening Comprehension	13.26	3.87	7–20	19 (100)	9.05	3.55	0–15	18 (95)
WIAT-III Listening Comprehension	11.42	1.90	8–16	19 (100)	4.95	4.13	0–15	18 (95)
Reading comprehension								
KABC Reading/Understanding	10.58	5.64	2–19	19 (100)	8.58	5.32	0–18	17 (89)
WRMT-III Passage Comprehension	13.32	3.73	9–22	19 (100)	8.68	4.41	2–17	19 (100)
TILLS Reading Comprehension	9.53	3.80	4–15	15 (79)	5.12	3.82	0–11	16 (84)
WIAT-III Reading Comprehension*	46.37	9.71	30–64	19 (100)	27.42	14.67	2–55	19 (100)

*Vertically scaled scores (not raw scores) reported for this measure, due to administration rules.
TD, typically developing; DS, Down syndrome; SD, standard deviation; WJ IV, Woodcock-Johnson IV (Schrank et al., 2014); KABC, Kaufman Assessment Battery for Children (Kaufman and Kaufman, 1983); WRMT-III, Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); TILLS, Test of Integrated Language and Literacy (Nelson et al., 2015); WIAT-III, Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009).

interclass correlation values as a measure of inter-rater reliability for each measure. The interclass correlation values for all measures were excellent (ICCs = 0.94–1.00).

Monotrait-heteromethod

The monotrait-heteromethod cells marked in yellow reflect the correlation between measures of the same trait using different measurement methods. Statistically significant and strong correlations ($rs = 0.77$–0.90, $p < 0.05$) were observed in half of the monotrait-heteromethod yellow cells, thus reflecting good convergent validity across the measures except for the passage-level with close-ended questions (TILLS) measures. Two notable exceptions of non-significant and weak correlations were observed, first, for the relation between the passage-level with close-ended questions (TILLS) listening comprehension measure ($rs = 0.05$–0.18, $p > 0.05$) and each of the other measures. Second, non-significant and weak correlations were observed for the relation between the passage-level with close-ended questions (TILLS) reading comprehension measure ($rs = 0.12$–0.40, $p > 0.05$) and each of the other measures. Given that the TILLS measures are not converging with other measures of the same construct, it appears that the TILLS measures are not tapping the construct that it's purporting to measure. It also may be the case that the TILLS measures are tapping a different dimension of the construct when compared with the other measurement methods. Because of the questionable construct validity of the TILLS (shaded cells in **Figure 1**), we will hone in on the cells reflecting only the associations of the remaining measures from this point forward.

Heterotrait-monomethod

The heterotrait-monomethod cells marked in purple reflect the correlations between listening comprehension and reading comprehension measures using the same method. Statistically significant and strong correlations ($rs = 0.69$–0.79, $p < 0.05$) were observed between the two traits—listening comprehension and reading comprehension—for three out of the four measurement methods (i.e., non-verbal response, cloze procedure, and passage-level with open-ended questions). This pattern of strong correlations (Cohen, 1998) demonstrates shared method variance, that measuring different constructs with the same methods results in correlated measures. The values in monotrait-heteromethod (yellow) cells were not significantly stronger than the values in the heterotrait-monomethod (purple) cells, as evidenced by the overlapping confidence intervals.

Heterotrait-heteromethod

Lastly, the heterotrait-heteromethod cells marked in green reflect the correlation between different traits measured using different methods. When excluding the associations related to the TILLS measures, statistically significant and strong correlations ($rs = 0.64$–0.75, $p < 0.05$) were observed in the remaining heterotrait-heteromethod green cells. Based on the overlapping confidence intervals, the values in the heterotrait-heteromethod (green) cells were not all significantly weaker than the values in the heterotrait-monomethod or the monotrait-heteromethod cells.

Summary of construct validity in Down syndrome group

In summary, the results for the DS group provide some support that these various measurement methods for listening comprehension and reading comprehension, with the exception of the TILLS measures, are measuring the same constructs. Within the DS group, however, the results do not provide evidence of discriminant validity. In other words, the listening comprehension and reading comprehension constructs are not separable.

Evaluating construct validity in typical development group

The TD group MTMM is displayed in **Figure 2**. Similar to the DS group, the interclass correlation values for all measures ranged from good to excellent (ICCs = 0.80–1.00) as shown in the reliability diagonal marked in blue.

Monotrait-heteromethod

The monotrait-heteromethod cells marked in yellow reflect the correlation between measures of the same trait using different measurement methods. Statistically significant and strong ($r = 0.5$; Cohen, 1998) correlations ($rs = 0.62$–0.88, $p < 0.05$) were observed in 75% of the monotrait-heteromethod yellow cells, thus reflecting good convergent validity across all measures. Similar to the DS group, though only for the passage-level with close-ended questions (TILLS) listening comprehension measure, non-significant and weak correlations ($rs = 0.15$–0.39, $p > 0.05$) were observed between this measure and each of the other measures. Given that this listening comprehension measure was not converging with other measures of the same construct, it appears that the TILLS listening comprehension measure is not tapping the construct that it purports to measure for the TD group. Because of the questionable construct validity of the TILLS listening comprehension measure (shaded cells in **Figure 2**), we again hone in on the cells reflecting only the associations of the remaining measures from this point forward.

Heterotrait-monomethod

The heterotrait-monomethod cells marked in purple reflect the correlations between listening comprehension and reading comprehension measures using the same method. Statistically significant and strong correlations ($rs = 0.53$–0.58, $p < 0.05$) were observed between the two traits of interest—listening comprehension and reading comprehension—for three out of the four measurement methods (i.e., cloze procedure, passage-level with close-ended questions, and passage-level with open-ended questions). However, the two traits of interest were not significantly correlated ($r = 0.41$, $p > 0.05$) for the non-verbal response (KABC Reading/Understanding and WJ IV TOL Oral Comprehension) measurement method. This pattern of strong correlations (Cohen, 1998) demonstrates shared method variance, that measuring different constructs with the same methods results in correlated measures. It is also important to note that the values in monotrait-heteromethod (yellow) cells were not significantly stronger than the values in the heterotrait-monomethod (purple) cells, as evidenced by the overlapping confidence intervals.

Heterotrait-hetermethod

Lastly, the heterotrait-heteromethod cells marked in green reflect the correlation between different traits measured using different methods. Statistically significant and strong correlations ($rs = 0.49$–0.66, $p < 0.05$) were observed in 58% of the heterotrait-heteromethod green cells, with the remaining cells reflecting moderate correlations ($rs = 0.38$–0.44, $p > 0.05$). Given this range of values and the overlapping confidence intervals, the values in the heterotrait-heteromethod (green) cells were not all significantly weaker than the values in the heterotrait-monomethod or the monotrait-heteromethod cells.

Summary of construct validity in typical development group

In summary, the results in the TD group provide some support that these measures, with the exception of the TILLS Listening Comprehension measure, are measuring the same constructs. Within the TD group, the results do not provide evidence of discriminant validity. In other words, listening comprehension and reading comprehension constructs are not separable.

Construct validity group comparisons

Regression analyses were performed to test whether the associations of interest within the MTMM (excluding the reliability diagonal) varied according to group. Only five associations were significantly different, all but one of which were within heterotrait cells. See **Figure 3** and **Table 6** for the regression results. Four associations reflected that correlations were slightly stronger in the DS group, although all correlations for both groups ranged from moderate to strong ($rs = 0.38$–0.79). Further, these correlations did not yield a meaningful interpretation given that they all index associations between indices purported to tap different constructs (i.e., associations moderated by group were all in heterotrait cells). The remaining three correlations moderated by group suggest that associations were attenuated in the DS group compared with the TD group. However, two of these correlations were expected to be small given that they were values contained in heterotrait-heteromethod cells. That is, they reflect correlations between variables that were purported to tap different constructs using different methods. In summary, only a few associations were moderated by group and thus the moderated associations on the whole do not suggest variable construct validity for these measures in the TD group compared with the DS group. The pattern of results was not moderated in any way that is meaningful for interpretation within the MTMM.

Discussion

In this pilot study, we assessed the construct validity of four parallel measures of listening comprehension and reading comprehension for individuals with DS and their peers with TD. Evaluation of psychometric properties is important to validate the use of commonly used norm-referenced measures for various clinical populations. Though establishing measures as demonstrating strong reliability and validity is essential in development, it is also essential to determine whether those characteristics hold true for each research sample of interest. Further, given that there is a variety of methods for assessing listening comprehension and reading comprehension, it is important to determine whether measures of the same traits using different methods demonstrate convergent validity (Cutting and Scarborough, 2006; Keenan et al., 2008). Researchers can make informed decisions regarding assessment and outcome measure selection based on the empirical evidence regarding feasibility and psychometric properties of commonly used measures.

Demonstrating construct validity

The MTMM approach proposed by Campbell and Fiske (1959) was chosen for the assessment and interpretation of construct validity. We were interested in evaluating whether measures of the same construct that use different methods were more strongly correlated than (a) measures of different constructs that use the same method and (b) measures of different constructs that use different methods. In other words, we were interested in evaluating whether the monotrait-heteromethod (yellow) associations were more strongly correlated compared with the heterotrait-monomethod (purple) and heterotrait-heteromethod (green) associations. Inspection of the MTMMs for both groups revealed that monotrait-heteromethod associations were not significantly different when compared with the heterotrait-monomethod and heterotrait-heteromethod associations, as evidenced by overlapping confidence intervals. Thus, the results indicate that the listening comprehension and reading comprehension constructs may not be separable or cannot be meaningfully differentiated for the study groups in this developmental period using these particular measures.

The current preliminary findings are consistent with the broader literature in which researchers have suggested that listening comprehension and reading comprehension are highly intercorrelated in readers (e.g., Sticht et al., 1974; Sinatra, 1990; Nation and Snowling, 2004). In a study of concurrent and longitudinal predictors of reading comprehension, Nation and Snowling (2004) examined reading development in 72 children at 8.5 and 13 years of age. Based on concurrent analyses at Time 1, they found that even after controlling for non-verbal cognition, phonological awareness, semantics, and expressive vocabulary, listening comprehension was the strongest contributor to reading comprehension, accounting for 31% of the unique variance. Based on longitudinal analyses, they found that even after controlling for Time 1 non-verbal cognition, reading comprehension, non-word reading, phonological awareness, semantics, and expressive vocabulary, listening comprehension accounted for an additional 14% of the unique variance in reading comprehension at Time 2. Further, Ebert and Scott (2016) found statistically significant and strong correlations between listening comprehension and reading comprehension among younger ($r = 0.47$, $p < 0.05$; aged 6.0–8.11) and older ($r = 0.47$, $p < 0.05$; aged 9.1–16.7) school-aged children with TD. Listening comprehension and reading comprehension have been found to be highly intercorrelated ($rs = 0.41–56$, $p < 0.05$) in studies with children with DS as well, with some stronger correlations between listening comprehension and reading comprehension observed in children with DS compared with children with TD (Roch and Levorato, 2009; Laws et al., 2016). Our results indicate a similar pattern with stronger correlations between listening comprehension and reading comprehension measured using the same method observed in the DS group ($rs = 0.69–0.79$) compared with the TD group ($rs = 0.41–0.58$).

It is important to consider the possible influences of development when interpreting these findings. Based on the simple view of reading model, it is not surprising that we observed strong intercorrelations between listening comprehension and reading comprehension given that the participants in this study had achieved some level of proficiency with word recognition. However, the strength of the relation between listening comprehension and reading comprehension is likely to vary across development. Thus, as other researchers have suggested, measuring listening comprehension earlier in development may be useful in predicting future reading comprehension (Nation and Snowling, 2004; Ebert and Scott, 2016). Capturing the predictive power of listening comprehension may be particularly important when children are developing reading skills or may be considered emergent readers, which may be a prolonged process for individuals with DS. Additionally, despite establishing strong correlations between listening comprehension and reading comprehension, the constructs are not perfectly correlated, thus some unexplained variance remains. Although beyond the scope of this project, future research must evaluate how other variables, such as those illustrated in Scarborough's (2001) reading rope model, contribute to reading comprehension for individuals with DS.

TABLE 6 Regression coefficients predicting each listening comprehension and reading comprehension measure.

Variable	Non-verbal response						Cloze procedure						Passage-level/ Multiple-choice questions						Passage-level/ Open-ended questions					
	LC: WJ TOL understanding directions			RC: KABC reading/ understanding			LC: WJ TOL oral comprehension			RC: WRMT passage comprehension			LC: TILLS listening comprehension			RC: TILLS reading comprehension			LC: WIAT listening comprehension			RC: WIAT reading comprehension		
	Est.	*SE*	*P*	Est.	*SE*	*P*	Est.	*SE*	*P*	Est.	*SE*	*P*	Est.	*SE*	*P*	Est.	*SE*	*P*	Est.	*SE*	*P*	Est.	*SE*	*P*
Predicting non-verbal response LC: WJ TOL understanding directions																								
Measure × Group				1.22	0.39	**0.00***	0.63	0.51	0.22	1.05	0.52	**0.05***	−0.33	0.84	0.69	−1.29	0.79	0.11	0.30	0.86	0.73	0.17	0.23	0.46
Predicting non-verbal response RC: KABC reading/understanding																								
Measure × Group	−0.12	0.22	0.60				−0.04	0.48	0.93	−0.03	0.29	0.91	−0.16	0.50	0.75	−0.70	0.45	0.13	−0.13	0.70	0.86	−0.22	0.12	0.07
Predicting Cloze Procedure LC: WJ TOL Oral Comprehension																								
Measure × Group	0.09	0.12	0.46	0.52	0.22	**0.02***				0.37	0.27	0.18	−0.06	0.39	0.87	−0.60	0.37	0.12	0.20	0.35	0.57	0.08	0.10	0.45
Predicting cloze procedure RC: WRMT passage comprehension																								
Measure × Group	−0.15	0.17	0.39	0.14	0.19	0.45	−0.36	0.35	0.32				−0.19	0.39	0.62	−0.36	0.36	0.33	−0.32	0.53	0.56	−0.12	0.09	0.21
Predicting passage-Level/MC LC: TILLS listening comprehension																								
Measure × Group	−0.028	0.21	0.20	−0.08	0.27	0.78	−0.55	0.41	0.19	−0.23	0.32	0.48				−0.43	0.34	0.22	−0.28	0.58	0.63	−0.18	0.12	0.16
Predicting passage-level/MC RC: TILLS reading comprehension																								
Measure × Group	−0.54	0.20	**0.01***	−0.38	0.26	0.15	−1.06	0.40	**0.01***	−0.39	0.32	0.23	−0.44	0.36	0.23				−0.89	0.58	0.13	−0.21	0.11	0.07
Predicting passage-level/OE LC: WIAT listening comprehension																								
Measure × Group	0.10	0.13	0.44	0.40	0.20	0.06	0.33	0.21	0.12	0.40	0.25	0.11	0.13	0.33	0.70	−0.23	0.32	0.49				0.10	0.09	0.30
Predicting passage-level/OE LC: WIAT reading comprehension																								
Measure × Group	−0.18	0.52	0.72	0.42	0.58	0.47	−0.16	0.98	0.87	0.48	0.67	0.48	−0.60	1.10	0.59	−0.39	0.96	0.69	−0.37	1.42	0.80			

*$p > 0.05$.
WJ TOL, Woodcock-Johnson IV Test of Oral Language (Schrank et al., 2014); WIAT, Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009); TILLS, Test of Integrated Language and Literacy (Nelson et al., 2015); KABC, Kaufman Assessment Battery for Children (Kaufman and Kaufman, 1983); WRMT, Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); Est., Estimate; MC, Multiple-Choice; OE, Open-Ended.

Combined Groups MTMM

			Nonverbal response: LC (WJ IV TOL Understanding Directions)	Nonverbal response: RC (KABC Reading/ Understanding)	Cloze procedure: LC (WJ IV TOL Oral Comprehension)	Cloze procedure: RC (WRMT-III Passage Comprehension)	Passage-level/ close-ended questions: LC (TILLS LC Comprehension)	Passage-level/ close-ended questions: RC (TILLS RC Comprehension)	Passage-level/ open-ended questions: LC (WIAT-III LC Comprehension)	Passage-level/ open-ended questions: RC (WIAT-III RC Comprehension)
Nonverbal response	LC	WJ IV TOL Understanding Directions								
	RC	KABC Reading/ Understanding	.54*							
Cloze procedure	LC	WJ IV TOL Oral Comprehension	.91	.53*						
	RC	WRMT-III Passage Comprehension	.74	.78	.75					
Passage-level/ close-ended questions	LC	TILLS LC	.42	.35	.51	.51				
	RC	TILLS RC	.44	.43	.43*	.57	.50			
Passage-level/ open-ended questions	LC	WIAT-III LC	.89	.49*	.92	.71	.45	.43		
	RC	WIAT-III RC	.78*	.73*	.80	.86	.50*	.67	.80	

FIGURE 3
Combined groups MTMM. Multitrait-multimethod matrix for Down syndrome group. Monotrait-monomethod cells (reliability diagonal) marked in blue, monotrait-heteromethod cells (validity diagonal) marked in yellow, heterotrait-monomethod cells marked in purple, and heterotrait-heteromethod cells marked in green. LC, Listening Comprehension; RC, Reading Comprehension; WJ IV TOL, Woodcock-Johnson Test of Oral Language IV (Schrank et al., 2014); KABC, Kaufman Assessment Battery for Children (Kaufman and Kaufman, 1983); WRMT-III, Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); WIAT-III, Wechsler Individual Achievement Test-Third Edition (Wechsler, 2009). *$p > 0.05$.

Inspection of the MTMMs provides additional information regarding the construct validity of the measurement methods (e.g., non-verbal response, cloze response, passage-level text with closed-ended response, and passage-level text with open-ended response) as well as the specific measures evaluated in this study. For the TD and DS groups, the results reflect high inter-rater reliability (ICC > 0.5) which is a precursor to further evaluate validity. For the DS group, the pattern of results reflecting high monotrait-heteromethod correlations ($r = 0.5$) were strongly in favor of convergent validity for three listening comprehension and reading comprehension measures. Three of the measurement methods converged on single listening comprehension and reading comprehension constructs: non-verbal response, cloze procedure, passage-level with open-ended questions. These results could also be interpreted as demonstrating that the three specific measures administered for each of these methods converged on single listening comprehension and reading comprehension constructs. For the TD group, the pattern of results reflecting high monotrait-heteromethod correlations ($r = 0.5$) were strongly in favor of convergent validity for all but one measure. Three of the measurement methods converged on a single listening comprehension construct: non-verbal response, cloze procedure, passage-level with open-ended questions and all of the measurement methods converged on a single reading comprehension construct.

Although the passage-level with close-ended questions (TILLS) listening comprehension and reading comprehension measures reflected a high degree of inter-rater reliability, inspection of the MTMM provides no evidence of construct validity, with one exception in the TD group. The correlations related to indices that were derived for the TILLS measures did not reflect strong construct validity for the DS group (TILLS Listening Comprehension and Reading Comprehension subtests) and was replicated in the TD group (TILLS Listening Comprehension subtest). Next, we speculate about some possible explanations for why the construct validity of the TILLS measure may be compromised in our study sample.

One possible explanation relates to the response format of the TILLS—passage-level text with close-ended questions. Anecdotally, some participants appeared to randomly select a response given the close-ended or forced choice comprehension questions which in turn is not necessarily a true reflection of their listening comprehension or reading comprehension. Another possible explanation related to the response format is the presence of "maybe" as a potential answer choice for the close-ended questions. Being able to consider "maybe" to a comprehension question reflects a certain degree of abstraction, which participants may not have fully understood, despite completing trial items with instructional feedback provided for the "maybe" response, if needed. We acknowledge that we did not account for a number of other variables that may influence performance on these measures. For instance, perhaps the TILLS measures, when used with children in these age and developmental ranges, are influenced more by verbal working memory or place greater demands on the decoding skills of participants when compared with the other measurement methods. For example, in the TD group, one possible explanation for why differential evidence of construct validity was observed may be that the TILLS Listening Comprehension measure was more heavily influenced by verbal working memory compared with the TILLS Reading Comprehension measure. The higher memory load required for the TILLS passage-level text measures may also explain why this measure did not converge with the other listening comprehension and reading comprehension measures in the DS group. Additional research evaluating the construct validity of the TILLS for individuals with DS is also warranted.

Group comparisons

In comparing the DS and TD groups, we aimed to evaluate the extent to which the evidence of construct validity was moderated by group. Within the MTMMs, a similar pattern of construct validity was demonstrated between groups, and thus only 7 or 25% of associations were moderated by group, though not in any particularly meaningful way. We only evaluated whether the associations were moderated by group because we hypothesized that group membership would capture any other potential moderators given that group differences on any other variables (e.g., non-verbal cognition and grammar comprehension) would be accounted for by group membership.

Limitations

The results of this study should be interpreted with the following limitations in mind. First, the MTMM approach that we used involves a primarily logical rather than analytical approach to guide interpretation of construct validity. Despite this limitation, analyzing and reporting the confidence intervals within the MTMMs enabled us to evaluate the extent to which the various classes of cells differed from one another according to Campbell and Fiske's (1959) guidelines. Further analysis using confirmatory factor analysis was not possible in this pilot study, though this approach could be used to evaluate construct validity with a larger participant sample. Second, we acknowledge that not all individuals with DS demonstrate sufficient word-level reading and reading comprehension skills necessary to complete the reading and language-related literacy tasks included in this study. In the current study, six individuals with DS (10 to 18 years of age) did not meet the eligibility criteria to participate in the study due to limited word-level reading abilities. These individuals may have presented with some pre-reading skills, though they did not demonstrate sufficient reading given the established eligibility criteria and thus we cannot draw any conclusions regarding the reading abilities for those non-readers or emergent readers who were not included in the study. Further, the study results should be viewed as a minimal estimate of the construct validity of the measures given that convenience sampling was used and due to the small sample size. In summary, the study results are specific to a particular subset of individuals with DS and the degree to which these results can be generalized for a broader and more representative sample in individuals with DS is unknown.

Finally, we did not control for a number of variables that are known to contribute to listening comprehension and reading comprehension in the analyses evaluating whether the evidence of construct validity was moderated by group. As mentioned previously, we hypothesized that any between-group differences that may have moderated the evidence of construct validity would be accounted for in analyzing the effect of group membership. To evaluate this hypothesis, future research should explore the extent to which participant characteristics such as working memory, background knowledge, verbal reasoning, and executive functioning influence the construct validity of various measurement methods across clinical populations (e.g., Perfetti et al., 1996; Snow, 2002; Kintsch and Kintsch, 2005; Miller et al., 2013).

Implications

The methods employed in this study address many of the challenges to valid assessment of listening comprehension and reading comprehension for individuals with DS. Although challenges with how listening comprehension and reading comprehension are defined will likely persist, we have clearly operationalized these two constructs as well as selected measures that align with those definitions. Reading comprehension was operationalized as constructing meaning from written text, and listening comprehension was operationalized as constructing meaning from read-aloud written text. In addition, the various measurement methods included in this study reflect a comprehensive range of methods for assessing listening comprehension and reading comprehension. Lastly, we also considered the DS phenotype in selecting which measures to evaluate. Given that floor effects are often observed when using norm-referenced assessments with individuals with DS, we included measures with a range of text formats so that participants, specifically those with limited word-level reading, would be able to complete at least some initial test items. Further, because individuals with DS often have limited speech intelligibility, we included measures with a range of response formats to limit the impact of poor speech intelligibility on examiner understanding and scoring of responses. We demonstrated a high degree of inter-rater reliability for the measurement methods that required a non-verbal or minimal verbal response. Not surprisingly we observed the lowest degree of inter-rater reliability in both groups, though still an acceptable value, for the most verbally robust (passage-level text with open-ended question) measurement method.

Consistent with the rationale for conducting this study, the results provide guidance on potentially valid measures for assessing listening comprehension and reading comprehension for individuals with DS and their peers with TD. Overall, the results demonstrate the inter-rater reliability of all the measures evaluated. However, strong evidence of convergent validity was only observed for three out of the four measurement methods, with no evidence of discriminant validity for the listening comprehension and reading comprehension constructs. The construct validity results are of critical importance in regards to using psychometrically sound assessment measures. Further, for examiners who may not have experience assessing and making accuracy judgements for individuals with limited speech intelligibility, it may be important to consider the response format alongside the evidence presented herein. Similarly, for test takers with limited reading proficiency, it also may be important to consider the text format. Taken together, the study results can guide listening comprehension and reading comprehension assessment selection for individuals with DS and their peers with TD. By establishing the inter-rater reliability and construct validity of multiple listening comprehension and reading comprehension measurement methods, researchers and clinicians can have greater confidence in using these measures to quantify skills and characterize patterns of strengths and weaknesses.

Future directions

This initial measurement investigation lays the foundation for developing and evaluating individualized reading interventions for individuals with DS. The current study results provide preliminary evidence of construct validity of multiple measurement methods, as well as identified the optimal methods of assessment, inform the outcome measure selection for studies of reading intervention for individuals with DS. Future analyses of the data from this investigation will apply generalizability (G) theory to conduct a decision (D) study to determine the number of measures needed to obtain stable estimates of listening comprehension and reading comprehension based on the current data of individuals with DS. The results of a G and D study will extend the current findings and enable researchers to ascertain how many of the evaluated (valid) measures should be administered to adequately capture an individual's listening comprehension and reading comprehension skills. As mentioned previously, additional research is needed for a larger sample replication and with individuals across different stages of reading development. Future research should also further evaluate the construct validity of the passage-level with close-ended questions (TILLS) listening comprehension and reading comprehension measures.

Conclusion

The current study contributes to the evidence base regarding the reliability and validity of commonly used measures of listening comprehension and reading comprehension in terms of their utility for individuals with DS. Key findings include strong evidence of reliability and construct validity for three of four measurement methods (non-verbal response, cloze procedure, and passage-level with open-ended questions). These results support the use of these measurement methods in clinical practice and future studies of reading comprehension in individuals with DS.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Vanderbilt University Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AP: conceptualization, formal analysis, investigation, project administration, and writing and preparing manuscript for sending. CS: conceptualization, supervision, reviewing, and editing manuscript. Both authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research, completed as part of AP's dissertation at Vanderbilt University, was supported by the US Department of Education Preparation of Leadership Personnel grant (H325D140087) and in part by Vanderbilt CTSA award No. UL1TR000445 from the National Center for Advancing Translational Sciences and a Russell G. Hamilton Graduate Leadership Institute Dissertation Enhancement Grant, Vanderbilt University Graduate School.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. Department of Education, the National Center for Advancing Translational Sciences, or the Russell G. Hamilton Graduate Leadership Institute.

References

Afacan, K., Wilkerson, K. L., and Ruppar, A. L. (2018). Multicomponent reading interventions for students with intellectual disability. *Remed. Special Educ.* 39, 229–242. doi: 10.1177/0741932517702444

American Speech-Language-Hearing Association (ASHA) (2022). *Childhood Hearing Screening.* Available online at: https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589935406&section=Key_Issues. (accessed April 1, 2022)

Buckley, S. (2001). *Reading and Writing for Individuals With Down Syndrome: An Overview*. Portmouth, UK: The Down Syndrome Educational Trust.

Byrne, A., Buckley, S., MacDonald, J., and Bird, G. (1995). Investigating the literacy, language and memory skills of children with down syndrome. *Down Synd. Res. Pract.* 3, 53–58. doi: 10.3104/reports.51

Byrne, A., MacDonald, J., and Buckley, S. (2002). Reading, language and memory skills: A comparative longitudinal study of children with Down syndrome and their mainstream peers. *Br. J. Educ. Psychol.* 72, 513–529. doi: 10.1348/00070990260377497

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56, 81–105. doi: 10.1037/h0046016

Carrow-Woolfolk, E. (2014). *Test for Auditory Comprehension of Language (TACL-4).* Austin, TX: Pro-Ed.

Catts, H., and Kamhi, A. (1986). The linguistic basis of reading disorders: Implications for the speech-language pathologist. *Lang. Speech Hearing Serv. Schools* 17, 329–341. doi: 10.1044/0161-1461.1704.329

Catts, H., and Kamhi, A. G. (1999). "Defining reading disabilities," in *Language and Reading Disabilities*, eds H. Catts and A. Kamhi (Toronto: Allyn & Bacon), 50–72.

Catts, H. W., Adlof, S. M., and Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *J. Speech Lang. Hear. Res.* 49, 278–293. doi: 10.1044/1092-4388(2006/023)

Catts, H. W., Herrera, S., Nielsen, D. C., and Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. *Read. Writ.* 28, 1407–1425. doi: 10.1007/s11145-015-9576-x

Chapman, R. S., and Hesketh, L. J. (2000). Behavioral phenotype of individuals with down syndrome. *Mental Retard. Dev. Disabili. Res. Rev.* 6, 84–95. doi: 10.1002/1098-277920006:23.0.CO;2-P

Cohen, J. (1998). *Statistical power analysis for behavioral sciences.* New York, NY: Routledge.

Cossu, G., Rossini, F., and Marshall, J. C. (1993). When reading is acquired but phonemic awareness is not: A study of literacy in Down's syndrome. *Cognition* 46, 129–138. doi: 10.1016/0010-0277(93)90016-O

Cutting, L., and Scarborough, H. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Sci. Stud. Read.* 10, 277–299. doi: 10.1207/s1532799xssr1003_5

Dunn, L. M., and Markwardt, F. C. (1970). *Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Service.

Ebert, K. D., and Scott, C. M. (2016). Bringing the simple view of reading to the clinic: Relationships between oral and written language skills in a clinical sample. *J. Commun. Dis.* 62, 147–160. doi: 10.1016/j.jcomdis.2016.07.002

Elleman, A. M., and Compton, D. L. (2017). Beyond comprehension strategy instruction: What's next? *Lang. Speech Hear. Serv. Schools* 48, 84–91. doi: 10.1044/2017_LSHSS-16-0036

Faul, F., Erdfelder, E., Buchner, A., and Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Fidler, D. (2006). The emergence of a syndrome-specific personality profile in young children with down syndrome. *Down Synd. Res. Pract.* 10, 53–60. doi: 10.3104/reprints.305

Fidler, D., and Philofsky, A. (2009). The down syndrome behavioral phenotype: Taking a developmental approach. *Down Synd. Res. Pract.* 12, 37–44. doi: 10.3104/reviews/2069

Fidler, D. J. (2005). The emerging down syndrome behavioral phenotype in early childhood: Implications for practice. *Infants Young Child. Interdiscipli. J. Early Child. Int.* 18, 86–103. doi: 10.1097/00001163-200504000-00003

Francis, D. J., Fletcher, J. M., Catts, H. W., and Tomblin, J. B. (2005). "Dimensions affecting the assessment of reading comprehension," in *Children's Reading Comprehension and Assessment*, eds S. Paris and S. Stahl (Mahwah, NJ: Erlbaum Associates), 369–394.

Fuchs, D., Hendricks, E., Walsh, M. E., Fuchs, L. S., Gilbert, J. K., Zhang, T. W., et al. (2018). Evaluating a multidimensional reading comprehension program and reconsidering the lowly reputation of tests of near-transfer. *Learn. Disabili. Res. Pract.* 33, 11–23. doi: 10.1111/ldrp.12162

Fudala, J., and Stegall, S. (2017). *Arizona Articulation Proficiency Scale–Fourth Revision (Arizona–4).* Los Angeles, CA: Western Psychological Services.

García, J. R., and Cain, K. K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Rev. Educ. Res.* 84, 74–111. doi: 10.3102/0034654313499616

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remed. Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104

Gray, S., Catts, H., Logan, J., and Pentimonti, J. (2017). Oral language and listening comprehension: Same or different constructs? *J. Speech Lang. Hear. Res.* 60, 1273–1284. doi: 10.1044/2017_JSLHR-L-16-0039

Groen, M. A., Laws, G., Nation, K., and Bishop, D. V. (2006). A case of exceptional reading accuracy in a child with down syndrome: Underlying skills and the relation to reading comprehension. *Cogn. Neuropsychol.* 23, 1190–1214. doi: 10.1080/02643290600787721

Hessling, A. (2020). *The Not So Simple View Of Reading In Down syndrome Ph. D, Thesis*. ProQuest Dissertations Publishing.

Hogan, T. P., Adlof, S. M., and Alonzo, C. N. (2014). On the importance of listening comprehension. *Int. J. Speech Lang. Pathol.* 16, 199–207. doi: 10.3109/17549507.2014.904441

Invernizzi, M., Meier, J. D., Swank, K., and Juel, C. (1997). *Phonological Awareness Literacy Screening (PALS)*. Charlottesville, VA: University Printing Services.

Kaufman, A. S. (2004). *Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., and Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service, doi: 10.1037/t27677-000

Keenan, J., Betjemann, R., and Olson, R. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Sci. Stud. Read.* 12, 281–300. doi: 10.1080/10888430802132279

Kim, Y.-S., and Phillips, B. (2014). Cognitive correlates of listening comprehension. *Read. Res. Quart.* 49, 269–281. doi: 10.1002/rrq.74

Kintsch, E., and Kintsch, W. (2005). "Comprehension," in *Children's Reading Comprehension and Assessment*, eds S. Paris and S. Stahl (Mahwah, NJ: L. Erlbaum Associates), 369–394.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge: Cambridge University Press.

Language and Reading Research Consortium (2017). Oral language and listening comprehension: Same or different constructs? *J. Speech Lang. Hear. Res.* 60, 1273–1284. doi: 10.1044/2017_JSLHR-L-16-0039

Lauren, L., and Caldwell, J. (2001). *Qualitative Reading Inventory-3*. New York: Addison Wesley Longman.

Laws, G., Brown, H., and Main, E. (2016). Reading comprehension in children with down syndrome. *Read. Writ.* 29, 21–45. doi: 10.1007/s11145-015-9578-8

Lemons, C. J., King, S. A., Davidson, K. A., Puranik, C. S., Al Otaiba, S., Fulmer, D., et al. (2017). Developing an early reading intervention aligned with the down syndrome behavioral phenotype. *Focus Autism Other Dev. Disabili.* 32, 176–187. doi: 10.1177/1088357615618941

MacGinitie, W. H., MacGinitie, R. K., Maria, K., and Dreyer, L. G. (2000). *Gates–MacGinitie Reading Tests*. Itasca, IL: Riverside.

Martin, G. E., Klusek, J., Estigarribia, B., and Roberts, J. E. (2009). Language characteristics of individuals with Down syndrome. *Topics Lang. Disord.* 29:112. doi: 10.1097/TLD.0b013e3181a71fe1

Martin, N., and Brownell, R. (2011b). *Receptive One-Word Picture Vocabulary Test, (ROWPVT-4)*. Novato, CA: Academic Therapy Publications.

Martin, N., and Brownell, R. (2011a). *Expressive One-Word Picture Vocabulary Test, (EOWPVT-4)*. Novato, CA: Academic Therapy Publications.

Mervis, C. B., and Klein-Tasman, B. P. (2004). Methodological issues in group-matching designs: α levels for control variable comparisons and measurement characteristics of control and target variables. *J. Autism Dev. Dis.* 34, 7–17. doi: 10.1023/B:JADD.0000018069.69562.b8

Miller, B., Cutting, L. E., and McCardle, P. (2013). *Unraveling Reading Comprehension*. Baltimore, MD: Brookes Publishing.

Nash, H., and Heath, J. (2011). The role of vocabulary, working memory and inference making ability in reading comprehension in down syndrome. *Res. Dev. Disabili.* 32, 1782–1791. doi: 10.1016/j.ridd.2011.03.007

Nation, K., and Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *J. Res. Read.* 27, 342–356. doi: 10.1111/j.1467-9817.2004.00238.x

Nelson, N., Plante, E., Helm-Estabrooks, N., and Hotz, G. (2015). *Test of Integrated Language and Literacy Skills (TILLS)*. Baltimore, MD: Brookes Publishing.

Paris, S. G., and Stahl, S. A. (2005). *Children's Reading Comprehension and Assessment*. Mahwah, NJ: L. Erlbaum Associates, doi: 10.4324/9781410612762

Perfetti, C. A., Marron, M. A., and Foltz, P. W. (1996). "Sources of comprehension failure: Theoretical perspectives and case studies," in *Reading Comprehension Difficulties: Processes and Intervention*, eds C. Cornoldi and J. Oakhill (Mahwah, NJ: Routledge), 137–165.

Petersen, D. B., and Stoddard, A. (2018). Psychometric requirements of oral and written language progress monitoring assessments. *Perspect. ASHA Special Int. Groups* 3, 180–197. doi: 10.1044/persp3.SIG1.180

Prahl, A., and Schuele, C. M. (2022). Reading and listening comprehension in individuals with down syndrome and word reading–matched typically developing children. *Am. J. Speech Lang. Pathol.* 31, 359–374. doi: 10.1044/2021_AJSLP-21-00208

Roch, M., and Levorato, M. C. (2009). Simple view of reading in Down's syndrome: The role of listening comprehension and reading skills. *Int. J. Lang. Commun. Dis.* 44, 206–223. doi: 10.1080/13682820802012061

Roizen, N. J., Wolters, C., Nicol, T., and Blondis, T. A. (1993). Hearing loss in children with down syndrome. *J. Pediatr.* 123, S9–S12. doi: 10.1016/S0022-3476(05)81588-4

Scarborough, H. S. (2001). "Connecting early language and literacy to later reading (DIS) abilities: Evidence, theory, and practice," in *Handbook for research in early literacy*, eds S. Neuman and D. Dickinson (New York, NY: Guilford Press), 97–110.

Schrank, F. A., Mather, N., and McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Oral Language*. Rolling Meadows, IL: Riverside.

Sinatra, G. M. (1990). Convergence of listening and reading processing. *Read. Res. Quart.* 25, 115–130. doi: 10.2307/747597

Snow, C. (2002). *Reading for Understanding: Toward An R&D Program in Reading Comprehension*. Washington, DC: Rand Corporation.

Sticht, T., Beck, L., Hauke, R., Kleiman, G., and James, J. A. (1974). *Reading: A Developmental Model*. Alexandria, VA: Human Resources Research Organization, doi: 10.1037/e457482004-001

Stoel-Gammon, C. (1997). Phonological development in down syndrome. *Mental Retard. Dev. Disabili. Res. Rev.* 3, 300–306. doi: 10.1002/(SICI)1098-277919973:43.0.CO;2-R

Wechsler, D. (1992). *Wechsler Individual Achievement Test: WIAT*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2009). *Wechsler Individual Achievement Test–Third Edition: WIAT-III*. San Antonio, TX: Psychological Corporation.

Wiederholt, J. L., and Bryant, B. R. (1992). *Gray Oral Reading Tests: GORT-3*. Austin, TX: Pro-Ed.

Wishart, J. G. (1996). "Avoidant learning styles and cognitive development in young children," in *New Approaches to Down Syndrome*, eds B. Stratford and P. Gunn (London: Cassell), 173–205.

Woodcock, R. W. (2011). *Woodcock reading mastery tests (WRMT-III)*. Circle Pines, MN: Pearson, doi: 10.1037/t15178-000

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.

BRIEF RESEARCH REPORT
published: 12 August 2022
doi: 10.3389/fpsyg.2022.929433

Inclusion of Individuals With Neurodevelopmental Disorders in Norm-Referenced Language Assessments

Susan J. Loveall[1], Marie Moore Channell[2], Laura J. Mattie[2] and Alexandria E. Barkhimer[2]*

[1] Department of Special Education and Communication Disorders, University of Nebraska–Lincoln, Lincoln, NE, United States, [2] Department of Speech and Hearing Science, University of Illinois at Urbana-Champaign, Champaign, IL, United States

Standardized, norm-referenced language assessment tools are used for a variety of purposes, including in education, clinical practice, and research. Unfortunately, norm-referenced language assessment tools can demonstrate floor effects (i.e., a large percentage of individuals scoring at or near the lowest limit of the assessment tool) when used with some groups with neurodevelopmental disorders (NDDs), such as individuals with intellectual disability and neurogenetic syndromes. Without variability at the lower end of these assessment tools, professionals cannot accurately measure language strengths and difficulties within or across individuals. This lack of variability may be tied to poor representation of individuals with NDDs in normative samples. Therefore, the purpose of this study was to identify and examine common standardized, norm-referenced language assessment tools to report the representation of individuals with NDDs in normative samples and the range of standard/index scores provided. A systematic search identified 57 assessment tools that met inclusion criteria. Coding of the assessment manuals identified that most assessment tools included a "disability" or "exceptionality" group in their normative sample. However, the total number of individuals in these groups and the number of individuals with specific NDDs was small. Further, the characteristics of these groups (e.g., demographic information; disability type) were often poorly defined. The floor standard/index scores of most assessment tools were in the 40s or 50s. Only four assessment tools provided a standard score lower than 40. Findings of this study can assist clinicians, educators, and researchers in their selections of norm-referenced assessment tools when working with individuals with NDDs.

Keywords: language assessment, neurodevelopmental disorders (NDDs), norm-referenced assessments, language, standardized assessment

Edited by:
Marisa Filipe,
Faculdade de Letras da Universidade de Lisboa, Portugal

Reviewed by:
Andreia Salé Veloso,
University of Porto, Portugal
Ana Paula Vale,
University of Trás-os-Montes and Alto Douro, Portugal

****Correspondence:***
Susan J. Loveall
sloveall-hague2@unl.edu

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *26 April 2022*
Accepted: *09 June 2022*
Published: *12 August 2022*

Citation:
Loveall SJ, Channell MM, Mattie LJ and Barkhimer AE (2022) Inclusion of Individuals With Neurodevelopmental Disorders in Norm-Referenced Language Assessments. Front. Psychol. 13:929433. doi: 10.3389/fpsyg.2022.929433

INTRODUCTION

Because the development of language is critical to meeting the demands of everyday life, accurate assessment of language is critical for diagnosing primary language disorders, identifying secondary language difficulties across other neurodevelopmental disorders (NDDs), and ultimately developing effective, targeted intervention and treatment plans that include monitoring progress over time. However, most commonly used assessment tools were not developed specifically for individuals with NDDs, and professionals who work with this population are often left without much guidance as to which tools to select. Therefore, the purpose of this study is to identify and examine common standardized, norm-referenced assessment tools of language to report the representation of individuals with NDDs in normative samples and the range of standard/index scores provided. This information can assist clinicians, educators, and researchers in their selections of norm-referenced assessment tools when working with individuals with NDDs.

Neurodevelopmental disorders are common in the United States, with birth cohort data ($n > 3.3$ million children) reporting that by 8 years of age, 23.9% of publicly insured children and 11% of privately insured children had a diagnosis of one or more NDDs (Straub et al., 2022). NDDs include a range of conditions resulting from either a genetic or multifactorial etiology (i.e., a combination of genetic and environmental factors) that occur during the developmental period and that are characterized by delays in cognition, communication, behavior, and/or motor skills (American Psychiatric Association [APA], 2013; Van Herwegen et al., 2015; World Health Organization [WHO], 2019). These conditions impact personal, social, academic, and/or occupational functioning (American Psychiatric Association [APA], 2013; Van Herwegen et al., 2015; World Health Organization [WHO], 2019). Specific NDDs include intellectual disability, communication disorders, autism, attention-deficit/hyperactivity disorder (ADHD), neurodevelopmental motor disorders (e.g., developmental coordination disorder, stereotypic movement disorder, and tic disorders), specific learning disorders, and some neurogenetic syndromes (e.g., Down syndrome, fragile X syndrome, and Williams syndrome). Different NDDs often co-occur; for example, individuals with autism may also have an intellectual disability or ADHD, and individuals with Down syndrome typically also have an intellectual disability (American Psychiatric Association [APA], 2013; Van Herwegen et al., 2015).

Many individuals with NDDs experience difficulties with language, though the exact pattern of these difficulties varies across diagnoses and individuals (e.g., Luyster et al., 2011). Some individuals have an NDD in which the primary diagnosis is specific to language. For example, developmental language disorder (under the umbrella of communication disorders) is linked to difficulties with pragmatics and structural aspects of language (e.g., Reed, 2018). Other individuals may have a different primary NDD but still also have language difficulties. For example, ADHD is often associated with secondary language difficulties in pragmatics (e.g., Geurts and Embrechts, 2008; Hawkins et al., 2016; Helland et al., 2016). Individuals with intellectual disability and neurogenetic syndromes also experience a range of difficulties in spoken language (Abbeduto et al., 2016; McDuffie et al., 2017), but the exact patterns of strength and difficulty often vary across different etiologies. For example, individuals with Down syndrome typically have relative strengths in vocabulary but more significant difficulties in grammar and syntax, whereas individuals with Williams syndrome tend to have relative strengths in concrete vocabulary but difficulties with relational vocabulary and pragmatics (Abbeduto et al., 2019).

One of the most common ways to measure language abilities is *via* standardized, norm-referenced language assessment tools. Norm-referenced assessment tools refer to those that have been tested (i.e., "normed") on a large group of individuals meant to represent the age and demographic makeup of those for whom the test is intended to be used (Peña et al., 2006). When the assessment is administered in a standardized way, as outlined in the administration manual, an individual's performance can then be compared to that of the normative sample to see how the individual compares to peers of a similar age and demographic makeup. However, the exact makeup of the normative sample can influence the scores of a norm-referenced assessment tool and its outcomes for the individual who is assessed (Peña et al., 2006; Spaulding et al., 2006). Thus, which norm-referenced language assessment tool a professional should use depends on the purpose of the assessment and the individual being assessed.

Two primary purposes of language assessment are to (1) diagnose language disorders and (2) describe language profiles. When the primary purpose of a language assessment is for diagnosis, a professional may want to select a norm-referenced assessment tool that did not include individuals with disabilities in the normative sample. Individuals with a primary language disorder may exhibit subtle, yet meaningful, language delays in which scores fall close to the diagnostic cut-off. In these cases, if individuals with disorders were included in the normative sample of the assessment tool being used, the normative group mean would be lower, with an increased standard deviation, resulting in decreased classification accuracy for identifying language impairment (i.e., a missed diagnosis), as demonstrated in a simulation study by Peña et al. (2006). On the other hand, for individuals with NDDs whose primary diagnosis is something other than a communication disorder (e.g., intellectual disability), the purpose of a language assessment is not typically for diagnosis but rather to describe their language profile and/or to identify their areas of strength or difficulty. This information can be used to guide intervention and academic planning. In these instances, it is important that norm-referenced assessment tools are not only reliable and valid for use in this population but that they also capture a wide range of skill levels, including at the lower-performing end where individuals with intellectual disability often fall.

Unfortunately, many standardized, norm-referenced assessment tools are not normed beyond three or four standard deviations below the normative mean, causing many participants with NDDs, such as individuals with intellectual disability, to score at the floor on standard/index scores (e.g., cf. Spaulding et al., 2006; Kasari et al., 2013; DiStefano et al., 2020). Floor effects occur when a large percentage of individuals have standard scores at or near the lowest limit of an assessment tool because its measurement range does not extend low enough to capture low levels of skills/performance (Hessling et al., 2004; McBee, 2010; Zhu and Gonzalez, 2017). Floor effects limit variability or separation in standard scores at the lower end of the assessment tool, and information regarding true differences across individuals is lost. These compressed scores, in turn, prevent researchers, clinicians, and educators from accurately capturing language strengths and difficulties within or across individuals and from tracking if individuals make clinically meaningful change/gains over time (Hessl et al., 2009; Sansone et al., 2014; Esbensen et al., 2017).

This issue of compressed scores is reflected in recent calls for the development of appropriate outcome measures for individuals with intellectual disability and neurogenetic syndromes (Esbensen et al., 2017; Hendrix et al., 2020). Floor effects have even been linked to recent failed pharmacological clinical trials for individuals with neurogenetic syndromes (Berry-Kravis et al., 2013; Budimirovic et al., 2017; Esbensen et al., 2017; see Abbeduto et al., 2020 for an overview; Baumer et al., 2022). Thus, many researchers, clinicians, and educators working with individuals with NDDs are pushing to develop more sensitive measures for use with these populations. Although there has been research addressing floor effects in cognitive/IQ tests (Hessl et al., 2009; Sansone et al., 2014), this line of research has not yet been extended to norm-referenced language assessment tools. At the same time, there is, and will continue to be, a need to use norm-referenced language assessment tools with this population, especially in clinical practice. Therefore, professionals who are assessing individuals who have an NDD that is *not* a primary language disorder and who are likely to score at the lower level of the assessment (e.g., intellectual disability) should select a norm-referenced assessment tool that has a low floor, to improve their ability to identify areas of strength and difficulty.

Given the variability in language skills across and within individuals with NDDs, and the various purposes of norm-referenced language assessment tools, researchers, educators, and clinicians need to be able to make informed decisions to select assessment tools that best meet their needs. Some may need norm-referenced assessment tools that did not include individuals with NDDs in their normative samples for better classification/diagnostic accuracy. Others may need norm-referenced assessment tools that have included individuals with NDDs in their normative samples, or at least that demonstrate variability at lower-performing ends of the assessment tool. Unfortunately, information on normative samples and psychometric properties is often not easily accessible before purchase, making it difficult to identify if a specific norm-referenced assessment tool meets one's needs. Therefore, the purpose of this study was to:

1) Identify common standardized, norm-referenced assessment tools of language.
2) Report the representation of individuals with NDDs in their normative samples and the range of standard/index scores available.

MATERIALS AND METHODS

Inclusion Criteria

To be included in our review, language assessment tools had to be a direct measure of oral language (e.g., the assessment tool could focus on any aspect of oral language, including phonology, but could not focus exclusively on articulation/speech or mostly on academics), have been published in the last 20 years (i.e., in or after 2002), and have been normed in the United States for individuals 22 years or younger (i.e., covers the developmental period; Schalock et al., 2021). In addition, the measure had to be published in English and commercially available for purchase by a main publishing house in the United States. Five main publishing houses in the United States were identified for review: Brookes Publishing, PARInc, Pearson, ProEd, and WPS. Screeners and caregiver-, teacher-, or self-report measures were not included.

Procedures

Identification of Assessment Tools

Each of the five publishing houses' websites was reviewed by two independent research assistants. The research assistants reviewed all assessment tools listed or tagged on the website as "speech and language" (or similar). Using the search function, they also searched for each of the following terms: "language," "grammar," "syntax," "morphology," "vocabulary," "phonology," "pragmatics," "listening comprehension," and "auditory processing." Research assistants excluded any assessment tools that clearly did not meet the inclusion criteria but defaulted to including any language assessment tools that were unclear as to whether or not they met the study's inclusion criteria. The first three authors made the final decisions on which assessment tools to include in situations of discrepancies across reviewers or when all reviewers were unsure. This process resulted in the identification of 55 assessment tools.

The assessment tool list was then reviewed by one university speech-language clinic director and one speech-language pathology clinical assistant professor with expertise in school-age language disorders. The clinicians were asked to review the list of assessment tools to determine if any language assessment tools were missed in the review. This process resulted in the inclusion of two additional assessment tools for a total of 57 assessment tools.

Coding of Assessment Tools

Following the identification of assessment tools, each assessment tool's administration or technical manual was independently reviewed and coded by two research assistants for the variables listed below. Discrepancies were identified and resolved by the first and fourth author, with assistance from research assistants, by consulting the assessment tool manual.

Variables

Full Normative Samples

The full normative sample of each assessment tool was coded for the total sample size and demographic information, including sex/gender, race/ethnicity, and geographic region. Each assessment tool was also coded for the chronological ages it was normed for and if the socioeconomic status of its normative sample was considered/reported.

Standard/Index Scores

We also documented the minimum and maximum standard/index scores provided by each assessment tool.

Inclusion of Individuals With Disabilities and Specific Neurodevelopmental Disorders in the Normative Sample

Many assessment tools included individuals with "disabilities" or "exceptionalities" in their normative samples without clearly differentiating disability type. For this reason, each assessment tool was coded for the total number of individuals with disabilities included in the normative sample. When possible, this information was also reported by disability type, including specific NDDs (e.g., number with intellectual disability, autism spectrum disorder, ADHD, learning disability). Demographic information was also coded for disability groups.

RESULTS

Full Normative Samples

Demographic information on the full normative samples is reported in **Table 1**. A majority (n = 45/57) of assessment tools had normative sample sizes of over 1,000 individuals with relatively equal numbers of males and females. These samples included individuals from all regions of the United States, though five assessment manuals did not specify where their participants were from, and one did not have participants representing all regions of the United States. Sample diversity (defined in terms of race and ethnicity) was reported for all but one assessment tool [i.e., the Bilingual English-Spanish Assessment, BESA (Peña et al., 2018)] and varied across assessment tools. Most assessment tools (n = 49/57) considered some aspect of socioeconomic status (e.g., maternal education, income, and/or percentage receiving free or reduced lunch).

Standard/Index Scores

The floor standard/index score of most assessment tools was in the 40s (n = 27/57) or 50s (n = 23/57). Three assessment tools had floor scores in the 60s [i.e., Clinical Assessment of Pragmatics, CAPs (Lavi, 2019); Communication and Symbolic Behavior Scales, Normed Edition, CSBS (Wetherby and Prizant, 2002); Listening Comprehension Test, LCT-2 (Bowers et al., 2006)]. Only four measures from our list provide a standard score lower than 40. The Phonological Awareness Test, Second Edition: Normative Update (PAT-2:NU; Robertson and Salter, 2018) provides standard scores down to 39. The WORD Test 3 Elementary (WORD-3; Bowers et al., 2014) provides scores down to −9. The Test of Adolescent and Adult Language (TOAL-4; Hammill et al., 2007) provides scores down to 34, and the Test of Early Communication and Emerging Language (TECEL; Huer and Miller, 2011) provides standard scores down to 25.

Inclusion of Individuals With Disabilities and Specific Neurodevelopmental Disorders in the Normative Sample

Number of Individuals With Disabilities and Neurodevelopmental Disorders in Normative Samples

Of the 57 assessment tools, 52 indicated that they included individuals with disabilities in the normative sample (numbers and percentages of individuals with disabilities and specific NDDs are reported in **Table 2**). However, five assessment tools did not include or report on any individuals with disabilities in their normative sample: the BESA (Peña et al., 2018), the Test of Integrated Language and Literacy Skills (TILLS; Nelson et al., 2016), the Test of Phonological Awareness, Second Edition Plus (TOPA-2+; Torgeson and Bryant, 2004), the Test of Semantic Skills Primary (TOSS-P; Huisingh et al., 2002), and the Vocabulary Assessment Scales – Expressive (VAS-E) and Receptive (VAS-R; Gerhardstein Nader, 2013). These assessment tools are therefore not included in **Table 2**.

Nine assessment tools indicated that they may have included some individuals with disabilities, or alternatively did not exclude all individuals with disabilities. However, they did not track and/or specify if/how many individuals with disabilities were included. To be as inclusive as possible, these assessment tools are reported in **Table 2**.

For the remaining 43 assessment tools that clearly included individuals with disabilities in their normative samples, the total number varied across assessment tools. However, in most cases, this was a low percentage of the normative sample (ranging from 3 to <26%). Only six assessment tools had normative samples in which 20% or more of the normative sample had a disability or an "exceptionality status": Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5; Wiig et al., 2013), Comprehensive Receptive and Expressive Vocabulary Test, Third Edition (CREVT-3; Wallace and Hammill, 2013), Khan-Lewis Phonological Assessment, Third Edition (KLPA-3; Khan and Lewis, 2015), Social Language Development Test – Adolescent: Normative Update (SLDT-A:NU; Bowers et al., 2017a), Test of Language Development – Intermediate: Fifth Edition (TOLD-I:5; Hammill and Newcomer, 2019), Test of Pragmatic Language, Second Edition (TOPL-2; Phelps-Terasaki and Phelps-Gunn, 2007). Another 21 assessment tools had normative samples in which 10–19% had a disability. Fifteen assessment tools had normative samples in which less than 10% had disabilities, and one assessment tool [i.e., the Auditory Processing Abilities Test, APAT (Ross-Swain and Long, 2004)] had between 9 and 16%, though the exact percentage was unclear. Further, the overall sample size (n) of individuals with disabilities was not reported in all assessment tools. When possible, we estimated the overall percentage of individuals with disabilities based on the available information (e.g., reported n's of specific NDDs). This method does not account for dual-diagnoses, though, so the reported number may be smaller than estimated.

Descriptions and Demographic Information of Individuals With Disabilities and Neurodevelopmental Disorders in Normative Samples

The makeup (i.e., disability type and demographic information) of individuals with disabilities was often poorly defined for these 43 assessment tools. Ten assessment manuals did not specify what type(s) of disabilities were represented in their normative sample (i.e., the number of individuals with specific disabilities or NDDs such as intellectual disability or learning disabilities). Another 10 assessment tools only

TABLE 1 | Normative samples in norm-referenced language assessments.

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
A Language Processing Skills Assessment, Fourth Edition (TAPS-4)	Martin et al., 2018	5:0–21:11	55–145	2023	Females (55), Males (45)	Race: White/Caucasian (80), Black/African American (9), Asian American (4), American Indian/Alaska Native (1), Native Hawaiian/Pacific Islander (0.3), Two or more Ethnicities (6); Ethnicity: Hispanic (20), Non-Hispanic (80), Not reported (0.1)	Northeast (15), Midwest (26), South (34), West (24)	Yes
Arizona Articulation and Phonology Scale, Fourth Revision (Arizona-4)	Fudala and Stegall, 2017	1:6–21:11	≤ 50– ≥ 125	3192	Females (52), Males (48)	White (56), Black/African American (17), Asian (2), American Indian/Alaska Native (0.4), Native Hawaiian/Pacific Islander (0.3), Other (4), Hispanic Origin (20)	Northeast (13), Midwest (25), South (42), West (21)	Yes
Auditory Processing Abilities Test (APAT)	Ross-Swain and Long, 2004	5:0–12:11	55–145	1087	Females (51), Males (49)	White (72), Black (12), Asian (1), Other (3), Hispanic (12)	Northeast (17), Midwest (19), South (41), West (24)	No
Bankson Expressive Language Test - Third Edition (BELT-3)	Bankson et al., 2018	3:0–6:11	49–156	684	Females (50), Males (50)	Race: White (72), Black/African American (15), Asian/Pacific Islander (5), American Indian/Eskimo/Aleut (1), Two or More (7); Ethnicity: Hispanic (24), Non-Hispanic (76)	Northeast (18), Midwest (22), South (37), West (23)	Yes
Bankson-Bernthal Test of Phonology, 2nd edition (BBTOP-2)	Bankson and Bernthal, 2020	3:0–9:11	40–128	770	Females (50), Males (50)	Race: White (73), Black/African American (16), Asian/Native Hawaiian/Other Pacific Islander (3), American Indian/Alaska Native (< 1), Two or more (7); Ethnicity: Hispanic (22), Non-Hispanic (78)	Northeast (16), Midwest (23), South (37), West (24)	Yes
Bilingual English-Spanish Assessment (BESA)	Peña et al., 2018	4:0–6:11	< 55– > 145	756	Females (47), Males (42), Not reported (11)	Not specified	Northeast (29), South (47), West (24)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Clinical Assessment of Articulation and Phonology, Second Edition (CAAP-2)	Secord and Donohue, 2014	2:6–11:11	55–124	1486	Females (51), Males (49)	Race White (81) African American (13) Other (6) Ethnicity Non-Hispanic (84) Hispanic (16)	Northeast (17), Midwest (21), South (38), West (24)	Yes
Clinical Assessment of Pragmatics (CAPs)	Lavi, 2019	7:0–18:11	64–136	914	Females (50), Males (50)	Race White (77), Black/African American (11), Asian (4), Other (7) Ethnicity Hispanic (14)	Northeast (19), Midwest (24), South (30), West (27)	Yes
Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5)	Wiig et al., 2013	5:0–21:11	40–160	2380	Females (50), Males (50)	White (57), African American (14), Asian (4), Other (6), Hispanic (20)	Northeast (15), Midwest (24), South (37), West (24)	Yes
Clinical Evaluation of Language Fundamentals Preschool, Third Edition (CELF-P3)	Wiig et al., 2020	3:0–6:11	40–160	700	Females (50), Males (50)	White (56) African American (13) Asian (2) Other (7) Hispanic (23)	Northeast (16), Midwest (23), South (44), West (18)	Yes
Communication and Symbolic Behavior Scales, Developmental Profile, First Normed Edition (CSBS), behavior sample	Wetherby and Prizant, 2002	1:0–2:0 Can be used up to 6:0 if developmental level is younger than 24 months	65–135	337	Females (49), Males (51)	Race White (87), Black (9), Asian (3), Other (0.9) Ethnicity Hispanic (6), Non-Hispanic (94)	Not specified	Yes
Comprehensive Assessment of Spoken Language, Second Edition (CASL-2)	Carrow-Woolfolk, 2017	3:0–21:11	40–160	2394	Females (51), Males (49)	White (57), Black/African American (14), Asian (3), American Indian/Alaska Native (0.4), Native Hawaiian/Pacific Islander (0.3), Other (3), Hispanic Origin (22)	Northeast (23), Midwest (16), South (37), West (25)	Yes
Comprehensive Receptive and Expressive Vocabulary Test, Third Edition (CREVT-3)	Wallace and Hammill, 2013	5:0–89:11	45–155	1535	Females (49), Males (51)	Race White (80), Black/African American (14), Asian/Pacific Islander (4), American Indian/Eskimo/Aleut (< 1), Two or more (2) Ethnicity Hispanic (15), Non-Hispanic (85)	Northeast (18), Midwest (20), South (37), West (25)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)	Wagner et al., 2013	4:0–24:11	43–165	1900	Females (51), Males (49)	Ethnicity White (76), Black/African American (14), Asian/Pacific Islander (2), Two or More (4), Other (4) Hispanic Hispanic (16), Non-Hispanic (84)	Northeast (18), Midwest (23), South (35), West (24)	Yes
Diagnostic Evaluation of Articulation and Phonology (DEAP)	Dodd et al., 2006	3:0–8:11	55–145	650	Females (50), Males (50)	White (60), African American (15), Asian (4), Other (3), Hispanic (20)	Northeast (17), Midwest (21), South (37), West (26)	Yes
Emerging Literacy & Language Assessment (ELLA)	Wiig and Secord, 2006	4:6–9:11	< 55–163	1267	Females (53), Males (47)	White (60), African American (20), Other (9), Hispanic (11), Not specified (0.2)	Northeast (28), Midwest (20), South (35), West (16)	Yes
Expressive Language Test - Second Edition: Normative Update (ELT-2:NU)	Bowers et al., 2018b	5:0–11:11	46–149	1007	Females (49), Males (51)	Race White (77), Black/African American (17), Asian/Pacific Islander (6) Ethnicity Hispanic (24), Non-Hispanic (76)	Northeast (18), Midwest (23), South (38), West (22)	No
Expressive One-Word Picture Vocabulary Test, Fourth Edition – English (EOWPVT-4)	Martin and Brownell, 2011a	2:0–70:0 +	< 55– > 145	2394	Females (56), Males (44)	Caucasian (63), African American (13), Asian American (3), Native American (1), Other (0.3), Hispanic (18), Not reported (1)	Northeast (24), Midwest (18), South (32), West (27)	Yes
Expressive One-Word Picture Vocabulary Test, Fourth Edition – Spanish - Bilingual Edition (EOWPVT-4:SBE)	Martin, 2013	2:0–70:0 +	< 55– > 145	1260	Females (55), Males (45)	Ethnicity Hispanic (94) White/Caucasian (4), African American (0.6), Native American (0.5), Asian American (0.1), Other (1), Not Reported (0.2) Hispanic Origin Mexico (62), Puerto Rico (10), South America (6), Central America (6), Cuba (5), Dominican Republic (3), Haiti (0.6)	Northeast (8), Midwest (15), South (36), West (42)	Yes
Expressive Vocabulary Test, Third Edition (EVT-3)	Williams, 2018	2:6–90:0 +	40–160	2720	Not specified	White (62), African American (14), Asian (3), Other (5), Hispanic (17)	Northeast (13), Midwest (21), South (49), West (17)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Hodson Assessment of Phonological Patterns, Third Edition (HAPP-3)	Hodson, 2004	3:0–7:11	< 55–114	886	Females (49), Males (51)	Race White (76), Black (16), Other (8); Ethnicity Hispanic (10), Non-Hispanic (90)	Northeast (19), Midwest (28), South (35), West (18)	Yes
Khan-Lewis Phonological Assessment, Third Edition (KLPA-3) & Goldman-Fristoe Test of Articulation, Third Edition (GFTA-3)	Khan and Lewis, 2015; Goldman and Fristoe, 2015	2:0–21:11	40–140	1500	Females (50), Males (50)	White (57), African American (11), Asian (2), Other (7), Hispanic (22)	Northeast (13), Midwest (24), South (41), West (23)	Yes
Language Processing Test 3: Elementary (LPT 3)	Richard and Hanner, 2005	5:0–11:11	< 40 - 150	1313	Females (50), Males (50)	Caucasian (61), African American (17), Asian American and Others (4), Hispanic (18)	Not specified	Yes
Listening Comprehension Test - Adolescent: Normative Update (LCT-A: NU)	Bowers et al., 2018a	12:0–17:11	48–136	1008	Females (50), Males (50)	Race White (77), Black/African American (16), Asian/Pacific Islander (4), Other (3); Ethnicity Hispanic (21), Non-Hispanic (79)	Northeast (17), Midwest (24), South (36), West (23)	No
Listening Comprehension Test, Second Edition (LCT-2)	Bowers et al., 2006	6:0–11:11	< 62–159	1504	Females (50), Males (50)	Caucasian (63), African American (15), Hispanic (17), Asian American and Others (5)	Not specified	Yes
Montgomery Assessment of Vocabulary Acquisition (MAVA) - Expressive Vocab Test	Montgomery, 2008a	3:0–12:11	< 55– > 145	1248	Females (49), Males (52)	White (63), African American (16), Other (6), Hispanic (15)	Northeast (18), Midwest (25), South (36), West (21)	Yes
Montgomery Assessment of Vocabulary Acquisition (MAVA) - Receptive Vocab Test	Montgomery, 2008b	3:0–12:11	< 55– > 145	1373	Females (48), Males (52)	White (62), African American (17), Other (5), Hispanic (16)	Northeast (16), Midwest (23), South (40), West (20)	Yes
Oral and Written Language Scales, Second Edition (OWLS-II)	Carrow-Woolfolk, 2011	3:0–21:11	40–160	2123	Females (51), Males (49)	White (55), Black/African American (18), Asian (2), Native American (0.5), Native Hawaiian/Pacific Islander (0.4), Two or more races (4), Other (1), Hispanic origin (any race) (19)	Northeast (22), Midwest (22), South (37), West (19)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (n)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Oral Passage Understanding Scale (OPUS)	Carrow-Woolfolk and Klein, 2017	5:0–21:11	40–160	1517	Females (51), Males (49)	White (55), Black/African American (15), Asian (3), American Indian/Alaska Native (0.6) Native Hawaiian/Pacific Islander (0.4) Two or More/Other (4) Hispanic (22)	Northeast (20), Midwest (20), South (35), West (24)	Yes
Peabody Picture Vocabulary Test, Fifth Edition (PPVT-5)	Dunn, 2019	2:6–90:11 +	40–160	2720	Not specified	White (62), African American (14), Asian (3), Other (5), Hispanic (17)	Northeast (13), Midwest (21), South (49), West (17)	Yes
Phonological and Print Awareness Scale (PPA Scale)	Williams, 2014	3:6–8:11	< 50– > 130	1104	Females (51), Males (49)	Race White (76), Black/African American (16), Asian (4), Native American (0.4), Native Hawaiian/Pacific Islander (0.3), Other (3) Ethnicity Hispanic (25), Non-Hispanic (76)	Northeast (28), Midwest (16), South (35), West (21)	Yes
Phonological Awareness Test, Second Edition: Normative Update (PAT-2: NU)	Robertson and Salter, 2018	5:0–9:11	39–123	1193	Females (49), Males (51)	Race White (77), Black/African American (17), Asian/Pacific Islander (3), Other (2) Ethnicity Hispanic (23), Non-Hispanic (77)	Northeast (19), Midwest (24), South (37), West (21)	No
Preschool Language Assessment Instrument - Second Edition (PLAI-2)	Blank et al., 2003	3:0–5:11	49–160	463	Females (49), Males (51)	Race White (79), Black (15), Other (6) Ethnicity African American (15), Hispanic American (13), Asian American (3), Native American (2), Other (67)	Northeast (18), Midwest (24), South (35), West (23)	Yes
Preschool Language Scales, Fifth Edition (PLS-5)	Zimmerman et al., 2011	birth–7:11	50 –150	1400	Females (50), Males (50)	White (54), African American (14), Asian (4), Hispanic (24), Other (4)	Northeast (20), Midwest (20), South (36), West (24)	Yes
Receptive One-Word Picture Vocabulary Test, Fourth Edition (ROWPVT-4)	Martin and Brownell, 2011b	2:0–80 +	< 55– > 145	2394	Females (56), Males (44)	Caucasian (63), African American (13), Asian American (3), Native American (1), Other (0.3), Hispanic (18), Not reported (1)	Northeast (24), Midwest (18), South (32), West (27)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Receptive, Expressive and Social Communication Assessment - Elementary (RESCA-E)	Hamaguchi and Ross-Swain, 2015	5:0–12:11	55– > 145	825	Females (50), Males (50) Not Reported (0.8)	Ethnicity White/Caucasian (77), Black/African American (13), Asian American (4), Native Hawaiian/Pacific Islander (0.4), American Indian/Alaska Native (0.1), Two or More (5), Not reported (0.2) Hispanic Origin Hispanic (27), Non-Hispanic (73), Not reported (0.1)	Northeast (18), Midwest (24), South (32), West (27)	Yes
Social Language Development Test – Adolescent: Normative Update (SLDT-A: NU)	Bowers et al., 2017a	12:0–17:11	45–160	868	Females (50), Males (50)	Race White (77), Black/African American (18), Asian/Pacific Islander (4), Other (< 1) Ethnicity Hispanic (21), Non-Hispanic (79)	Northeast (17), Midwest (22), South (38), West (23)	No
Social Language Development Test-Elementary: Normative Update (SLDT-E: NU)	Bowers et al., 2017b	6:0–11:11	47–160	1002	Females (49), Males (51)	Race White (78), Black/African American (15), Asian/Pacific Islander (5), Other (2) Ethnicity Hispanic (19), Non-Hispanic (81)	Northeast (17), Midwest (23), South (38), West (22)	No
Structured Photographic Expressive Language Test, Third Edition (SPELT-3)	Dawson et al., 2003	4:0–9:11	< 40–142	1580	Females (48), Males (52)	Caucasian (66), African American (16), Hispanic (11), Other (7)	Northeast (19), Midwest (39), South (23), West (19)	Yes
The WORD Test, Third Edition: Elementary (WORD-3)	Bowers et al., 2014	6:0–11:11	–9–143	1302	Females (49), Males (51)	Caucasian (66), African American (13), Hispanic (17), Asian American and others (4)	Not specified	Yes
Test for Auditory Comprehension of Language, Fourth Edition (TACL-4)	Carrow-Woolfolk, 2014	3:0–12:11	45–160	1142	Females (49), Males (51)	Race White (78), Black/African American (14), Asian/Pacific Islander (3), American Indian/Eskimo/Aleut (1), Two or More (4) Ethnicity Hispanic (20), Non-Hispanic (80)	Northeast (17), Midwest (22), South (36), West (25)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Test of Adolescent & Adult Language, Fourth Edition (TOAL-4)	Hammill et al., 2007	12:0–24:11	34–168	1671	Females (51), Males (49)	Ethnicity White (81), Black/African American (11), Asian/Pacific Islander (3), American Indian/Eskimo (2), Two or More (2), Other (1) Hispanic Status Hispanic (12), Non-Hispanic (88)	Northeast (21), Midwest (22), South (35), West (22)	Yes
Test of Early Communication and Emerging Language (TECEL)	Huer and Miller, 2011	2 weeks–2:0	25–160	558	Females (50), Males (50)	Ethnicity White (78), Black/African American (10), Asian/Pacific Islander (5), American Indian/Eskimo (1), Two or More (6), Other (1) Hispanic Status Hispanic (16), Non-Hispanic (84)	Northeast (15), Midwest (19), South (34), West (32)	Yes
Test of Early Language Development, Fourth Edition (TELD-4)	Hresko et al., 2018	3:0–7:11	50–155	1074	Females (50), Males (50)	Race White (75), Black/African American (14), Asian/Pacific Islander (5), American Indian/Eskimo (1), Two or more (5), Ethnicity Hispanic (23), Non-Hispanic (77)	Northeast (17), Midwest (22), South (37), West (24)	Yes
Test of Expressive Language (TEXL)	Carrow-Woolfolk and Allen, 2014	3:0–12:11	47–159	1205	Females (51), Males (49)	Race White (78), Black/African American (15), Asian/Pacific Islander (4), American Indian/Eskimo/Aleut (1), Two or More Races (2) Ethnicity Hispanic (21), Non-Hispanic (79)	Northeast (16), Midwest (22), South (36), West (26)	Yes
Test of Integrated Language and Literacy Skills (TILLS)	Nelson et al., 2016	6:0–18:11	40–145	1262	Females (50), Males (50)	White (Non-Hispanic) (73), African American (10), Asian (5), Native American (1) Other (1), Hispanic (Any Race) (10)	Northeast (5), Midwest (50), South (16), West (29)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Test of Language Development – Intermediate: Fifth Edition (TOLD-I:5)	Hammill and Newcomer, 2019	8:0–17:11	40 –167	1012	Females (51), Males (49)	Race: White (74), Black/African American (14), Asian/Pacific Islander (4), American Indian/Alaska Native (2), Two or More (6); Ethnicity: Hispanic (25), Non-Hispanic (75)	Northeast (17), Midwest (21), South (36), West (26)	Yes
Test of Language Development – Primary, Fifth Edition (TOLD-P:5)	Newcomer and Hammill, 2019	4:0–8:11	41–165	1007	Females (47), Males (53)	Race: White (71), Black/African American (13), Asian/Pacific Islander (6), American Indian/Alaska Native (3), Two or More (7); Ethnicity: Hispanic (25), Non-Hispanic (75)	Northeast (16), Midwest (20), South (36), West (28)	Yes
Test of Narrative Language, Second Edition (TNL-2)	Gilliam and Pearson, 2017	4:0–15:11	50–155	1310	Females (50), Males (50)	Race: White (78), Black/African American (14), Asian/Pacific Islander (5), American Indian/Eskimo/Aleut (< 1) Two or More (2); Ethnicity: Hispanic (22), Non-Hispanic (78)	Northeast (16), Midwest (21), South (38), West (25)	Yes
Test of Phonological Awareness, Second Edition Plus (TOPA-2 +)	Torgeson and Bryant, 2004	5:0–8:11	49 –143	2085	Females (48), Males (52)	Race: White (71), Black (16), Other (14); Ethnicity: White/European American (65), Black/African American (16), Asian/Pacific Islander (5), Native American/Eskimo/Aleut (2), Hispanic (15)	Northeast (20), Midwest (24), South (34), West (23)	Yes
Test of Pragmatic Language, Second Edition (TOPL-2)	Phelps-Terasaki and Phelps-Gunn, 2007	6:0–18:11	55–139	1136	Females (51), Males (49)	Race: White (79), Black/African American (13), Asian/Pacific Islander (4), Native American (1), Two or More (2), Other (1); Ethnicity: Hispanic (13), Non-Hispanic (87)	Northeast (19), Midwest (23), South (35), West (23)	Yes

(Continued)

TABLE 1 | (Continued)

Assessment	References	Ages normed	Range of standard/index scores	Sample size (*n*)	Sex/gender (%)	Race/ethnicity (%)[a]	Region (%)	SES considered
Test of Preschool Vocabulary (TOPV)	Mathews and Miller, 2015	2:0–5:11	54–160	1190	Females (49), Males (51)	Race: White (74), Black/African American (13), Asian/Pacific Islander (5) American Indian/Eskimo/Aleut (1), Two or More (7) Ethnicity: Hispanic (23), Non-Hispanic (77)	Northeast (16), Midwest (22), South (38), West (24)	Yes
Test of Semantic Reasoning (TOSR)	Lawrence and Seifert, 2016	7:0–17:11	<55– > 145	1117	Females (49), Males (51)	Race: White/Caucasian (71), Black/African American (19), Asian American (3), American Indian/Alaska Native (0.7), Native Hawaiian/Pacific Islander (0.4), Two or More (5), Not reported (0.5) Ethnicity: Hispanic (21), Non-Hispanic (79)	Northeast (13), Midwest (20), South (38), West (29)	Yes
Test of Semantic Skills - Intermediate: Normative Update (TOSS-I:NU)	Huisingh et al., 2019	9:0–13:11	52–143	1234	Females (49), Males (51)	Race: White (77), Black/African American (18), Asian American/Native Hawaiian/Other Pacific Islander (4), American Indian/Alaska Native (2), Other (< 1) Ethnicity: Hispanic (20), Non-Hispanic (80)	Northeast (16), Midwest (24), South (36), West (24)	No
Test of Semantic Skills - Primary (TOSS-P)	Huisingh et al., 2002	4:0–8:11	<45–179	1510	Females (51), Males (49)	Caucasian (62), African American (17), Asian American and Others (5), Hispanic-American (16)	Not specified	No
Test of Word Finding, Third Edition (TWF-3)	German, 2014	4:6–12:11	45–132	1283	Females (49), Males (51)	Race: White (77), Black/African American (14), Asian/Pacific Islander (4), American Indian/Eskimo/Aleut (1), Two or More (4) Ethnicity: Hispanic (19), Non-Hispanic (81)	Northeast (18), Midwest (25), South (35), West (22)	Yes
Vocabulary Assessment Scales – Expressive & Receptive (VAS-E & VAS-R)	Gerhardstein Nader, 2013	2:6–95:0	50–150	2678	Females (50), Males (50)	Caucasian (63), African American (13), Other (5), Hispanic (19)	Northeast (17), Midwest (19), South (52), West (13)	Yes

[a] *Race and ethnicity categories for each assessment are reported as they were presented in each manual.*

TABLE 2 | Individuals with disabilities in normative samples.

Assessment	Disability Total Sample Size (*n*)[a]	% Normative Sample[b]	Diagnosis (*n*)[c]	% Normative Sample by Diagnosis	Sample Description
A Language Processing Skills Assessment (TAPS-4)	179	9	Specific Learning Disability/Dyslexia	2	Disability *Note: "Any Disability" represents the total number of individuals in the sample reporting one or more disability status categories.*
			Attention-Deficit Hyperactivity Disorder	2	
			Auditory Processing Disorder	0.4	
			Specific Language Impairment	3	
			Cochlear Implant/Hearing Impairment	2	
			Any Disability	9	
Arizona Articulation and Phonology Scale, Fourth Revision (Arizona - 4)	Not specified	7	Speech/Language Impairment	3	Individuals with severe disabilities (e.g., intellectual disability, moderate to severe autism spectrum disorder) were excluded from the standardization sample, whereas those with mild disabilities were included as long as they spent most of their day in a general education (not gifted or special education) classroom... 7% of the standardization sample had a diagnosed disability (3% speech/language impairment and 4% other diagnoses, including learning disability, developmental disability, intellectual disability, hearing/vision impairment, autism spectrum disorder, emotional disturbance, or other physical/health impairment; these other diagnoses each occurred with a frequency of 1% or less).
			Learning Disability	<1	
			Developmental Disability	<1	
			Intellectual Disability	<1	
			Autism Spectrum Disorder	<1	
			Hearing/Vision Impairment	<1	
			Emotional Disturbance	<1	
			Other Physical/Health Impairment	<1	
Auditory Processing Abilities Test (APAT)	Not specified	9-16[d]	Learning Disability	2	Disability Status *Note: 84% of the sample listed as having "no disability."*
			Speech-Language Disorder	3	
			Other	4	
Bankson Expressive Language Test - Third Edition (BELT-3)	Not specified	<8	Intellectual Disability	<1	Exceptionality Type
			Attention-Deficit/Hyperactivity Disorder	1	
			Developmental Delay	1	
			Speech-Language Impairment	3	
			Learning Disability	<1	
			Autism Spectrum Disorder	<1	
Bankson-Bernthal Test of Phonology, 2nd edition (BBTOP-2)	Not specified	<9	Speech-Language Impaired	3	Exceptionality Status
			Attention-Deficit/Hyperactivity Disorder	2	
			Learning Disabled	1	
			Developmentally Delayed	1	
			Intellectually Disabled	<1	
			Autism Spectrum Disorder	1	
Clainical Assessment of Articulation and Phonology, Second Edition (CAAP-2)	Not specified	7	Not specified	Not specified	Seven percent of the standardization subjects were receiving speech and language services.
Clinical Assessment of Pragmatics (CAPs)	137	15	Autism Spectrum Disorder	2	Clinical Groups
			Specific Language Impairment	3	
			Other	10	

(Continued)

TABLE 2 | (Continued)

Assessment	Disability Total Sample Size (*n*)[a]	% Normative Sample[b]	Diagnosis (*n*)[c]	% Normative Sample by Diagnosis	Sample Description
Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5)	Not specified	<26	Attention Deficit Disorders	5	Of the students in the standardization sample, 11% reported the following diagnoses: 5% attention deficit disorders (inattentive, hyperactive, and combined); 1% learning disability; 1% intellectual disability, pervasive developmental disorder, Down syndrome, or developmental delay; and less than 1% each emotional disturbance, cerebral palsy, color blindness, central auditory processing disorder, visual impairment, autistic spectrum disorder, or other diagnoses not specified. Approximately 3% of the sample was receiving occupational or physical therapy. Approximately 12% of the sample was diagnosed with a speech and/or language disorders; of those, 7.2% reported diagnoses of language disorder (including receptive/expressive language disorder or pragmatics impairment), 4.2% reported articulation or phonological disorder, and less than 1% reported fluency and voice disorder.
			Learning Disability	1	
			Intellectual Disability, Pervasive Developmental Disorder, Down syndrome, and Developmental Delay	1	
			Emotional Disturbance	<1	
			Cerebral Palsy	<1	
			Color blindness	<1	
			Central Auditory Processing Disorder	<1	
			Visual Impairment	<1	
			Autistic Spectrum Disorder	<1	
			Other Diagnosis Not Specified	<1	
			Language Disorder	7	
			Articulation or Phonological Disorder	4	
			Fluency and Voice Disorder	<1	
			Occupational or Physical Therapy	3	
Clinical Evaluation of Language Fundamentals Preschool, Third Edition (CELF-P3)	Not specified	7	Occupational or Physical Therapy	1	According to the inclusion and exclusion specifications for the normative sample, the children included did not meet the diagnosis criteria for a language impairment, a learning disorder in reading or writing, or a hearing impairment... To reflect the variability in learning needs that naturally occur in the general population, a limited number of children with special education placement were included in the normative sample. Approximately 8% of children in the sample were reported as receiving special services: less than 1% for gifted and talented, 1.3% for occupational or physical therapy, 2% early childhood or other services, and an overlapping 4% received services for both speech and language.
			Early Childhood and Other Services	2	
			Services in Both Speech and Language	4	
Communication and Symbolic Behavior Scales, Developmental Profile, First Normed Edition (CSBS)	Not specified	Not specified	Not specified	Not specified	Children with known developmental delays or who qualified for Part C early intervention services were excluded from the standardization sample... Because of the extent of the under-identification of children with developmental delays from birth-24 months of age, it is presumed that at least 10% of the standardization sample has developmental delays or disabilities, although children with severe disabilities are likely not included in this sample.
Comprehensive Assessment of Spoken Language, Second Edition (CASL-2)	Not specified	Not specified	Not specified	Not specified	Individuals with severe disabilities (e.g., intellectual disability, moderate to severe autism spectrum disorder) were excluded from the standardization sample, while those with mild disabilities were included as long as they spent most of their school day in a general education classroom (not gifted or special education), at a grade level appropriate to the child's chronological age.
Comprehensive Receptive and Expressive Vocabulary Test, Third Edition (CREVT-3)	Not specified	25	Learning Disability	5	Exceptionality Status
			Articulation Disorder	5	
			Language Impaired	5	
			Attention-Deficit Disorder	4	
			Other	6	
Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)	Not specified	<7	Specific Learning Disabilities	1	Exceptionality Status
			Intellectual Disability	<1	
			Hearing Impairment	<1	
			Other Health Impairment	<1	
			Attention-Deficit Disorder	2	
			Other Disability	1	

(Continued)

TABLE 2 | (Continued)

Assessment	Disability Total Sample Size (n)[a]	% Normative Sample[b]	Diagnosis (n)[c]	% Normative Sample by Diagnosis	Sample Description
Diagnostic Evaluation of Articulation and Phonology (DEAP)	52	8	Disability (subgroup percentages not reported)	8	The DEAP norms were developed on a sample that included 650 children; 94.3% had no speech or language disorder and 5.7% had diagnosed articulation, phonological, or oral motor disorders... Eight percent of children in the normative sample were reported by parents and examiners to be diagnosed with one or more of the following: receptive and/or expressive language disorder, articulation disorder, phonological disorder, oral motor disorder, Attention-deficit/hyperactivity disorder, cerebral palsy, developmental delay, fluency disorder, learning disability, orthopedic handicap, visual impairment, and other health impairment. Less than 1 percent were identified as gifted.
Emerging Literacy & Language Assessment (ELLA)	Not specified	Not specified	Not specified	Not specified	Some children with language disorders, learning disabilities, and some children receiving remediation services in reading (but not special education services) were included in the standardization sample.
Expressive Language Test - Second Edition: Normative Update (ELT-2:NU)	Not specified	14	Language Impairment	4	Exceptionality Status *Note: 86% of the sample listed as having "no disability."*
			Other Disability	10	
Expressive One-Word Picture Vocabulary Test, Fourth Edition -English (EOWPVT-4)	298	12	Any Disability	12	Information regarding disability status is from the U.S. Department of Education (2000).
Expressive One-Word Picture Vocabulary Test, Fourth Edition - Spanish-Bilingual Edition (EOWPVT-4:SBE)	131	10	Any Disability	10	Not specified
Expressive Vocabulary Test, Third Edition (EVT-3)	Not specified	4	Attention-Deficit Disorders	0.8	According to the inclusion and exclusion specifications for the study, the individuals included in this sample did not meet the diagnosis criteria for language disorder, learning disorder, or hearing impairment. Review of each individual's test results indicated expected performance for individuals without language and/or learning disability... To reflect the variability in learning needs that naturally occur in the general population, a limited number of individuals with special education placement were included in the normative sample. Of the sample, 0.8% were reported with an educational placement for gifted or talented. In addition, 3.7% of the individuals in the normative sample reported an educational diagnosis: approximately 0.8% attention deficit disorders (inattentive, hyperactive, and combined); 0.7% autism spectrum disorder; 0.2% developmental delay; 0.2% hearing impairment; 0.6% learning disability in reading and/or writing; 0.3% speech and/or language delay; and 0.9% speech and/or language disorder.
			Autism Spectrum Disorder	0.7	
			Developmental Delay	0.2	
			Hearing Impairment	0.2	
			Learning Disability in Reading and/or Writing	0.6	
			Speech and/or Language Delay	0.3	
			Speech and/or Language Disorder	0.9	
Hodson Assessment of Phonological Patterns, Third Edition (HAPP-3)	Not specified	3	Phonological Impairment	2	Disability Status *Note: 97% of the sample listed as having "no disability."*
			Other Disability	1	
Khan-Lewis Phonological Assessment, Third Edition (KLPA-3) & Goldman-Fristoe Test of Articulation, Third Edition (GFTA-3)	Not specified	20	Speech and/or Language Disorder	8	20% were reported with the following diagnoses: approximately 8% speech and/or language disorder; 4% attention deficit disorders (inattentive, hyperactive, and combined); 3% learning disability; 2% intellectual disability, pervasive developmental disorder, Down syndrome, or developmental delay; and less than 1% each emotional disturbance, cerebral palsy, central auditory processing disorder, visual impairment, autistic spectrum disorder, or other diagnoses not specified.
			Attention Deficit Disorders	4	
			Learning Disability	3	
			Intellectual Disability, Pervasive Developmental Disorder, Down syndrome, or Developmental Delay	2	
			Emotional Disturbance	<1	
			Cerebral Palsy	<1	
			Central Auditory Processing Disorder	<1	
			Visual Impairment	<1	
			Autistic Spectrum Disorder	<1	
			Other diagnoses not specified	<1	
Language Processing Test 3: Elementary (LPT-3)	Not specified	Not specified	Not specified	Not specified	The sample included normal subjects and subjects with language-learning disorders. Subjects previously identified as having hearing impairment, mental disabilities, emotional disabilities, or limited English proficiency were excluded from the standardization sample.
Listening Comprehension Test - Adolescent: Normative Update (LCT-A: NU)	Not specified	12	Specific Language Impairment	4	Exceptionality status: *Note: Other/Special education consisted of students receiving special education services for a variety of conditions.*
			Other/Special Education	8	

(Continued)

TABLE 2 | (Continued)

Assessment	Disability Total Sample Size (*n*)[a]	% Normative Sample[b]	Diagnosis (*n*)[c]	% Normative Sample by Diagnosis	Sample Description
Listening Comprehension Test, Second Edition (LCT-2)	Not specified	Not specified	Not specified	Not specified	Subjects from regular education; special education ... were included in the study. In addition, subjects with IEPs for special services (e.g., articulation disorder, remedial reading) but who attend regular education classes were included. Subjects excluded from the study included those who were not able to use English proficiently at school, were non-verbal, had any degree of hearing loss, or who resided outside of the United States.
Montgomery Assessment of Vocabulary Acquisition (MAVA) - Expressive Vocab Test	Not specified	10	Not Specified	Not specified	Ten percent of this population included children in special education with known vocabulary deficits.
Montgomery Assessment of Vocabulary Acquisition (MAVA)- Receptive Vocab Test	Not specified	10	Not Specified	Not specified	Ten percent of this population included children in special education with known vocabulary deficits.
Oral and Written Language Scales, Second Edition (OWLS - II)	Not specified	Not specified	Not specified	Not specified	Individuals with diagnosed disabilities were included in the standardization sample as long as they spent most of their school day in a regular classroom. The percentage of such individuals matched what is expected in the population.
Oral Passage Understanding Scale (OPUS)	Not specified	Not specified	Not specified	Not specified	Individuals with severe disabilities (e.g., intellectual disability, moderate to severe autism) were excluded from the standardization sample, while those with mild disabilities were included as long as they spent most of their day in a general classroom.
Peabody Picture Vocabulary Test, Fifth Edition (PPVT-5)	Not specified	4	Attention Deficit Disorders	0.8	According to the inclusion and exclusion specifications for the study, the individuals included in this sample did not meet the diagnosis criteria for language disorder, learning disorder, or hearing impairment. Review of each individual's test results indicated expected performance for individuals without language and/or learning disability... To reflect the variability in learning needs that naturally occur in the general population, a limited number of individuals with special education placement were included in the normative sample. Of the sample, 0.8% were reported with an educational placement for gifted or talented. In addition, 3.7% of the individuals in the normative sample reported an educational diagnosis: approximately 0.8% attention deficit disorders (inattentive, hyperactive, and combined); 0.7% autism spectrum disorder; 0.2% developmental delay; 0.2% hearing impairment; 0.6% learning disability in reading and/or writing; 0.3% speech and/or language delay; and 0.9% speech and/or language disorder.
			Autism Spectrum Disorder	0.7	
			Developmental Delay	0.2	
			Hearing Impairment	0.2	
			Learning Disability in reading and/or writing	0.6	
			Speech and/or Language Delay	0.3	
			Speech and/or Language Disorder	0.9	
Phonological and Print Awareness Scale (PPA Scale)	Not specified	Not specified	Not specified	Not specified	Children with mild disabilities were included in the standardization sample as long as they spent most of their school day in a regular classroom.
Phonological Awareness Test, Second Edition: Normative Update (PAT-2: NU)	Not specified	15	Language Impairment	3	Exceptionality Status *Note: 85% of the sample listed as having "no disability."*
			Special Education	12	
Preschool Language Assessment Instrument - Second Edition (PLAI-2)	Not specified	11	Speech-Language Disorder	4	Disability Status *Note: 89% of the sample listed as having "no disability."*
			Intellectual Disability	0	
			Other Handicap	7	
Preschool Language Scales - Fifth Edition (PLS-5)	90	6	Speech Language Disorder	4	Educational Classification/Diagnosis *Note: Other includes hearing impairments, other health impairments, visual impairments, multiple disabilities, deaf-blindness, and traumatic brain injury.*
			Intellectual Disability	0.1	
			Developmental Delay	1	
			Attention-Deficit/Hyperactivity Disorder	0.2	
			Orthopedic/Motor Impairment	0.1	
			Other	0.9	
Receptive One-Word Picture Vocabulary Test, Fourth Edition (ROWPVT-4)	298	12	Any Disability	12	Information regarding disability status is from the U.S. Department of Education (2000).

(Continued)

TABLE 2 | (Continued)

Assessment	Disability Total Sample Size (*n*)[a]	% Normative Sample[b]	Diagnosis (*n*)[c]	% Normative Sample by Diagnosis	Sample Description
Receptive, Expressive and Social Communication Assessment - Elementary (RESCA-E)	126	15	Autism Spectrum Disorder	4	Disability *Note: "Any Disability" represents the total number of individuals in the sample reporting one or more disability status categories.*
			Attention Deficit/Hyperactivity Disorder	2	
			Developmental Disability	3	
			Emotional Disturbance	1	
			Intellectual Disability	0.7	
			Social Communication Disorder	1	
			Speech and Language Impairment	5	
			Learning Disability	4	
			Any Disability	15	
Social Language Development Test – Adolescent: Normative Update (SLDT-A: NU)	Not specified	20	Specific Language Impairment	6	Exceptionality Status *Note: Other/Special Education subgroup consisted of students receiving special education services for a variety of conditions.*
			Autism Spectrum Disorder	5	
			Other/Special Education	9	
Social Language Development Test-Elementary: Normative Update (SLDT-E: NU)	Not specified	17	Specific Language Impairment	5	Exceptionality Status *Note: Other/Special Education subgroup consisted of children receiving special education services for a variety of conditions.*
			Autism Spectrum Disorder	2	
			Other/Special Education	10	
Structured Photographic Expressive Language Test - Third Edition (SPELT-3)	Not specified	> 7	Not specified	Not Specified	Slightly more than 7% of the sample was identified as language impaired consistent with prevalence estimates of 7% in the population (Leonard, 1998) and 7.4% (Tomblin et al., 1997).
The WORD Test, Third Edition: Elementary (WORD-3)	Not specified	Not specified	Not specified	Not specified	In addition, subjects with Individualized Education Plans (IEPs) for special services (e.g., articulation disorder, remedial reading) but who attended regular education classes were included. Subjects excluded from this study included those who were not able to use English proficiently at school, were non-verbal, had any degree of hearing loss, or who resided outside of the United States.
Test for Auditory Comprehension of Language, Fourth Edition (TACL-4)	Not specified	18	Intellectual Disability	4	Exceptionality Type
			Deaf/Hard of Hearing	1	
			Language Impairment	4	
			Learning Disability	4	
			Attention-Deficit/Hyperactivity Disorder	3	
			Autism Spectrum Disorder	2	
Test of Adolescent & Adult Language, Fourth Edition (TOAL-4)	Not specified	15	Disabled	15	Exceptionality Status *Note: 85% of the sample listed as "not disabled." The data on exceptionality status represent students being served under the Individuals with Disabilities Education Act and does not include those who have a language disorder, who have attention-deficit/hyperactivity disorder, or who are gifted and talented.*
Test of Early Communication and Emerging Language (TECEL)	47	8	Not specified	Not specified	The TECEL was normed on a sample of 558 persons (47 with disabilities).
Test of Early Language Development, Fourth Edition (TELD-4)	Not specified	13	Intellectual Disability	1	Exceptionality Status
			Developmental Disability	2	
			Speech/Language Impairment	6	
			Learning Disability	2	
			Attention-Deficit/Hyperactivity Disorder	1	
			Autism Spectrum Disorder	1	

(Continued)

TABLE 2 | (Continued)

Assessment	Disability Total Sample Size (*n*)[a]	% Normative Sample[b]	Diagnosis (*n*)[c]	% Normative Sample by Diagnosis	Sample Description
Test of Expressive Language (TEXL)	Not specified	16	Intellectual Disability	2	Exceptionality Type
			Language Impairment	3	
			Articulation Disorder	3	
			Learning Disability	4	
			Attention-Deficit Hyperactivity Disorder	2	
			Autism Spectrum Disorder	1	
			Deaf/Hard of Hearing	1	
Test of Language Development – Intermediate: Fifth Edition (TOLD-I:5)	Not specified	<21	Intellectual Disability	1	Exceptionality Status
			Deaf/Hard of Hearing	<1	
			Attention-Deficit/Hyperactivity Disorder	4	
			Articulation Disorder	2	
			Asperger Syndrome/High-Functioning Autism	<1	
			Developmental Delay	<1	
			Emotional/Behavior Disorder	2	
			Specific Learning Disability	5	
			Language Impairment	2	
			Low-Functioning Autism	<1	
			Other Disability	<1	
Test of Language Development – Primary, Fifth Edition (TOLD-P:5)	Not specified	<20	Intellectual Disability	<1	Exceptionality Status
			Attention-Deficit/Hyperactivity Disorder	<1	
			Articulation Disorder	5	
			Asperger Syndrome/High-Functioning Autism	1	
			Developmental Delay	2	
			Behavior Disorder	1	
			Learning Disability	4	
			Language Impairment	3	
			Low-Functioning Autism	1	
			Other Disability	1	
Test of Narrative Language, Second Edition (TNL-2)	Not specified	<11	Specific Learning Disabilities	2	Exceptionality Status
			Intellectual Disability	3	
			Deaf/Hard of Hearing	<1	
			Other Health Impairments	<1	

(Continued)

TABLE 2 | (Continued)

Assessment	Disability Total Sample Size (n)[a]	% Normative Sample[b]	Diagnosis (n)[c]	% Normative Sample by Diagnosis	Sample Description
			Attention-Deficit/Hyperactivity Disorder	2	
			Physically Impaired	1	
			Other Disability	1	
Test of Pragmatic Language, Second Edition (TOPL-2)	Not specified	23	Behavioral Disorder	<1	Disability/Exceptionality Status *Note: 77% of the sample listed as having "no exceptionality/disability."*
			Developmental Delay	1	
			Asperger's Syndrome	1	
			Articulation Disorder	2	
			Learning Disability	5	
			Attention-Deficit/Hyperactivity Disorder	2	
			Intellectual Disability	<1	
			Autism	<1	
			Emotional Disturbance	3	
			Physical Impairment	<1	
			Speech-Language Impairments	2	
			Deaf/Hard of Hearing	<1	
			Blind/Visual Impairments	<1	
			Traumatic Brain Injury	<1	
Test of Preschool Vocabulary (TOPV)	Not specified	15	Intellectual Disability	1	Exceptionality Status *Note: 85% of the sample listed as having "no exceptionality."*
			Deaf/Hard of Hearing	1	
			Developmental Delay	6	
			Emotional Disturbance	<1	
			Behavioral Disorder	<1	
			Language Impairment	7	
			Autism Spectrum Disorder	2	
Test of Semantic Reasoning (TOSR)	114	10	Specific Language Impairment	3	Disability *Note: "Any Disability" is defined as total number of individuals in the sample reporting one or more disability status categories.*
			Learning Disability	3	
			Autism	2	
			Attention-Deficit/Hyperactivity Disorder	4	
			Any Disability	10	
Test of Semantic Skills - Intermediate: Normative Update (TOSS-I:NU)	Not specified	9	Language Impairment	3	Exceptionality Status *Note: Other/Special Education subgroup consisted of children receiving special education services for a variety of conditions.*
			Other/Special Education	6	
Test of Word Finding, Third Edition (TWF-3)	Not specified	14	Attention-Deficit/Hyperactivity Disorder	3	Exceptionality Status *Note: 86% of the sample listed as having "no disability."*
			Specific Learning Disability	2	

(Continued)

TABLE 2 | (Continued)

Assessment	Disability Total Sample Size (*n*)[a]	% Normative Sample[b]	Diagnosis (*n*)[c]	% Normative Sample by Diagnosis	Sample Description
			Speech or Language Impairment	3	
			Word Finding Problem	3	
			Intellectual Disability	<1	
			Other Disability	5	

Percentages were rounded to the nearest whole number unless they were less than 1.
[a]Sample size (n) is only reported here if the n was provided in the assessment manual.
[b]Some assessments did not report the overall sample size (n) or overall percentage of individuals with disabilities in their manual but instead reported the sample size (n) or percentages of individual disability groups (e.g., n's with learning disability, autism). When this happened, the overall percentage reported in the table was estimated based on the available information.
[c]Individual disability groups are reported as presented and labeled in each assessment manual. One exception is that the term "mental retardation" was updated to "intellectual disability".
[d]Under the category of "Disability Status", the APAT manual reported that 84% of the normative sample had "no disability", 1.7% had a "learning disability", 3.2% had a "speech-language disability", and 4.3% had an "other" disability. Even if there were no dual-diagnoses, these percentages do not account for 100% of the normative sample.

reported 2–3 disability groups (e.g., language impairment and "other/special education"). Further, no assessment tools reported race, ethnicity, or socioeconomic status information specifically for subgroups of individuals with disabilities in their normative samples.

DISCUSSION

The purpose of this study was to identify the number and characteristics of individuals with NDDs in commonly used and commercially available standardized, norm-referenced language assessment tools. Our findings indicate that many of these assessment tools, though not all, did include some individuals with "disabilities" in their normative sample. However, the number of individuals with specific types of disabilities or NDDs was often very low, and minimal demographic information was provided about groups with disabilities.

We identified 43 assessment tools that included individuals with disabilities in their normative samples. These "disability" groups typically included individuals with *any* disabilities, not just NDDs, and the groups were often not broken down by disability type. Therefore, the number of individuals with NDDs, specifically, in the normative samples was often unclear. There was high variability in the percentages of individuals with "disabilities" in the normative samples, ranging from 3% to <26%. These rates align with some available prevalence data on individuals with disabilities in the United States. For example, 2019 United States census data reveal that 4.3% of children under 18 have a disability (Young, 2021), and 2020–2021 United States special education data indicate that 15% of 3-to-21-year-olds receive services under the Individuals with Disabilities Education Act (National Center for Education Statistics, 2022). However, the percentage of children with NDDs reported from public or private insurance data is even higher [i.e., 23.9 and 11%, respectively (Straub et al., 2022)]. It would be helpful if our results could be easily interpreted within the available prevalence data, but similar to the reporting of disabilities within the assessment tools we reviewed, these data are also difficult to interpret and vary based on how disabilities are defined. This presents a barrier to the selection of standardized assessment tools for these populations.

Another barrier is the lack of demographic information (i.e., race, ethnicity, and socioeconomic status) provided about the individuals with disabilities or NDDs in the normative samples of the assessment tools. Without this information, the diversity of the individuals in these groups is unknown. It is possible, for example, that there were no Black individuals with autism included in some normative samples or no Hispanic individuals with intellectual disability. Thus, it is unknown if the individuals with disabilities who were included are representative of these groups as a whole, including across race, ethnicity, and socioeconomic status. Together with the lack of definition of "disability" provided by many of the assessment tools we reviewed, it is unclear if their normative samples are representative of the population of individuals with disabilities or NDDs in the United States.

Considerations for the Selection of Norm-Referenced Language Assessment Tools

There are several scenarios in which clinicians, educators, and researchers must use standardized, norm-referenced language assessment tools with individuals who have, or who are suspected of having, NDDs. The information extracted from this study can be used by these professionals to guide the selection of such assessment tools and while interpreting their scores.

The decision about which language assessment tools are most suitable depends on the specific population of interest and the intended purpose of the assessment. Professionals using norm-referenced assessments tools to identify if an individual has a *primary* diagnosis of a communication disorder may want to choose an assessment tool that does not include individuals with "disabilities" in the normative sample because they are trying to determine if an NDD (e.g., a communication disorder) is present or absent. To make this determination, an individual's score should be compared to a normative sample of peers who *do not* have an NDD. Relatedly, professionals using norm-referenced assessment tools to determine if clients qualify for services (e.g., special education services) may also wish to use assessment tools that do not include individuals with disabilities in their samples, because, as Peña et al. (2006) demonstrated, the presence of individuals with disabilities in the normative sample can lower the level of performance that falls within the average range. Consequently, it becomes less likely that an individual who has a disability will score below the average range and thus be eligible for services. This is particularly important when evaluating an individual with relatively mild delays. In our review, this included the BESA (Peña et al., 2018), the TILLS (Nelson et al., 2016), the TOPA-2+ (Torgeson and Bryant, 2004), the TOSS-P (Huisingh et al., 2002), and the VAS-E and VAS-R (Gerhardstein Nader, 2013).

In contrast, professionals working with individuals with NDDs may select a standardized assessment tool for the purpose of identifying areas of strength and need to support intervention and educational planning. This may be common when an individual has a *primary* diagnosis of an NDD other than a communication disorder, in which language is also affected (e.g., intellectual disability). In such cases, it is ideal to select an assessment tool that has been developed and normed with others who have a similar NDD and who are demographically similar to their client. This is especially important when working with clients who have more severe disabilities and who are at risk of performing at the floor level of an assessment tool. Thus, yet another important consideration is the range and floor level of the standard/index scores. Those with lower floors may allow for more separation between scores at the lower-performing end. This, in turn, allows for greater differentiation across specific skills or language domains. These assessment tools are also better options for professionals who are using norm-referenced assessment tools to monitor progress over time (e.g., clinical gains or intervention success). In our review, we identified that the floor score of some language assessment tools is in the 60s, while others include scores between 50 and 40, and only four had standard scores of 40 or lower. Those with floors below 40 were the PAT-2:NU (Robertson and Salter, 2018), the WORD-3 (Bowers et al., 2014), the TOAL-4 (Hammill et al., 2007), and the TECEL (Huer and Miller, 2011).

Similarly, researchers who are documenting patterns of strength and difficulty to inform the field about different NDD phenotypes should also consider selecting norm-referenced assessment tools that have included individuals with NDDs and that have a wide range of standard/index scores with lower floors. This allows for more nuance in understanding the variability among participant samples, especially at the lower-performing end. The ability to include participant samples with more diverse language profiles can lead to more precise phenotyping that can ultimately be applied to develop evidence-based language interventions. This could also improve the likelihood of intervention success because intervention studies and clinical trials often fail to demonstrate response to treatment due in part to poor outcome measures (see Esbensen et al., 2017; Abbeduto et al., 2020). If language assessment tools can better differentiate among different language profiles, it may be possible for researchers to specify who does and does not respond to certain interventions. When researchers select measures that *do not* include individuals with NDDs in the normative sample, the interpretation of skills and abilities is reduced to comparisons with neurotypical peers. Instead, if individuals with NDDs are compared to individuals with other NDDs (e.g., Down syndrome vs. intellectual disability), areas of unique strength and need can be identified and used in treatment planning.

Future Directions and Recommendations for Holistic Language Assessment for Individuals With Neurodevelopmental Disorders

Several researchers have noted the limited utility of standardized, norm-referenced assessment tools for individuals with certain NDDs (e.g., intellectual disability and neurogenetic syndromes) and have started developing more sensitive measures for these populations (e.g., Berry-Kravis et al., 2013; Budimirovic et al., 2017; Esbensen et al., 2017; Abbeduto et al., 2020; Baumer et al., 2022). For example, Brady et al. (2012, 2018, 2020) developed the Communication Complexity Scale to assess communication skills in individuals who have intellectual disabilities and are minimally speaking, and Abbeduto et al. (2020) and Thurman et al. (2021) developed an expressive language sampling procedure for use with individuals with intellectual disability and neurogenetic syndromes. These measures capture more variability in language and communication skills in individuals with intellectual and developmental disabilities, with demonstrated evidence of their usefulness as outcome measures. Thus, professionals working with individuals with NDDs should consider these measures when tracking progress over time. These assessment tools can also be used by professionals working with neurotypical individuals; for example, Channell et al. (2018) documented that expressive language sampling in the context of narration showed age-related increases in syntactic complexity and lexical diversity from 4 years up until 18.5 years. As these language assessment tools continue to be tested and examined, professionals may have more options in which to assess clients with NDDs.

In addition to these language/communication sampling assessment tools, there will continue to be a need for norm-referenced language assessment tools for use with individuals with NDDs. Thus, in the future, test developers should not only consider including a more representative number of individuals with NDDs in their normative samples but also as part of the iterative test development and standardization processes. Test developers should also consider the possibility of including separate norms for individuals with NDDs and/or who perform at the lower ends of their assessment tool (e.g., Hendrix et al., 2020). Importantly, test developers should better define the characteristics of individuals with disabilities who are included and seek to include diverse samples of individuals with disabilities.

Information about the normative sample composition is a critical part of assessment tool selection; therefore, the inclusion of this information would aid professionals when determining the best assessment tool for an individual client or student.

Until then, standardized, norm-referenced assessment tools that do not include individuals with disabilities broadly and/or NDDs specifically can still be used when working with this population. In particular, professionals can examine item-level performance and/or use growth or deviation scores to track change over time (e.g., Sansone et al., 2014). Professionals should also continue to use holistic approaches to assessment when working with individuals with NDDs by supplementing norm-referenced assessment tools with additional non-standardized assessment tools and dynamic assessment methods (Haywood and Tzuriel, 2002; Grigorenko, 2009).

Study Limitations

There are several limitations to note in the current study. First, this review focused on normative samples, specifically. Many of the reported assessment tools have conducted follow-up validity or clinical research studies to test their measure on small groups of individuals with disabilities or NDDs. Although these participants are not included in the normative sample, the information can still be helpful for understanding if an assessment tool is appropriate for use with individuals with NDDs (e.g., if it will capture variability at lower ends, if items are appropriate, and/or if it yields valid and reliable scores in these populations). Future studies could review these validation studies to provide a comprehensive summary of the additional testing that has been conducted. Another limitation of the current study was that it excluded norm-referenced academic assessment tools that include a language subtest, as well as screeners and caregiver-, teacher-, and self-report measures. Therefore, we are unable to comment on their normative samples. Similarly, our review was limited to language, and we therefore cannot comment on norm-referenced measures of speech, other communication skills, or cognition more broadly. Lastly, although all discrepancies were resolved, we did not track the percentage of agreement across reviewers for the identification and coding of assessment tools and therefore cannot report inter-rater reliability.

CONCLUSION

Researchers, clinicians, and educators who work with individuals with NDDs must often use standardized, norm-referenced language assessment tools. Unfortunately, many norm-referenced assessment tools have floor effects when used with individuals with intellectual disability or neurogenetic syndromes. We proposed that these floor effects may be due, in part, to the limited inclusion of individuals with NDDs in normative samples. However, even if some professionals wanted to use norm-referenced assessment tools that included individuals with NDDs in their normative samples, or at least that demonstrate variability at lower-performing ends of the assessment tool, this information can be difficult to access. Therefore, we reviewed and reported the representation of individuals with disabilities and NDDs in the normative samples of standardized, norm-referenced language assessment tools, as well as the range of standard/index scores provided. This information can be used to guide professionals' selections of assessment tools, based on the individual or sample of individuals they are working with and the purpose of the assessment.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SL, MC, and LM conceptualized the study and drafted and edited the manuscript. SL and AB reviewed and coded assessments and drafted and edited tables for the Results section. All authors contributed to the article and approved the submitted version.

FUNDING

This project was funded by startup funds from the Department of Special Education and Communication Disorders and the College of Education and Human Sciences at the University of Nebraska–Lincoln.

ACKNOWLEDGMENTS

We would like to thank Dr. Kristy Weissling and Ms. Jennifer Dahman for their reviews and recommendations for which assessments to include. We would also like to thank our research assistants: Hailey Droge, Claire Kubicek, Katelyn Pick, Anna Suppes, and Abbie Zoucha.

REFERENCES

Abbeduto, L., Berry-Kravis, E., Sterling, A., Sherman, S., Edgin, J. O., McDuffie, A., et al. (2020). Expressive language sampling as a source of outcome measures for treatment studies in fragile X syndrome: feasibility, practice effects, test-retest reliability, and construct validity. *J. Neurodev. Disord.* 12:10. doi: 10.1186/s11689-020-09313-6

Abbeduto, L., McDuffie, A., Thurman, A., and Kover, S. T. (2016). Language development in individuals with intellectual and developmental disabilities: from phenotypes to treatments. *Int. Rev. Res. Dev. Disabil.* 50, 71–118. doi: 10.1016/bs.irrdd.2016.05.006

Abbeduto, L., Thurman, A. J., Bullard, L., Nelson, S., and McDuffie, A. (2019). "Genetic syndromes associated with intellectual disabilities," in *Handbook of Medical Neuropsychology*, eds C. Armstrong and L. Morrow (Cham: Springer), 263–299. doi: 10.1007/978-3-030-14895-9_13

American Psychiatric Association [APA] (2013). *Neurodevelopmental Disorders Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, Vol. 5. Washington, DC: American Psychiatric Association. doi: 10.1176/appi.books.9780890425596

Bankson, N. W., and Bernthal, J. E. (2020). *Bankson-Bernthal Test of Phonology*, 2nd Edn. Austin, TX: Pro-Ed, Inc. [Examiner's Manual].

Bankson, N. W., Mentis, M., and Jagielko, J. R. (2018). *Bankson Expressive Language Test*, 3rd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Baumer, N. T., Becker, M. L., Capone, G. T., Egan, K., Fortea, J., Handen, B. L., et al. (2022). Conducting clinical trials in persons with Down syndrome: summary from the NIH INCLUDE Down syndrome clinical trials readiness working group. *J. Neurodev. Disord.* 14:22. doi: 10.1186/s11689-022-09435-z

Berry-Kravis, E., Hessl, D., Abbeduto, L., Reiss, A. L., Beckel-Mitchener, A., Urv, T. K., et al. (2013). Outcome measures for clinical trials in fragile X syndrome. *J. Dev. Behav. Pediatr.* 34, 508–522. doi: 10.1097/DBP.0b013e31829d1f20

Blank, M., Rose, S. A., and Berlin, L. J. (2003). *Preschool Language Assessment Instrument*, 2nd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Bowers, L. B., Huisingh, R., and LoGiudice, C. (2006). *Listening Comprehension Test*, 2nd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Bowers, L. B., Huisingh, R., and LoGiudice, C. (2017a). *Social Language Development Test – Adolescent: Normative Update*. Austin, TX: Pro-Ed. [Examiner's Manual].

Bowers, L. B., Huisingh, R., and LoGiudice, C. (2017b). *Social Language Development Test – Elementary: Normative Update*. Austin, TX: Pro-Ed. [Examiner's Manual].

Bowers, L. B., Huisingh, R., LoGiudice, C., and Orman, J. (2018b). *Expressive Language Test*, 2nd Edn Normative Update. Austin, TX: Pro-Ed. [Examiner's Manual].

Bowers, L. B., Huisingh, R., and LoGiudice, C. (2018a). *Listening Comprehension Test – Adolescent: Normative Update*. Austin, TX: Pro-Ed. [Examiner's Manual].

Bowers, L. B., Huisingh, R., LoGiudice, C., and Orman, J. (2014). *The WORD Test: Elementary*, 3rd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Brady, N. C., Fleming, K., Romine, R. S., Holbrook, A., Muller, K., and Kasari, C. (2018). Concurrent validity and reliability for the communication complexity scale. *Am. J. Speech Lang. Pathol.* 27, 237–246. doi: 10.1044/2017_AJSLP-17-0106

Brady, N. C., Fleming, K., Thiemann-Bourque, K., Olswang, L., Dowden, P., Saunders, M. D., et al. (2012). Development of the communication complexity scale. *Am. J. Speech Lang. Pathol.* 21, 16–28. doi: 10.1044/1058-0360(2011/10-0099)

Brady, N. C., Romine, R. E. S., Holbrook, A., Fleming, K. K., and Kasari, C. (2020). Measuring change in the communication skills of children with autism spectrum disorder using the communication complexity scale. *Am. J. Intellect. Dev. Disabil.* 125, 481–492. doi: 10.1352/1944-7558-125.6.481

Budimirovic, D. B., Berry-Kravis, E., Erickson, C. A., Hall, S. S., Hessl, D., Reiss, A. L., et al. (2017). Updated report on tools to measure outcomes of clinical trials in fragile X syndrome. *J. Neurodev. Disord.* 9:14. doi: 10.1186/s11689-017-9193-x

Carrow-Woolfolk, E. (2011). *Oral and Written Language Scales*, 2nd Edn. Torrance, CA: Western Psychological Services. [Manual].

Carrow-Woolfolk, E. (2014). *Test for Auditory Comprehension of Language*, 4th Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Carrow-Woolfolk, E. (2017). *Comprehensive Assessment of Spoken Language*, 2nd Edn. Torrance, CA: Western Psychological Services. [Manual].

Carrow-Woolfolk, E., and Allen, E. A. (2014). *Test of Expressive Language*. Pro-Ed: Austin, TX. [Examiner's Manual].

Carrow-Woolfolk, E., and Klein, A. M. (2017). *Oral Passage Understanding Scale*. Bloomington, MN: NCS Pearson. [Examiner's Manual].

Channell, M. M., Loveall, S. J., Conners, F. A., Harvey, D. J., and Abbeduto, L. (2018). Narrative language sampling in typical development: implications for clinical trials. *Am. J. Speech Lang. Pathol.* 27, 123–135. doi: 10.1044/2017_AJSLP-17-0046

Dawson, J. I., Stout, C. E., and Eyer, J. A. (2003). *Structured Photographic Expressive Language Test*, 3rd Edn. Dekalb, IL: Janelle Publications. [Manual].

DiStefano, C., Sadhwani, A., and Wheeler, A. C. (2020). Comprehensive assessment of individuals with significant levels of intellectual disability: challenges, strategies, and future directions. *Am. J. Intellect. Dev. Disabil.* 125, 434–448. doi: 10.1352/1944-7558-125.6.434

Dodd, B., Hua, Z., Crosbie, S., Holm, A., and Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology*. San Antonio, TX: Pearson Education.

Dunn, D. M. (2019). *Peabody Picture Vocabulary Test*, 5th Edn. Bloomington, MN: NCS Pearson. [Manual].

Esbensen, A. J., Hooper, S. R., Fidler, D., Hartley, S. L., Edgin, J., d'Ardhuy, X. L., et al. (2017). Outcome measures for clinical trials in Down syndrome. *Am. J. Intellect. Dev. Disabil.* 122, 247–281. doi: 10.1352/1944-7558-122.3.247

Fudala, J. B., and Stegall, S. (2017). *Arizona Articulation and Phonology Scale*, 4th Edn. Torrance, CA: Western Psychological Services. [Examiner's Manual].

Gerhardstein Nader, R. (2013). *Vocabulary Assessment Scales – Expressive & Receptive*. Lutz, FL: PARInc.

German, D. J. (2014). *Test of Word Finding*, 3rd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Geurts, H. M., and Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *J. Autism Dev. Disord.* 38, 1931–1943. doi: 10.1007/s10803-008-0587-1

Gilliam, R. B., and Pearson, N. A. (2017). *Test of Narrative Language*, 2nd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Goldman, R., and Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation*, 3rd Edn. Bloomington, MN: NCS Pearson. [Manual].

Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: two sides of one coin. *J. Learn. Disabil.* 42, 111–132. doi: 10.1177/0022219408326207

Hamaguchi, P., and Ross-Swain, D. (2015). *Receptive, Expressive, and Social Communication Assessment - Elementary*. Novato, CA: Academic Therapy Publications. [Technical and Administration Manuals].

Hammill, D. D., Brown, V., Larsen, S., and Wiederholt, L. (2007). *Test of Adolescent & Adult Language*, 4th Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Hammill, D. D., and Newcomer, P. L. (2019). *Test of Language Development – Intermediate*, 5th Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Hawkins, E., Gathercole, S., Astle, D., Calm Team, and Holmes, J. (2016). Language problems and ADHD symptoms: how specific are the links? *Brain Sci.* 6:50. doi: 10.3390/brainsci6040050

Haywood, H. C., and Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody J. Educ.* 77, 40–63. doi: 10.1207/S15327930PJE7702_5

Helland, W. A., Posserud, M. B., Helland, T., Heimann, M., and Lundervold, A. J. (2016). Language impairments in children with ADHD and in children with reading disorder. *J. Atten. Disord.* 20, 581–589. doi: 10.1177/1087054712461530

Hendrix, J. A., Amon, A., Abbeduto, L., Agiovlasitis, S., Alsaied, T., Anderson, H. A., et al. (2020). Opportunities, barriers, and recommendations in Down syndrome research. *Transl. Sci. Rare Dis.* 5, 99–129. doi: 10.3233/TRD-200090

Hessl, D., Nguyen, D. V., Green, C., Chavez, A., Tassone, F., Hagerman, R. J., et al. (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: the case of fragile X syndrome. *J. Neurodev. Disord.* 1, 33–45. doi: 10.1007/s11689-008-9001-8

Hessling, R. M., Schmidt, T. J., and Traxel, N. M. (2004). "Floor effect," in *Encyclopedia of Social Science Research Methods*, eds M. S. Lewis-Beck, A. Bryman, and T. F. T. Liao (Thousand Oaks, CA: SAGE Publications, Inc), 390–391.

Hodson, B. W. (2004). *Hodson Assessment of Phonological Patterns*, 3rd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Hresko, W., Reid, K., and Hammill, D. (2018). *Test of Early Language Development*, 4th Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Huer, M. B., and Miller, L. (2011). *Test of Early Communication and Emerging Language*. Austin, TX: Pro-Ed. [Examiner's Manual].

Huisingh, R., Bowers, L. B., LoGiudice, C., and Orman, J. (2002). *Test of Semantic Skills – Primary*. Austin, TX: Pro-Ed. [Examiner's Manual].

Huisingh, R., Bowers, L. B., LoGiudice, C., and Orman, J. (2019). *Test of Semantics Skills – Intermediate: Normative Update*. Austin, TX: Pro-Ed. [Examiner's Manual].

Kasari, C., Brady, N., Lord, C., and Tager-Flusberg, H. (2013). Assessing the minimally verbal school-aged child with autism spectrum disorder. *Autism Res.* 6, 479–493. doi: 10.1002/aur.1334

Khan, L. M., and Lewis, N. (2015). *Khan-Lewis Phonological Assessment*, 3rd Edn. Bloomington, MN: NCS Pearson. [Manual].

Lavi, A. (2019). *Clinical Assessment of Pragmatics (CAPs)*. Torrance, CA: Western Psychological Services. [Manual].

Lawrence, B., and Seifert, D. (2016). *Test of Semantic Reasoning*. Novato, CA: Academic Therapy Publications. [Manual].

Leonard, L. B. (1998). *Children with Specific Language Impairment*. Cambridge, MA: The MIT Press.

Luyster, R. J., Seery, A., Talbott, M. R., and Tager-Flusberg, H. (2011). Identifying early-risk markers and developmental trajectories for language impairment in neurodevelopmental disorders. *Dev. Disabil. Res. Rev.* 17, 151–159. doi: 10.1002/ddrr.1109

Martin, N., Brownell, R., and Hamaguchi, P. (2018). *A Language Processing Skills Assessment*, 4th Edn. Novato, CA: Academic Therapy Publications. [Manual].

Martin, N. A. (2013). *Expressive One-Word Picture Vocabulary Test*, 4th EdnSpanish-Bilingual. Novato, CA: Academic Therapy Publications. [Manual].

Martin, N. A., and Brownell, R. (2011a). *Expressive One-Word Picture Vocabulary Test*, 4th Edn. Novato, CA: Academic Therapy Publications. [Examiner's Manual].

Martin, N. A., and Brownell, R. (2011b). *Receptive One-Word Picture Vocabulary Test*, 4th Edn. Novato, CA: Academic Therapy Publications. [Examiner's Manual].

Mathews, S., and Miller, L. (2015). *Test of Preschool Vocabulary*. Austin, TX: Pro-Ed. [Examiner's Manual].

McBee, M. (2010). Modeling outcomes with floor or ceiling effects: an introduction to the Tobit model. *Gifted Child Q.* 54, 314–320. doi: 10.1177/0016986210379095

McDuffie, A., Thurman, A. J., Channell, M. M., and Abbeduto, L. (2017). "Language disorders in children with intellectual disability of genetic origin," in *Handbook of Child Language Disorders*, 2nd Edn, ed. R. Schwartz (Milton Park: Taylor & Francis), 52–81. doi: 10.4324/9781315283531-2

Montgomery, J. (2008a). *Montgomery Assessment of Vocabulary Acquisition* (Expressive Vocab Test). Austin, TX: Pro-Ed. [Examiner's Manual].

Montgomery, J. (2008b). *Montgomery Assessment of Vocabulary Acquisition* (Receptive Vocab Test). Austin, TX: Pro-Ed. [Examiner's Manual].

National Center for Education Statistics (2022). *Students with Disabilities. Condition of Education*. Washington, DC: U.S. Department of Education.

Nelson, N., Plante, E., Helm-Estabrooks, N., and Hotz, G. (2016). *Test of Integrated Language and Literacy Skills*. Baltimore, MD: Brookes Publishing Company. [Examiner's Manual].

Newcomer, P. L., and Hammill, D. D. (2019). *Test of Language Development – Primary*, 5th Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., and Bedore, L. M. (2018). *Bilingual English-Spanish Assessment*. Baltimore, MD: Brookes Publishing Company. [Test Manual].

Peña, E. D., Spaulding, T. J., and Plante, E. (2006). The composition of normative groups and diagnostic decision making: shooting ourselves in the foot. *Am. J. Speech Lang. Pathol.* 15, 247–254. doi: 10.1044/1058-0360(2006/023)

Phelps-Terasaki, D., and Phelps-Gunn, T. (2007). *Test of Pragmatic Language*, 2nd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Reed, R. (2018). "Toddlers and preschoolers with specific language impairment," in *An Introduction to Children with Language Disorders*, 5th Edn, ed. V. Reed (New York, NY: Pearson Education), 77–129.

Richard, G. J., and Hanner, M. A. (2005). *Language Processing Test 3: Elementary*. Austin, TX: Pro-Ed. [Examiner's Manual].

Robertson, C., and Salter, W. (2018). *Phonological Awareness Test*, 2nd Edn Normative Update. Austin, TX: Pro-Ed. [Examiner's Manual].

Ross-Swain, D., and Long, N. (2004). *Auditory Processing Abilities Test*. Novato, CA: Academic Therapy Publications. [Manual].

Sansone, S. M., Schneider, A., Bickel, E., Berry-Kravis, E., Prescott, C., and Hessl, D. (2014). Improving IQ measurement in intellectual disabilities using true deviation from population norms. *J. Neurodev. Disord.* 6:16. doi: 10.1186/1866-1955-6-16

Schalock, R. L., Luckasson, R., and Tassé, M. J. (2021).? *Intellectual Disability: Definition, Diagnosis, Classification, and Systems of Supports*. Silver Spring: American Association on Intellectual and Developmental Disabilities. doi: 10.1352/1944-7558-126.6.439

Secord, W. A., and Donohue, J. S. (2014). *Clinical Assessment of Articulation and Phonology*, 2nd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Spaulding, T. J., Plante, E., and Farinella, K. A. (2006). Eligibility criteria for language impairment: is the low end of normal always appropriate? *Lang. Speech Hear. Serv. Schl.* 37, 61–72. doi: 10.1044/0161-1461(2006/007)

Straub, L., Bateman, B. T., Hernandez-Diaz, S., York, C., Lester, B., Wisner, K. L., et al. (2022). Neurodevelopmental disorders among publicly or privately insured children in the United States. *JAMA Psychiatry* 79, 232–242. doi: 10.1001/jamapsychiatry.2021.3815

Thurman, A. J., Edgin, J. O., Sherman, S. L., Sterling, A., McDuffie, A., Berry-Kravis, E., et al. (2021). Spoken language outcome measures for treatment studies in Down syndrome: feasibility, practice effects, test-retest reliability, and construct validity of variables generated from expressive language sampling. *J. Neurodev. Disord.* 13:13. doi: 10.1186/s11689-021-09361-6

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., and O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *J. Speech Lang. Hear. Res.* 40, 1245–1260.

Torgeson, J. K., and Bryant, B. R. (2004). *Test of Phonological Awareness*, 2nd Edn Plus. Austin, TX: Pro-Ed. [Examiner's Manual].

U.S. Department of Education (2000). *Annual Report to Congress on the Implementation of IDEA*. Washington, DC: U.S. Government Printing Office.

Van Herwegen, J., Riby, D., and Farran, E. K. (2015). "Neurodevelopmental disorders: definitions and issues," in *Neurodevelopmental Disorders: Research Challenges and Solutions*, eds J. Van Herwegen and D. Riby (London: Psychology Press), 3–18. doi: 10.4324/9781315735313

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., and Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing*, 2nd Edn. Austin, TX: Pro-Ed. [Examiner's Manual]. doi: 10.1037/t52630-000

Wallace, G., and Hammill, D. D. (2013). *Comprehensive Receptive and Expressive Vocabulary Test*, 3rd Edn. Austin, TX: Pro-Ed. [Examiner's Manual].

Wetherby, A. M., and Prizant, B. M. (2002). *Communication and Symbolic Behavior Scales, Normed Edition*. Baltimore, MD: Brookes Publishing Company. [Manual]. doi: 10.1037/t11527-000

Wiig, E. H., and Secord, W. A. (2006). *Emerging Literacy & Language Assessment*. Greenville, SC: Super Duper Publications. [Manual].

Wiig, E. H., Secord, W. A., and Semel, E. (2020). *Clinical Evaluation of Language Fundamentals Preschool*, 3rd Edn. Bloomington, MN: NCS Pearson. [Manual].

Wiig, E. H., Semel, E., and Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals*, 5th Edn. Bloomington, MN: NCS Pearson. [Manual].

Williams, K. T. (2014). *Phonological and Print Awareness Scale*. Torrance, CA: Western Psychological Services. [Manual].

Williams, K. T. (2018). *Expressive Vocabulary Test*, 3rd Edn. Bloomington, MN: NCS Pearson. [Manual].

World Health Organization [WHO] (2019). *International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision)*. Geneva: World Health Organization.

Young, N. A. E. (2021). *"Childhood Disability in the United States: 2019," ACSBR-006, American Community Survey Briefs*. Washington, DC: U.S. Census Bureau.

Zhu, L., and Gonzalez, J. (2017). Modeling floor effects in standardized vocabulary test scores in a sample of low ses Hispanic preschool children under the multilevel structural equation modeling framework. *Front. Psychol.* 8:2146. doi: 10.3389/fpsyg.2017.02146

Zimmerman, I., Steiner, V., and Pond, R. (2011). *Preschool Language Scales*, 5th Edn. Bloomington, MN: NCS Pearson. [Manual]. doi: 10.1037/t15141-000

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 September 2022
DOI 10.3389/fpsyg.2022.936392

Check for updates

OPEN ACCESS

EDITED BY
Lénia Carvalhais,
Infante D. Henrique Portucalense University, Portugal

REVIEWED BY
Eleni Peristeri,
Aristotle University of Thessaloniki, Greece
Aparna Nadig,
McGill University, Canada
Bérengère G. Digard,
University of Edinburgh, United Kingdom

*CORRESPONDENCE
Rachel Reetzke
Reetzke@kennedykrieger.org

†PRESENT ADDRESS
Danika L. Pfeiffer,
Department of Speech-Language Pathology and Audiology, Towson University, Towson, MD, United States

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 05 May 2022
ACCEPTED 12 August 2022
PUBLISHED 06 September 2022

CITATION
Reetzke R, Singh V, Hong JS, Holingue CB, Kalb LG, Ludwig NN, Menon D, Pfeiffer DL and Landa RJ (2022) Profiles and correlates of language and social communication differences among young autistic children.
Front. Psychol. 13:936392.
doi: 10.3389/fpsyg.2022.936392

Profiles and correlates of language and social communication differences among young autistic children

Rachel Reetzke[1,2]*, Vini Singh[1], Ji Su Hong[1,2], Calliope B. Holingue[1,3], Luther G. Kalb[1,3,4], Natasha N. Ludwig[2,4], Deepa Menon[1,2], Danika L. Pfeiffer[1,2†] and Rebecca J. Landa[1,2]

[1]Center for Autism and Related Disorders, Kennedy Krieger Institute, Baltimore, MD, United States, [2]Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University School of Medicine, Baltimore, MD, United States, [3]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States, [4]Department of Neuropsychology, Kennedy Krieger Institute, Baltimore, MD, United States

Delays in early language development are characteristic of young autistic children, and one of the most recognizable first concerns that motivate parents to seek a diagnostic evaluation for their child. Although early language abilities are one of the strongest predictors of long-term outcomes, there is still much to be understood about the role of language impairment in the heterogeneous phenotypic presentation of autism. Using a person-centered, Latent Profile Analysis, we first aimed to identify distinct patterns of language and social communication ability in a clinic-based sample of 498 autistic children, ranging in age from 18 to 60 months (M = 33 mo, SD = 12 mo). Next, a multinomial logistic regression analysis was implemented to examine sociodemographic and child-based developmental differences among the identified language and social communication profiles. Three clinically meaningful profiles were identified from parent-rated and clinician-administered measures: Profile 1 (48% of the sample) "Relatively Low Language and Social Communication Abilities," Profile 2 (34% of the sample) "Relatively Elevated Language and Social Communication Abilities," and Profile 3 (18% of the sample) "Informant Discrepant Language and Relatively Elevated Social Communication Abilities." Overall, young autistic children from the lowest-resource households exhibited the lowest language and social communication abilities, and the lowest non-verbal problem-solving and fine-motor skills, along with more features of attention-deficit/hyperactivity disorder and atypical auditory processing. These findings highlight the need for effective community-based implementation strategies for young autistic children from low-resource households and underrepresented communities to improve access to individualized quality care.

KEYWORDS

autism, latent profile analysis, correlates, child-based factors, sociodemographic factors

Introduction

Autism is one of the most common and complex neurodevelopmental conditions, with an early-onset (Landa and Garrett-Mayer, 2006; Ozonoff et al., 2010) high heritability (heritability of 0.9; Tick et al., 2016), and increasing prevalence (1 in 44; Maenner, 2021). Language delays are characteristic of young autistic children (Landa and Garrett-Mayer, 2006; Ozonoff et al., 2010; Tager-Flusberg, 2016), and one of the most recognizable first concerns that motivates parents to seek a diagnostic evaluation for their child (Herlihy et al., 2015; Talbott et al., 2015). Although early language abilities (prior to age 6 years) are one of the strongest predictors of later academic performance, relationships, and quality of life (Petersen et al., 2013; Howlin and Magiati, 2017), there is still much to be understood about the role of language impairment in the heterogeneous phenotypic presentation of autism.

Just as autistic children exhibit significant heterogeneity across all levels of the phenotype (Jeste and Geschwind, 2014), they also show a range of language profiles. Studies have revealed that variability in language development is most pronounced prior to age 6 (Landa et al., 2013; Randall et al., 2016; Brignell et al., 2019) with more stable language profiles observed from prior to age 6 to 19 years (Pickles et al., 2014). Although many children develop spoken language, despite delays early in life, approximately 30% of autistic children use no to fewer than 20–30 spoken words beyond age 5 (Tager-Flusberg and Kasari, 2013; Tager-Flusberg, 2016). Differences in the use and understanding of language have been found to vary greatly (Ellis Weismer et al., 2010; Landa et al., 2013; Pickles et al., 2014; Whyte and Nelson, 2015). Some studies report greater difficulty in receptive language relative to expressive language (Charman et al., 2003; Luyster et al., 2008; Hudry et al., 2010), while others have found the opposite pattern (Luyster et al., 2008; Ellis Weismer et al., 2010; Kover et al., 2013). In support of the variability in receptive-expressive language profiles observed across studies, a meta-analysis examined discrepancies in receptive and expressive language ability among younger (Age: 1 to 5 years) and older (Age: 6 to 19 years) groups of autistic children and youth, and did not find evidence of a common receptive-expressive language profile in either age group (Kwok et al., 2015).

Social communication and interaction skills (hereafter, social communication) are intimately linked with language development in young autistic children. Based on the DSM-5 criteria, social communication skills broadly encompass behaviors related to social-emotional reciprocity (e.g., reduced sharing of emotions/affect/interests and difficulty initiating/responding to others), non-verbal communication (e.g., difficulty using and understanding gestures/body postures), and developing and maintaining relationships (e.g., lack of interest in others, difficulty making friends, and limited imaginative play) (American Psychiatric Association, 2013). Extant evidence suggests that there may be a bidirectional relationship between social communication and language development such that early social communication skills (i.e., initiation of and response to joint attention, gesture use, and imitation) are predictive of spoken language outcomes (Yoder et al., 2015; Delehanty et al., 2018; Pecukonis et al., 2019); and early expressive language skills are predictive of social communication outcomes (Anderson et al., 2009; Dillon et al., 2021). This can be seen in a retrospective study that examined predictors of language outcomes in a sample of 535, 8-year-old autistic children with significant language delay from the Simons Simplex Collection (Wodka et al., 2013). This study reported that higher levels of parent-reported social communication skills were associated with the acquisition of phrase and fluent speech, as well as an earlier age of language acquisition (even when non-verbal cognition was taken into account; Wodka et al., 2013). In contrast, other studies have found that social communication is not strongly associated with language development after other factors, such as non-verbal cognition, are considered (Sigman and McGovern, 2005; Thurm et al., 2015). One reason for the discrepancy in findings across studies may be due to differences in sample size and, as a result, statistical power. For example, Sigman and McGovern, 2005's study included 48 participants; Thurm et al., 2015's study included 70 participants; whereas Wodka et al., 2013's study included 535 participants. Another reason for the discrepancy in findings may be due to differences in sample sociodemographic factors (e.g., child age, sex, race, parental education) and/or child-based characteristics (e.g., developmental functioning, presence of co-occurring conditions). Given the heterogeneity of language and social communication profiles observed across young autistic children, it is important to identify factors that may account for such variability to improve the development of and access to individually tailored interventions aimed at improving language and social communication outcomes.

In terms of sociodemographic factors, lower parental education is one of the most consistently reported correlates of lower language abilities in autistic children (Anderson et al., 2007; Pungello et al., 2009; Warlaumont et al., 2014; Ellis Weismer and Kover, 2015; Olson et al., 2021). For example, Olson et al. (2021) examined the association between household- and neighborhood-level socioeconomic status variables and receptive and expressive language skills in autistic and neurotypical 15- to 64-month-olds. Lower maternal education was more strongly associated with lower receptive and expressive language skills across both participant groups, compared to income-based socioeconomic variables (Olson et al., 2021). Similarly, Ellis Weismer and Kover (2015) found that maternal education contributed to the correct classification of 80% of autistic children into high versus low language groups at age 5.5 years.

To date, no direct relationship between race and language development has been reported in the autism literature. However, there is evidence to suggest an indirect link between child/family race and language, and social communication outcomes. Recent reports indicate that children from underrepresented racial groups are less likely to receive an ASD evaluation compared to their non-Hispanic White counterparts by the child's third birthday (Maenner, 2021), which inevitably leads to delays in access to intervention services. Such intervention delays could hold negative consequences for language and social communication development, as history of intervention services has been associated with positive language and social communication outcomes (Mazurek et al., 2012).

Among child-based developmental factors, non-verbal cognition is one of the most consistently reported correlates of language outcomes (Mazurek et al., 2012; Wodka et al., 2013; Ellis Weismer and Kover, 2015; Thurm et al., 2015; Bal et al., 2020). Ellis Weismer and Kover (2015), found that non-verbal cognition at 2.5 years was a strong predictor of expressive language at 5.5 years. In two separate Simons Simplex Collection cohort studies, Mazurek et al. (2012) and Wodka et al. (2013) found that social communication intervention gains and acquisition of phrase or fluent speech were greatest among autistic children and adolescents with higher non-verbal cognition, respectively. Evaluating non-verbal cognition as a correlate of language outcome affords an estimation of general cognitive ability without the confound of language ability (DeThorne and Schaefer, 2004). As such, if a low level of non-verbal cognition is a correlate of similarly low levels of receptive and expressive language then such a profile may be indicative of a broad developmental delay, as opposed to a delay specific to language.

Beyond non-verbal cognition, auditory processing is another putative correlate of language outcomes in autism (Haesen et al., 2011; Boucher, 2012; Kujala et al., 2013; Matsuzaki et al., 2019). Here, one hypothesis is that difficulty in efficiently representing brief sounds and rapid auditory transitions (i.e., auditory temporal processing) may lead to further difficulty distinguishing phonemic contrasts (e.g., /ba/ vs. /ga/). Consistent with this hypothesis, children with lower language ability have shown difficulty in tasks that require the encoding of rapid spectrotemporal changes in auditory signals (Tallal and Gaab, 2006); and positive correlations have been found between auditory processing and later expressive language outcomes in infants at elevated likelihood for developing autism (Riva et al., 2018).

Both prospective (Bhat et al., 2012; LeBarton and Iverson, 2013; Iverson et al., 2019; LeBarton and Landa, 2019; Bal et al., 2020) and retrospective studies (Mody et al., 2017) have additionally shown a positive association between fine motor skills and later language and social communication development. Fine motor skills, the ability to coordinate the small muscles of the fingers and hands to reach, grasp, and manipulate objects, have been found to be more susceptible to delay in autism, when compared to gross motor skills, such as walking (Landa et al., 2012). This may explain some of the variability in language and social communication development among young autistic children, as evidence suggests that infants with more sophisticated object manipulation and exploration have increased opportunities to interact with their environment and learn from their caregivers. This in turn provides a rich scaffolding for language development (for a review see, Iverson, 2021).

Finally, features of other co-occurring neurodevelopmental and psychiatric conditions have also been associated with differences in language and social communication abilities among young autistic children. Indeed, autistic children with co-occurring attention deficit/hyperactivity disorder (ADHD) symptoms are often reported to have greater impairments in communication and socialization skills (Rao and Landa, 2014; Lyall et al., 2017; Yerys et al., 2019). Features of anxiety are also commonly observed in young autistic children and have been found to also interact with language and social communication abilities. However, unlike the inverse relationship observed between ADHD symptoms and language/social communication abilities, young autistic children with higher expressive language abilities tend to exhibit higher levels of anxiety symptoms, and vice versa (Davis et al., 2012; Vasa et al., 2016; Rodas et al., 2017).

To our knowledge, no study to date has characterized concurrent patterns of language *and* social communication abilities in a sample of autistic toddlers and preschoolers using a person-centered latent profile analytic (LPA) approach. LPA is particularly useful for describing the heterogeneity observed in a given sample by identifying subgroups (or latent profiles) with similar patterns of performance across multiple domains (Collins and Lanza, 2009; Lanza et al., 2013). In addition, no study has examined the extent to which sociodemographic and child-based developmental factors—commonly observed to interact with language and social communication development early in life—are associated with different profiles of language and social communication abilities among young autistic children. Identifying specific correlates of language and social communication profiles (beyond global non-verbal cognitive ability) is important to inform the development of individualized intervention targets (Bal et al., 2020; Saul and Norbury, 2020). To this end, we ask the following research questions: (1) How many qualitatively different profiles of language and social communication can be identified in a clinic-based sample of autistic toddlers and preschoolers using a person-centered (LPA) analytic approach?; (2) Are sociodemographic (i.e., race, parental education, medical insurance status, history of intervention) and child-based developmental factors (non-verbal problem solving skills, fine motor skills, auditory processing, as well as commonly co-occurring symptoms of ADHD and anxiety) differentially associated with identified language and social communication latent profiles?

Materials and methods

Participants

Data for this retrospective study were obtained from a sample of young autistic children who received a comprehensive autism spectrum disorder (ASD) evaluation at an urban, outpatient ASD specialty clinic located in the Mid-Atlantic region of the United States between June 2014 and December 2019. This research was approved by the Johns Hopkins Medicine Institutional Review Board.

Inclusion criteria for our analytic sample consisted of children who: (a) were between the ages of 18 to 60 months; (b) received an ASD diagnosis by a licensed, medical provider (e.g., psychiatrist, developmental behavioral pediatrician, or neurodevelopmental pediatrician) or licensed psychologist (clinical or neuro) based on the *Diagnostic and Statistical Manual of Mental Disorders, 5^{th} Edition* diagnostic criteria (American Psychiatric Association, 2013) and clinical judgment, informed by the *Autism Diagnostic Observation Schedule-Second Edition* (ADOS-2; Lord et al., 2012), medical, developmental and family history, as well as behavioral testing; and (c) completed all pre-appointment paperwork (completion rate = 62%) within 6 months (98% within a week) of their evaluation appointment.

The final analytic sample consisted of 498 autistic children, ranging in age from 18 to 58 months (*M* = 33 months; *SD* = 7 months). Children in this sample were predominantly male (80%), White (45%), and non-Hispanic (93%). 76.5% were diagnosed with ASD by a physician and 23.5% were diagnosed by a licensed psychologist. Parents who completed intake questionnaires and parent-report measures of language and social communication ability consisted of mothers (85%), with a college level education (52%), and private insurance (64%).

Measures

Language and social communication measures for latent profile analysis

Measures selected for the LPA included both parent-report and clinician-administered measures of language and social communication abilities (see **Table 2**).

Parent-rating of child language

During the clinic intake process, parents responded to five yes/no items and two multiple choice items about their child's receptive, expressive, and pragmatic language abilities on a clinic-specific Parent-rating of Child Language (PCL) questionnaire: (1) Are you worried about your child's language development? [Yes = 0; No = 1]; (2) Does your child have any problems with talking, hearing, being understood by others, or understanding what he/she is told? [Yes = 0; No = 1]; (3) Can your child speak phrases with at least 2- or 3-word combinations? [Yes = 3; No = 0]; (4) Can your child tell you about their day? [Yes = 3; No = 0]; (5) Can your child have a conversation? [Yes = 5; No = 0]; (6) How does your child usually communicate? [Babbling = 1; Single Words = 2; Short phrases = 3; Full sentences = 4]; (7) Does your child use sign language or any other communication device? [Sign language or PECS or Speech Generating Device = 1; No = 0]. The above codes were derived by weighted parent responses such that a higher code reflected a higher level of language ability. Codes were first created by the first author who is a certified and licensed speech-language pathologist (SLP). The codes were then discussed with and fine-tuned by an interdisciplinary team (which included two additional certified SLPs, two epidemiologists, a neuropsychologist, a psychiatrist, and a developmental behavioral pediatrician). For each participant, scores were summed for a total score ranging from 0 to 18, with higher scores reflecting higher language ability.

Mullen scales of early learning: Receptive and expressive language subscales

The Mullen Scales of Early Learning (MSEL; Mullen, 1995) is a clinician-administered standardized developmental assessment for children birth to 68 months. The Receptive Language (RL) and Expressive Language (EL) subscales were administered during the ASD diagnostic evaluation. Consistent with previous literature, developmental quotients (DQs) were calculated by dividing each MSEL subscale age-equivalent score by the child's chronological age and multiplying by 100 (Messinger et al., 2013). The RL and EL DQs were included in the LPA.

Child behavior checklist 1.5-5: Withdrawn subscale

The Child Behavior Checklist 1.5-5 (CBCL; Achenbach and Rescorla, 2000) is a parent-report, norm-referenced reliable and valid questionnaire developed to measure emotional, behavioral, and social limitations in young children. The checklist consists of 99 items describing the presence of a specific behaviors that are rated on a 3-point Likert scale (0 = not true, 1 = somewhat/sometimes true, 2 = very often true). Item scores are summed and converted to T-scores (M = 50; SD = 10) to derive "problem scores." The Withdrawn subscale T-score was included in the LPA to capture parent's ratings of child social communication functioning, as items of this scale include: "avoids looking others in the eye;" "doesn't answer when people talk to him/her;" "refuses to play active games;" "seems unresponsive to affection;" "shows little affection toward people;" "withdrawn, doesn't get involved with others."

Autism diagnostic observation schedule, second edition

The Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) is a clinician-administered, semi-structured standardized measure developed to assess the presence or absence of features of ASD related to communication, social interaction, play, and restricted, repetitive behaviors (Lord et al., 2012). The ADOS-2 consists of different modules, with module selection based on chronological age and language ability at the time of testing. Children were administered the ADOS-2 during their ASD diagnostic evaluation. The ADOS-2 was administered by a certified and licensed SLP or a licensed psychologist, clinically trained to administer the ADOS-2. Specifically, all clinicians completed a required ADOS-2 clinical training workshop with a certified ADOS-2 trainer prior to clinical administration of the ADOS-2. Clinicians had access to quarterly booster trainings and research-reliable ADOS-2 clinicians for consultation, if needed. The ADOS-2 Social Affect Calibrated Severity Score (ADOS-2-SA-CSS; score 1 to 10) was derived, reflecting the relative severity of social communication impairment and allowing comparisons across modules (Esler et al., 2015). Higher SA CSS scores reflect greater social communication limitations. The ADOS-2 SA CSS was included in the LPA to capture clinician's ratings of child social-communication functioning.

Correlates of language and social communication profiles

Correlates of language and social communication latent profiles included both sociodemographic and child-based developmental factors hypothesized to account for language and social communication heterogeneity among young autistic children.

Sociodemographic factors

Parents completed clinic-specific questionnaires upon initiating their child's intake process. This form was used for deriving sociodemographic variables. Questionnaires captured the following sociodemographic data: child age, sex, race/ethnicity, parent education, medical insurance, and history of intervention. Race was categorized as a four-level variable (Asian, Black, White, and Other). Other races included Native American, Pacific Islander, multiracial, and any other race. Prior to 2019, ethnicity was captured on this clinic-specific questionnaire as a racial category and therefore informants were unable to report both race and ethnicity for the majority of this study. Parental education was classified as No College Education vs. College Education. Medical Insurance was classified as public (reflecting Medical Assistance) vs. private (e.g., PPO) plans. History of intervention (i.e., parent report of speech therapy or general early intervention services) was derived as a binary variable (yes/no).

Auditory processing

Based on the child's age, auditory processing was measured using the auditory processing subscales from one of five different parent-report questionnaires: the Sensory Processing Measure-Preschool Home Form (SPM-P), the Toddler Sensory Profile-2 (TSP-2), the Child Sensory Profile-2 (CSP-2), the Infant Toddler Sensory Profile (ITSP), or the Short Sensory Profile (SSP). The SPM-P is a 75-item, reliable and valid questionnaire developed to assess seven different sensory processing categories in children aged 2–5 years (Glennon et al., 2011). The TSP-2 and the CSP-2 are part of the Sensory Profile 2, which has high internal consistency, interrater reliability, and test-retest stability (Dunn, 2014). The TSP-2 is a 54-item questionnaire developed to assess seven different sensory processing categories in children 7–35 months old. The CSP-2 is an 86-item questionnaire developed to assess six different sensory processing categories in children aged 3–14 years. The ITSP is a 36-item questionnaire developed to assess six different sensory processing categories. The ITSP has high internal consistency, test-retest reliability, and convergent validity (Dunn, 2002; Ben-Sasson et al., 2008; Beranova et al., 2017; Niedźwiecka et al., 2020). Finally, the SSP is a shortened 38-item version of the Sensory Profile 2 to assess sensory processing in children aged 3–17 years. A binary auditory processing variable was derived such that "typical auditory processing" was defined by a rating of *typical performance* (SPM, ITSP, SSP) or *just like the majority of the others* (CSP, TSP) and "atypical auditory processing" was defined by all other ratings. We did not differentiate between over and under processing since SPM (*Typical*, *Some Dysfunction*, *Definite dysfunction*) does not have this distinction and comprised 70% of the data.

Non-verbal problem-solving and fine motor skills

The MSEL's (described above) Visual Reception (VR) and Fine Motor (FM) subscales were administered during the ASD diagnostic evaluation. VR and FM DQs were derived and included as continuous variables in the multinomial logistic regression analysis.

Co-occurring features of attention deficit/hyperactivity disorder and anxiety

The CBCL, as described above, yields a total of five DSM-based condition scores. To include measures of ADHD and anxiety traits, respective subscale T-scores were included as continuous variables in the multinomial logistic regression analysis. A T-score of 70 and above reflects "clinically significant" features.

Statistical analysis

Distinct language and social communication profiles were derived using Latent Profile Analysis (LPA). LPA is a person-centered, mixture modeling approach for detecting and estimating underlying sample clustering. LPA divides participants into distinct subgroups, for the purposes of maximizing within group homogeneity, based on the posterior probability of continuous indicator responses (i.e., phenotypic characteristics). In the current study, parent-reported (PCL and CBCL–Withdrawn T-Score) and clinician-administered (MSEL EL-DQ, RL-DQ, and ADOS SA-CSS) measures served as model indicators. These measures were used to identify distinct language and social communication profiles among young autistic children.

Models with two to six class solutions were fitted to the data and compared in: (a) goodness of fit statistics, (b) proportion of participants classified within each profile, and (c) whether the profiles were clinically meaningful. Fit statistics included the Akaike Information Criterion (AIC), Bayesian information criterion (BIC; Schwarz, 1978), Integrated Completed Likelihood (ICL; Biernacki et al., 2000), Entropy, and Parametric Bootstrapped likelihood ratio test (BLRT; McLachlan and Rathnayake, 2014). Lower values of BIC, AIC and ICL indicate better model fit. Entropy values closer to 1 denote an improvement in the classification of participants. A significant p-value associated with the BLRT for a given model indicates an improvement in fit compared to K-1 classes. After determining the optimal number of classes, participants were assigned class membership based on their posterior probability. All variables were converted to a z-score distribution (M = 0, SD = 1), to provide a uniform metric for the LPA. Lower z-scores reflect greater impairment.

After determining class membership, differences in sociodemographic and child-based factors were examined using chi-square tests for categorical measures and one-way analysis of variance for continuous measures. Two separate multinomial regression models–one for sociodemographic characteristics and one for child-based developmental factors–were then used to find adjusted correlates of language and social communication profiles. The outcome for all models was odds of belonging to a particular latent profile and the correlates were those identified as significant ($p < 0.05$) from the bivariate analysis. We opted to use two separate models as we did not have *a priori* hypotheses regarding how different latent profiles and child-based developmental differences would interact with sociodemographic factors. In addition, we found that child-based factors were strongly associated with group membership leading to very little variability that could be explained by sociodemographic factors. Finally, we wanted to avoid potential collinearity between child age and the fine motor and non-verbal problem-solving DQs, as age is part of the calculation of all DQs. Therefore, sociodemographic and child-based factors were examined in separate models to obtain interpretable associations.

All analyses were completed in R Version 1.2.5033 (R Core Team, 2020) using packages mclust (Scrucca et al., 2016) and nnet (Ripley and Venables, 2022). All tests were 2-sided with p-values of < 0.05 considered statistically significant.

Results

Latent profile analysis

As shown in **Table 1**, AIC, BIC, and ICL values were the lowest for the 3- and 4-class models. Entropy values for the 3- and 4-class solutions suggested similar classification of participants (3-class entropy = 0.83 and 4-class entropy = 0.87). However, two classes in the 4-class model were nearly identical, with differences only observed on the ADOS SA CSS (6.39 vs. 9.34). As a result, the 4-class solution was deemed less interpretable, and the 3-class model was selected as the more clinically meaningful solution.

Figure 1 shows standardized means by profile and **Table 2** shows the actual means by profile.

Profile 1 "Relatively Low Language and Social Communication Abilities", included 48% of the sample (n = 237; 77% male). Children in Profile 1 were characterized by the lowest levels of language and social communication abilities on both clinician-administered and parent-report measures. These children also exhibited the lowest non-verbal problem solving (M = 54.1; SD = 17.0) and fine-motor skills (M = 61.1; SD = 14.9) and included the highest percentage of autistic children with atypical auditory processing (67%), and relatively more features of ADHD (M = 58.7; SD = 7.9) compared to Profile 2.

Profile 2 "Relatively Elevated Language and Social Communication Abilities", included 34% of the sample

TABLE 1 Model fit statistics and *n* (%) by class for latent profile models with two to six classes.

						Class *n* (%) based on the estimated model					
No. of classes	**AIC**	**BIC**	**ICL**	**Entropy**	**BLRT**	**1**	**2**	**3**	**4**	**5**	**6**
2	16201	16331	16468	0.89	< 0.001	276 (56)	222 (45)				
3	**16124**	**16300**	**16461**	**0.83**	**< 0.001**	**237 (48)**	**170 (35)**	**91 (19)**			
4	**16073**	**16297**	**16469**	**0.87**	**< 0.001**	**102 (21)**	**173 (35)**	**75 (15)**	**148 (30)**		
5	16114	16383	16548	0.79	0.86	48 (10)	193 (39)	52 (11)	118 (24)	87 (18)	
6	16086	16402	16598	0.76	–	93 (19)	178 (36)	10 (2)	91 (19)	46 (9)	80 (16)

AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; ICL, Integrated Completed Likelihood; BLRT, parametric bootstrapped likelihood ratio test. Entropy closer to 1 reflects a good classification of participants. A significant BLRT *p*-value indicates that the model with a greater number of classes fit the data better relative to a fewer number of classes. Bold denotes the two best-fitting models.

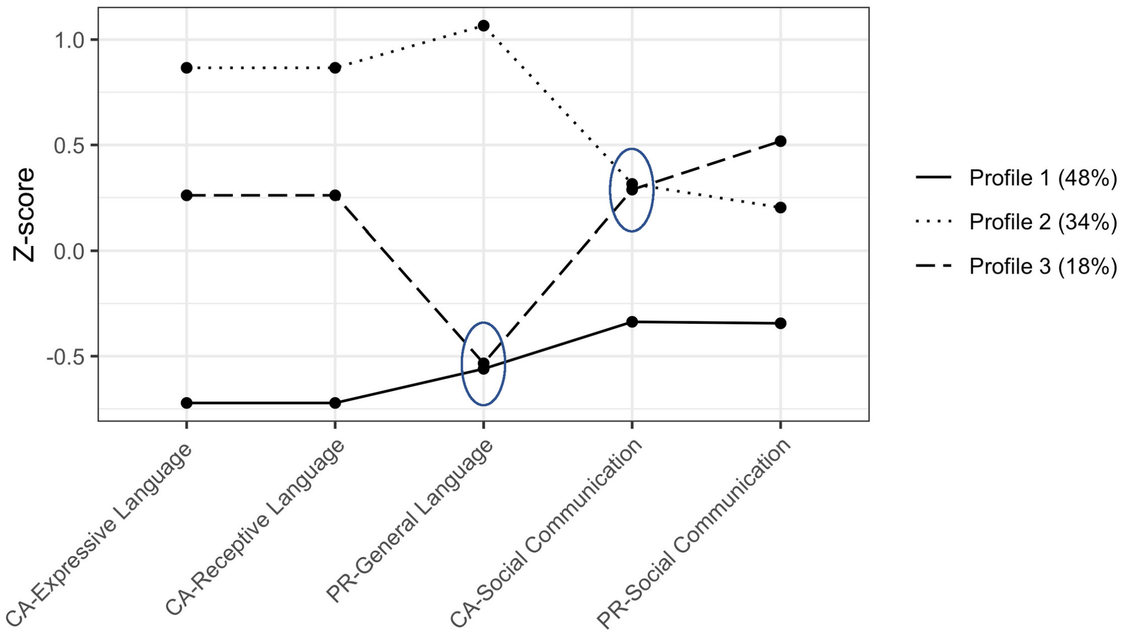

FIGURE 1
Language and social communication variable z-scored means for the three-profile solution. Ellipses encompass means that are not significantly different. CA, clinician-administered; PR, parent-rated. Profile 1, "Relatively Low Language and Social Communication Abilities;" Profile 2, "Relatively Elevated Language and Social Communication Abilities;" Profile 3, "Informant Discrepant Language and Relatively Elevated Social Communication Abilities."

TABLE 2 Means and standard deviations of language and social communication variables by latent profile.

			Profile 1	Profile 2	Profile 3
Assessment variable	**Assessor**	**Score**	**(*n* = 237)**	**(*n* = 170)**	**(*n* = 91)**
ADOS-2 SA CSS (Social Communication)	Clinician	CSS (range = 1–10)	8.10 (1.71)	6.75 (2.14)[a]	6.80 (2.24)[a]
CBCL-Withdrawn (Social Communication)	Parent	T Score (M = 50; SD = 10)	73.7 (10.6)	67.7 (11.1)	64.2 (8.16)
MSEL EL DQ (Expressive Language)	Clinician	DQ (M = 100; SD = 15)	30.5 (10.4)	68.8 (23.9)	54.2 (13.8)
MSEL RL DQ (Receptive Language)	Clinician	DQ (M = 100; SD = 15)	29.1 (9.75)	63.1 (25.1)	54.7 (18.8)
PCL (General Language)	Parent	Composite (range = 0–26)	3.21 (0.67)[a]	7.28 (2.55)	3.27 (0.87)[a]

CSS, Calibrated Severity Score; CBCL, Child Behavior Checklist; MSEL EL DQ, Mullen Scales of Early Learning – Expressive Language Developmental Quotient; MSEL RL DQ, Mullen Scales of Early Learning – Receptive Language Developmental Quotient; PCL, Parent rating of Child Language (Custom score, see text for details).
[a]Denotes non-significant difference between marked groups ($p > 0.05$); All other differences are significant at $p < 0.05$.

(n = 170; 82% male). Children in Profile 2 were characterized by the highest levels of language and social communication ability on both clinician and parent reported measures. These children also exhibited the highest non-verbal problem solving skills (M = 76.0; SD = 22.8), which were significantly elevated compared to Profile 1, and the most features of anxiety (M = 56.7; SD = 9.81), which was significantly elevated compared to Profile 3, but not Profile 1.

Profile 3 "Informant Discrepant Language and Relatively Elevated Social Communication Abilities", included 18% of the sample (n = 91; 86% male). Children in Profile 3 were characterized by elevated levels of social communication ability on both clinician-administered and parent-report measures but moderate (clinician-administered) and low (parent-report) levels of language ability. These children exhibited the highest fine motor skills (M = 78.8; SD = 24.5) compared to both Profiles 1 and 2. Only in comparison to Profile 1, Profile 3 included the lowest percentage of children with atypical auditory processing (51%) and fewer features of ADHD (M = 56.3.0; SD = 7.14). Finally, compared to Profile 2 only, Profile 3 showed significantly fewer features of anxiety (M = 54.0; SD = 6.01).

Correlates of language and social communication profiles

Bivariate relationships between latent class membership and sociodemographic and child-based correlates are presented in Table 3. Among the sociodemographic variables, child age, race, parental education and medical insurance status were significantly different ($p < 0.05$) across groups. The multinomial logistic regression included these sociodemographic variables while controlling for child age (Table 4). Children with college-educated parents were more likely to be in Profile 2 (OR = 2.61; $p < 0.001$) and Profile 3 (OR = 1.99; $p < 0.02$), compared to Profile 1. Compared to White children, Black children were less likely to be in Profile 2 (OR = 0.52; p = 0.04), relative to Profile 1, and more likely to be in Profile 3 (OR = 2.88; p = 0.03) compared to Profile 2. Insurance and other race categories were not significant after adjusting for parent education.

As shown in Table 5, the child-based variables associated with group membership were VR DQ (non-verbal problem solving skills), FM DQ (fine motor skills), CBCL ADHD (features of ADHD), CBCL Anxiety (features of Anxiety) and auditory processing. Compared to Profile 1, higher fine motor, non-verbal problem-solving skills, and typical auditory processing were associated with increased likelihood of being in Profile 2 [non-verbal problem solving: odds ratio (OR) = 1.06; auditory processing: OR = 1.99; all $ps < 0.05$] or Profile 3 (non-verbal problem solving: OR = 1.04, fine motor skills: OR = 1.04; auditory processing skills: OR = 2.07; all $p < 0.05$). Higher ADHD symptomatology were significantly associated with lower likelihood of being in Profile 3 (OR = 0.96; p = 0.04) but not Profile 2 as compared to Profile 1. Higher non-verbal problem-solving skills (OR = 1.02), lower fine motor skills (OR = 0.96), and higher levels of anxiety (OR = 1.05) were significantly associated with being in Profile 2 compared to Profile 3.

Discussion

The current study aimed to understand the heterogeneity in language and social communication profiles, and their sociodemographic and child-based developmental correlates, within a large clinic-based sample of young autistic children. Using LPA, a person-centered approach to statistical modeling,

TABLE 3 Sociodemographic and child characteristics by latent profile.

N (%)	Profile 1	Profile 2	Profile 3	*P*-value
	237 (48%)	**170 (34%)**	**91 (18%)**	
Age at evaluation (M, SD)	2.65 (0.58)	3.03 (0.59)	2.46 (0.56)	< 0.001
Male sex (n,%)	181 (76.7)	139 (82.2)	78 (85.7)	
Race (n,%)				0.02
Asian	9 (3.81)	16 (9.47)	4 (4.40)	
Black	77 (32.6)	31 (18.3)	28 (30.8)	
White	100 (42.4)	86 (50.9)	39 (42.9)	
Other	50 (21.2)	36 (21.3)	20 (22.0)	
Ethnicity (n,%)				0.72
Hispanic	18 (7.79)	10 (6.10)	5 (5.68)	
Not Hispanic	213 (92.2)	154 (93.9)	83 (94.3)	
Parent education (n,%)				< 0.001
No College Degree	143 (61.1)	54 (32.1)	39 (42.9)	
College Degree	91 (38.9)	114 (67.9)	52 (57.1)	
Medical insurance (n,%)				< 0.001
Medicaid/Public	108 (45.8)	40 (23.7)	32 (35.2)	
Private	128 (54.2)	129 (76.3)	59 (64.8)	
History of intervention (n,%)	195 (84.8)	145 (85.8)	73 (84.9)	0.96
MSEL VR DQ (M, SD)	54.1 (17.0)	76.0 (22.8)	73.5 (17.5)	< 0.001
MSEL FM DQ (M, SD)	61.1 (14.9)	73.5 (15.4)	78.8 (24.5)	< 0.001
Auditory processing (n,%)				0.02
Atypical	156 (67.0)	96 (57.1)	46 (51.1)	
Typical	77 (33.0)	72 (42.9)	44 (48.9)	
CBCL, ADHD (M, SD)	58.7 (7.85)	57.4 (7.91)	56.3 (7.14)	0.03
CBCL, Anxiety (M, SD)	55.1 (7.19)	56.7 (9.81)	54.0 (6.01)	0.02

M, mean; SD, standard deviation; MSEL VR DQ, Mullen Scales of Early Learning Visual Reception Developmental Quotient; MSEL FM DQ, Mullen Scales of Early Learning Fine Motor Developmental Quotient; CBCL, Child Behavior Checklist; ADHD, attention deficit/hyperactivity disorder.

TABLE 4 Parameter estimates of the sociodemographic factors multinomial logistic regression model.

	Profile 2 vs. 1	Profile 3 vs. 1	Profile 3 vs. 2	Wald	Pairwise
	OR (CI)	OR (CI)	OR (CI)		
Age at evaluation	3.53 (2.37–5.26)[c]	0.53(0.33–0.85)[b]	0.12[c](0.06–0.21)	72.73[c]	2 > 1 > 3
Race					
White	REF	REF	REF		
Black	0.52 (0.29–0.93)[a]	1.34 (0.7–2.56)	2.88 (1.28–6.59)[a]	6.42[a]	3 > 2, 1 > 2
Asian	0.89 (0.51–1.56)	0.94 (0.49–1.82)	0.95 (0.43–2.03)	0.08	–
Other	1.48 (0.58–3.77)	0.82 (0.23–2.91)	0.55 (0.13–1.89)	1.23	–
Parent education					
No college	REF	REF	REF		
College degree	2.62 (1.59–4.32)[c]	2.00 (1.11–3.58)[c]	1.00 (0.48–2.10)	16.30[c]	2 > 1, 3 > 1
Medical Insurance					
Medicaid/Public	REF	REF	REF		
Private	1.46 (0.84–2.55)	1.18 (0.62–2.26)	0.67 (0.29–1.53)	1.74	–

OR = Odds Ratio; CI = Confidence Interval; [a]$p < 0.05$; [b]$p < 0.01$; [c]$p < 0.001$.

TABLE 5 Parameter estimates of the child-based factors multinomial logistic regression model.

	Profile 2 vs. 1	Profile 3 vs. 1	Profile 3 vs. 2	Wald	Pairwise
	OR (CI)	OR (CI)	OR (CI)		
MSEL VR DQ	1.06 (1.04–1.08)[c]	1.04 (1.02–1.06)[c]	0.98 (0.96 – 0.99)[a]	41.63[c]	2 > 3 > 1
MSEL FM DQ	1.01 (0.98–1.03)	1.04 (1.02–1.07)[b]	1.04 (1.01–1.06)[b]	11.93[b]	3 > 1,3 > 2
Auditory processing					
Atypical	REF	REF	REF		
Typical	1.99 (1.16–3.41)[a]	2.07 (1.13–3.78)[a]	0.97 (0.53–1.78)	8.08[a]	2 > 1,3 > 1
CBCL, ADHD problems	0.96 (0.93–0.99)[a]	0.97 (0.93–1.01)	1.01 (0.97–1.06)	4.64	1 > 2
CBCL, Anxiety problems	1.04 (0.99–1.07)	0.99 (0.95–1.04)	0.95 (0.91–0.99)[a]	5.88[a]	2 > 3

OR, Odds Ratio; CI, Confidence Interval; [a]$p < 0.05$; [b]$p < 0.01$; [c]$p < 0.001$.

three meaningful profiles of language and social communication abilities were identified from parent-report and clinician-administered measures: Profile 1 "Relatively Low Language and Social Communication Abilities", Profile 2 "Relatively Elevated Language and Social Communication Abilities", Profile 3 "Informant Discrepant Language and Relatively Elevated Social Communication Abilities." Slightly less than half of the children were in Profile 1, whereas a third were in Profile 2 and the remaining eighteen percent were in Profile 3. Overall, profiles were distinguished by different levels of language and social communication ability (e.g., high, medium, and low) and discrepant parent-report and clinician measurement of language (i.e., Profile 2 vs. 3). Significant differences were found in the sociodemographic and child-based developmental correlates of these profiles, including level of parental education, race, non-verbal problem-solving skills, fine motor skills, auditory processing, and co-occurring child mental health characteristics (i.e., ADHD and anxiety).

Despite previous reports suggesting that autistic children present with a unique language profile such that use of language (i.e., expressive language) exceeds the ability to understand language (i.e., receptive language; Charman et al., 2003; Luyster et al., 2008; Hudry et al., 2010), we did not observe such a pattern in any of the three profiles identified in the current study. Instead, our results are consistent with findings from a meta-analysis, revealing no specific receptive-expressive profile among young autistic children (Kwok et al., 2015). Taken together, our findings suggest that a receptive-expressive language discrepancy is not a common profile in autistic toddlers or preschoolers. However, differences may emerge as a child develops language, beyond toddlerhood and preschool age, if abilities take divergent trajectories.

Characterizing almost half of the sample, autistic children in Profile 1 exhibited the lowest scores across clinician-administered measures of receptive language, expressive language, and social communication as well as the lowest scores on a parent-rated measure of social communication abilities. Perhaps not surprisingly, children in Profile 1 also exhibited the lowest scores in non-verbal problem-solving, fine motor skills, auditory processing, and high levels of ADHD symptoms. These results are consistent with literature indicating that autistic children with co-occurring ADHD symptoms often have greater impairments in communication and social communication skills (Rao and Landa, 2014; Lyall et al., 2017; Yerys et al., 2019), and consistent with an extant and growing body of evidence suggesting that children with ASD plus high ADHD symptoms tend to present with generally lower developmental functioning (Karalunas et al., 2018; Antshel and Russo, 2019; Miller et al., 2020; Hong et al., 2021; Reetzke et al., 2021a).

Children from the lowest-resource households (defined by lower levels of parent education and a higher percentage of families with a public medical insurance) tended to be in Profile 1. This group also included a higher percentage of Black children, which may be suggestive of potential racial disparity in services. Indeed, recent reports indicate that children from underrepresented racial groups are less likely to receive an ASD evaluation compared to their non-Hispanic White counterparts by three years (Maenner, 2021), which inevitably leads to delays in access to intervention services. However, our findings indicated that there was no significant difference regarding history of intervention for the children in Profile 1. This finding is consistent with reports that early intervention (e.g., Birth-to-3 speech-language therapy) may not be predictive of language growth in autistic children from 2.5 to 5.5 years of age (Ellis Weismer and Kover, 2015). The low level of language, social communication, and developmental functioning may instead reflect the *quality and quantity* of intervention services children in Profile 1 received compared to children in the other groups. Even if these children did receive a comparable quantity of services, which could not be identified given the dichotomous classification, the intervention received may have not met the specific needs of the child and family. Unfortunately, the current data are limited in being able to pinpoint the exact mechanism underlying this disparity. These findings highlight the need for effective community-based implementation strategies for autistic children from low-resource households and underrepresented communities to improve access to individualized, quality care. For example, the Exploration, Preparation, Implementation, Sustainment (EPIS; Aarons et al., 2011), an implementation framework focused on influential contextual factors (e.g., service environment, policies, family cultural characteristics, etc.), can be used in partnership with community stakeholders to identify potential barriers that may hinder the uptake of community-based early interventions (Stahmer et al., 2019).

Children in Profile 2 exhibited relatively elevated language and social communication abilities, as well as non-verbal problem-solving skills and fewer parent-endorsed features of ADHD. Children in this group included a higher percentage of children from White, college-educated families. Children in Profile 2 were uniquely characterized by a higher level of parent-endorsed features of anxiety, consistent with the previous literature showing young autistic children with higher expressive language abilities tend to exhibit higher levels of anxiety symptoms, and vice versa (Davis et al., 2012; Vasa et al., 2016; Rodas et al., 2017). This may be a function of measurement limitations given that most parent-report based measures of anxiety, like the CBCL, are highly reliant on a child's ability to verbally express their anxiety. Thus, our present findings, and the broader literature, may underestimate the presence of anxiety symptoms in young autistic children until reliable and valid measures are developed for individuals at all levels of language ability (i.e., non-speaking to speaking).

Profile 3 had similarly low parent-rated language abilities as Profile 1, yet relatively moderate language abilities per clinician-administered direct measures of receptive and expressive language, reflecting an informant discrepancy in language abilities. This finding contrasts with Profiles 1 and 2 as well as extant and emerging literature (Miller et al., 2017; Reetzke et al., 2021b), showing that parent report of language abilities does not significantly differ from clinician assessment of receptive and expressive language skills. To better understand why this pattern of informant discrepancy was only observed for Profile 3, we examined whether there were specific language differences between Profiles 3 and 1 which might not have been captured by our clinic-based, parent-report measure of child language.

Parents of 32- to 36-month-old children categorized as Profiles 3 and 1 reported similar concerns about expressive and receptive language skills, indicating that although their children used some single words they were not yet using two- or three-word phrases to communicate. However, parents of children in Profile 1 reported that their children typically babbled to communicate, while parents of children in Profile 3 indicated that their children typically used single words to communicate. While our clinic-based parent-report measure of child language was not sensitive to these differences, the MSEL receptive and expressive language scales were. For example, children in Profile 3 used single words to label objects and pictures in their environment on the MSEL, while children in Profile 1 were not yet able to do this. Overall, these findings highlight the importance of collecting multiple types of parent-report and clinician-administered measures to estimate a child's language ability during the clinic evaluation process.

In terms of general patterns observed across correlates of language profiles, our finding that children from households with higher parental education were in the relatively elevated language and social communication Profile 2 and children with lower parental education tended to be in the relatively low language and social communication Profile 1 is consistent with the literature showing strong association between parental education and language abilities in autistic children (Anderson et al., 2007; Pungello et al., 2009; Warlaumont et al., 2014; Ellis Weismer and Kover, 2015; Olson et al., 2021). In addition, our findings are aligned with the majority of extant literature which has found a strong positive association between non-verbal cognitive ability and language abilities (Anderson et al., 2007; Wodka et al., 2013; Thurm et al., 2015; Yoder et al., 2015; Brignell et al., 2019). Even after consideration of other correlates of language and social communication abilities, non-verbal problem-solving skills were associated rather robustly with language and social communication profile membership. In addition, consistent with previous findings (Bhat et al., 2012; LeBarton and Iverson, 2013; LeBarton and Landa, 2019; Bal et al., 2020; Iverson, 2021), we found a significant association between fine motor skills and latent profile group membership. Here, Profile 3 showed relative strengths in fine motor skills compared to the other two profiles. These findings provide further support for the role of non-verbal problem-solving and fine motor skills in the development of receptive and expressive language; and highlight the importance of early intervention focused on these developmental domains as potential pathways for improving language outcomes in young autistic children.

Limitations

The results of this study should be interpreted within the context of several limitations. First, with only one time point, we are unable to examine the temporal stability of the sociodemographic and child-based developmental correlates of language and social communication profiles identified in the current study. Research is needed to replicate the findings of the present study in a longitudinal cohort to examine whether the correlates of language and social communication profiles in young autistic children are predictive of more developmentally downstream language and social communication outcomes. Second, our findings may not be representative of the general autistic population, as our participants were recruited from a single autism specialty clinic, limiting the generalizability of these findings. Third, our clinic-based, parent-rated measure of child language ability was a relatively short, omnibus measure of language ability which differed from our clinician-administered standardized measure of receptive and expressive language ability. Although our parent-rated measure of child language ability was strongly correlated with the MSEL receptive and expressive language subscales, it is possible that our LPA may have yielded different results with more comparable parent-rated and clinician-administered measures of receptive and expressive language ability. Fourth, the measures included in the current study were limited to what was available in the patient medical records (e.g., history of intervention captured as a binary variable), introducing potential measurement bias. Future work should aim to capture both the quantity and quality of intervention services to better understand how disparities in access to intervention services may be associated with variability in language and social communication profiles among young autistic children.

Conclusion

An LPA identified three language and social communication profiles based on parent-report and clinician-administered measures of language and social communication ability. Children from the lowest-resource households exhibited the lowest language and social-communication abilities, and the lowest non-verbal problem solving and fine-motor skills, along with more features of attention-deficit/hyperactivity disorder and atypical auditory processing. These findings highlight the need for effective community-based implementation strategies for autistic children from low-resource households to improve access to individualized quality care.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Johns Hopkins Medicine Institutional Review Board. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

RR and RL conceived, outlined, and supervised the original study. JH and VS acquired the data. VS organized the database and performed the statistical analyses. RR and VS wrote the first draft of the manuscript. All authors contributed to the refinement of the research questions and methodology, contributed to the interpretation of the results, critical revision of the manuscript, read, and approved the submitted version.

Funding

This study was supported in part by the Kennedy Krieger Institute Intellectual and Developmental Disabilities Research Center (P50 HD103538).

Acknowledgments

We acknowledge the children and their families who participated in this study, as well as the clinicians who administered the standardized assessments, and the staff who provided administrative support, collected data, and processed data. We appreciate Jacqueline Liu, Stephanie Choi, and Quing Lu who abstracted the MSEL and sensory measures from the clinic database under the supervision of JH. We also thank Jessica Morrel for her assistance with manuscript formatting and preparation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aarons, G. A., Hurlburt, M., and Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Adm. Policy Ment. Health* 38, 4–23. doi: 10.1007/s10488-010-0327-7

Achenbach, T. M., and Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles*. Burlington, VT: University of Vermont.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders. Fifth Edition (DSM-5)*. Arlington, VA: American Psychiatric Association.

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., et al. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *J. Consult Clin. Psychol.* 75:594. doi: 10.1037/0022-006X.75.4.594

Anderson, D. K., Oti, R. S., Lord, C., and Welch, K. (2009). Patterns of growth in adaptive social abilities among children with autism spectrum disorders. *J. Abnorm. Child Psychol.* 37, 1019–1034. doi: 10.1007/s10802-009-9326-0

Antshel, K. M., and Russo, N. (2019). Autism Spectrum Disorders and ADHD: Overlapping Phenomenology, Diagnostic Issues, and Treatment Considerations. *Curr. Psychiatry Rep.* 21:34. doi: 10.1007/s11920-019-1020-5

Bal, V. H., Fok, M., Lord, C., Smith, I. M., Mirenda, P., Szatmari, P., et al. (2020). Predictors of longer-term development of expressive language in two independent longitudinal cohorts of language-delayed preschoolers with autism spectrum disorder. *J. Child Psychol. Psychiatry* 61, 826–835. doi: 10.1111/jcpp.13117

Ben-Sasson, A., Cermak, S. A., Orsmond, G. I., Tager-Flusberg, H., Kadlec, M. B., and Carter, A. S. (2008). Sensory clusters of toddlers with autism spectrum disorders: Differences in affective symptoms. *J. Child Psychol. Psychiatry* 49, 817–825. doi: 10.1111/j.1469-7610.2008.01899.x

Beranova, S., Stoklasa, J., Dudova, I., Markova, D., Kasparova, M., Zemankova, J., et al. (2017). A possible role of the Infant/Toddler Sensory Profile in screening for autism: A proof-of-concept study in the specific sample of prematurely born children with birth weights <1,500 g. *Neuropsychiatr. Dis. Treat.* 13, 191–200. doi: 10.2147/NDT.S123066

Bhat, A. N., Galloway, J. C., and Landa, R. J. (2012). Relation between early motor delay and later communication delay in infants at risk for autism. *Infant Behav. Dev.* 35, 838–846. doi: 10.1016/j.infbeh.2012.07.019

Biernacki, C., Celeux, G., and Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 719–725. doi: 10.1109/34.865189

Boucher, J. (2012). Research review: Structural language in autistic spectrum disorder - characteristics and causes. *J. Child Psychol. Psychiatry* 53, 219–233. doi: 10.1111/j.1469-7610.2011.02508.x

Brignell, A., May, T., Morgan, A. T., and Williams, K. (2019). Predictors and growth in receptive vocabulary from 4 to 8 years in children with and without autism spectrum disorder: A population-based study. *Autism* 23, 1322–1334. doi: 10.1177/1362361318801617

Charman, T., Drew, A., Baird, C., and Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *J. Child Lang.* 30, 213–236. doi: 10.1017/s0305000902005482

Collins, L. M., and Lanza, S. T. (2009). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences.* Hoboken, NJ: John Wiley & Sons, Inc, doi: 10.1002/9780470567333

Davis, T. E. III, Moree, B. N., Dempsey, T., Hess, J. A., Jenkins, W. S., Fodstad, J. C., et al. (2012). The effect of communication deficits on anxiety symptoms in infants and toddlers with autism spectrum disorders. *Behav. Ther.* 43, 142–152. doi: 10.1016/j.beth.2011.05.003

Delehanty, A. D., Stronach, S., Guthrie, W., Slate, E., and Wetherby, A. M. (2018). Verbal and nonverbal outcomes of toddlers with and without autism spectrum disorder, language delay, and global developmental delay. *Autism Dev. Lang. Impairments* 3:239694151876476. doi: 10.1177/2396941518764764

DeThorne, L. S., and Schaefer, B. A. (2004). A Guide to Child Nonverbal IQ Measures. *Am. J. Speech Lang. Pathol.* 13, 275–290. doi: 10.1044/1058-0360(2004/029)

Dillon, E., Holingue, C., Herman, D., and Landa, R. J. (2021). Psychometrics of the pragmatic rating scale for school-age children with a range of linguistic and social communication skills. *J. Speech Lang. Hear.* 64, 3477–3488. doi: 10.1044/2021_JSLHR-20-00753

Dunn, W. (2002). *Infant/Toddler Sensory Profile manual.* New York, NY: Psychological Corporation.

Dunn, W. (2014). *Sensory Profile 2. User's manual.* Bloomington, MN: Pearson.

Ellis Weismer, S., and Kover, S. T. (2015). Preschool language variation, growth, and predictors in children on the autism spectrum. *J. Child Psychol. Psychiatry* 56, 1327–1337. doi: 10.1111/jcpp.12406

Ellis Weismer, S., Lord, C., and Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *J. Autism Dev. Disord.* 40, 1259–1273. doi: 10.1007/s10803-010-0983-1

Esler, A. N., Bal, V. H., Guthrie, W., Wetherby, A., Ellis Weismer, S., and Lord, C. (2015). The Autism Diagnostic Observation Schedule, Toddler Module: Standardized Severity Scores. *J. Autism Dev. Disord.* 45, 2704–2720. doi: 10.1007/s10803-015-2432-7

Glennon, T. J., Miller Kuhaneck, H., and Herzberg, D. (2011). The Sensory Processing Measure–Preschool (SPM-P)—Part One: Description of the Tool and Its Use in the Preschool Environment. *J. Occup. Ther. Sch. Early Interv.* 4, 42–52. doi: 10.1080/19411243.2011.573245

Haesen, B., Boets, B., and Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Res. Autism Spectr. Disord.* 5, 701–714. doi: 10.1016/j.rasd.2010.11.006

Herlihy, L., Knoch, K., Vibert, B., and Fein, D. (2015). Parents' first concerns about toddlers with autism spectrum disorder: Effect of sibling status. *Autism* 19, 20–28. doi: 10.1177/1362361313509731

Hong, J. S., Singh, V., and Kalb, L. (2021). Attention Deficit Hyperactivity Disorder Symptoms in Young Children with Autism Spectrum Disorder. *Autism Res.* 14, 182–192. doi: 10.1002/aur.2414

Howlin, P., and Magiati, I. (2017). Autism spectrum disorder: Outcomes in adulthood. *Curr. Opin. Psychiatry* 30, 69–76.

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., et al. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *Int. J. Lang. Commun. Disord.* 45, 681–690. doi: 10.3109/13682820903461493

Iverson, J. M. (2021). Developmental Variability and Developmental Cascades: Lessons from Motor and Language Development in Infancy. *Curr. Dirr. Psychol. Sci.* 30, 228–235. doi: 10.1177/0963721421993822

Iverson, J. M., Shic, F., Wall, C. A., Chawarska, K., Curtin, S., Estes, A., et al. (2019). Early motor abilities in infants at heightened versus low risk for ASD: A Baby Siblings Research Consortium (BSRC) study. *J. Abnorm. Psychol.* 128:69. doi: 10.1037/abn0000390

Jeste, S. S., and Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat. Rev. Neurol.* 10, 74–81. doi: 10.1038/nrneurol.2013.278

Karalunas, S. L., Hawkey, E., Gustafsson, H., Miller, M., Langhorst, M., Cordova, M., et al. (2018). Overlapping and Distinct Cognitive Impairments in Attention-Deficit/Hyperactivity and Autism Spectrum Disorder without Intellectual Disability. *J. Abnorm. Child Psychol.* 46, 1705–1716. doi: 10.1007/s10802-017-0394-2

Kover, S. T., McDuffie, A. S., Hagerman, R. J., and Abbeduto, L. (2013). Receptive vocabulary in boys with autism spectrum disorder: Cross-sectional developmental trajectories. *J. Autism Dev. Disord.* 43, 2696–2709. doi: 10.1007/s10803-013-1823-x

Kujala, T., Lepistö, T., and Näätänen, R. (2013). The neural basis of aberrant speech and audition in autism spectrum disorders. *Neurosci. Biobehav. Rev.* 37, 697–704. doi: 10.1016/j.neubiorev.2013.01.006

Kwok, E. Y., Brown, H. M., Smyth, R. E., and Cardy, J. O. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Res. Autism Spectr. Disord.* 9, 202–222. doi: 10.1016/j.rasd.2014.10.008

Landa, R. J., Gross, A. L., Stuart, E. A., and Bauman, M. (2012). Latent class analysis of early developmental trajectory in baby siblings of children with autism. *J. Child Psychol. Psychiatry* 53, 986–996. doi: 10.1111/j.1469-7610.2012.02558.x

Landa, R. J., Stuart, E. A., Gross, A. L., and Faherty, A. (2013). Developmental Trajectories in Children With and Without Autism Spectrum Disorders: The First 3 Years. *Child Dev.* 84, 429–442. doi: 10.1111/j.1467-8624.2012.01870.x

Landa, R., and Garrett-Mayer, E. (2006). Development in infants with autism spectrum disorders: A prospective study. *J. Child Psychol. Psychiatry* 47, 629–638. doi: 10.1111/j.1469-7610.2006.01531.x

Lanza, S. T., Bray, B. C., and Collins, L. M. (2013). An introduction to latent class and latent transition analysis. *Handb. Psychol.* 2, 691–716. doi: 10.1002/9781118133880.hop202024

LeBarton, E. S., and Iverson, J. M. (2013). Fine motor skill predicts expressive language in infant siblings of children with autism. *Dev. Sci.* 16, 815–827. doi: 10.1111/desc.12069

LeBarton, E. S., and Landa, R. J. (2019). Infant motor skill predicts later expressive language and autism spectrum disorder diagnosis. *Infant Behav. Dev.* 54, 37–47. doi: 10.1016/j.infbeh.2018.11.003

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., and Bishop, S. (2012). *Autism Diagnostic Observation Schedule*, (Second Edn). Los Angeles, CA: Western Psychological Services.

Luyster, R. J., Kadlec, M. B., Carter, A., and Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *J. Autism Dev. Disord.* 38, 1426–1438. doi: 10.1007/s10803-007-0510-1

Lyall, K., Schweitzer, J. B., Schmidt, R. J., Hertz-Picciotto, I., and Solomon, M. (2017). Inattention and hyperactivity in association with autism spectrum disorders in the CHARGE study. *Res. Autism Spectr. Disord.* 35, 1–12. doi: 10.1016/j.rasd.2016.11.011

Maenner, M. J. (2021). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR Surveill. Summ.* 70, 1–16. doi: 10.15585/mmwr.ss7011a1

Matsuzaki, J., Kuschner, E. S., Blaskey, L., Bloy, L., Kim, M., Ku, M., et al. (2019). Abnormal auditory mismatch fields are associated with communication impairment in both verbal and minimally verbal/nonverbal children who have autism spectrum disorder. *Autism Res.* 12, 1225–1235. doi: 10.1002/aur.2136

Mazurek, M. O., Kanne, S. M., and Miles, J. H. (2012). Predicting improvement in social–communication symptoms of autism spectrum disorders using retrospective treatment data. *Res. Autism Spectr. Disord.* 6, 535–545. doi: 10.1016/j.rasd.2011.07.014

McLachlan, G. J., and Rathnayake, S. (2014). On the number of components in a Gaussian mixture model. *Wires Data Min. Knowl. Discov.* 4, 341–355. doi: 10.1002/widm.1135

Messinger, D., Young, G. S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., et al. (2013). Beyond Autism: A Baby Siblings Research Consortium Study of High-Risk Children at Three Years of Age. *J. Am. Acad. Child Adolesc. Psychiatry* 52, 300–308.e1. doi: 10.1016/j.jaac.2012.12.011

Miller, L. E., Perkins, K. A., Dai, Y. G., and Fein, D. A. (2017). Comparison of Parent Report and Direct Assessment of Child Skills in Toddlers. *Res. Autism Spectr. Disord.* 41, 57–65. doi: 10.1016/j.rasd.2017.08.002

Miller, M., Austin, S., Iosif, A.-M., De La Paz, L., Chuang, A., Hatch, B., et al. (2020). Shared and distinct developmental pathways to ASD and ADHD phenotypes among infants at familial risk. *Dev. Psychopathol.* 32, 1323–1334. doi: 10.1017/S0954579420000735

Mody, M., Shui, A., Nowinski, L., Golas, S., Ferrone, C., O'Rourke, J., et al. (2017). Communication Deficits and the Motor System: Exploring Patterns of Associations in Autism Spectrum Disorder (ASD). *J. Autism Dev. Disord.* 47, 155–162. doi: 10.1007/s10803-016-2934-y

Mullen, E. M. (1995). *Mullen scales of early learning (AGS ed.).* Circle Pines, MN: American Guidance Service Inc.

Niedźwiecka, A., Domasiewicz, Z., Kawa, R., Tomalski, P., and Pisula, E. (2020). Sensory processing in toddlers with autism spectrum disorders. *Eur. J. Dev. Psychol.* 17, 527–555. doi: 10.1080/17405629.2019.1676715

Olson, L., Kinnear, M., Chen, B., Reynolds, S., Ibarra, C., Wang, T., et al. (2021). Socioeconomic Factors Account for Variability in Language Skills in Preschoolers with Autism Spectrum Disorders. *J. Dev. Behav. Pediatr.* 42, 101–108. doi: 10.1097/DBP.0000000000000870

Ozonoff, S., Iosif, A.-M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., et al. (2010). A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *J. Am. Acad. Child Adolesc. Psychiatry* 49, 256–266. doi: 10.1016/j.jaac.2009.11.009

Pecukonis, M., Plesa Skwerer, D., Eggleston, B., Meyer, S., and Tager-Flusberg, H. (2019). Concurrent Social Communication Predictors of Expressive Language in Minimally Verbal Children and Adolescents with Autism Spectrum Disorder. *J. Autism Dev. Disord.* 49, 3767–3785. doi: 10.1007/s10803-019-04089-8

Petersen, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., et al. (2013). Language ability predicts the development of behavior problems in children. *J. Abnorm. Psychol.* 122, 542–557. doi: 10.1037/a0031963

Pickles, A., Anderson, D. K., and Lord, C. (2014). Heterogeneity and plasticity in the development of language: A 17-year follow-up of children referred early for possible autism. *J. Child Psychol. Psychiatry* 55, 1354–1362. doi: 10.1111/jcpp.12269

Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., and Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Dev. Psychol.* 45, 544–557. doi: 10.1037/a0013917

R Core Team (2020). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing.

Randall, M., Sciberras, E., Brignell, A., Ihsen, E., Efron, D., Dissanayake, C., et al. (2016). Autism spectrum disorder: Presentation and prevalence in a nationally representative Australian sample. *Aust. N. Z. J. Psychiatry* 50, 243–253. doi: 10.1177/0004867415595287

Rao, P. A., and Landa, R. J. (2014). Association between severity of behavioral phenotype and comorbid attention deficit hyperactivity disorder symptoms in children with autism spectrum disorders. *Autism* 18, 272–280. doi: 10.1177/1362361312470494

Reetzke, R., Iosif, A., Hatch, B., de la Paz, L., Chuang, A., Ozonoff, S., et al. (2021a). Patterns of objectively measured motor activity among infants developing ASD and concerns for ADHD. *J. Child Psychol. Psychiatry* 63, 663–673. doi: 10.1111/jcpp.13504

Reetzke, R., Pfeiffer, D., Kalb, L. G., Holingue, C., Zetlmeisl, C., Hong, J. S., et al. (2021b). Informant (Dis) Agreement on Ratings of Challenging Behaviors and Social Communication in Preschool Children With Autism Spectrum Disorder. *J. Speech Lang. Hear. Res.* 64, 2766–2775. doi: 10.1044/2021_JSLHR-20-00707

Ripley, B., and Venables, W. (2022). *Package 'nnet.' R package version, 7.3-17*. Available online at: http://www.stats.ox.ac.uk/pub/MASS4/ (accessed June 25, 2022).

Riva, V., Cantiani, C., Mornati, G., Gallo, M., Villa, L., Mani, E., et al. (2018). Distinct ERP profiles for auditory processing in infants at-risk for autism and language impairment. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-017-19009-y

Rodas, N. V., Eisenhower, A., and Blacher, J. (2017). Structural and Pragmatic Language in Children with ASD: Longitudinal Impact on Anxiety and Externalizing Behaviors. *J. Autism Dev. Disord.* 47, 3479–3488. doi: 10.1007/s10803-017-3265-3

Saul, J., and Norbury, C. (2020). Does phonetic repertoire in minimally verbal autistic preschoolers predict the severity of later expressive language impairment? *Autism* 24, 1217–1231. doi: 10.1177/1362361319898560

Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Statist.* 6, 461–464. doi: 10.1214/aos/1176344136

Scrucca, L., Fop, M., Murphy, T. B., and Raftery, A. E. (2016). mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *R J* 8, 289–317.

Sigman, M., and McGovern, C. W. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *J. Autism Dev. Disord.* 35, 15–23. doi: 10.1007/s10803-004-1027-5

Stahmer, A. C., Dababnah, S., and Rieth, S. R. (2019). Considerations in implementing evidence-based early autism spectrum disorder interventions in community settings. *Pediatr. Med.* 2:18. doi: 10.21037/pm.2019.05.01

Tager-Flusberg, H. (2016). Risk Factors Associated with Language in Autism Spectrum Disorder: Clues to Underlying Mechanisms. *J. Speech Lang. Hear. Res.* 59, 143–154. doi: 10.1044/2015_JSLHR-L-15-0146

Tager-Flusberg, H., and Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Res.* 6, 468–478. doi: 10.1002/aur.1329

Talbott, M. R., Nelson, C. A., and Tager-Flusberg, H. (2015). Diary Reports of Concerns in Mothers of Infant Siblings of Children with Autism Across the First Year of Life. *J. Autism Dev. Disord.* 45, 2187–2199. doi: 10.1007/s10803-015-2383-z

Tallal, P., and Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends Neurosci.* 29, 382–390. doi: 10.1016/j.tins.2006.06.003

Thurm, A., Manwaring, S. S., Swineford, L., and Farmer, C. (2015). Longitudinal study of symptom severity and language in minimally verbal children with autism. *J. Child Psychol. Psychiatry* 56, 97–104. doi: 10.1111/jcpp.12285

Tick, B., Bolton, P., Happé, F., Rutter, M., and Rijsdijk, F. (2016). Heritability of autism spectrum disorders: A meta-analysis of twin studies. *J. Child Psychol. Psychiatry* 57, 585–595. doi: 10.1111/jcpp.12499

Vasa, R. A., Mazurek, M. O., Mahajan, R., Bennett, A. E., Bernal, M. P., Nozzolillo, A. A., et al. (2016). Assessment and Treatment of Anxiety in Youth with Autism Spectrum Disorders. *Pediatrics* 137, S115–S123. doi: 10.1542/peds.2015-2851J

Warlaumont, A. S., Richards, J. A., Gilkerson, J., and Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychol. Sci.* 25, 1314–1324. doi: 10.1177/0956797614531023

Whyte, E. M., and Nelson, K. E. (2015). Trajectories of pragmatic and nonliteral language development in children with autism spectrum disorders. *J. Commun. Disord.* 54, 2–14. doi: 10.1016/j.jcomdis.2015.01.001

Wodka, E. L., Mathy, P., and Kalb, L. (2013). Predictors of phrase and fluent speech in children with autism and severe language delay. *Pediatrics* 131, e1128–e1134. doi: 10.1542/peds.2012-2221

Yerys, B. E., Bertollo, J. R., Pandey, J., Guy, L., and Schultz, R. T. (2019). Attention-Deficit/Hyperactivity Disorder Symptoms Are Associated With Lower Adaptive Behavior Skills in Children With Autism. *J. Am. Acad. Child Adolesc. Psychiatry* 58, 525–533. doi: 10.1016/j.jaac.2018.08.017

Yoder, P., Watson, L. R., and Lambert, W. (2015). Value-Added Predictors of Expressive and Receptive Language Growth in Initially Nonverbal Preschoolers with Autism Spectrum Disorders. *J. Autism Dev. Disord.* 45, 1254–1270. doi: 10.1007/s10803-014-2286-4

frontiers | Frontiers in Psychology

TYPE Mini Review
PUBLISHED 14 September 2022
DOI 10.3389/fpsyg.2022.934490

OPEN ACCESS

EDITED BY
David Saldaña,
Seville University, Spain

REVIEWED BY
Sara Onnivello,
University of Padua, Italy
Vesna Stojanovik,
University of Reading, United Kingdom

*CORRESPONDENCE
Marisa G. Filipe
labfon@letras.ulisboa.pt

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 02 May 2022
ACCEPTED 15 August 2022
PUBLISHED 14 September 2022

CITATION
Filipe MG, Cruz S, Veloso AS and Frota S (2022) Early predictors of language outcomes in Down syndrome: A mini-review. *Front. Psychol.* 13:934490. doi: 10.3389/fpsyg.2022.934490

Early predictors of language outcomes in Down syndrome: A mini-review

Marisa G. Filipe[1]*, Sara Cruz[2], Andreia S. Veloso[3] and Sónia Frota[1]

[1]Center of Linguistics, School of Arts and Humanities, University of Lisbon, Lisbon, Portugal, [2]Universidade Lusófona do Porto, Porto, Portugal, [3]Faculty of Psychology and Education Sciences, Center for Psychology at University of Porto, University of Porto, Porto, Portugal

As children with Down syndrome (DS) typically manifest significant delays in language development, the research has pointed out the predictors of later language skills for this clinical population. The purpose of this study was to systematically explore the evidence for early predictors of language outcomes in infants and toddlers with DS from studies published between 2012 and 2022. After the search, nine studies met the inclusion criteria. The results indicated that maternal educational level, adaptive level of functioning, cognitive function, attention skills, communicative intent of the child, early vocalizations, gestures, baby signs, parents' translation of their children's gestures into words, and vocabulary level are significant predictors of language outcomes in children with DS. These findings provide a timely and warranted summary of published work that contributes to current understanding of the development of language and communication in DS. They are therefore useful to researchers, clinicians, and families.

KEYWORDS

Down syndrome, systematic review, early predictors, language, impairments

Introduction

As language is crucial for learning and academic achievement (Johnson et al., 2010; Conti-Ramsden et al., 2018; Eadie et al., 2021), the development of language skills is essential to meet the increasing demands of modern societies (Duncan et al., 2007). Indeed, research has shown that children with low language abilities are at high risk of difficulties with literacy, academic achievement, and social-emotional and behavioral adjustment (Voci et al., 2006; Zubrick et al., 2007; Tromblin, 2008; Durkin and Conti-Ramsden, 2010; Johnson et al., 2010; Conti-Ramsden et al., 2018). Thus, research on language development is particularly useful.

Previous studies have identified typical trajectories for language development. For example, at the age of 10–12 months, children can discriminate phonemes in their native language (for a review, refer to Kuhl, 2010), begin to understand and utter words, and produce representational and deictic gestures (Fenson et al., 1994; Caselli et al., 2012). At 18 months, typically developing (TD) children reach a lexical repertoire of approximately 50 words and use gesture–word combinations frequently

(Fenson et al., 1994; Capirci and Volterra, 2008). Between 20 and 24 months, they increase expressive vocabulary and start to combine words (Fenson et al., 1994; Capirci and Volterra, 2008). Children at the age of 3 years have been found to produce a more complex lexicon, as well as utterances that are grammatically more accurate and richer (for a review, see Guasti, 2017).

Identification of these typical language trajectories is important as many children can experience language delays (Reilly et al., 2007; Zubrick et al., 2007), as a result of biological, cognitive, and environmental factors (Kuhl, 2010; Perani et al., 2011; Riva et al., 2017). In fact, several children diagnosed with neurodevelopmental disorders have language specificities and may later be diagnosed with language impairments. For example, children with Down syndrome (DS; which results from a partial or complete duplication of chromosome 21; Epstein, 1986) display a complex neurocognitive profile including particular patterns of language skills that are characterized by relative strengths and weaknesses. On the one hand, receptive vocabulary (Laws et al., 2015) and the use of gestures (Iverson et al., 2003) appear as relative strengths in the language profile of children with DS. But on the other hand, children with DS frequently display severe language difficulties (Abbeduto et al., 2007a) and are less likely to accompany prelinguistic communicative gestures with vocalizations when compared to TD peers matched by their sensorimotor development (Greenwald and Leonard, 1979). Children with this clinical condition also tend to produce their first words at approximately 21 months (Stoel-Gammon, 2001) in line with their cognitive abilities (Miller, 1999), and expressive language abilities can be delayed when compared to receptive language and non-verbal skills (Chapman and Hesketh, 2000; Abbeduto et al., 2007b). Furthermore, in DS, the development of word segmentation competencies is seriously compromised (Mason-Apps et al., 2018), infants with DS do not use prosody as a facilitator for word segmentation unlike TD infants (Frota et al., 2020), reduced speech intelligibility is common (Kumin, 1994; Kent and Vorperian, 2013), and more substantial delays in expressive syntax than in expressive vocabulary have been reported (Kover et al., 2012). Longitudinal studies have also suggested that vocabulary development in DS is slower compared to the language development of TD peers, which, in turn, seems to be related to general cognitive abilities (Cuskelly et al., 2016; Kaat-van den Os et al., 2017).

To understand how different variables impact development and predict which children are most likely to have language impairments, researchers are identifying early predictors of language trajectories in different subgroups of community cohorts (McKean et al., 2017). In fact, several environmental and child-related factors associated with language delays or impairments have been found, such as male gender, prematurity, low birth weight, perinatal disorders, low income, and low parental education (Nelson et al., 2006; Sansavini et al., 2010; Snowling et al., 2016; Bishop et al., 2017). Other variables may also predict language outcomes in typical and atypical development. For instance, non-verbal requesting is a longitudinal predictor of expressive language development (e.g., Mundy et al., 1995) and prelinguistic communication reveals children's readiness to acquire language while eliciting language-facilitating responses from parents (Yoder and Warren, 1993; Yoder et al., 1998). Auditory and visual processing in early speech perception has also been shown to be crucial to language outcomes (Friederici, 2006; Kuhl et al., 2008; Ellis et al., 2015), affecting speech segmentation, word learning, and phrase-level processing.

Regarding DS, it is often suggested that the same environmental and child-related predictors found for TD children apply to children with this condition (e.g., Deckers et al., 2019). Indeed, previous studies found that gestures predict language development in children with typical development and DS (Capone and McGregor, 2004; Rowe et al., 2008; Zampini and D'Odorico, 2009). In addition, (i) the use of gestures at 24 and 36 months of age has been shown to predict future vocabulary growth (Zampini and D'Odorico, 2009), (ii) early prosodic development predicted lexical development in similar ways for infants and toddlers with typical development, at-risk for language impairments, or with DS (Sousa et al., 2022), (iii) babbling correlated with later language development (Locatelli et al., 2021) in line with previous studies on TD children (e.g., Lang et al., 2019), and (iv) the relationship between motor and language development was found to become stronger as the age of children increases (Yamauchi et al., 2019), a pattern that is also consistent with findings for TD children (e.g., Alcock and Krawczyk, 2010).

However, research has also suggested that different variables might predict language development in children with DS. Mason-Apps et al. (2018) showed that (i) non-verbal mental skills were the significant longitudinal predictors of language for infants with DS but not for TD infants, (ii) speech segmentation abilities only predicted language outcomes in the TD group, and (iii) while initiating joint attention was critical for TD participants, response to joint attention was more predictive of language scores in infants with DS than in TD participants. Indeed, research has shown important differences in early visual attention abilities and audiovisual speech processing in infants with DS compared to typically developing infants (D'Souza et al., 2016; Pejovic et al., 2021).

As several predictors of language outcomes have been reported in children with DS, the aim of this study was to systematically review the articles that focus on early precursors of language in infants with this genetic condition. We will focus on early predictors that appear before 30 months of age, given the potential of early screening to identify children at risk of developing language difficulties in the first 2 years of life (Määttä et al., 2012). Understanding these early predictors of language variability is important to determine the factors that explain why some children with DS acquire language before others

(Sameroff and Chandler, 1975). This could also contribute to the development of an early intervention that facilitates language learning in young children, which is strongly recommended due to the link between language skills and later development (e.g., Luyster et al., 2007).

Methods

This study adopted the method of a systematic review, as required by the Cochrane Collaboration and the PRISMA framework (Moher et al., 2009). In March 2022, using EBSCOhost, the following databases were searched: Academic Search Complete, APA PsyArticles, ERIC, MEDLINE, ScienceDirect, and Psychology and Behavioral Sciences Collection. The keywords *language* AND *longitudinal* OR *prospective* AND *down syndrome* OR *trisomy 21* OR *down's syndrome* were used to conduct the search. The following filters were applied: (i) publication date from 2012 to 2022, (ii) academic journals, and (iii) peer-reviewed. All titles/abstracts identified in the electronic databases were independently screened for eligibility by two authors (MF and SC), according to the following inclusion criteria:

- The study was published in a peer-reviewed journal from 2012 to 2022.
- Participants were followed for a period of 3 months or more in a prospective cohort study.
- The study design was experimental or observational.
- The report presented at least one early (collected before the first 30 months of age) and a later language measure/outcome.

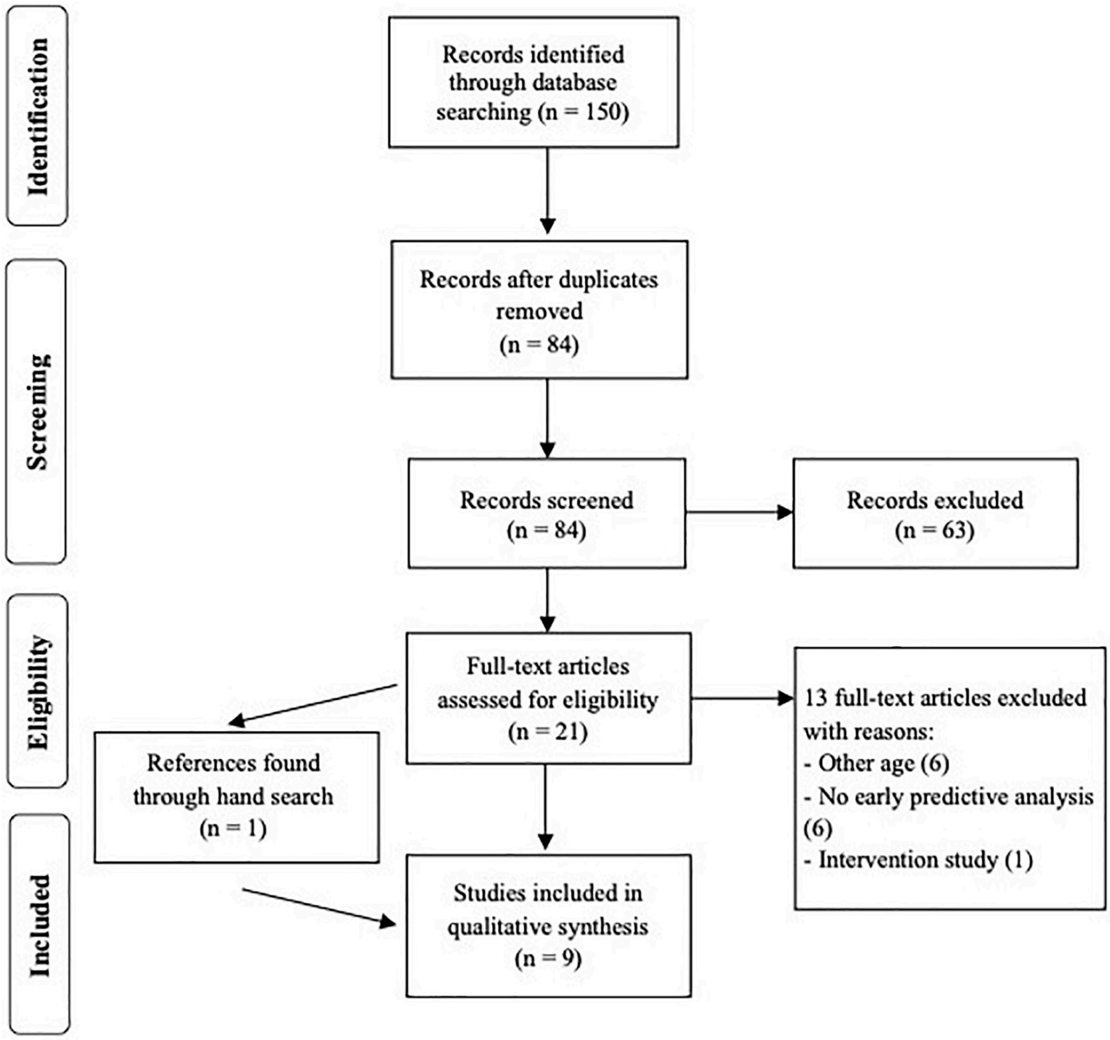

FIGURE 1
Preferred reporting items in systematic reviews and meta-analyses (PRISMA) flow diagram of the study selection procedure.

- The subsequent result(s) should include at least one quantitative measure to compare the findings across the studies.
- The report was written in English.

The search identified 150 articles. After the removal of duplicates, if the title and abstract suggested that the study may be appropriate for inclusion, the full-text article was evaluated according to the previously established inclusion criteria. A total of 21 articles were selected for full-text review. Hand searches, which included checking the reference lists of the included journal articles, identified another paper which was also read in full. A total of nine studies were included in the mini-review. Percentage agreement on the selection of included studies was 95.51%. Percentage agreement after consensus building was 100%. The selection of studies is depicted in **Figure 1** in a PRISMA flow diagram (Moher et al., 2009). The list of excluded studies along with reasons for exclusion are presented in **Supplementary material**.

From each eligible study, the following data were extracted: first author name, publication date, study location, primary language, number of participants, age at intake, time to follow-up, language predictors, language predictor measures, language outcomes, language outcome measures, main findings, and effect sizes.

Results

A summary of each study characteristics is presented in **Table 1**. The sample sizes of children with DS ranged from 5 to 26 participants. Almost half of the studies were conducted in the United States, and the remaining studies were carried out in different countries including the Netherlands, Sweden, Italy, and the United Kingdom. Age at intake varied between 10 and 84 months, with follow-ups conducted 6–53 months later.

Several predictors of language outcomes were evaluated, namely, socioeconomic status, general cognitive function, developmental level, adaptative level of functioning, auditory working memory, attention skills, joint attention, behavioral and emotional problems, temperament, auditory discrimination, number of communication partners, level of communicative intent, book reading experiences, parents' translations of child gestures, gestures, signs, initiation of behavioral requests, speech segmentation, consonant use, vocabulary, and phonological/phonemic awareness (cf. **Table 1**).

The following language outcomes were evaluated: consonant production, functional intelligibility, auditory comprehension, expressive communication, referents later expressed in speech, receptive and expressive vocabulary, vocabulary growth, and receptive and expressive language. Language measures varied between the studies. The MacArthur Communicative Development Inventories and the Communication Play Protocol were the most common language measures employed (cf. **Table 1**).

The results showed that most of the language outcomes were related to vocabulary. Regarding language predictors, adaptive level of functioning, vocabulary skills, maternal educational status, level of communicative intent of the child, attention skills, phonological/phonemic awareness (Deckers et al., 2019), parents' translation of their children's gestures into words (Dimitrova et al., 2016), baby signs (Özçaliskan et al., 2016), general cognitive function (Kaat-van den Os et al., 2017; including non-verbal mental ability: Mason-Apps et al., 2018), and joint attention (Zampini et al., 2015; responding to join attention: Mason-Apps et al., 2018) were the significant predictors of vocabulary skills. Non-verbal mental ability and responding to join attention were also the predictives of auditory comprehension (Mason-Apps et al., 2018). Furthermore, a significant positive correlation was found between the age at which a child expressed referents uniquely in gesture and the mean age they were expressed later in speech (Özçalışkan et al., 2017). Finally, a high number of different true consonants at early ages was associated with a higher consonant production measured at follow-up (Nyman et al., 2021).

Discussion

This review contributed to a better understanding of early predictors (before 30 months of age) of language outcomes in children with DS. This enhances our theoretical understanding of language development by revealing the factors that underpin language acquisition. Identifying language predictors is critical to promote the early identification of individuals with language impairments. In general, the studies included in the review show that most children with DS make positive language gains that are evident in vocabulary measurements. Although it is difficult to draw strong conclusions based on the limited evidence available, it is becoming increasingly clear that early predictors of later language development may be present in the first 30 months of life. Based on the results of this review, the predictors of language outcomes in DS will be discussed in the following paragraphs.

Between 2 and 7 years of age, maternal educational level appears to be a predictor of later expressive vocabulary in DS (Deckers et al., 2019). Indeed, previous research had also suggested that mothers of TD children from a higher socioeconomic status used longer utterances and a more diverse vocabulary when talking to their toddlers, which was associated with greater vocabulary growth (Hoff, 2003).

Evidence for the adaptive level of functioning (i.e., the child's level of participation in daily tasks involving conceptual, social, and practical skills) was also found as an early predictor

TABLE 1 Characteristics of the included studies.

Authors (year) location, language	Participants (n)	Age at intake	Age at follow-up	Measures and predictors of language development	Measures and language outcomes	Main findings
Deckers et al. (2019) Netherlands, Dutch	DS: *N* =20	2.0–7.0 years	+ 1.6 years or 18 months	**Measure:** Vineland Screener **Predictor:** Adaptive level of functioning **Measure:** Subscale Working Memory from the Behavior Rating Inventory of Executive functions—Preschool version **Predictor:** Working memory **Measure:** Child Behavior Checklist 1.5–5 **Predictor:** Behavioral and emotional problems, attention distractibility and temperament **Measure:** The Bridge: Emergent literacy skills **Predictor:** Book reading experiences and phonological/phonemic awareness **Measure:** Social Networks Questionnaire **Predictor:** Number of communication partners **Measure:** Sociodemographic Questionnaire **Predictor:** Socioeconomic status, chronological age of the child, siblingship size, educational level, and involvement of the child **Measure:** Receptive One-word Picture Vocabulary Test **Predictor:** Receptive vocabulary **Measure:** Auditory Discrimination Task **Predictor:** Auditory discrimination **Measure:** Auditory Working Memory Test **Predictor:** Auditory working memory **Measure:** Communicative Intentive Onderzoek **Predictor:** Communicative intent, joint attention and parental support and responsiveness	**Measure:** MacArthur Communicative Development Inventories **Outcome:** Expressive vocabulary **Measure:** Receptive One-word picture Vocabulary Test **Outcome:** Receptive vocabulary	• Expressive vocabulary development was best predicted by the adaptive level of functioning (R^2 = 0.80; p = 0.01), receptive vocabulary (R^2 = 0.73; p = 0.001), maternal educational level (R^2 = 0.42; p = 0.01), level of communicative intent of the child (R^2 = 0.53; p = 0.01), attention skills (R^2 = 0.63; $p < 0.05$), and phonological/phonemic awareness (R^2 = 0.69; p = 0.01). • Receptive vocabulary development was best predicted by the adaptive level of functioning (R^2 = 0.88; p = 0.001) and early receptive vocabulary skills (R^2 = 0.84; p = 0.001).
Dimitrova et al. (2016) USA, English	TD: n = 23 Autism: n = 23 DS: n = 23	TD: 18–30 months Autism: 31–43 months DS: 30–45 months	± 12 months	**Measure:** Communication Play Protocol **Predictor:** Parents' translations of child gesture	**Measure:** Communication Play Protocol **Outcome:** Expressive vocabulary development	• Parents translate a high percentage of their children's gestures into words, and this input was beneficial for children in each group as they acquire more words for the translated gestures than the not translated ones. Translation: $F(1, 63) = 5.97$, p = 0.02, v^2_p = 0.09 . Group: $F(2, 63) = 8.01$, p = 0.001, v^2_p = 0.20 . Group × Translation: $F(2, 63)$ = 0.05, p = 0.95 • This benefit on child vocabulary development was particularly evident for children who show evidence of vocabulary growth over time. . Translation: $F(1, 45) = 6.63$, p = 0.013, v^2_p= 0.13 . Group: $F(2, 45) = 6.54$, p = 0.003, v^2_p = 0.23 . Group × Translation: $F(2, 45)$ = 0.30, p = 0.743

(Continued)

TABLE 1 (Continued)

Authors (year)location, language	Participants (n)	Age at intake	Age at follow-up	Measures and predictors of language development	Measures and language outcomes	Main findings
						• The use of these spoken labels had the same facilitative effect on vocabulary development for children with TD and DS.
Kaat-van den Os et al. (2017) Netherlands, Dutch	DS: $N = 26$	18–24 months	Monthly assessments over an 18-month period	**Measure:** Cognition Scale of the Bayley Scales of Infant and Toddler Development, Third Edition **Predictor:** General cognitive function	**Measure:** Lexi Questionnaire **Outcome:** Expressive vocabulary growth and modality (gesture and/or verbal production)	• Three patterns of vocabulary growth were identified: children with a marginal vocabulary growth, children with an increase in vocabulary without a growth spurt, and children who showed a vocabulary growth spurt. • All groups significantly differed in the rate of vocabulary growth. . Growth spurt (GS): $M = 56.2$, $SD = 52.9$. Without growth spurt (WGS): $M = 3.9$, $SD = 2.9$. Marginal growth pattern (MGP): $M = 1.1$, SD = 0.6 - GS vs. WGS: $p < 0.05$ - WGS vs. MGP: $p < 0.01$ • The general cognitive function of the children with a marginal growth pattern was significantly lower than that of the children in the groups with a substantial increase in vocabulary or vocabulary spurt. . GS: $M_{age} = 19$. WGS: $M_{age} = 18.5$. MGP: $M_{age} = 15.9$ - GS vs. MGP: $p < 0.05$ - WGS vs. MGP: $p < 0.05$ • The general cognitive function of the groups with or without a growth spurt did not differ significantly. • Correlation showed that the rate of vocabulary growth was significantly correlated with the general cognitive function ($r = 0.44$, $p < 0.05$).
Mason-Apps et al. (2018) United Kingdom, English	DS: $n = 14$ TD: $n = 35$	10–19 months	Measures collected at two time points, approximately 6 and 12 months apart from intake	**Measure:** Mullens Scales of Early Learning **Predictor:** Non-verbal mental ability **Measures:** Strong-Weak Task (to assess infants' ability to segment bisyllabic words with a strong-weak stress pattern) and Weak-Strong Task (to assess the ability to segment bisyllabic words with a weak-strong stress pattern) **Predictor:** Speech segmentation skills **Measure:** Early Social Communication Scales **Predictor:** Social communication skills (initiating and responding to joint attention; initiating behavioral requests)	**Measure:** Preschool Language Scales-4 **Outcome:** Auditory comprehension and expressive communication **Measure:** Reading Communicative Development Inventory **Outcome:** Receptive and expressive vocabulary	• In the TD group, speech segmentation and initiating joint attention were the strongest predictors of later language. . Speech segmentation (SS; T1) × expressive communication (EC; T2): $r = 0.701$, $p \leq 0.001$. SS (T1) × expressive vocabulary (EV; T2): $r = 0.553$, $p \leq 0.01$. Initiating joint attention (IJA; T1) × expressive communication (EC; T2): $r = 0.490$, $p \leq 0.05$. IJA (T1) × EV (T2): $r = 0.402$, $p \leq 0.05$ - Regression analysis (EC, SS, IJA, age): $F(4, 15) = 18.17$, $p < 0.001$, Adj$R^2 = 0.783$ - Regression analysis (EV, SS, IJA, age): $F(3, 18) = 5.68$, $p = 0.006$, Adj$R^2 = 0.401$

(Continued)

TABLE 1 (Continued)

Authors (year)location, language	Participants (n)	Age at intake	Age at follow-up	Measures and predictors of language development	Measures and language outcomes	Main findings
						. SS (T1) × auditory comprehension (AC; T3): $r = 0.498$, $p \leq 0.05$. SS (T1) × EC (T3): $r = 0.685$, $p \leq 0.001$. SS (T1) × receptive vocabulary (RV; T3): $r = 0.565$, $p \leq 0.05$. SS (T1) × EV (T3): $r = 0.827$, $p \leq 0.001$. IJA (T1) × EV (T3): $r = 0.413$, $p \leq 0.05$ - Regression analysis (EC, SS, age): $F(3, 17) = 7.04$, $p = 0.003$, Adj$R^2 = 0.475$ • In the DS group, non-verbal mental ability and responding to joint attention were the strongest predictors of later language. . Non-verbal mental ability (NVMA; T1) × AC (T2): $r = 0.862$, $p \leq 0.001$. NVMA (T1) × Receptive vocabulary (RV; T2): $r = 0.855$, $p \leq 0.01$. Non-verbal mental ability (NVMA; T1) × RV (T3): $r = 0.871$, $p \leq 0.001$. Responding to JA (RJA; T1) × AC (T3): $r = 0.614$, $p \leq 0.01$. RJA (T1) × EC (T3): $r = 0.812$, $p \leq 0.001$. RJA (T1) × RV (T3): $r = 0.629$, $p \leq 0.05$. RJA (T1) × EV (T3): $r = 0.656$, $p \leq 0.05$ - Regression analysis (NVMA, RJA, RV, age): $F(4, 7) = 12.662$, $p = 0.003$, Adj$R^2 = 0.809$ - Regression analysis (EC, RJA, age): $F(1, 10) = 11.906$, $p = 0.002$, Adj$R^2 = 0.645$ • Non-verbal mental skills were a significant longitudinal predictor of language for infants with DS but not for TD infants, speech segmentation abilities only predicted language outcomes in the TD group, and while initiating joint attention was critical for TD participants, response to joint attention was more predictive of language scores in infants with DS than in TD participants.
Nyman et al. (2021) Sweden, Swedish	DS: $n = 5$ Cerebral palsy (CP): $n = 4$ Chromosomal deletion syndromes: $n = 2$	12–22 months	4:11–5:4 years	**Measure:** Audio-video recordings of parent–child interaction, using a standardized procedure and set of toys. A babbling observation was performed, and the occurrence of different babbling variables was noted using an observation form containing a list of all 18 Swedish consonant sounds. **Predictor:** Consonant use	**Measure:** Test for Reception of Grammar-2 or Reynell Developmental Language Scales-III **Outcome:** Receptive language **Measure:** The five longest utterances for each child were identified based on all spontaneous communication. Mean maximum utterance length was calculated by taking the five longest utterances, adding up the number of words and dividing it by five **Outcome:** Expressive language **Measure:** Expressive Vocabulary and Sentence Recall from the Clinical Evaluation of Language Fundamentals-4 **Outcome:** Expressive language **Measure:** Swedish Communicative Development Inventory III or Swedish Communicative Development—words and gestures **Outcome:** Number of words the child understands and produces	• Children with DS performed lower than participants with other types of neurological disabilities on two consonant production measures of the Swedish Articulation and Nasality Test. . Percentage of consonants correct (PCC): DS vs. CP: $U = 0$, $p = 0.016$. Number of established consonants: DS vs. CP: $U = 1.5$, $p = 0.032$ • However, participants with DS who used a high number of different true consonants at the first assessment also had higher consonant production measured at the follow-up. . Correlation (n true consonants at T1 × PCC at T2): $r_s = 0.553$, $p = 0.077$. Correlation (n true consonants at T1 × PCC at T2 – DS subgroup analysis): $r_s = 0.894$, $p = 0.041$

(Continued)

TABLE 1 (Continued)

Authors (year)location, language	Participants (n)	Age at intake	Age at follow-up	Measures and predictors of language development	Measures and language outcomes	Main findings
					Measure: Swedish Articulation and Nasality Test **Outcome:** Consonant Production **Measure:** Presence of motor speech disorder was assessed based on the audio and video recorded articulation test **Outcome:** Presence of motor speech disorder **Measure:** Intelligibility in Context Scale **Outcome:** Functional intelligibility	
Özçalışkan et al. (2017) USA, English	DS: $n = 23$ TD: $n = 23$ Autism (ASD): $n = 23$	DS: 30 months TD: 18 months ASD: 30 months	5 times over a year	**Measure:** Communication Play Protocol **Predictor:** Referents expressed uniquely in gesture	**Measure:** Communication Play Protocol **Outcome:** Referents later expressed in speech	• A significant positive correlation was found between the age at which a child expressed referents uniquely in gesture and the mean age they were expressed later in speech across the three groups and within each group. . Correlation (across all groups): $r = 0.93$, $p < 0.001$. Correlation (ASD): $r = 0.87$, $p < 0.001$. Correlation (DS): $r = 0.81$, $p < 0.001$ • Most of the referents conveyed uniquely in gesture entered children's spoken vocabularies as words for both TD children and children with autism within a year. This pattern was less pronounced for children with DS, who differed significantly from both groups. . Modality shift from gesture to speech: $F(1, 63) = 4.46$, $p = 0.04$, $\eta^2_p = 0.07$. Interaction between group and modality shift: $F(2, 63) = 6.45$, $p = 0.003$, $\eta^2_p = 0.17$ • The time interval from when a referent was observed in gesture and its observation in speech was longer for DS compared to TD. . Timing of the modality shift from gesture to speech: - modality: $F(1, 48) = 427.92$, $p < 0.001$, $\eta^2_p = 0.90$ - group: $F(2, 48) = 92.36$, $p < 0.001$ $\eta^2_p = 0.79$ - interaction between group and modality: $F(2, 48) = 9.52$, $p < 0.001$, $\eta^2_p = 0.28$
Özçaliskan et al. (2016) USA, English	DS: $n = 23$ TD: $n = 23$	DS: 2.6 TD: 1.6	+ 12 months	**Measure:** Communication Play Protocol **Predictor:** Gestures and signs (deictic, conventional, iconic)	**Measure:** Previously transcribed transcripts **Outcome:** Spoken vocabulary **Measure:** Expressive Vocabulary Test **Outcome:** Vocabulary size	• For children with DS, the production of baby signs predicted expressive vocabulary size 1 year later (*Spearman's rho* = 0.60, $p = 0.005$). Neither deictic nor conventional gestures produced by children with DS had a significant relation to later spoken vocabulary. • Deictic gestures reliably predicted expressive vocabulary size for TD children (*Spearman's rho* = 0.64, $p = 0.002$), while baby signs were positively related to later vocabulary of children with DS.

(Continued)

TABLE 1 (Continued)

Authors (year)location, language	Participants (n)	Age at intake	Age at follow-up	Measures and predictors of language development	Measures and language outcomes	Main findings
Zampini and D'Odorico (2013) Italy, Italian	DS: $N = 18$	Ten 2-year-old children Eight 3-year-old children	2-year-old children were followed for a 2-year period 3-year-old children were followed for a 1-year period	**Measure:** MacArthur–Bates Communicative Development Inventories (production checklist) **Predictor:** Vocabulary size **Measure:** Brunet–Lézine Scale of Psychomotor Development **Predictor:** Developmental level	**Measure:** MacArthur–Bates Communicative Development Inventories (production checklist) **Outcome:** Lexical outcomes	• Only at 36 and 42 months could vocabulary size explain individual differences on subsequent lexical development at 48 months, and only at 42 and 48 months could developmental age explain the variability in children's lexical outcomes. . Lexical outcomes at 48 months and first stages of vocabulary acquisition: - 36 months × low outcome group × medium outcome group × high outcome group: $K = 12.97$, $p = 0.002$ - 42 months × low outcome group × medium outcome group × high outcome group: $K = 15.05$, $p = 0.001$. Individual differences in children's developmental ages and children's lexical outcomes: - 42 months × low outcome group × medium outcome group × high outcome group: $K = 7.67$, $p = 0.022$ - 48 months outcome group × low outcome group × medium outcome group × high outcome group: $K = 9.08$, $p = 0.011$
Zampini et al. (2015) Italy, Italian	DS: $N = 18$	24 months	30 months	**Measure:** Semi-structured free-play sessions in interaction with their mothers **Predictor:** Joint attention	**Measure:** MacArthur-Bates Communicative Development Inventory **Outcome:** Vocabulary development (both receptive and expressive)	• The children's behavior of proposing a joint attention focus to their communicative partners appeared to be a significant predictor of the children's vocabulary comprehension skills as assessed 6 months later. . Total amount of time spent in joint attention and word comprehension: $r = 0.577$, $p = 0.024$. Regressions: - Word comprehension at 24 months: $F(1, 16) = 60.11$, $p < 0.001$, $R^2 = 0.79$, Adj$R^2 = 0.78$ - Word comprehension at 24 months + joint attention propose + joint attention follow: $F(2, 15) = 41.07$, $p < 0.001$, $R^2 = 0.85$, Adj$R^2 = 0.83$

of expressive and receptive vocabulary in children with DS, between 2 and 7 years of age (Deckers et al., 2019). This highlights that language development and the adaptative level of functioning might be interrelated. Indeed, previous studies with individuals with DS highlighted stronger skills in daily living activities and socialization compared with the relative weaknesses in motor and communication skills (e.g., Van Duijn et al., 2010). Probably, children with DS who are more likely to show social competence will elicit more reactions from communication partners, experience different social contexts, and learn more different words, while the use of language to communicate may in turn increase the ability to manage social situations. However, it is important to highlight that language development and the adaptative level of functioning are interrelated and growth in one skill might affect the functioning of the other.

We also found that, before 2 years of age, cognitive domains such as non-verbal mental ability (Mason-Apps et al., 2018), play skills, information processing, memory, habituation skills, and reasoning abilities (named by the authors as general cognitive function; Kaat-van den Os et al., 2017) predict vocabulary growth in DS. Furthermore, non-verbal mental ability was also found to predict auditory comprehension (Mason-Apps et al., 2018). Although research has shown that language outcomes in DS are not merely a result of a cognitive disability (e.g., Dodd and Thompson, 2001), the studies included in this review highlighted that several cognitive skills predicted language outcomes showing a clear link between cognitive skills and language learning. This is not surprising since domain-general abilities apply across different kinds of tasks (Federmeier et al., 2020).

A finding that is also evident in the present review is that attention skills found to predict language outcomes in TD children were also visible in children with DS. Namely, at 19 and 24 months of age (respectively, Zampini et al., 2015; Mason-Apps et al., 2018), joint attention predicted language outcomes, and between 2 and 7 years of age, attention skills predicted expressive vocabulary (Deckers et al., 2019). These results are in line with the previous studies for typically developing children. For instance, in TD 1-year-olds, the effect of maternal education and warm parenting on vocabulary growth was found to be mediated by attention skills and parent–child book reading when the children completed 3 years of age (Farrant and Zubrick, 2012). Furthermore, in TD individuals, higher attention demands negatively affect the aspects of spoken vocabulary (Hula et al., 2007). Thus, attention skills are important for language development in TD and in DS, probably because children with greater attention skills may be more likely to experience more opportunities for language learning.

We also found that, between 2 and 7 years of age, the level of communicative intent could be a predictor of later expressive vocabulary for children with DS (Deckers et al., 2019). Indeed, previous research has reported a result along similar lines for TD toddlers, showing that the level of communicative intent is a predictor of later language outcomes (Wetherby et al., 2002). Higher rates of communication could increase the opportunities for interaction and shape communication development (McCathren, 2000). For example, Yoder et al. (1994) showed that mothers provided more verbal modeling when children have a higher communicative intent.

Our findings also highlighted that consonant measures might be useful in evaluating toddlers with DS, namely, the number of true consonants assessed from 12 to 22 months of age might predict later consonant production (Nyman et al., 2021). A continuity between early vocalizations and language outcomes in atypical and typical development has been suggested in the literature. For instance, canonical babbling (which consists of consonant-vowel-syllables with a rapid transition between them) is commonly used in the study of early vocalizations in children at risk of language difficulties (Nyman et al., 2021). For TD, the early consistent use of consonants has also been associated with better expressive vocabulary (McGillion et al., 2017).

Children with DS are as likely as TD children to point to and request objects using gestures prior to using words, and our review highlighted that, at 30 months of age, the onset of referents expressed uniquely in gestures could predict the onset of similar spoken words (Özçaliskan et al., 2016). Also, at 1 year of age, parents' translation of children's gestures into words might predict later vocabulary development (Dimitrova et al., 2016). This is in line with what previous findings have suggested that parents gather information from the gestures their children produce and tailor their verbal responses to the communicative interests of the child (Goldin-Meadow et al., 2007; Olson and Masur, 2011). These parents' translations of child gestures could help the child to map the word to the object of interest conveyed in gesture. Thus, children's gestures probably provide cues to the parents about the child's readiness to learn a particular word (Dimitrova et al., 2016).

Also related to gestures, an important finding is that baby signs (i.e., iconic or arbitrary signs intentionally taught by adults) at 2.5 years of age may be positively related to later vocabulary outcomes in children with DS (Özçaliskan et al., 2016). Baby signs are learned in the everyday context when a parent produces signs to refer to a particular object. The use of these repeated signed symbols might create a state symbol stand for objects (DeLoache, 2004) that could help children with DS to move from a repertoire of signed symbols to a repertoire of words. Thus, findings from this review seem especially significant considering current knowledge about the importance of early non-verbal communicative skills for the prediction of later language outcomes.

Finally, our results showed that a particularly important behavioral domain is the use of vocabulary skills as a key precursor to language development. Deckers et al. (2019) found that receptive vocabulary, between 2 and 7 years of age, was a predictor of later expressive and receptive vocabulary. A similar conclusion was reached by other studies. For instance, in children with DS, early receptive vocabulary skills tend to be a predictor of receptive and expressive vocabulary (Chapman et al., 2000; Chiat and Roy, 2008). However, Zampini and D'Odorico (2013) assessed children with DS from 2 years of age and showed that individual differences at 48 months could be explained by vocabulary size only at 36 and 42 months.

It seems that some predictors had the same facilitative effect for TD children and children with DS, such as the parents' translation of gestures into words (Dimitrova et al., 2016). However, our review also emphasized that early predictors of language outcomes might be different for the two groups: (i) the time interval from when a referent was observed in gesture and its observation in speech was longer for DS compared to TD (Özçalışkan et al., 2017); (ii) deictic gestures reliably predicted expressive vocabulary size for TD children, but it was baby signs (and not deictic gestures) that predicted expressive vocabulary development for children with DS (Özçaliskan et al., 2016); (iii) non-verbal mental skills predicted language for infants with DS but not for TD children (Mason-Apps et al., 2018); (iv) speech segmentation abilities predicted language outcomes only for TD children (Mason-Apps et al., 2018); and (v) response to joint attention was more predictive of language outcomes in children with DS than in TD peers (Mason-Apps et al., 2018).

This review offers systematic information for researchers, families, and clinicians on language development over time and on language outcomes for individuals with DS. Further research should focus on the yet to be fully studied early predictors of language impairments, and the association between early and later outcomes in DS must be confirmed in larger cohorts. Furthermore, to attain the goal of identifying predictors of language and communication impairments in DS, future studies should combine a set of innovative features, as proposed, for example, within the Predictors of Language Outcomes Project (PLOs)[1]: (1) inclusion of early measures and later assessments of language abilities for several at-risk groups for language impairments enabling cross-group comparisons; (2) multimethodology approach to a set of potential early predictors of later language outcomes, which combines quantitative and qualitative measures but also other non-invasive methods such as eye gaze, eye tracking, and brain measures; and (3) examination of several language domains at the word and phrase levels (e.g., stress discrimination, word learning, and intonation). This will offer a timely opportunity to promote more effective methods of screening, prevention, early intervention, and diagnosis of language impairments.

In sum, this systematic review shows that there are only a few comprehensive studies that have explored key early predictors of later language acquisition in DS. Although it is difficult to draw strong conclusions based on the relatively limited evidence available, it is becoming increasingly clear that predictors of later language development could be evident in the 5 years of life. Overall, this review confirms that both child-related factors (e.g., maternal education) and prelinguistic communication could predict later language for infants with DS. One important behavioral domain that has received particular attention as a key precursor to language for this clinical population is non-verbal communicative skills such as gestures and signs, together with early vocabulary measures. Furthermore, domain-general processes such as non-verbal cognitive skills have been shown to account for some variations in later language outcomes. However, more studies are needed to identify which factors are the most robust predictors of language development for children with DS, and whether these predictors differ between different clinical populations. A better understanding of the developmental factors that underlie, facilitate, and predict language acquisition in DS would shed light on the nature of this disorder and allow the refinement of targeted early interventions. Such an endeavor would be very relevant for policymakers and service providers to support individuals with DS throughout their lives.

1 http://labfon.letras.ulisboa.pt/babylab/PLOS/en/

Author contributions

MF and SF contributed to the conception and design of the work. MF and SC reviewed the abstracts and the manuscript. MF and AV obtained the data from the selected articles. MF prepared the first draft of the manuscript. SF and AV revised the manuscript critically for important intellectual content. MF, AV, and SF revised the last version of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the Portuguese Foundation for Science and Technology (2020.01866.CEECIND; PTDC/LLT-LIN/1115/2021) in conjunction with the European Regional Development Fund from the EU, Portugal 2020, and Lisboa 2020 (PTDC/LLT-LIN/29338/2017).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.934490/full#supplementary-material

References

Abbeduto, L., Brady, N., and Kover, S. T. (2007a). Language development and Fragile X Syndrome: Profiles, syndrome-specificity, and within-syndrome differences. *Ment. Retard. Dev. Disabil. Res. Rev.* 13, 36–46. doi: 10.1002/mrdd.20142

Abbeduto, L., Warren, S. F., and Conners, F. A. (2007b). Language development in Down Syndrome: From the prelinguistic period to the acquisition of literacy. *Ment. Retard. Dev. Disabil. Res. Rev.* 13, 247–261. doi: 10.1002/mrdd.20158

Alcock, K. J., and Krawczyk, K. (2010). Individual differences in language development: Relationship with motor skill at 21 months. *Dev. Sci.* 13, 677–691. doi: 10.1111/j.1467-7687.2009.00924.x

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., and Greenhalgh, T. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *J. Child Psychol. Psychiatry* 58, 1068–1080. doi: 10.1111/jcpp.12721

Capirci, O., and Volterra, V. (2008). Gesture and speech: The emergence and development of a strong and changing partnership. *Gesture* 8, 22–44. doi: 10.1075/gest.8.1.04cap

Capone, N. C., and McGregor, K. K. (2004). Gesture development: A review for clinical and research practices. *J. Speech Lang. Hear. Res.* 47, 173–186. doi: 10.1044/1092-4388(2004/015)

Caselli, M. C., Rinaldi, P., Stefanini, S., and Volterra, V. (2012). Early action and gesture "vocabulary" and its relation with word comprehension and production. *Child Dev.* 83, 526–542. doi: 10.1111/j.1467-8624.2011.01727.x

Chapman, R. S., and Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down Syndrome. *Ment. Retard. Dev. Disabil. Res. Rev.* 6, 84–95. doi: 10.1002/1098-277920006:2<84::AID-MRDD2<3.0.CO;2-P

Chapman, R. S., Seung, H. K., Schwartz, S. E., and Bird, E. K. R. (2000). Predicting language production in children and adolescents with Down Syndrome: The role of comprehension. *J. Speech Lang. Hear. Res.* 43, 340–350. doi: 10.1044/jslhr.4302.340

Chiat, S., and Roy, P. (2008). Early phonological and sociocognitive skills as predictors of later language and social communication outcomes. *J. Child Psychol. Psychiatry* 49, 635–645. doi: 10.1111/j.1469-7610.2008.01881.x

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., and Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *Int. J. Lang. Commun. Disord.* 53, 237–255. doi: 10.1111/1460-6984.12338

Cuskelly, M., Povey, J., and Jobling, A. (2016). Trajectories of development of receptive vocabulary in individuals with Down Syndrome. *J. Policy Pract. Intell. Disabil.* 13, 111–119. doi: 10.1111/jppi.12151

Deckers, S. R. J. M., Van Zaalen, Y., Van Balkom, H., and Verhoeven, L. (2019). Predictors of receptive and expressive vocabulary development in children with Down Syndrome. *Int. J. Speech Lang. Pathol.* 21, 10–22. doi: 10.1080/17549507.2017.1363290

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends Cogn. Sci.* 8, 66–70. doi: 10.1016/j.tics.2003.12.004

Dimitrova, N., Özçalışkan, Ş., and Adamson, L. B. (2016). Parents' translations of child gesture facilitate word learning in children with Autism, Down Syndrome and typical development. *J. Autism Dev. Disord.* 46, 221–231. doi: 10.1007/s10803-015-2566-7

Dodd, B., and Thompson, L. (2001). Speech disorder in children with Down's syndrome. *J. Intell. Disabil. Res.* 45, 308–316. doi: 10.1046/j.1365-2788.2001.00327.x

D'Souza, D., D'Souza, H., Johnson, M. H., and Karmiloff-Smith, A. (2016). Audio-visual speech perception in infants and toddlers with Down Syndrome, Fragile X Syndrome, and Williams Syndrome. *Infant Behav. Dev.* 44, 249–262. doi: 10.1016/j.infbeh.2016.07.002

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Dev. Psychol.* 43, 1428–1446. doi: 10.1037/0012-1649.43.6.1428

Durkin, K., and Conti-Ramsden, G. (2010). Young people with specific language impairment: A review of social and emotional functioning in adolescence. *Child Lang. Teach. Ther.* 26, 105–121. doi: 10.1177/0265659010368750

Eadie, P., Bavin, E. L., Bretherton, L., Cook, F., Gold, L., Mensah, F., et al. (2021). Predictors in infancy for language and academic outcomes at 11 years. *Pediatrics* 147:e20201712. doi: 10.1542/peds.2020-1712

Ellis, E. M., Borovsky, A., Elman, J. L., and Evans, J. L. (2015). Novel word learning: An eye-tracking study. Are 18-month-old late talkers really different from their typical peers? *J. Commun. Disord.* 58, 143–157. doi: 10.1016/j.jcomdis.2015.06.011

Epstein, C. J. (1986). *The consequences of chromosome imbalance*. New York, NY: Cambridge University Press.

Farrant, B. M., and Zubrick, S. R. (2012). Early vocabulary development: The importance of joint attention and parent-child book reading. *First Lang.* 32, 343–364. doi: 10.1177/0142723711422626

Federmeier, K. D., Jongman, S. R., and Szewczyk, J. M. (2020). Examining the role of general cognitive skills in language processing: A window into complex cognition. *Curr. Direct. Psychol. Sci.* 29, 575–582. doi: 10.1177/0963721420964095

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., et al. (1994). Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* 59, 1–173. doi: 10.2307/1166093

Friederici, A. D. (2006). The neural basis of language development and its impairment. *Neuron* 52, 941–952. doi: 10.1016/j.neuron.2006.12.002

Frota, S., Pejovic, J., Severino, C., and Vigário, M. (2020). Looking for the edge: Emerging segmentation abilities in atypical development. *Proc. Speech Prosody* 2020, 814–818. doi: 10.21437/SpeechProsody.2020-166

Goldin-Meadow, S., Goodrich, W., Sauer, E., and Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Dev. Sci.* 10, 778–785. doi: 10.1111/j.1467-7687.2007.00636.x

Greenwald, C. A., and Leonard, L. B. (1979). Communicative and sensorimotor development of Down's Syndrome children. *Am. J. Ment. Defic.* 84, 296–303.

Guasti, M. T. (2017). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Dev.* 74, 1368–1378. doi: 10.1111/1467-8624.00612

Hula, W. D., McNeil, M. R., and Sung, J. E. (2007). Is there an impairment of language-specific attentional processing in aphasia?. *Brain Lang.* 103, 240–241. doi: 10.1016/j.bandl.2007.07.023

Iverson, J. M., Longobardi, E., and Caselli, M. C. (2003). Relationship between gestures and words in children with Down's syndrome and typically developing children in the early stages of communicative development. *Int. J. Lang. Commun. Disord.* 38, 179–197. doi: 10.1080/1368282031000062891

Johnson, C. J., Beitchman, J. H., and Brownlie, E. B. (2010). Twenty-year follow-up of children with and without speech-language impairments: Family, educational, occupational, and quality of life outcomes. *Am. J. Speech Lang. Pathol.* 19, 51–65. doi: 10.1044/1058-0360(2009/08-0083)

Kaat-van den Os, D. T., Volman, C., Jongmans, M., and Lauteslager, P. (2017). Expressive vocabulary development in children with Down Syndrome: A longitudinal study. *J. Policy Pract. Intell. Disabil.* 14, 311–318. doi: 10.1111/jppi.12212

Kent, R. D., and Vorperian, H. K. (2013). Speech impairment in Down Syndrome: A review. *J. Speech Lang. Hear. Res.* 56, 178–210. doi: 10.1044/1092-4388(2012/12-0148)

Kover, S. T., McDuffie, A., Abbeduto, L., and Brown, W. T. (2012). Effects of sampling context on spontaneous expressive language in males with Fragile X Syndrome or Down Syndrome. *J. Speech Lang. Hear. Res.* 55, 1022–1038. doi: 10.1044/1092-4388(2011/11-0075)

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron* 67, 713–727. doi: 10.1016/j.neuron.2010.08.038

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. B Biol. Sci.* 363, 979–1000. doi: 10.1098/rstb.2007.2154

Kumin, L. (1994). Intelligibility of speech in children with Down Syndrome in natural settings: Parents' perspective. *Percept. Mot. Skills* 78, 307–313. doi: 10.2466/pms.1994.78.1.307

Lang, S., Bartl-Pokorny, K. D., Pokorny, F. B., Garrido, D., Mani, N., Fox-Boyer, A. V., et al. (2019). Canonical Babbling: A marker for earlier identification of late detected developmental disorders? *Curr. Dev. Disord. Rep.* 6, 111–118. doi: 10.1007/s40474-019-00166-w

Laws, G., Briscoe, J., Ang, S. Y., Brown, H., Hermena, E., and Kapikian, A. (2015). Receptive vocabulary and semantic knowledge in children with SLI and children with Down syndrome. *Child Neuropsychol.* 21, 490–508. doi: 10.1080/09297049.2014.917619

Locatelli, C., Onnivello, S., Antonaros, F., Feliciello, A., Filoni, S., Rossi, S., et al. (2021). Is the Age of Developmental Milestones a Predictor for Future Development in Down Syndrome?. *Brain Sci.* 11:655. doi: 10.3390/brainsci11050655

Luyster, R., Qiu, S., Lopez, K., and Lord, C. (2007). Predicting outcomes of children referred for Autism using the MacArthur–Bates communicative development inventory. *J. Speech Lang. Hear. Res.* 50, 667–681. doi: 10.1044/1092-4388(2007/047)

Määttä, S., Laakso, M. L., Tolvanen, A., Ahonen, T., and Aro, T. (2012). Developmental trajectories of early communication skills. *J. Speech Lang. Hear. Res.* 55, 1083–1096. doi: 10.1044/1092-4388(2011/10-0305)

Mason-Apps, E., Stojanovik, V., Houston-Price, C., and Buckley, S. (2018). Longitudinal predictors of early language in infants with Down Syndrome: A preliminary study. *Res. Dev. Disabil.* 81, 37–51. doi: 10.1016/j.ridd.2017.12.021

McCathren, R. B. (2000). Teacher-implemented prelinguistic communication intervention. *Focus Autism Other Dev. Disabil.* 15, 21–29. doi: 10.1177/108835760001500103

McGillion, M., Herbert, J. S., Pine, J., Vihman, M., dePaolis, R., Keren-Portnoy, T., et al. (2017). What Paves the Way to Conventional Language? The Predictive Value of Babble, Pointing, and Socioeconomic Status. *Child Dev.* 88, 156–166. doi: 10.1111/cdev.12671

McKean, C., Wraith, D., Eadie, P., Cook, F., Mensah, F., and Reilly, S. (2017). Subgroups in language trajectories from 4 to 11 years: The nature and predictors of stable, improving and decreasing language trajectory groups. *J. Child Psychol. Psychiatry* 58, 1081–1091. doi: 10.1111/jcpp.12790

Miller, J. F. (1999). "Profiles of language development in children with Down Syndrome," in *Improving the communication of people with Down Syndrome*, eds J. Miller, M. Leddy, and L. A. Leavitt (Baltimore, US: Brookes Publishing), 11–40.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Prisma Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* 6:e1000097. doi: 10.1371/journal.pmed.1000097

Mundy, P., Kasari, C., Sigman, M., and Ruskin, E. (1995). Nonverbal communication and early language acquisition in children with Down Syndrome and in normally developing children. *J. Speech Lang. Hear. Res.* 38, 157–167. doi: 10.1044/jshr.3801.157

Nelson, H. D., Nygren, P., Walker, M., and Panoscha, R. (2006). Screening for speech and language delay in preschool children: Systematic evidence review for the US preventive services task force. *Pediatrics* 117, e298–e319. doi: 10.1542/peds.2005-1467

Nyman, A., Strömbergsson, S., Lindström, K., Lohmander, A., and Miniscalco, C. (2021). Speech and language in 5-year-olds with different neurological disabilities and the association between early and later consonant production: Speech/language in neurological disabilities. *Dev. Neurorehabil.* 24, 408–417. doi: 10.1080/17518423.2021.1899327

Olson, J., and Masur, E. F. (2011). Infants' gestures influence mothers' provision of object, action and internal state labels. *J. Child Lang.* 38, 1028–1054. doi: 10.1017/S0305000910000565

Özçaliskan, S., Adamson, L. B., Dimitrova, N., Bailey, J., and Schmuck, L. (2016). Baby sign but not spontaneous gesture predicts later vocabulary in children with Down Syndrome. *J. Child Lang.* 43, 948–963. doi: 10.1017/S030500091500029X

Özçalışkan, Ş, Adamson, L. B., Dimitrova, N., and Baumann, S. (2017). Early gesture provides a helping hand to spoken vocabulary development for children with Autism, Down Syndrome, and typical development. *J. Cogn. Dev.* 18, 325–337. doi: 10.1080/15248372.2017.1329735

Pejovic, J., Cruz, M., Severino, C., and Frota, S. (2021). Early visual attention abilities and audiovisual speech processing in 5–7 month-old Down Syndrome and typically developing infants. *Brain Sci.* 11:939. doi: 10.3390/brainsci11070939

Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., et al. (2011). Neural language networks at birth. *Proc. Natl. Acad. Sci. U. S. A.* 108, 16056–16061. doi: 10.1073/pnas.1102991108

Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., et al. (2007). Predicting language at 2 years of age: A prospective community study. *Pediatrics* 120, e1441–e1449. doi: 10.1542/peds.2007-0045

Riva, V., Cantiani, C., Dionne, G., Marini, A., Mascheretti, S., Molteni, M., et al. (2017). Working memory mediates the effects of gestational age at birth on expressive language development in children. *Neuropsychology* 31, 475–485. doi: 10.1037/neu0000376

Rowe, M. L., Özçalışkan, Ş., and Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Lang.* 28, 182–199. doi: 10.1177/0142723707088310

Sameroff, A., and Chandler, M. (1975). "Reproductive risk and the continuum of caretaking casualty," in *Review of child development research*, eds M. Horowitz, M. Hetherington, S. Scarr-Salapatek, and G. Siegel (Pennsylvania, US: University Park Press), 187–244.

Sansavini, A., Guarini, A., Justice, L. M., Savini, S., Broccoli, S., Alessandroni, R., et al. (2010). Does preterm birth increase a child's risk for language impairment? *Early Hum. Dev.* 86, 765–772. doi: 10.1016/j.earlhumdev.2010.08.014

Snowling, M. J., Duff, F. J., Nash, H. M., and Hulme, C. (2016). Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *J. Child Psychol. Psychiatry* 57, 1360–1369. doi: 10.1111/jcpp.12497

Sousa, R., Silva, S., and Frota, S. (2022). Early prosodic development predicts Lexical development in typical and atypical language acquisition. *Proc. Speech Prosody* 2022, 387–391.

Stoel-Gammon, C. (2001). Down Syndrome phonology: Developmental patterns and intervention strategies. *Down Syndr. Res. Pract.* 7, 93–100. doi: 10.3104/reviews.118

Tromblin, B. (2008). "Validating diagnostic standards for specific language impairment using adolescent outcomes," in *Understanding developmental language disorders*, eds C. F. Norbury, J. B. Tomblin, and D. V. M. Bishop (London: Psychology Press), 93–114.

Van Duijn, G., Dijkxhoorn, Y., Scholte, E. M., and Van Berckelaer-Onnes, I. A. (2010). The development of adaptive skills in young people with Down Syndrome. *J. Intell. Disabil. Res.* 54, 943–954. doi: 10.1111/j.1365-2788.2010.01316.x

Voci, S. C., Beitchman, J. H., Brownlie, E. B., and Wilson, B. (2006). Social anxiety in late adolescence: The importance of early childhood language impairment. *J. Anxiety Disord.* 20, 915–930. doi: 10.1016/j.janxdis.2006.01.007

Wetherby, A. M., Allen, L., Cleary, J., Kublin, K., and Goldstein, H. (2002). Validity and reliability of the communication and symbolic behavior scales developmental profile with very young children. *J. Speech Lang. Hear. Res.* 45, 1202–1218. doi: 10.1044/1092-4388(2002/097)

Yamauchi, Y., Aoki, S., Koike, J., Hanzawa, N., and Hashimoto, K. (2019). Motor and cognitive development of children with Down syndrome: The effect of acquisition of walking skills on their cognitive and language abilities. *Brain Dev.* 41, 320–326. doi: 10.1016/j.braindev.2018.11.00

Yoder, P. J., and Warren, S. F. (1993). "Can developmentally delayed children's language development be enhanced through prelinguistic intervention?," in *Enhancing children's communication: Research foundations for intervention*, eds A. P. Kaiser and D. B. Gray (Baltimore, US: Paul H. Brookes Publishing Co., Inc), 35–61.

Yoder, P. J., Warren, S. F., Kim, K., and Gazdag, G. E. (1994). Facilitating prelinguistic communication skills in young children with developmental delay II: Systematic replication and extension. *J. Speech Lang. Hear. Res.* 37, 841–851. doi: 10.1044/jshr.3704.841

Yoder, P. J., Warren, S. F., and McCathren, R. B. (1998). Determining spoken language prognosis in children with developmental disabilities. *Am. J. Speech Lang. Pathol.* 7, 77–87. doi: 10.1044/1058-0360.0704.77

Zampini, L., and D'Odorico, L. (2009). Communicative gestures and vocabulary development in 36-month-old children with Down's Syndrome. *Int. J. Lang. Commun. Disord.* 44, 1063–1073. doi: 10.1080/13682820802398288

Zampini, L., and D'Odorico, L. (2013). Vocabulary development in children with Down syndrome: Longitudinal and cross-sectional data. *J. Intell. Dev. Disabil.* 38, 310–317. doi: 10.3109/13668250.2013.828833

Zampini, L., Salvi, A., and D'Odorico, L. (2015). Joint attention behaviours and vocabulary development in children with Down Syndrome. *J. Intell. Disabil. Res.* 59, 891–901. doi: 10.1111/jir.12191

Zubrick, S. R., Taylor, C. L., Rice, M. L., and Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *J. Speech Lang. Hear. Res.* 50, 1562–1592. doi: 10.1044/1092-4388(2007/106)

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 03 October 2022
DOI 10.3389/fpsyg.2022.934826

Check for updates

OPEN ACCESS

EDITED BY
Lénia Carvalhais,
Infante D. Henrique Portucalense University, Portugal

REVIEWED BY
Christelle Declercq,
Université de Reims Champagne-Ardenne, France
Nuria Sagarra,
Rutgers, The State University of New Jersey, United States
Vsevolod Kapatsinski,
University of Oregon, United States

*CORRESPONDENCE
Natalia Arias-Trejo
nariast@unam.mx

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 03 May 2022
ACCEPTED 13 September 2022
PUBLISHED 03 October 2022

CITATION
Angulo-Chavira AQ, Castellón-Flores AM, Barrón-Martínez JB and Arias-Trejo N (2022) Word prediction using closely and moderately related verbs in Down syndrome.
Front. Psychol. 13:934826.
doi: 10.3389/fpsyg.2022.934826

Word prediction using closely and moderately related verbs in Down syndrome

Armando Q. Angulo-Chavira[1],
Alejandra M. Castellón-Flores[1], Julia B. Barrón-Martínez[2] and
Natalia Arias-Trejo[1]*

[1]Laboratorio de Psicolingüística, Facultad de Psicología, Universidad Nacional Autónoma de México, Mexico City, Mexico, [2]Facultad de Estudios Superiores Zaragoza, Universidad Nacional Autónoma de México, Mexico City, Mexico

People with Down syndrome (DS) have several difficulties in language learning, and one of the areas most affected is language production. Theoretical frameworks argue that prediction depends on the production system. Yet, people with DS can predict upcoming nouns using semantically related verbs. Possibly, prediction skills in people with DS are driven by their associative mechanism rather than by the prediction mechanism based on the production system. This study explores prediction mechanisms in people with DS and their relationship with production skills. Three groups were evaluated in a preferential-looking task: young adults, children with DS, and a typically developing control group paired by sex and mental age. Participants saw two images, a target and a distractor. They also heard a sentence in one of the three conditions: with a verb that was closely related to the object (e.g., "The woman read the book"), with a verb that was moderately related to the object (e.g., "My uncle waited for the bus"), or with a verb that was unrelated to the object (e.g., "My sister threw a broom"). Their productive vocabulary was then measured. In the young adult and typically developing groups, the results showed prediction in sentences with highly and moderately related verbs. Participants with DS, however, showed prediction skills only in the highly related context. There was no influence of chronological age, mental age, or production on prediction skills. These results indicate that people with DS base prediction mainly on associative mechanisms and they have difficulty in generating top-down predictions.

KEYWORDS
Down syndrome, prediction, verb restriction, association strength, productive vocabulary

Introduction

Lexical prediction allows people to anticipate information based on the top-down pre-activation of potential word candidates (Schoknecht, 2022) and respond rapidly and assertively to linguistic information (Federmeier, 2007; Fine et al., 2013; Kuperberg and Jaeger, 2016). Children can use different kinds of cues to make predictions, including the transitional probability between words (Pelucchi et al., 2009), phonological forms, prosody (Ito and Speer, 2008; DeLong et al., 2015), morphology (Martin and Ellis, 2012; Arias-Trejo et al., 2013; Huettig and Brouwer, 2015), syntaxis (Federmeier and Kutas, 1999), and sentence context (Campanelli et al., 2018).

There is extensive evidence about prediction using semantically related verbs (Altmann and Kamide, 1999; Mani and Huettig, 2012, 2014). Altmann and Kamide (1999) showed that adults anticipate referents based on the semantic attributes of a verb. On hearing *eat*, they looked more at the image of an edible object than a non-edible one. There are also several studies on prediction during language comprehension in younger populations, such as infants and school children with typical development (TD), that demonstrate prediction skills as early as 24 months of age (Borovsky et al., 2012; Mani and Huettig, 2012; Lukyanenko and Fisher, 2016; Mani et al., 2016; Gambi et al., 2018). However, little is known about people with genetic syndromes that lead to different developmental trajectories (Arias-Trejo et al., 2019).

Down syndrome (DS) is a genetic disorder caused by all or part of an extra copy of chromosome 21 (Lubec and Engidawork, 2002). One in every thousand babies born presents DS. It is the most frequent biological cause of intellectual disability (Dierssen, 2012), resulting in cognitive development that falls behind chronological age (van Gameren-Oosterom et al., 2011). One area of disadvantage in children with DS is language production. Part of this disadvantage is related to physical abnormalities in the vocal apparatus (Kumin et al., 1994), including a small oral cavity, irregular teeth, a large tongue, and abnormalities in the facial muscles. They are usually affected by hearing loss and otitis media, which affect not only comprehension but also perception of speech during oral production (Dodd and Thompson, 2001).

Word comprehension scores in children with DS are similar to those of their TD peers matched by mental age (see Næss et al., 2011). Their comprehension of nouns and verbs, for instance, is remarkably well preserved (Michael et al., 2012), but they experience problems with grammar (Witecy and Penke, 2017), use of contextual cues (Hsu, 2019), and syntax (Iverson et al., 2003). In general, word production is lower in children with DS than in those with TD (Roberts et al., 2007; Martin et al., 2013).[1]

They usually present delayed speech (Roberts et al., 2005) and speech errors (Rosin, 1988). In a longitudinal study, Næss (2022) evaluated children with and without DS with similar non-verbal mental age, auditory memory, oral motor skills, and receptive vocabulary and found a slower growth of expressive vocabulary in people with DS than those with TD. People with DS also experience morphosyntactic difficulties in production: problems with gender and number agreement between articles and nouns (Eadie et al., 2002), and errors in grammatical categories, including verbs, in spontaneous speech (Chapman et al., 2000; Chapman, 2006). They also have problems producing grammar, morphemes, and syntax (Yoder et al., 2006), and problems with semantic processing (Laws et al., 2015; Andreou and Katsarou, 2016; cf. Barrón-Martínez and Arias-Trejo, 2020, 2022). Andreou and Katsarou (2016) evaluated the semantic performance of adolescents with DS with a mental age of 3.5–6.5 years, using tests that measured receptive and expressive semantic skills. They found that the group with DS had lower performance on all tests compared to a TD group matched by mental age, and those with DS performed lower on expressive than receptive semantic skills. In sum, the evidence showed generalized language problems in people with DS, and particular weakness in production and semantic processing.

Models of prediction in language comprehension (Dell and Chang, 2014; Pickering and Gambi, 2018) postulate that the production system is highly important to make predictions during comprehension. Experimental evidence has shown that using the production system during language comprehension makes predictive processing difficult; people are not able to make predictions while they produce syllables during comprehension tasks (Martin et al., 2018). Verbal fluency is also related to prediction skills (Rommers et al., 2015). In a correlational study with 2-year-old German toddlers, Mani and Huettig (2012) found a positive correlation between their ability to predict a target object using a semantically related verb and their productive vocabulary: high-scoring producers predicted the target, but low-scoring producers did not. However, prediction skills did not correlate with comprehension scores (Borovsky et al., 2012; Mani et al., 2016). These results suggest the need for a well-developed production system to make predictions during language comprehension.

If people with DS have production problems, they should therefore also have problems with prediction; however, there is evidence for some prediction skills in this population. In a preferential-looking task using an eye-tracker, Arias-Trejo et al. (2019) reported that children with DS (mental age: 5.48 years), as well as their TD peers matched by mental age, used the semantic information contained in a verb (e.g., *eat*) to anticipate an edible target (e.g., *cake*) in preference to a non-edible distractor. Thus, the question is, if people with DS have problems with production, how do they make predictions about upcoming linguistic information? Understanding this predictive

1 Studies have found, however, that in the early stages of development, production is equivalent in groups matched by mental age (Galeote et al., 2011; Checa et al., 2016).

processing in people with DS is essential to understanding their language difficulties because children learn a language using prediction and prediction errors (Dell and Chang, 2014; Reuter et al., 2019).

Two mechanisms cooperate to create predictions during language comprehension: prediction-by-association and prediction-by-production (Huettig, 2015; Pickering and Gambi, 2018). Prediction-by-association mechanism is a bottom-up mechanism of automatic spreading activation based on representations shared between words. This mechanism has been described extensively in priming studies (Collins and Loftus, 1975; Dell, 1986; Anderson, 2013); it is essentially predictive since the activation spreads among concepts before the presentation of the target word (Huettig, 2015; Pickering and Gambi, 2018). The word *dog*, for example, pre-activates the word *bone* because these words occur together in speech and the environment. The prediction-by-association mechanism is inefficient because the activation spreads freely through all related[2] concepts, regardless of the context. For instance, in the sentence "My dog is chasing a cat," the activation from *dog* can pre-activate *cat* and other incongruent but related words like *bone*. Nevertheless, the cognitive load is low, and the pre-activation is virtually instantaneous. The prediction-by-production mechanism is more efficient because it considers contextual information, both linguistic and non-linguistic, to make predictions, but these predictions are slower and require more cognitive processing. Pickering and Gambi (2018) argue that the top-down predictions generated by this second mechanism are based on the production system: to make predictions during the comprehension process, the production system predicts the concept of the word based on linguistic and non-linguistic context. Notably, these two mechanisms, prediction-by-production and prediction-by-association are complementary: the extent to which predictions rely on one system or the other depends on the availability of information, resources, and time (Pickering and Gambi, 2018).

According to prediction theory, the production problems of people with DS should result in difficulties in creating top-down predictions using contextual information, but not bottom-up automatic predictions. Recent studies have shown that children with DS may use pre-activation mechanisms based on the association between concepts (Barrón-Martínez and Arias-Trejo, 2020, 2022). Barrón-Martínez and Arias-Trejo (2022) evaluated children with and without DS in a preferential-looking task using an eye-tracker. In half of the trials, participants were exposed to pairs of words (prime and target) that were related, and in the other half to pairs that were unrelated. The participants looked more at a named target image preceded by a related prime than one preceded by an unrelated prime. This finding suggests that the prediction-by-association mechanism is preserved in children with DS.

Arias-Trejo et al. (2019) demonstrate that children with DS can use the verb information to make predictions. Verbs provide information about the action and important semantic and grammatical information about the agent and patient of the action. These thematic roles are verb-specific concepts (McRae et al., 1997; Ferretti et al., 2001). For example, the verb *gamble* activates information about the location of the action (e.g., *casino*) and possible participants in the action (e.g., *gambler*).

The information provided by verbs can be used in making predictions (Altmann and Kamide, 1999); however, these predictions are not tied to the verb itself but to the event in which verbs occur together with agents and patients. Kamide et al. (2003) evaluated young adults using the visual world paradigm. Participants were presented with an array of images, including several objects: a motorcycle, a carousel, a man, and a girl. When participants heard the sentence "The man will ride the...," they looked at the motorcycle, but not the carousel; when they heard the sentence "The girl will ride...," they looked at the carousel. Thus, although verbs can be linked to specific noun concepts, the elicited link depends on the context.

Stefaniak et al. (2021) found developmental differences using contextual information. In Experiment 2, school-age children and adults performed a grammatical judgment task, including both typical and unusual (but grammatically correct) patients for verbs. They found that both groups showed better performance with typical than with unusual patients; however, younger children showed lower performance in judging unusual patients. The authors interpreted these results based on the declarative/procedural model: in the processing of typical patients, the declarative memory assigns a meaning, and the procedural memory evaluates whether the patient can be used with the verb; with unusual patients, the declarative memory does not generate meaning, and the procedural memory does not evaluate the patient.

Stefaniak et al. (2021) argue that spreading activation has little influence on the verb-patient typicality effect since there was no variation from the free association norms, and the task relied more on the syntactic cues and the thematic roles. This interpretation is congruent with our theoretical framework: the prediction-by-association system always generates predictions, and the prediction-by-production system uses contextual information such as syntax or grammar to generate predictions requiring more information. For example, the verb *read* is highly associated with the patient noun *book*; in this case, the prediction relies more on the prediction-by-association system. However, the lower degree of association between the verb *wash* and the patient *bucket* relies less on prediction-by-association and more on prediction-by-production because the verb *wash* can be applied to different objects. In the latter

2 Throughout this text we will use the words *association* and *relationship* as synonyms.

case, additional contextual information is needed to formulate a correct prediction.

The present study examines whether children with DS anticipate a referent in the same way as their mental age-matched peers in two different contexts: when there are higher and lower levels of association between verbal cues. We hypothesized that in a predictive sentence with a high degree of association between the verb and the target noun (e.g., *read—book*), children with and without DS would look at the target image before it was named because they would rely on prediction-by-association. However, in a sentence with a lower degree of association between the verb and the noun (e.g., *wash—bucket*), participants with TD, but not those with DS, would look at the target image before it was named. Here, they need to rely more on the prediction-by-production mechanism; thus, problems with production in people with DS would affect this mechanism. We also hypothesized that vocabulary production would modulate prediction in DS, as participants with higher vocabulary production scores would use their greater production skills in predictive sentences with low associations between the verb and the target noun.

Materials and methods

Participants

The study was carried out online because of the COVID-19 pandemic. We evaluated 21 participants with DS with a mean chronological age of 20.784 years (SD = 5.754, range: 11.460–29.563) and a mean verbal mental age of 5.524 years (SD = 2.363, range: 3.5–13.83). Five were non-verbal and therefore did not produce any language. All participants with DS lived in a monolingual environment, according to their parents or primary guardian. We also evaluated a control group of children with TD paired by mental age and sex with the participants with DS (see **Table 1**). This group included 21 participants with a mean chronological age of 5.524 years (SD = 2.363, range: 3.25–13.58) and a mean verbal mental age of 5.829 years (SD = 2.418, range: 3.25–13.83). Another 21 participants with TD were excluded because they had a mental age greater than their chronological age and could not be paired with participants with DS. All were monolingual Spanish speakers. According to parental reports, all participants had a normal or corrected-to-normal hearing and vision and had no neurological/psychiatric problems. An additional group of 39 adults was assessed (M = 23.87 years, SD = 2.48, range: 18–28, 22 male) to test the functioning of our experimental manipulation. Three adults were excluded from this group because of failures in calibration. All participants, or, in the case of minors, their parents or guardian, provided informed consent. The study was approved by the research ethics committee of the Facultad de Psicología, Universidad Nacional Autónoma de México (Approval No. FCPE_13092021_H_AC).

Instruments

Mental age: Receptive vocabulary assessment

Participants' verbal mental age was evaluated to match participants from the TD and DS groups and to determine whether cognitive development affected linguistic prediction skills. Verbal mental age was measured with remote administration of the Receptive One-Word Picture Vocabulary Test: SBE (Martin, 2010b). Participants were presented with four images and asked to match a word they heard to the correct image. The test was suspended after four consecutive errors or failure to respond to six stimuli. The raw score was calculated by subtracting the number of errors from the total number of items reached and converted to mental age using standardized tables (Martin, 2010b). The approximate duration of the test was 20 min. For younger children and participants with DS who had difficulty verbally indicating the image, parents or guardians were asked to indicate the images the child had pointed to, even if they were incorrect. The test administrator corroborated the answers by noting the part of the screen the participant pointed to. We used this mental age evaluation because our experimental task measures prediction during language comprehension; it is thus an appropriate measure for pairing participants with similar comprehension skills since receptive vocabulary is a good predictor of general comprehension (Ricketts et al., 2007; Stolt et al., 2016; Cheung et al., 2022). This evaluation also has two methodological advantages: there are normative values for the Mexican Spanish-speaking population and it can be performed online.

Productive vocabulary assessment

Participants' expressive vocabulary was evaluated to determine the effect of production skills on language prediction. The Expressive One-Word Picture Vocabulary Test: SBE (Martin, 2010a) was administered remotely. Participants were presented with one image and asked to name it. The test was suspended after four consecutive errors. In this evaluation, some participants with DS scored zero points; they were non-verbal according to their parents. However, we assumed that participants understood the task because they followed the instructions for the mental age evaluation and the experimental task. The raw score was calculated by subtracting the number of errors from the total number of items reached. The test was suspended if participants failed to respond to six stimuli (Martin, 2010a). The approximate duration of the test was 20 min. Parents were asked to avoid interaction with participants while they performed the test.

TABLE 1 Socio-demographic data.

		TD	DS	*P*-value
Age	M (SD)	5.524 (2.363)	20.936 (5.765)	<0.001
Sex	N (male/female)	11 /10	11 /10	–
Mental age	M (SD)	5.829 (2.418)	5.773 (2.482)	0.941
Productive vocabulary	M (SD)	51.904 (22.248)	43.600 (27.400)	0.185

P-value corresponds to an independent sample test between the two groups. TD, typical development; DS, Down syndrome.

Materials

Three types of sentences were created: predictable sentences with a closely related verb (CV; e.g., "The woman read the book"), a moderately related verb (MV; e.g., "My uncle waited for the bus"), and unpredictable sentences with an unrelated verb (UV; e.g., "The woman lost the sock"). A total of 56 sentences were created, 14 for the CV condition, 14 for the MV condition, and 28 for the UV condition. All words used in the sentences were familiar to children (Alva-Canto, 2001). Verbs and direct objects in the CV condition had a high association strength and those in the MV condition had a lower association strength, according to the validation studies described below. The UV sentences used the same target nouns as the predictive sentences but with unrelated verbs. **Supplementary Appendix 1** shows the experimental sentences, the targets, and distractors in the CV condition, and their corresponding UV sentences. **Supplementary Appendix 2** shows the experimental sentences, the targets, and distractors in the MV condition, and their corresponding UV sentences.

The sentences were audio recorded in a female, child-directed voice, with no specific emphasis on any part of the sentence, in a quiet room (a basement with low noise levels), using a Shure MV51 microphone at 44,100 Hz and 16-bits. They were edited in Adobe Audition CS6 with noise reduction, normalization, and sound amplification. The lists of sentences were recorded four times in different orders. First, they were recorded in ascending order (from the beginning to the end of the list) and then in descending order, and then the sequence was repeated.

Two objects were presented visually as competitors: a target and a distractor (**Supplementary Table 1**). The target was the noun that appropriately completed the sentence for the closely and moderately related grammatical constructions. The UV condition used the same target and distractor as the CV and MV conditions. The visual stimuli were realistic photographs of the targets and distractors. The images were edited in Adobe Photoshop CS6 and adjusted to 600 × 600 pixels. Individual images were placed on a gray background (RGB: 225, 225, 225; 1920 × 1080 pixels). The visual and auditory stimuli were then embedded in AVI videos created with Adobe Flash CS6 and Adobe Premiere Pro.

Sentence validation studies

Two pilot studies were carried out to determine the plausibility of each sentence and the degree of association between the verb and the expected noun. The first evaluated whether the sentences would be likely to be heard in an everyday context (**Supplementary Appendix 3**). Thirty undergraduates ($M_{age} = 25.3$, 17 male) evaluated the plausibility of the sentence with the target (e.g., "The woman read the book") and with the distractor (e.g., "The woman read the sock"). Kruskal–Wallis tests found differences between conditions in the target ($X^2 = 34.707$, $p < 0.001$) and the distractor ($X^2 = 24.996$, $p < 0.001$). *Post-hoc* analysis with a Mann–Whitney *U*-test showed that the CV sentences have more plausibility with the target ($z = 2.987$, $p = 0.002$) but less plausibility with the distractor ($z = 4.412$, $p < 0.001$) than the MV sentences. They also have more plausibility with the target ($z = 4.896$, $p < 0.001$) but less plausibility with the distractor ($z = 4.483$, $p < 0.001$) than the UV sentences. The MV sentences have more plausibility with the target ($z = 4.243$, $p < 0.001$) than the UV sentences; however, there are no differences between MV and UV sentences with the distractor ($z = 0.480$, $p = 0.644$). These results confirm that our predictable sentences are considered more natural than the non-predictable ones, which is expected because regularity generates prediction in language.

The second validation was an association strength task. A "restricted" association task was performed to determine the association levels between the verbs and the nouns in the sentences. A total of 30 university students from Mexico City participated ($M_{age} = 26.2$, range: 18–30; 19 male). The pilot experiment was created on the Cognition platform (Cognition Run, 2021), and it lasted approximately 10 min. Participants were asked to write a verb in response to the noun stimulus in this task. The instructions to the participants were: "Next, you will see a series of nouns; please write the first VERB that comes into your mind when reading the noun. Answer as quickly as you can." **Table 2** shows the association strength of the experimental stimuli.

A Kruskal–Wallis test revealed significant differences between conditions in the association strength between the verb and the target ($X^2 = 49.353$, $p < 0.001$) but not between the verb and the distractor ($X^2 = 3.526$, $p = 0.172$). The exploration of significance analysis showed that the CV condition had a greater

TABLE 2 Association strength between targets and distractors.

ID	CV		UV		ID	MV		UV	
	Target	Distractor	Target	Distractor		Target	Distractor	Target	Distractor
1	80	0	0	0	15	23.33	0	0	0
2	56.66	0	0	0	16	16.66	0	0	0
3	70	0	0	6.66	17	20	0	0	0
4	73.33	0	0	6.66	18	6.66	0	0	0
5	60	0	0	0	19	13.33	0	0	13.33
6	66.66	0	0	3.33	20	23.33	0	0	6.66
7	66.66	0	0	0	21	16.66	0	0	0
8	46.66	0	0	0	22	10	0	16.66	0
9	63.33	0	0	0	23	20	6.66	0	6.66
10	53.33	0	0	0	24	20	0	0	0
11	53.33	0	0	0	25	6.66	0	0	0
12	46.66	0	0	0	26	13.33	3.33	0	0
13	73.33	0	0	0	27	20	0	0	0
14	56.66	0	0	0	28	3.33	0	0	0

The ID corresponds to the sentences presented in **Supplementary Appendices 2 and 3**. CV, closely related verb; MV, moderately related verb; UV, unrelated verb.

association strength between the verb and the noun than the MV ($z = 4.491$, $p < 0.001$) and the UV ($z = 6.103$, $p < 0.001$); the association was greater in the MV than the UV ($z = 5.555$, $p < 0.001$). These results corroborate changes in the association strength between conditions.

Finally, we found a strong positive correlation between the association strength and the plausibility values ($r = 0.639$, $p < 0.001$), implying a relationship between the verbal association and sentence plausibility.

An additional validation study was performed after the review process. This validation study was not used in the stimulus selection, but it is important to corroborate the association strength in both directions between the verb and the expected noun. This additional validation used the same procedure as the original verb association validation, except that the verb was presented as a cue, and participants were asked to provide the first word that came to mind when they saw it. We compared only the CV and MV conditions because all the values in the UV condition were zero. A Mann–Whitney *U*-test showed that the CV condition had higher association values than the MV condition ($Z = 3.37$, $p < 0.001$). The values for the MV condition were very close to zero ($M = 3.225$; $SD = 5.93$), suggesting that the verb does not elicit the expected noun (see **Supplementary Appendix 4**), and participants probably need the visual context to create a prediction.

Procedure

The participants were recruited through informational posts on social media and specialized care foundations for people with DS. Parents who contacted us were told about the procedures and objectives of the study and then formalized their participation by signing the informed consent. A socio-demographic questionnaire was first administered to participants' parents on a Zoom video call to verify that they met the inclusion criteria.

The gaze of the participants was recorded remotely using the RealEye.io online platform. This platform is a webcam-based eye-tracker with a maximum sample rate of 60 Hz; it calculates the gaze position when participants look at their personal computers with an accuracy of approximately 100 px (∼1.5 cm) and with a visual angle error of ∼ 4.17 degrees (RealEye, 2020). This accuracy is appropriate for a two-image visual display and fixation analysis.

Two calibration processes were performed with RealEye. Participants first tracked points using the computer mouse and then performed standard calibrations in which they looked at four different points on the screen. The platform does not store the participant's image, sound, or location data, but only their gaze position.

Each participant heard 28 sentences: 7 CV sentences, 7 MV sentences, and 14 UV sentences. The sentences were counterbalanced across subjects in four different orders so that each pair of images was presented only once to each participant. Each trial had a duration of 8,000 ms. From 0 to 1,000 ms, a fixation point was presented on the screen. The images of the two competitors were presented from 1,000 to 7,000 ms (see **Figure 1**). The sentence (e.g., "The woman read the book") was presented as follows: the subject (e.g., "The woman") was presented in a pre-verb window from 1,000 to 3,000 ms and the verb from 3,000 to 5,000 ms. The verb (e.g., "read") was presented at 3,000 ms, followed by the determiner (e.g., "the") and then a period of silence. Then, in the noun window, the

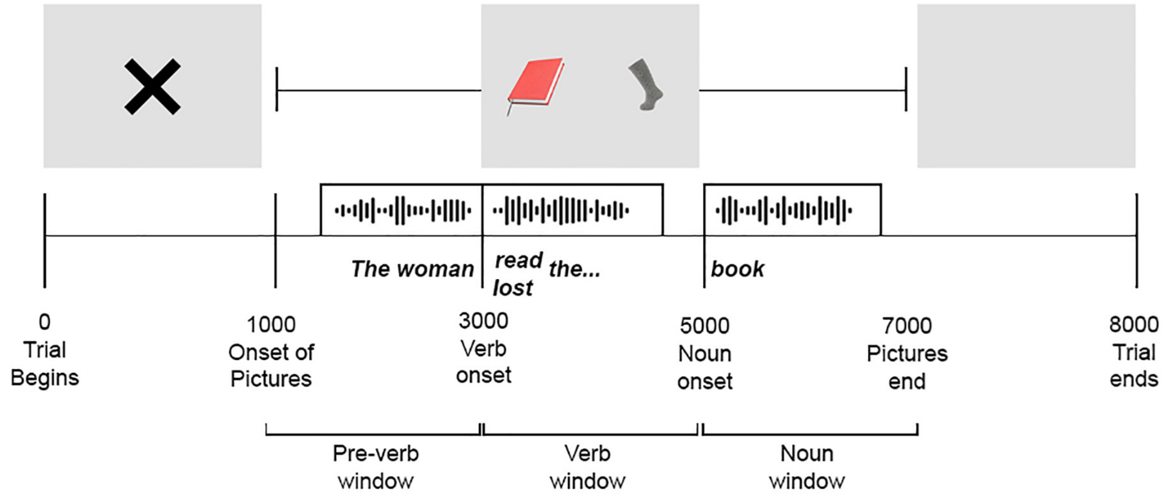

FIGURE 1
Example of an experimental trial. The image used a creative commons zero (cc0) license and it is the image of public domain.

direct object (e.g., "book") was presented from 5,000 to 7,000 ms. Finally, the screen was blank from 7,000 to 8,000 ms.[3]

The receptive and expressive tests were administered in a Zoom video call with support from the participant's parent or guardian to manage the practical details. The total duration of the evaluations was approximately 90 min. The results of the scales were delivered to the parents in a report that also contained suggestions for educational intervention.

Data processing

As noted earlier, the data quality of the webcam eye-tracker is lower than that usually employed in an experimental laboratory, but enough for our experimental design. The raw fixation signal was thus interpolated and filtered to enhance robustness and precision (Wass et al., 2014). A Gaussian filter ($\sigma = 5$) was applied to reduce high-frequency noise and enhance the precision of the data. A linear interpolation was applied to reconstruct the missing data and standardize the sample rate across participants. To obtain a better reconstruction, signal segments with >150 ms of missing data were not interpolated (Wass et al., 2014). We also adjusted all signals to a sample period of 20 ms (50 Hz); the maximum sample period of the webcam eye-tracker is ~16 ms (60 Hz). This standardization of the sample rate allowed us to compare changes in temporality between groups and conditions; otherwise, comparing the average looking time over the trial could produce type II errors. Since we hypothesize that participants with DS had weaker prediction skills, avoiding this type of error is important.

Since participants performed the experiments on their computers, there was variation in the size and location of competitors on screens. We thus adjusted the areas of interest by modifying them in proportion to the screen size. The original areas of interest measured 960 × 1,080 pixels and were embedded on a 1,920 × 1,080 background. If, for example, the participant's screen measured 1,600 × 1,200 pixels, the areas of interest should measure 800 × 1,200 pixels. Changes in height and width were independent to enhance the adjustment. A similar process was applied to the location of the areas of interest. The original location of the upper–left competitor was at 480 × 540 pixels; in the same example, the new location would be at 400 × 600 pixels.

The fixations on the two areas of interest were coded as 1 when the gaze signal coordinates were located inside the area of interest; otherwise, they were coded as 0. Each trial thus had two binary time series indicating when participants looked at any specific competitor. Since participants could only fixate on one competitor at a time, an increased fixation on one competitor implies a decreased fixation on other competitors. To reduce the autocorrelation (temporal dependence between samples) of the fixation signals, we binned the data by averaging it every 100 ms (Mirman, 2014).

Trials were excluded in which participants looked <25% of the time when the competing pictures were present (0–6,000 ms, relative to the picture presentation).

3 The auditory sentences were manipulated by adding pauses so that the verbs and expected nouns in all of the sentences were heard at the same positions on the timeline, as in previous studies with children (Mani and Huettig, 2012, 2014; Arias-Trejo et al., 2019). This presentation is advantageous for subjects with processing speed problems, and it also provides a specific period for prediction. We thus expected the predictive look in the verb window for the CV and MV conditions, and the look to the named noun in the noun window for all conditions.

Statistical analysis

Growth curve analysis (Barr, 2008; Mirman, 2014) and a cluster-based non-parametric test (Maris and Oostenveld, 2007) were used to analyze the prediction over the time course of the trial. The growth curve analysis compared the temporal dynamic among conditions and groups from the verb presentation until the end of the picture presentation (4,000–6,000 ms relative to the picture presentation). This analysis window allowed for modeling the predictive and non-predictive responses using low-order polynomials. The analysis was performed in R version 4.1.1 (R Core Team, 2019) using the glmmPQL function of the mass package. We used a mixed-effects binomial logistic regression because the fixations are binary variables (fixated or not). The dependent variable was the log odds ratio of the fixation computed as follows (Barr, 2008): $log\frac{F}{N-F}$, where F is the sum of fixations in a specific bin and N is the total number of fixations in the bin. The time was modeled using third-order orthogonal polynomials (Mirman, 2014). The fixed effects of the model were all-time terms (linear, quadratic, and cubic), condition (CV, MV, or UV), and group (TD or DS). For the random effect, we used the maximal random structure that allowed convergence (Barr, 2008); for all analyses, the maximal random structure was the slope of all time terms on the subject and the intercept of the trials. The categorical variables were dummy coded using the UV condition and the TD group as a reference.

The cluster-based non-parametric test better describes the temporality of prediction effects (beginning, duration, and end); these were evaluated from the onset to the end of the picture presentation (0–6,000 ms). To compare conditions, we used paired t-tests contrasting CV and MV conditions against the UV condition, independently for each group. We also compared each condition against chance level (0.5) using a one-sample t-test; this comparison was performed independently for each group and condition. Clusters were created by summing the adjacent t-values higher than the critical value for $\alpha = 0.05$ (adults: 2.02; TD and DS groups: 2.08). The permuted distribution (100,000 iterations) was created by shuffling the data randomly between conditions for paired tests and shuffling the mathematical sign for the one-sample test. In each iteration, we took only the maximum permuted cluster. A cluster was significant if its value was less than 5% of the total values of the permuted distribution.

Using the model comparison approach, we also evaluated the effect of chronological age, mental age, production, and association strength on prediction in the DS group. Chronological age was used to assess the influence of language experience, mental age was used to evaluate the effect of cognitive development, and production was used to determine the effect of preservation of the productive system on prediction skills. The association strength between the verb and the expected noun was used to evaluate whether the participants with DS had better predictions when there was a high degree of association. The fixation data were aggregated from 2,500 to 4,000 ms, relative to the picture presentation: the period in which participants could predict the upcoming noun. All continuous variables were min-max normalized (−0.5 to 0.5) to improve the convergence of the model. The categorical variables were dummy coded using the UV condition as a reference. Binomial mixed effect models (the glmer function) were compared using the change in log-likelihood (−2 times) with a chi-squared distribution. Thus, we first created a reference model including only the condition as fixed effects and the subjects on the slope of the condition, and the intercept of the items as a random effect. The demographic variables were then included independently in the reference model. We also computed the Bayes factor using the package bayestestR (Makowski et al., 2019) to provide evidence for the null or alternative hypothesis. If the Bayes factor was <0.33 (Wetzels et al., 2011), we assumed that the variable was not relevant to the explanation of the predictive effect.

Results

Adults

All trials were analyzed for 39 adults (see section "Data processing"); however, three adults were excluded from the final sample because of calibration problems. The upper panels of **Figure 2** show the probability of fixation in each condition (left) and the modeled data (right). Preliminary examination revealed an increase in fixation in the CV and MV conditions after the verb presentation and the UV condition after the presentation of the noun.

Table 3 presents the statistical values of the growth curve analysis. The results were significant for the CV and MV conditions, indicating that participants looked more in these conditions than in the UV condition.

According to the interaction of both predictive conditions with the quadratic term, participants had a sharper fixation pattern in both predictive conditions (CV and MV) than in the non-predictive one (UV). Finally, the interaction of the MV condition with the cubic term suggests that participants looked more at the target and looked away faster in the MV than in the UV condition.

The cluster-based permutation analysis revealed that participants looked more in the CV than in the UV condition from 3,200 to 4,800 ms ($t_{cluster} = 53.615$, $t_{max} = 4.109$, $p < 0.001$), and more in the MV than in the UV condition from 3,200 to 3,900 ms ($t_{cluster} = 59.380$, $t_{max} = 4.240$, $p < 0.001$). They also looked more at the target than chance level in the CV condition from 2,000 to 6,000 ms ($t_{cluster} = 117.235$, $t_{max} = 4.808$, $p < 0.001$). Adults looked more at the target than chance level in the MV condition in two time clusters: from

10.3389/fpsyg.2022.934826

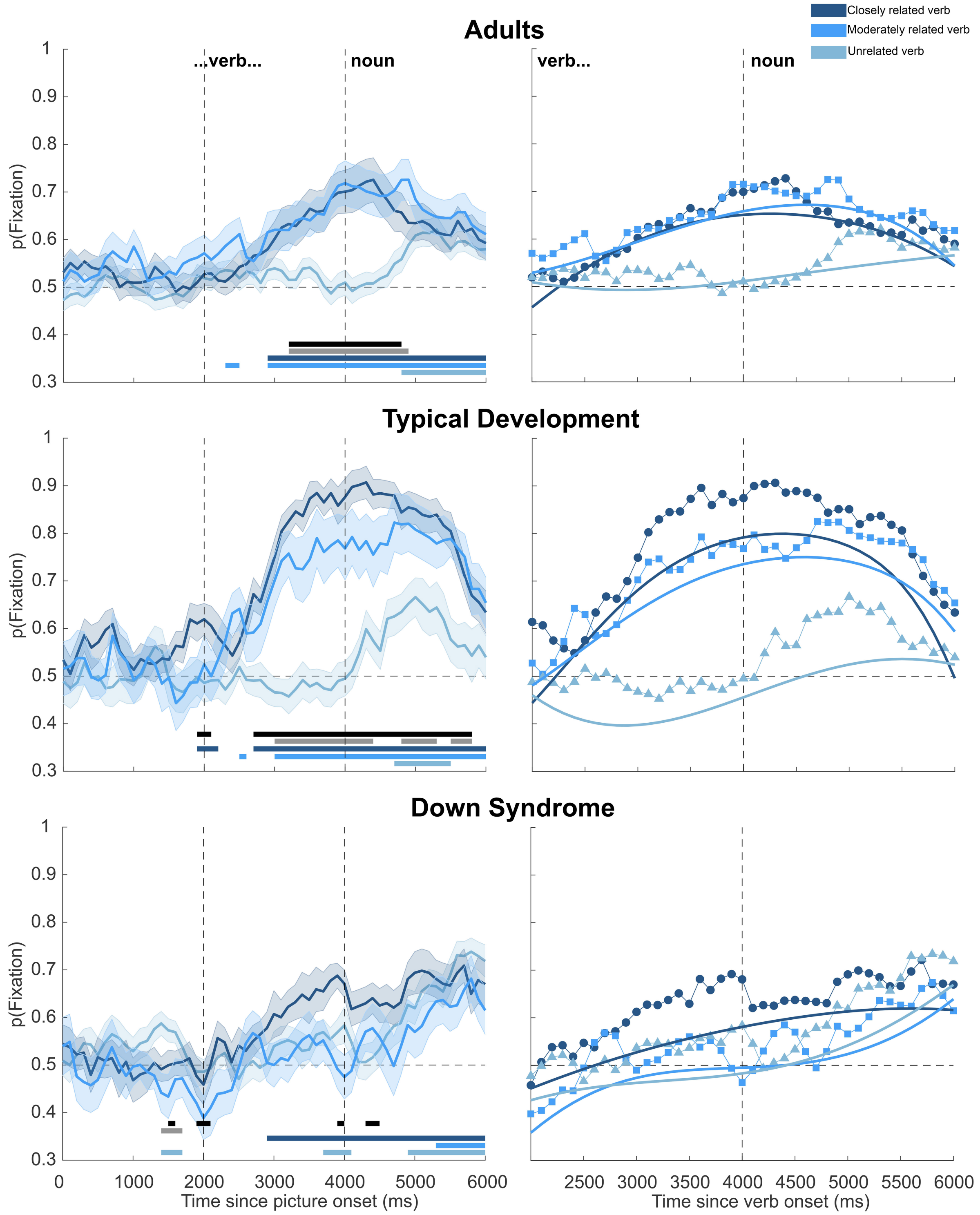

FIGURE 2
Probability of fixation and fitted lines for all groups. Left panels lines represent the average probability of fixation. Shaded areas show standard error. The horizontal dashed line indicates chance level; the vertical dashed lines the presentation of the verb and the noun. Horizontal bars in the lower part of each plot indicate the significant clusters. Closely related and unrelated verb differences are shown in black, and moderately related and unrelated verb differences are shown in gray. Blue lines represent the difference with chance level (0.5), and the colored lines correspond to the colors of the conditions. Right panels lines with markers show the average probability of fixation. Solid lines indicate the fitted line from the growth curve analysis. The horizontal dashed line shows the chance level, and the vertical dashed line the presentation of the noun. Note that the time shown in this plot begins with the presentation of the verb. The image used a creative commons zero (cc0) license and it is the image of public domain.

2,300 to 2,500 ms ($t_{cluster} = 7.737$, $t_{max} = 3.144$, $p = 0.019$) and from 2,900 to 6,000 ms ($t_{cluster} = 125.750$, $t_{max} = 5.237$, $p < 0.001$). Finally, they looked more at the target than chance level in the UV condition from 4,800 to 6,000 ms ($t_{cluster} = 48.407$, $t_{max} = 4.418$, $p < 0.001$).

Typical development and Down syndrome groups

Approximately 95% of the trials (559 of 588) with the TD group and 97% (573 of 588) with the DS group were analyzed. No participants were excluded for missing data or calibration problems in both youngest groups. Table 4 presents the statistical values of the growth curve analysis.

The middle and lower panels of Figure 2 show the probability of fixation in each condition (left) and the modeled data (right) for the TD and DS groups, respectively. Preliminary examination revealed that the TD group had more fixations in the CV and MV conditions after the verb presentation and in the UV condition after the noun presentation. The DS group presented an increase in fixation after the verb in the CV condition and after the noun in the MV and UV conditions.

The growth curve analysis revealed that in the TD group, the CV and MV had a greater and sharper increase in looking (seen in interaction with the quadratic term) than the UV condition. The positive slope of the interaction of the groups with the cubic term suggests that in the UV condition, the DS group looked more at the target and looked away faster than the TD group. The interaction of the groups with both conditions (CV and MV) indicated that in the control group, there was a greater difference between the predictive and the non-predictive conditions than in the DS group; this pattern of looking was sharper in the TD than in the DS group (seen in interaction with the quadratic term). Finally, the interaction of the groups with the linear term suggests that the difference between the MV and UV conditions increases faster over time in the control group than in the DS group.

The cluster-based permutation analysis revealed that children with TD looked more in the CV than in the UV condition in two time clusters: from 1,800 to 2,000 ms ($t_{cluster} = 7.384$, $t_{max} = 2.747$, $p = 0.025$) and from 2,600 to 5,700 ms ($t_{cluster} = 168.421$, $t_{max} = 9.571$, $p < 0.001$). They also looked more in the MV than in the UV condition in three time clusters: from 2,900 to 4,300 ms ($t_{cluster} = 57.854$, $t_{max} = 5.806$, $p < 0.001$), from 4,700 to 5,200 ms ($t_{cluster} = 17.835$, $t_{max} = 7.735$, $p < 0.001$), and from 5,400 to 5,700 ms ($t_{cluster} = 14.468$, $t_{max} = 4.616$, $p < 0.001$). The TD group looked more than chance level in the CV condition in two time clusters: from 1,800 to 2,100 ms ($t_{cluster} = 12.211$, $t_{max} = 3.514$, $p = 0.003$) and from 2,600 to 6,000 ms ($t_{cluster} = 322.759$, $t_{max} = 14.883$, $p < 0.001$). They also looked more than chance level in the MV condition in two clusters: from 2,400 to 2,500 ms ($t_{cluster} = 6.236$, $t_{max} = 3.124$, $p = 0.039$) and from 2,900 to 600 ms ($t_{cluster} = 146.690$, $t_{max} = 7.964$, $p < 0.001$). The TD

TABLE 3 Growth curve analysis for adults.

Fixed effects	β	*SE*	*df*	*t*	*p*
Intercept	0.0856	0.107	43591	0.793	0.427
Linear	**0.643**	**0.266**	**43591**	**2.414**	**0.015**
Quadratic	0.457	0.294	43591	1.555	0.119
Cubic	−0.104	0.172	43591	−0.603	0.546
CV	**0.441**	**0.131**	**1049**	**3.371**	**<0.001**
MV	**0.588**	**0.131**	**1049**	**4.474**	**<0.001**
Linear: CV	0.299	0.162	43591	1.849	0.064
Linear: MV	0.274	0.162	43591	1.684	0.092
Quadratic: CV	**−2.259**	**0.164**	**43591**	**−13.725**	**<0.001**
Quadratic: MV	**−1.898**	**0.164**	**43591**	**−11.516**	**<0.001**
Cubic: CV	0.128	0.163	43591	0.791	0.428
Cubic: MV	**−0.490**	**0.163**	**43591**	**−2.990**	**0.002**

Formula: log odds (fixations) ∼ (Linear + Quadratic + Cubic) × Condition + [(Linear + Quadratic + Cubic)| Subject] + (1| Item). Conditions: Unrelated verbs (UV), closely related verbs (CV), moderately related verbs (MV). *SE*, standard error; *df*, degrees of freedom. Bold values indicate significant effects.

TABLE 4 Growth curve analysis for DS and TD groups.

Fixed effects	β	*SE*	*df*	*t*	*P*
Intercept	−0.249	0.179	45262	−1.392	0.163
Linear	**1.524**	**0.46**	**45262**	**3.306**	**<0.001**
Quadratic	0.600	0.368	45262	1.628	0.103
Cubic	**−0.580**	**0.233**	**45262**	**−2.487**	**0.012**
CV	**1.468**	**0.142**	**1086**	**10.304**	**<0.001**
MV	**1.326**	**0.145**	**1086**	**9.109**	**<0.001**
Group	0.301	0.251	40	1.197	0.238
Linear: CV	0.204	0.234	45262	0.873	0.382
Linear: MV	0.625	0.245	45262	2.543	0.011
Quadratic: CV	**−4.844**	**0.25**	**45262**	**−19.354**	**<0.001**
Quadratic: MV	**−3.255**	**0.256**	**45262**	**−12.715**	**<0.001**
Cubic: CV	−0.287	0.244	45262	−1.174	0.240
Cubic: MV	−0.094	0.253	45262	−0.373	0.709
Linear: Group	0.583	0.65	45262	0.897	0.369
Quadratic: Group	0.266	0.518	45262	0.514	0.607
Cubic: Group	**1.023**	**0.326**	**45262**	**3.134**	**0.001**
CV: Group	**−1.149**	**0.2**	**1086**	**−5.738**	**<0.001**
MV: Group	**−1.386**	**0.202**	**1086**	**−6.839**	**<0.001**
Linear: CV: Group	−0.459	0.325	45262	−1.412	0.157
Linear: MV: Group	**−0.701**	**0.333**	**45262**	**−2.106**	**0.035**
Quadratic: CV: Group	**3.59**	**0.338**	**45262**	**10.616**	**<0.001**
Quadratic: MV: Group	**2.626**	**0.341**	**45262**	**7.683**	**<0.001**
Cubic: CV: Group	−0.114	0.334	45262	−0.342	0.732
Cubic: MV: Group	0.481	0.34	45262	1.413	0.157

Formula: log-odds ∼ (Linear + Quadratic + Cubic) × Condition × Group + [(Linear + Quadratic + Cubic)| Subject] + (1| Item). Conditions: Unrelated verbs (UV), closely related verbs (CV), moderately related verbs (MV). TD, typical development; DS, Down syndrome; *SE*, standard error; *df*, degrees of freedom. Bold values indicate significant effects.

group looked more than chance level in the UV condition from 4,600 to 5,400 ms ($t_{cluster} = 27.553$, $t_{max} = 4.082$, $p < 0.001$).

Participants with DS looked more in the UV than in the CV condition from 1,400 to 1,500 ms ($t_{cluster} = 5.589$, $t_{max} = 2.113$, $p = 0.006$), but more in the CV than in the UV condition from 3,800 to 3,900 ms ($t_{cluster} = 6.340$, $t_{max} = 3.402$, $p = 0.002$), and from 4,200 to 4,400 ms ($t_{cluster} = 7.923$, $t_{max} = 2.808$, $p < 0.001$). They also looked more in the UV than in the MV condition from 1,300 to 1,500 ms ($t_{cluster} = 8.562$, $t_{max} = 3.346$, $p < 0.001$). They looked more at the target than chance level in the CV condition from 2,800 to 6,000 ms ($t_{cluster} = 126.342$, $t_{max} = 7.0733$, $p < 0.001$), and more at the target than chance level in the MV condition from 5,200 to 6,000 ms ($t_{cluster} = 24.491$, $t_{max} = 3.665$, $p < 0.001$). They looked more at the target than chance level in the UV condition in three time clusters: from 1,300 to 1,600 ms ($t_{cluster} = 12.014$, $t_{max} = 3.397$, $p < 0.001$), from 3,600 to 4,000 ms ($t_{cluster} = 11.827$, $t_{max} = 2.542$, $p < 0.001$), and from 4,800 to 6,000 ms ($t_{cluster} = 70.208$, $t_{max} = 8.033$, $p < 0.001$).

Factors influencing prediction

The binomial mixed-effect analysis showed that the reference model replicated the main results of the temporal analysis (Table 5); the CV condition, but not the MV condition, had more predictive looks than the UV condition. The fixation probability was higher in the predictive conditions (CV and MV) than in the non-predictive ones (UV).

The model comparison found that including the factors of chronological age or production did not improve the fit (Table 6). Furthermore, all Bayes factors were <0.001, suggesting that the null hypothesis should be accepted. Thus, neither chronological age nor production were related to the prediction effect.

In contrast, mental age significantly improved the fit of the model (Table 6). Further exploration of the mental age model showed a significant interaction between the MV condition and mental age, indicating that the differences between the UV and MV conditions increase with mental age.

Notably, the slope of the results was negative (Table 7), indicating that participants looked less in the MV than in the UV condition as mental age increased. This result should be taken with caution because the Bayes factor provides evidence in favor of the null hypothesis (BF < 0.001).

The association strength between the verb and the expected noun also improved the fit of the model (Table 6), and it provided strong evidence in favor of the alternative hypothesis (BF = 2.98e + 10). Exploration of the model including association strength showed an interaction between association strength and both the CV and MV conditions, with a positive slope (Table 8). This result indicates that both predictive conditions showed more predictive looks than the UV condition with higher association strength.

TABLE 5 Model of the average prediction window for the DS group.

Fixed effects	β	*SE*	*z*	*P*
Intercept	−0.135	0.116	−1.167	0.243
CV	**0.458**	**0.115**	**3.986**	**<0.001**
MV	−0.119	0.161	−0.738	0.460

Formula: log odds ∼ Condition + (Cond| Subject) + (1| Item). Conditions: Unrelated verbs (UV), high-related verbs (CV), low-related verbs (MV). DS, Down syndrome; *SE*, standard error; *df*, degrees of freedom. Bold values indicate significant effects.

TABLE 6 Fit comparison of demographic models for the DS group.

Fixed effect structure	Ln(L)	X^2	*p*
Condition	−28103		
Condition × Chronological age	−28100	6.745	0.080
Condition × Mental age	**−28097**	**12.717**	**0.005**
Condition × Production	−28102	2.098	0.552
Condition × Association strength	**−28065**	**75.608**	**<0.001**

All models were compared directly with the reference model ($df = 3$). The dependent variable was the log odds ratio of fixation. The random structures were the subject and the slope of the condition, and the intercept of the Item. Condition: unrelated verb, closely related verb, moderately related verb. DS, Down syndrome. Ln(L), −2 times log-likelihood. Bold values indicate significant effects.

Discussion

We tested the prediction ability of young people with DS and a control group of children with TD, paired by verbal mental age of around 5 years, based on the relationship between a heard verb and a depicted pair of images representing target and distractor nouns. We also tested a group of adults with TD to corroborate the prediction effect expected in the other two groups. We presented three types of relationships between verbs and nouns embedded in sentences: closely related verb (CV; e.g., *to read—book*), moderately related verb (MV; e.g., *to wait—bus*), and unrelated verb (UV; e.g., *to arrive—dog*). We

TABLE 7 Model for mental age exploration in the DS group.

Fixed effects	β	*SE*	*z*	*p*
Intercept	−0.042	0.158	−0.266	0.789
CV	**0.263**	**0.168**	**1.568**	**0.116**
MV	**−0.656**	**0.198**	**−3.315**	**<0.001**
Mental age	0.333	0.393	0.848	0.396
CV: Mental age	−0.697	0.456	−1.525	0.127
MV: Mental age	**−1.916**	**0.538**	**−3.563**	**<0.001**

Formula: log odds ∼ Condition × Mental age + (Cond| Subject) + (1| Item). *SE*, standard error; *df*, degrees of freedom. Bold values indicate significant effects.

TABLE 8 Model for association strength exploration in the DS group.

Fixed effects	β	*SE*	*z*	*p*
Intercept	−0.153	0.114	−1.341	0.179
CV	0.191	0.132	1.446	0.148
MV	**0.707**	**0.199**	**3.553**	**<0.001**
Association Strength	**−0.434**	**0.197**	**−2.202**	**0.027**
CV: Association Strength	**1.019**	**0.216**	**4.716**	**<0.001**
MV: Association Strength	**2.424**	**0.34**	**7.115**	**<0.001**

Formula: log odds ∼ Condition × Association Strength + (Cond| Subject) 1| Item). Conditions: High-related verbs (CV), low-related verbs (MV). *SE*, standard error; *df*, degrees of freedom. Bold values indicate significant effects.

hypothesized that adults and children with TD would predict the intended target in both closely and moderately related sentences but not in unrelated pairs. In the case of participants with DS, we expected to capture prediction only with closely related verbs but not with moderately related or unrelated verbs. Finally, we expected that vocabulary production would play a significant role in prediction by participants with DS.

Our results corroborate our hypothesis for adults and children with TD. Both groups could anticipate the target before it was named, based on the level of relationship between the verb and the noun. In the case of young people with DS, we found an ability to predict only in closely related sentences, confirming their need for a high degree of relationship between verbs and nouns. We also found that their ability to predict was slow compared to children with TD: they took about 200 ms longer to anticipate the target noun. In all cases, preference for the labeled noun at the end of the noun window confirmed that participants followed the task. Our last hypothesis, positing a relationship between the level of the productive vocabulary of people with DS and their predictive ability, was not confirmed.

Our results show that participants with DS could anticipate the subsequent noun only in sentence constructions with a closely related verb, not in those with a moderately related verb, while the TD group showed linguistic anticipation skills with both closely and moderately related verbs. These results support the idea that different factors are involved in prediction, depending on the degree of relationship between the context and the upcoming word, as proposed by the theoretical prediction models (Pickering and Gambi, 2018). A higher degree of association between the verb and the noun makes the generation of linguistic predictions more likely, even though both sentence constructions are possible at the grammatical level and also predictable. The development of these differential factors associated with moderately related verbs could be delayed or impaired in DS participants but not in those with TD.

In their prediction theory, Pickering and Gambi (2018) postulate two prediction mechanisms: prediction-by-association and prediction-by-production. Although both predictive mechanisms are involved in the experimental condition, the sentences were designed to require different uses of each mechanism. The prediction-by-association mechanism is based on spreading activation between related concepts, is automatic, and uses fewer cognitive resources. Sentence prediction with a close verb-noun relationship is assumed to be supported mainly by this mechanism because the activation spreads strongly from the verb to the noun. The activation of the target may also produce lateral inhibition in unrelated elements of the lexicon (Chow et al., 2016; Angulo-Chavira and Arias-Trejo, 2021): in this case, the distractor. Our results in the closely related condition show that these mechanisms are relatively preserved in participants with DS, which is consistent with previous studies showing spreading activation between related nouns in this population (Barrón-Martínez and Arias-Trejo, 2020; Barrón-Martínez et al., 2020). Nevertheless, people with DS seem to present weak connections between related concepts: the magnitude and velocity of the predictions are less in the DS group than in the TD group. This explanation is plausible, at least for the connection between verbs and nouns, because children with DS present a similar spreading activation between nouns as their mental age peers (Barrón-Martínez and Arias-Trejo, 2020; Barrón-Martínez et al., 2020).

By contrast, prediction-by-production is efficient because it uses linguistic and non-linguistic contextual information and interaction with the speaker to make inferences about their intentions. This system is slow, uses a high level of cognitive resources, and is optional (Pickering and Gambi, 2018). We assume that our moderately related condition depends on prediction-by-production because participants needed to rely more on visual information to predict the target.[4] For example, there is more variability in the possible direct objects connected to the verb *fix* than to the verb *sweep*, which is closely related to *broom*; participants are thus forced to look for a *fixable* object and discard all *unfixable* objects based on the picture displayed (e.g., *washing machine* vs. *watermelon*). Adults and children with TD predicted the moderately related verb condition; however, in the TD group, they did so less in this condition than in the closely related condition, suggesting that the moderately related condition is harder to process, in line with the prediction-by-production hypothesis. Note that adults did not present a clear difference between the two, indicating that prediction-by-production improves during development, at least for a syntactically simple sentence with common words. It is possible that the prediction in the closely related and moderately related conditions behaved asymptotically, as in associative learning models (Plaut and Booth, 2000; Kapatsinski, 2021). This asymptotic behavior contributes to maintaining a degree of

4 Prediction-by-association could be involved, since there is a degree of association; however, its involvement should be small, since our free association task asked for a verb. It is possible that in typical association norms participants do not even mention these verbs.

uncertainty (Kapatsinski, 2021) and when the associations are weak (Plaut and Booth, 2000), there is a delay in approaching the asymptotic point, as seen in the TD children and adults in our study. Participants with DS possibly did not have enough resources to predict the moderately related condition. Since prediction-by-production is optional (Huettig, 2015; Huettig and Mani, 2016; Pickering and Gambi, 2018), the language system prioritizes the comprehension of bottom-up information over top-down prediction.

The question then arises as to what resources are necessary for people with DS to use prediction-by-production. To answer this question, we explored variables that could explain the individual difference in prediction, particularly chronological age, mental age, association strength, and production. We measured the influence of chronological age in prediction skills because older participants have more experience with language than younger ones; however, it seems that the ability to predict a referent in highly or less highly semantically related environments does not underlie this factor. This result does not mean that prediction is not dependent on experience in people with DS; in fact, the prediction of highly semantically related information indicates that they need very common word pairs to make predictions. The sentences with close relationships were also those that our plausibility study found to have higher probabilities of being heard. The association strength between the verb and the noun also facilitates prediction regardless of the condition. Thus, the frequency of the sentences and the frequency of the relationships may contribute to the prediction of a noun. Less common combinations of verbs and nouns also diminish predictive ability in young people with DS.

We also found that mental age influences prediction in people with DS in an unexpected direction: participants with DS with greater mental age looked less at the target in the MV condition. This result is contrary to that of Arias-Trejo et al. (2019), who found a positive correlation between mental age and the predictive ability of people with DS. Differences in the mental age evaluation might explain this discrepancy. Arias-Trejo et al. (2019) computed mental age by evaluating verbal and non-verbal cognitive domains. In the present study, mental age was based on comprehension ability. In other words, if comprehension skills do not determine the ability of people with DS to predict the upcoming noun, then more general cognitive skills may do.

Associative models show that an increase in vocabulary produces difficulties in word recognition because of the competition and addition of weak associations to the lexicon (Ramscar et al., 2014). This difficulty might be present in the predictive recovery of words. It is possible that people with DS had such difficulties related to the addition of new words to the lexicon. One mechanism that helps overcome competition problems is inhibition (e.g., McClelland and Elman, 1986), a mechanism developed in early childhood (Chow et al., 2016, 2019). People with DS may suppress weak associations to avoid the interference produced by the competition that increases with cognitive development. This interpretation is supported by the observation that people with DS predict better when the association strength is higher. Nevertheless, this is a speculative interpretation and should be taken with caution, not only because our experiment was not designed to prove this point but also because the Bayes factor provides evidence against the influence of verbal mental age on prediction skills.

It is hypothesized that predictions are made by the production system (Dell and Chang, 2014; Huettig, 2015; Huettig and Janse, 2016; Pickering and Gambi, 2018). For example, participants who scored better on productive vocabulary tests were those who also presented better linguistic anticipation skills (Mani and Huettig, 2012; Mani et al., 2016). In the present study, we found no influence of production, either in the closely related or moderately related verb conditions, in any group of participants.

The lack of a relationship between production and prediction could be interpreted as the production system not being involved in the generation of top-down predictions; however, this is unlikely in light of previous evidence (Martin et al., 2013; Dell and Chang, 2014; Huettig, 2015; Huettig and Janse, 2016; Pickering and Gambi, 2018). A second explanation is in the use of resources in prediction-by-production: people with DS have several cumulative factors that can hinder top-down predictions. Working memory problems and processing speed in people with DS are likely to interfere with the ability to predict upcoming linguistic information (Huettig and Janse, 2016; Ito et al., 2018). For example, Huettig and Janse (2016) found more predictive eye movements in the visual world paradigm in people with better working memory and faster processing speed. Participants with DS tended to have poor reading skills, which could hinder their ability to predict, as reported by Mishra (2012) for adults with low literacy and (Huettig and Brouwer, 2015) for Dutch adults with dyslexia. Thus, the lack of a relationship between production and top-down predictions in people with DS may be better explained by limitations in general cognition.

Limitations and future studies

The present study describes some prediction processes in people with DS; however, it is important to consider some of the study's limitations. First, the sample of participants is small; it is difficult to generalize our results to all populations with DS since there is a high degree of variability in their cognitive profiles. The sample size also affects the fixation data. There are unexpected but significant differences across the trial: a slight preference for the unrelated condition in the pre-verb and verb windows in the DS group and a slight preference in the pre-verb window for the closely related verb condition. We assume that these differences result from the small sample because there is no consistency in the presentation of these clusters across groups or conditions; it can thus be interpreted as a random preference created by the high variability of our data.

Another limitation is that we evaluated only receptive and expressive vocabulary because the COVID-19 pandemic required us to administer the assessments online. Future studies must explore more general cognitive skills, such as working memory and processing speed, to better explain the factors underlying DS prediction. Receptive vocabulary as a measure of verbal mental age could also be insufficient; further research should measure additional language skills or general cognitive development.

People with DS also have a high prevalence of nystagmus, which affects ocular control (Mathan et al., 2022). Given the online nature of our study, we relied on parents for information about possible problems with vision and hearing. Although this bias would be a constant in the within-subject comparisons, it is necessary to consider the problem in the between-subject comparisons and also consider more robust measures of ocular problems in the population with DS.

Conclusion

This study evaluated prediction skills in people with DS using a preferential-looking task. It provides evidence that young people with DS can anticipate upcoming information based on the semantic relatedness between a verb and a noun. Participants with DS predicted nouns in closely related verb-noun pairs but not in pairs that were only moderately related and in which they needed visual context to generate the prediction. These effects are not explained by chronological age, mental age, or productive vocabulary. These results suggest that in people with DS, prediction is driven by association; this offers clues about how people in this group process and extract information from speech and in context. By studying the mechanism that allows this, we can better understand how this population uses it to learn more rapidly in situations varying in context and how established predictions can be used to promote learning. Our findings support an ecological and feasible evaluation tool for the systematic measurement of lexical prediction in people with DS, useful for understanding the cognitive mechanisms of lexical prediction and how these mechanisms can be strengthened through the implementation of stimulation programs.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Comité de Ética, Facultad de Psicología, UNAM. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AA-C: conception of the original idea, experimental design, data analysis, data processing, stimulus creation, revision, and critical writing of the manuscript. AC-F: experimental design, conceptualization of the final experimental design, creation of stimuli, data processing, search, and attention of participants, and writing of the manuscript. JB-M: search and attention of participants, and writing of the manuscript. NA-T: funding acquisition, administration and supervision of the project, conceptualization of experimental design, revision, and critical writing of the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank the PAPIIT project, UNAM IN303221 "Efecto de la restricción oracional y la similitud de palabras en la actividad electroencefalográfica anticipatoria."

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.934826/full#supplementary-material

References

Altmann, G. T. M., and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73, 247–264. doi: 10.1016/S0010-0277(99)00059-1

Alva-Canto, E. A. (2001). Cómo usan los niños las palabras. El uso de los derivados de las palabras en el lenguaje espontáneo de los niños en interacción libre entre iguales. *Facultad de Psicología, Univ. Nacional Autónoma de México* 1, 19–173.

Anderson, J. R. (2013). A spreading activation theory of memory. *Read. Cogn. Sci. Perspect. Psychol. Artifi. Intellig.* 2013, 137–154. doi: 10.1016/B978-1-4832-1446-7.50016-9

Andreou, G., and Katsarou, D. (2016). *Semantic Processing In Children With Down Syndrome*. 59–66. Available online at: https://doi.org/10.26262/istal.v21i0.5208 (accessed January 23, 2022).

Angulo-Chavira, A. Q., and Arias-Trejo, N. (2021). Mediated semantic priming interference in toddlers as seen through pupil dynamics. *J. Exp. Child Psychol.* 208:105146. doi: 10.1016/j.jecp.2021.105146

Arias-Trejo, N., Angulo-Chavira, A. Q., and Barrón-Martínez, J. B. (2019). Verb-mediated anticipatory eye movements in people with down syndrome. *Int. J. Lang. Commun. Dis.* 54, 756–766. doi: 10.1111/1460-6984.12473

Arias-Trejo, N., Falcón, A., and Alva-Canto, E. A. (2013). The gender puzzle: Toddlers' use of articles to access noun information. *Psicologica* 34, 1–23.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *J. Memory Lang.* 59, 457–474. doi: 10.1016/j.jml.2007.09.002

Barrón-Martínez, J. B., and Arias-Trejo, N. (2020). Perceptual similarity effect in people with down syndrome. *Int. J. Dev. Disabili.* 68, 182–189. doi: 10.1080/20473869.2020.1729016

Barrón-Martínez, J. B., and Arias-Trejo, N. (2022). Perceptual similarity effect in people with down syndrome. *Int. J. Dev. Disabili.* 68, 182–189.

Barrón-Martínez, J. B., Arias-Trejo, N., and Salvador-Cruz, J. (2020). Associative lexical relationships in children with down syndrome. *Int. J. Disabili. Dev. Educ.* 2020, 1–13. doi: 10.1016/j.jcomdis.2020.105975

Borovsky, A., Elman, J. L., and Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *J. Exp. Child Psychol.* 112, 417–436. doi: 10.1016/j.jecp.2012.01.005

Campanelli, L., van Dyke, J. A., and Marton, K. (2018). *CUNY Academic Works The Modulatory Effect of Expectations on Memory Retrieval During Sentence Comprehension How Does Access To This Work Benefit You? Let Us Know! The Modulatory Effect of Expectations on Memory Retrieval During Sentence Comprehension*. New York City, NY: City University New York Graduated Center.

Chapman, R. S. (2006). Language learning in down syndrome: The speech and language profile compared to adolescents with cognitive impairment of unknown origin. *Down Syndrome Res. Pract. J. Sarah Duffen Centre Univ. Portsmouth* 10, 61–66. doi: 10.3104/reports.306

Chapman, R. S., Seung, H. K., Schwartz, S. E., and Bird, E. K. (2000). Predicting language production in children and adolescents with down syndrome: The role of comprehension. *J. Speech Lang. Hear. Res.* 43, 343–350. doi: 10.1044/jslhr.4302.340

Checa, E., Galeote, M., and Soto, P. (2016). The Composition of early vocabulary in Spanish children with down syndrome and their peers with typical development. *Am. J. Speech Lang. Pathol.* 25, 605–619. doi: 10.1044/2016_AJSLP-15-0095

Cheung, R. W., Hartley, C., and Monaghan, P. (2022). Receptive and expressive language ability differentially support symbolic understanding over time: Picture comprehension in late talking and typically developing children. *J. Exp. Child Psychol.* 214:105305. doi: 10.1016/j.jecp.2021.105305

Chow, J., Aimola Davies, A. M., Fuentes, L. J., and Plunkett, K. (2016). Backward semantic inhibition in toddlers. *Psychol. Sci.* 27, 1312–1320. doi: 10.1177/0956797616659766

Chow, J., Aimola Davies, A. M., Fuentes, L. J., and Plunkett, K. (2019). The vocabulary spurt predicts the emergence of backward semantic inhibition in 18-month-old toddlers. *Dev. Sci.* 22:12754. doi: 10.1111/desc.12754

Cognition Run (2021). *Cogntion Run Cognitive Experiments Online*. Available online at: https://www.cognition.run/ (accessed October 10, 2021).

Collins, A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428. doi: 10.1037/0033-295X.82.6.407

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283

Dell, G. S., and Chang, F. (2014). The p-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philos. Trans. R. Soc. Biol. Sci.* 369:394. doi: 10.1098/rstb.2012.0394

DeLong, C., Nessler, C., Wright, S., and Wambaugh, J. (2015). Semantic feature analysis: Further examination of outcomes. *Am. J. Speech Lang. Pathol.* 24, 1–14. doi: 10.1044/2015_AJSLP-14-0155

Dierssen, M. (2012). Down syndrome: The brain in trisomic mode. *Nat. Rev. Neurosci.* 13, 844–858. doi: 10.1038/nrn3314

Dodd, B., and Thompson, L. (2001). Speech disorder in children with Down's syndrome. *J. Intellect. Disabili. Res.* 45, 308–316. doi: 10.1046/j.1365-2788.2001.00327.x

Eadie, P. A., Fey, M. E., Douglas, J. M., and Parsons, C. L. (2002). Profiles of grammatical morphology and sentence imitation in children with specific language impairment and down syndrome. *J. Speech Lang. Hear. Res.* 45, 720–732. doi: 10.1044/1092-4388(2002/058)

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology* 44, 491–505. doi: 10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., and Kutas, M. (1999). A rose by any other name: long-term memory structure and sentence processing. *J. Memory Lang.* 41, 469–495. doi: 10.1006/jmla.1999.2660

Ferretti, T. R., McRae, K., and Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *J. Memory Lang.* 44, 516–547. doi: 10.1006/jmla.2000.2728

Fine, A. B., Jaeger, T. F., Farmer, T. A., and Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS One* 8:77661. doi: 10.1371/journal.pone.0077661

Galeote, M., Sebastián, E., Checa, E., Rey, R., and Soto, P. (2011). The development of vocabulary in Spanish children with down syndrome: Comprehension, production, and gestures. *J. Intellect. Dev. Disabili.* 36, 184–196. doi: 10.3109/13668250.2011.599317

Gambi, C., Gorrie, F., Pickering, M. J., and Rabagliati, H. (2018). The development of linguistic prediction: Predictions of sound and meaning in 2- to 5-year-olds. *J. Exp. Child Psychol.* 173, 351–370. doi: 10.1016/j.jecp.2018.04.012

Hsu, C. F. (2019). Contextual effects on semantic grouping in individuals with down syndrome. *Int. J. Dev. Disabilit.* 65, 65–72. doi: 10.1080/20473869.2017.1353659

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Res.* 1626, 118–135. doi: 10.1016/j.brainres.2015.02.014

Huettig, F., and Brouwer, S. (2015). Delayed anticipatory spoken language processing in adults with dyslexia - evidence from eye-tracking. *Dyslexia* 21, 97–122. doi: 10.1002/dys.1497

Huettig, F., and Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Lang. Cogn. Neurosci.* 31, 80–93. doi: 10.1080/23273798.2015.1047459

Huettig, F., and Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Lang. Cogn. Neurosci.* 31, 19–31. doi: 10.1080/23273798.2015.1072223

Ito, A., Corley, M., and Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism* 21, 251–264. doi: 10.1017/S1366728917000050

Ito, K., and Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *J. Memory Lang.* 58, 541–573. doi: 10.1016/j.jml.2007.06.013

Iverson, J. M., Longobardi, E., and Caselli, M. C. (2003). Relationship between gestures and words in children with Down's syndrome and typically developing children in the early stages of communicative development. *Int. J. Lang. Commun. Dis.* 38, 179–197. doi: 10.1080/1368282031000062891

Kamide, Y., Altmann, G. T. M., and Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *J. Memory Lang.* 49, 133–156. doi: 10.3389/fpsyg.2016.00150

Kapatsinski, V. (2021). Learning fast while avoiding spurious excitement and overcoming cue competition requires setting unachievable goals: Reasons for using the logistic activation function in learning to predict categorical outcomes. *Lang. Cogn. Neurosci.* [Preprint]. doi: 10.1080/23273798.2021.1927120

Kumin, L., Councill, C., and Goodman, M. (1994). A longitudinal study of the emergence of phonemes in children with down syndrome. *J. Commun. Disord.* 27, 293–303. doi: 10.1016/0021-9924(94)90019-1

Kuperberg, G. R., and Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Lang. Cogn. Neurosci.* 31, 32–59. doi: 10.1080/23273798.2015.1102299

Laws, G., Briscoe, J., Ang, S. Y., Brown, H., Hermena, E., and Kapikian, A. (2015). Receptive vocabulary and semantic knowledge in children with SLI and children with down syndrome. *Child Neuropsychol.* 21, 490–508. doi: 10.1080/09297049.2014.917619

Lubec, G., and Engidawork, E. (2002). The brain in down syndrome (trisomy 21). *J. Neurol.* 249, 1347–1356. doi: 10.1007/s00415-002-0799-9

Lukyanenko, C., and Fisher, C. (2016). Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition* 146, 349–370. doi: 10.1016/j.cognition.2015.10.012

Makowski, D., Ben-Shachar, M., and Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *J. Open Source Softw.* 4:1541. doi: 10.21105/joss.01541

Mani, N., and Huettig, F. (2012). Prediction during language processing is a piece of cake-but only for skilled producers. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 843–847. doi: 10.1037/a0029284

Mani, N., and Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *J. Exp. Child Psychol.* 126, 264–279. doi: 10.1016/j.jecp.2014.05.004

Mani, N., Daum, M. M., and Huettig, F. (2016). 'Proactive' in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. *Quart. J. Exp. Psychol.* 69, 2189–2201. doi: 10.1080/17470218.2015.1111395

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024

Martin, C. D., Branzi, F. M., and Bar, M. (2018). Prediction is production: The missing link between language production and comprehension. *Sci. Rep.* 8, 1–9. doi: 10.1038/s41598-018-19499-4

Martin, G. E., Losh, M., Estigarribia, B., Sideris, J., and Roberts, J. (2013). Longitudinal profiles of expressive vocabulary, syntax and pragmatic language in boys with fragile X syndrome or down syndrome. *Int. J. Lang. Commun. Dis.* 48, 432–443. doi: 10.1111/1460-6984.12019

Martin, K. I., and Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Stud. Second Lang. Acquisit.* 34, 379–413. doi: 10.1017/S0272263112000125

Martin, N. A. (2010b). *ROWPVT-4: Receptive One-Word Picture Vocabulary Test*. Novato, CA: ATP: Assesment a division of Academic Therapy of Publications.

Martin, N. A. (2010a). *EOWPVT-4: Expressive One-Word Picture Vocabulary Test*. Novato, CA: ATP: Assesment a division of Academic Therapy of Publications.

Mathan, J. J., Simkin, S. K., Gokul, A., and McGhee, C. N. J. (2022). Down syndrome and the eye: Ocular characteristics and ocular assessment. *Survey Ophthalmol.* [Preprint]. doi: 10.1016/j.survophthal.2022.03.006

McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Psychol.* 18, 1–86. doi: 10.1016/0010-0285(86)90015-0

McRae, K., Ferretti, T. R., and Amyote, L. (1997). Thematic roles as verb-specific concepts. *Lang. Cogn. Proc.* 12, 137–176. doi: 10.1080/016909697386835

Michael, S. E., Ratner, N. B., and Newman, R. (2012). Verb comprehension and use in children and adults with down syndrome. *J. Speech Lang. Hear. Res.* 55, 1736–1749. doi: 10.1044/1092-4388(2012/11-0050)

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. Taylor & Francis: Chapman & Hall.

Mishra, R. K. (2012). Spoken language-mediated anticipatory eyemovements are modulated by reading ability - evidence from Indian low and high literates. *J. Eye Move. Res.* 5, 1–10. doi: 10.16910/jemr.5.1.3

Næss, K. A. B. (2022). A randomized trial of the digital down syndrome languageplus (DSL+) vocabulary intervention program. *Remed. Special Educ.* [Preprint]. doi: 10.1177/07419325211058400

Næss, K. A. B., Lyster, S. A., Hulme, C., and Melby-Lervåg, M. (2011). Language and verbal short-term memory skills in children with down syndrome: A meta-analytic review. *Res. Dev. Disabili.* 32, 2225–2234. doi: 10.1016/j.ridd.2011.05.014

Pelucchi, B., Hay, J. F., and Saffran, J. R. (2009). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition* 113, 244–247. doi: 10.1016/j.cognition.2009.07.011

Pickering, M. J., and Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychol. Bull.* 144, 1002–1044. doi: 10.1037/bul0000158

Plaut, D. C., and Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychol. Rev.* 107, 786–823. doi: 10.1037/0033-295x.107.4.786

R Core Team (2019). *R: A Language And Environment For Statistical Computing, Reference Index Version 3.5.3*.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., and Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Top. Cogn. Sci.* 6, 5–42. doi: 10.1111/tops.12078

RealEye (2020). *RealEye Eye-Tracking System Technology Whitepaper*. Available online at: https://d48dl6cezsco5.cloudfront.net/static/doc/RealEye+-+Technical+Whitepaper+v1.1.0.pdf (accessed July 27, 2022).

Reuter, T., Borovsky, A., and Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Dev. Psychol.* 55, 1656–1665. doi: 10.1037/dev0000754

Ricketts, J., Nation, K., and Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Sci. Stud. Read.* 11, 235–257. doi: 10.1080/10888430701344306

Roberts, J., Long, S. H., Malkin, C., Barnes, E., Skinner, M., Hennon, E. A., et al. (2005). A comparison of phonological skills of boys with fragile X syndrome and down syndrome. *J. Speech Lang. Hear. Res.* 48, 980–995. doi: 10.1044/1092-4388(2005/067)

Roberts, J., Martin, G. E., Moskowitz, L., Harris, A. A., Foreman, J., and Nelson, L. (2007). Discourse skills of boys with fragile X syndrome in comparison to boys with down syndrome. *J. Speech Lang. Hear. Res.* 50, 475–492. doi: 10.1044/1092-4388(2007/033)

Rommers, J., Meyer, A. S., and Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attent. Percept. Psychophys.* 77, 720–730. doi: 10.3758/s13414-015-0873-x

Rosin, M. M. (1988). Communication profiles of adolescents with down syndrome. *Commun. Dis. Quart.* 12, 49–64. doi: 10.1177/152574018801200105

Schoknecht, P. (2022). The interaction of predictive processing and similarity-based retrieval interference: An ERP study. *Lang. Cogn. Neurosci.* 2022, 1–19. doi: 10.1080/23273798.2022.2026421

Stefaniak, N., Baltazart, V., and Declercq, C. (2021). Processing verb meanings and the declarative/procedural model: A developmental study. *Front. Psychol.* 12:714523. doi: 10.3389/fpsyg.2021.714523

Stolt, S., Lind, A., Matomäki, J., Haataja, L., Lapinleimu, H., and Lehtonen, L. (2016). Do the early development of gestures and receptive and expressive language predict language skills at 5;0 in prematurely born very-low-birth-weight children? *J. Commun. Dis.* 61, 16–28. doi: 10.1016/j.jcomdis.2016.03.002

van Gameren-Oosterom, H. B. M., Fekkes, M., Buitendijk, S. E., Mohangoo, A. D., Bruil, J., and Van Wouwe, J. P. (2011). Development, problem behavior, and quality of life in a population based sample of eight-year-old children with down syndrome. *PLoS One* 6:16–20. doi: 10.1371/journal.pone.0021879

Wass, S. V., Forssman, L., and Leppänen, J. (2014). Robustness and precision: How data quality may influence key dependent variables in infant eye-tracker analyses. *Infancy* 19, 427–460. doi: 10.1111/infa.12055

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., and Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspect. Psychol. Sci.* 6, 291–298. doi: 10.1177/1745691611406923

Witecy, B., and Penke, M. (2017). Language comprehension in children, adolescents, and adults with down syndrome. *Res. Dev. Disabili.* 62, 184–196. doi: 10.1016/j.ridd.2017.01.014

Yoder, P. J., Camarata, S., Camarata, M., and Williams, S. M. (2006). Association between differentiated processing of syllables and comprehension of grammatical morphology in children with down syndrome. *Am. Assoc. Mental Retardat.* 111, 138–152. doi: 10.1352/0895-8017(2006)111[138:ABDPOS]2.0.CO;2

frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 18 October 2022
DOI 10.3389/fcomm.2022.977724

Check for updates

OPEN ACCESS

EDITED BY
Leonard Abbeduto,
University of California, Davis,
United States

REVIEWED BY
Meagan Talbott,
University of California, Davis,
United States
Laura J. Mattie,
University of Illinois at
Urbana-Champaign, United States
Angela John Thurman,
University of California, Davis,
United States

*CORRESPONDENCE
Shruthi Ravi
Shruthi.Ravi1@utdallas.edu

†These authors share senior authorship

SPECIALTY SECTION
This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

RECEIVED 24 June 2022
ACCEPTED 23 September 2022
PUBLISHED 18 October 2022

CITATION
Ravi S, Bradshaw A, Abdi H, Meera SS, Parish-Morris J, Yankowitz L, Paterson S, Dager SR, Burrows CA, Chappell C, St.John T, Estes AM, Piven J, Swanson MR and the IBIS Network (2022) Are early social communication skills a harbinger for language development in infants later diagnosed autistic?—A longitudinal study using a standardized social communication assessment. *Front. Commun.* 7:977724.
doi: 10.3389/fcomm.2022.977724

Are early social communication skills a harbinger for language development in infants later diagnosed autistic?—A longitudinal study using a standardized social communication assessment

Shruthi Ravi[1]*, Allison Bradshaw[1], Hervé Abdi[1], Shoba Sreenath Meera[2], Julia Parish-Morris[3], Lisa Yankowitz[4], Sarah Paterson[5], Stephen R. Dager[6], Catherine A. Burrows[7], Chad Chappell[8], Tanya St.John[6], Annette M. Estes[6], Joseph Piven[8†], Meghan R. Swanson[1†] and the IBIS Network

[1]Department of Psychology, The University of Texas at Dallas, Richardson, TX, United States, [2]Department of Speech Pathology & Audiology, National Institute of Mental Health and Neurosciences, Bangalore, India, [3]Center for Autism Research, Children's Hospital of Philadelphia, Philadelphia, PA, United States, [4]Boston Children's Hospital, Harvard Medical School, Boston, MA, United States, [5]The James S. McDonnell Foundation, St. Louis, MO, United States, [6]Department of Radiology, University of Washington, Seattle, WA, United States, [7]Department of Pediatrics, University of Minnesota, Minneapolis, MN, United States, [8]Department of Psychiatry, University of North Carolina, Chapel Hill, NC, United States

The early emergence of social communication challenges and their impact on language in infants later diagnosed with autism has sparked many early intervention programs that target social communication skills. While research has consistently shown lower scores on social communication assessments in the first year of life, there is limited research at 12-months exploring associations between different dimensions of social communication and later language. Understanding associations between early social communication skills and language would enhance our ability to choose high priority intervention goals that will impact downstream language skills. The current study used a standardized assessment to profile social communication skills across 516 infants with a high (HL) or low likelihood (LL-Neg) for autism (84% White, 60% Male), based on the presence of a sibling with autism in the family. The primary aim of the study was to profile social communication skill development in the second year of life and to evaluate associations between social communication skills and later language. HL infants who met criteria for autism (HL-ASD, $N = 81$) demonstrated widespread reductions in social communication skills at 12-months compared to HL infants who did not meet criteria for autism (HL-Neg, $N = 277$) and LL-Neg ($N = 158$) infants. Across all infants in the study, those with better social communication skills at 12-months had better language at 24-months. However, within group analyses indicated that infants who met criteria for autism did not show this developmental coupling until 24-months-of-age at which point social communication was positively associated with downstream language skills. The cascading pattern of reduced social communication skills as well as overall significant positive associations with later language provide further evidence for the need to support developing social communication skills prior to formal autism diagnosis, a goal that could possibly be reached through pre-emptive interventions.

KEYWORDS

autism, language, social communication, longitudinal, infancy

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition, characterized by restricted, repetitive patterns of behavior and challenges in social communication (DSM-5; American Psychiatric Association, 2013). By 9- to 12-months of age, infants, who are later diagnosed autistic score lower on social communication assessments (Zwaigenbaum et al., 2005; Landa et al., 2007; Ozonoff et al., 2010; Bradshaw et al., 2020). During the early pre-diagnostic period, infants later diagnosed with autism remain consistently low, or make fewer gains on social communication skills, when compared to infants who do not meet criteria for autism (Bradshaw et al., 2021). For most children, these early observable behavioral differences in social communication consolidate into a diagnosable behavioral phenotype of autism around 24- to 36-months-of-age (Piven et al., 2017; Grzadzinski et al., 2021).

Early social communication skills are a harbinger for later language development (Wetherby et al., 2007; Delehanty et al., 2018). The cascading pattern of social communication challenges in the first two years of life experienced by infants who meet criteria for autism highlights the need to strengthen these skills (Bradshaw et al., 2021). However, current interventions for autism are typically provided following formal diagnoses. While these interventions have demonstrated some promise with improving social communication skills in autistic toddlers (Sandbank et al., 2020), there is a growing need to provide interventions that focus on early developmental trajectories rather than as a reaction to an autism diagnosis (Green et al., 2022).

Pre-emptive interventions

Interventions provided before a diagnosis are referred to as pre-emptive interventions. Pre-emptive interventions have been reported to be feasible and acceptable by families (Green et al., 2013). Recent evidence suggests that parents can effectively implement strategies taught through pre-emptive interventions, and parent fidelity in turn leads to better child social communication outcomes (Hampton and Rodriguez, 2021; Yoder et al., 2021). Pre-symptomatic interventions have been provided for infants who have a higher familial likelihood for autism (i.e., infants who have an older autistic sibling) or for those who show early symptoms of autism (Green, 2020). These interventions have focused on social communication skills such as social play, joint attention, symbolic play, and infant vocalizations (Hampton and Rodriguez, 2021). Research has also demonstrated that parent-mediated social communication interventions have significant overall positive effects on autism symptoms during the pre-diagnostic period (Green et al., 2017), as well as autism symptom severity following a formal diagnosis (Whitehouse et al., 2021). Together, this body of research suggests that working on early social communication skills during the pre-diagnostic period might shift developmental trajectories for autistic infants.

While preemptive interventions are effective in teaching parents strategies to improve social communication outcomes, it is important to consider the impact of these interventions on language development trajectories. Fostering better language skills has been identified as a priority for interventions by the autism community (Kapp, 2020; Green et al., 2022). A few studies have explored long-term language outcomes following pre-emptive interventions, and results have been mixed (Green et al., 2017; Whitehouse et al., 2021; Yoder et al., 2021). However, proximal significant effects on parent fidelity have been found to mediate language outcomes (Watson et al., 2017; Yoder et al., 2021). More recently, a pre-emptive intervention program focused on early social communication skills resulted in significant improvements in parent-reported measures of vocabulary (Whitehouse et al., 2021). In summary, social communication skills are often targeted in pre-emptive intervention programs as they have been found to relate to better downstream language skills (Yoder et al., 2015; Delehanty et al., 2018). However, results from preemptive intervention studies suggest that long-term effects on language are mixed.

Associations between social communication skills and downstream language in autistic toddlers

Autistic toddlers with better social communication skills measured in the second year of life also have better later language skills (Toth et al., 2006; Wetherby et al., 2007; Yoder et al., 2015; Delehanty et al., 2018). More specifically, studies have reported significant positive associations between social communication skills such as joint attention and goal-directed (i.e., intentional communication) communicative acts and later receptive and expressive language (Toth et al., 2006; Yoder et al., 2015; Delehanty et al., 2018). Intentional communication, and frequency of intentional communicative acts were positively associated with receptive and expressive language abilities (Delehanty et al., 2018). Superior communication skills related to speech (i.e., number of consonants and words used in communicative acts) in the second year of life were found to be associated with better expressive language abilities (but not receptive language abilities) at 3 years of age (Delehanty et al., 2018). However, non-verbal communication abilities (i.e., use of gestures) was found to be positively associated with receptive *and* expressive language abilities (Delehanty et al., 2018). Further, better symbolic play (i.e., using an object to represent something else such as using a block to represent a phone) and receptive vocabulary skills between 18- and 24-months-of-age were associated with lower receptive and expressive language scores a year later (Delehanty et al., 2018).

While research has reported significant positive associations between social communication skills measured in the first year of life and later language, most of the existing studies focused on a limited number of specific skills, such as joint attention (Bottema-Beutel, 2016) and use of gestures (Choi et al., 2020). Much of the research that has explored associations to later language across multiple dimensions of social communication have primarily focused on associations in the second year of life (Toth et al., 2006; Wetherby et al., 2007; Yoder et al., 2015; Delehanty et al., 2018). The current study aims to measure associations between seven specific categories of social communication skills, and later language. A fine-grained analysis will inform the selection of intervention targets that will have the most impact on later language. In addition, the current study aims to explore associations to language skills measured at 24-and-36-months-of-age. Previous research has been based on language outcomes measured at a single time point (Wetherby et al., 2007; Yoder et al., 2015; Delehanty et al., 2018). The additional language data at will enable us to measure changes in social communication and language associations over time. Overall, the present study aims to examine the timing and nature of associations between specific features of social communication and language in infants who go on to receive and autism diagnosis.

The present study is part of two multisite Infant Brain Imaging (IBIS) Network studies that prospectively followed three groups of infants: (a) typically developing infants with a low likelihood for developing autism (LL-Neg), (b) infants who have a family-history of autism but do not develop autism themselves (HL-Neg), and (c) infants who have a family-history of autism and who go on to have autism (HL-ASD). This study design provides the opportunity to prospectively explore social communication skills and their association to later language. The purposes of this study were: (a) to explore developmental trajectories of social communication skills across the three groups of infants, (b) to identify differences in social communication skills at 12-months and 24-months across the three groups of infants, (c) to explore temporal relationships between social communication skills measured at 12-and-24-months-of-age and later language, measured at 24-and-36-months-of-age and (d) to understand how social communication skills predict autism and language diagnostic outcomes.

Methods

Participants

This study included 516 infants form two IBIS studies. Data for the IBIS 1 study were collected between 2007 and 2012; and data for the IBIS 2 study were collected between 2012 and 2018. Data were collected at four sites: University of North Carolina at Chapel Hill; University of Washington; The Children's Hospital of Philadelphia; and Washington University in St. Louis Data for the current study were collected between January 10th, 2008, and February 19th, 2018. Procedures for this study were approved by local Institutional Review Boards. Written informed consent was obtained from parents prior to participation.

All participants were screened, and exclusions were made for the following reasons: (1) genetic conditions or syndromes, (2) medical/neurological conditions affecting growth, development, or cognition (e.g., seizure disorder) or significant sensory impairments (e.g., vision or hearing loss), (3) birth weight <2000 g and/or gestational age <36 weeks or significant perinatal adversity and/or exposure in utero to neurotoxins, (4) contraindication for MRI, (5) predominant home language other than English, (6) adopted children or half siblings, (7) first-degree relative with psychosis, schizophrenia, or bipolar disorder (Family Interview for Genetic Studies (FIGS; Maxwell, 1992), and (8) twins.

For the IBIS 1 study, data were collected when infants were 6, 12, and 24-months-of-age. A detailed description of the IBIS 1 data collection protocol can be found in Estes et al. (2015). IBSI 2 had a variable visit schedule, where infants were seen at four of the following time points: 3, 6, 9, 12, 15, and 24. Infants from these two studies were included in the current study if they contributed at least one social communication data point and had diagnostic data at 24-months-of-age. We used 24-month diagnostic classification because a 36-month time point was not conducted across the full sample. Previous research has suggested strong diagnostic stability for autism from 24 to 36 months of age (Lord, 1995; Lord et al., 2006; e.g., Chawarska et al., 2009; Corsello et al., 2013; Guthrie et al., 2013; Shen et al., 2013; Barbaro and Dissanayake, 2017).

Infants were classified as autistic at 24-months if they met DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders, edition IV, Text Revision; American Psychiatric Association, 2000) criteria for autistic disorder or PDD-NOS. A clinical best-estimate diagnosis of autism was made using the DSM-IV-TR criteria by expert clinicians. Clinicians used all available developmental, clinical, and parent reported measures available at 24-months to determine the diagnostic classification for each participant. These measures included the Autism Diagnostic Observation Scale (ADOS; Lord et al., 2000), The Mullen Scales of Early Learning (MSEL; Mullen, 1995), and the Vineland Adaptive Behavior Scales, Second Edition (Vineland; Sparrow et al., 2005).

Infant participants were assigned to three groups based on familial history status and diagnostic outcome. Infants who had a high likelihood for autism, by virtue of having an older sibling with autism, and met criteria for autism were assigned to the HL-ASD group ($N = 81$). Infants who had a high likelihood for autism and did not meet criteria for autism were assigned to the HL-Neg group ($N = 277$). Infants with a low likelihood for autism, who did not meet criteria for autism were assigned to the LL-Neg group ($N = 158$).

TABLE 1 Overview of social communication skills measures in the CSBS composites.

Composite	Cluster	Skills measured
Social	Emotion and eye gaze	Shifting gaze between object and communicative partner Sharing positive affect with communicative partner Responding to joint attention (RJA)
	Communication	Frequency and variety of intentional communicative acts (e.g., requesting, refusing, seeking comfort)
	Gestures	Use of conventional gestures (e.g., pointing, nodding, showing) Use of distal gestures (i.e., gestures that do not involve touching an object or person)
Speech	Sounds	Use of sounds in communicative acts
	Words	Use of words in communicative acts
Symbolic	Understanding	Comprehension of object names, body part names, and person names (i.e., receptive vocabulary)
	Object Use	Use of objects during symbolic play Constructive play (i.e., stacking blocks)

Procedures and measures

Communication and symbolic behavior scales, developmental profile

The behavior sample of the CSBS (Wetherby and Prizant, 2002) was administered at 12, 15, and 24-months-of-age to assess early social communication skills. This standardized assessment uses attractive manipulatives to enable direct observations of natural play. The CSBS uses the following strategies to elicit social communication skills: communicative temptations, book sharing, pretend play, and constructive play. Administrations were videotaped and coded based on the CSBS manual. CSBS weighted raw scores for each cluster were extracted for the analyses. Raw scores were used to avoid floor effects in standard scores. The cluster scores included: emotion and eye gaze, communicative acts, use of gestures, use of sounds, use of words, understanding, and object use. Table 1 describes skills measured under each cluster of the CSBS.

All coders were trained based on guidelines described in the CSBS manual. Coders first reviewed coded practice videos with a trained coder. Next, they coded practice videos independently until they achieved 80% reliability with gold standard scoring. The gold-standard coding video was originally coded by a clinician with expertise in CSBS coding. Approximately 5% of the videos ($N = 25$) were double coded, and the coders had 86 % agreement on average ($SD = 3.83$).

Through the coding process, administration errors were identified in the symbolic and social composites, which impacted the following clusters: emotion and eye gaze, communication, gestures, understanding, and object use. Videos with administration errors were excluded from analyses for the clusters with incorrectly administered composite(s), resulting in different data sets for each CSBS score (see Tables S1 and S2 in Supplementary methods for additional information on administration errors).

Developmental and language measures

The MSEL (Mullen, 1995) is a standardized, direct assessment of cognitive functioning. It was administered at 12 and 24 months-of-age. A subset of infants participated in the MSEL when they were 36 months old. Subscale raw scores for receptive language (RL) and expressive language (EL) were extracted at the 24- and 36-month timepoints. Raw scores were used to avoid floor effects in standardized scores. Mullen *T*-scores were used as follows to generate two groups within the HL sample based on if the infants showed signs of early language delay (HL-Language Delay vs. HL-No Delay). These groups were made irrespective of ASD diagnostic status. Infants were determined to have signs of early language delay if their RL or EL *T-scores* fell 1.5 standard deviations below the mean (i.e. *T-scores* <= 35) (Swanson et al., 2017; Marrus et al., 2018). The Non-Verbal Developmental Quotient (NVDQ) was computed by averaging the age equivalent scores from the fine motor and visual reception subscales at 24-months to measure non-verbal cognitive skills. The Early Learning Composite (MSEL-ELC) score was not used in analyses but is reported to provide an overall description of developmental functioning across the sample at 12, 24, and 36-months-of-age.

A second assessment, the Vineland (Sparrow et al., 2005) was used to measure receptive and expressive language. The Vineland, a standardized measure of adaptive functioning, was used to provide a parent-reported measure of language abilities. It was administered at 12- and 4-months-of-age *via* parent interviews. The Vineland evaluates adaptive functioning across the following domains: communication, daily living skills, socialization, and motor. Overall adaptive behavior was evaluated at 12- and 24-months of age using the Adaptive Behavior Composite (ABC). The ABC was not included in the analyses, but is reported to provide a description of adaptive functioning across the groups. EL and RL raw scores were derived at 24-months from the communication domain. While the receptive and expressive language subtests of the Vineland and MSEL measure the same construct, data from both the assessments were used in order to provide a comprehensive picture of the infant's language abilities across different settings (lab setting vs. home environment).

Statistical analysis plan

Table 1 includes the number of participants by group at each time point. All analyses were performed using R, version 4.1. Group differences in the trajectory of CSBS scores from 12- to 24-months-of-age were examined using the nlme package (Pinheiro et al., 2021). The mixed linear model was used to analyze the effects of group and age (in months) for all CSBS scores. The interaction effect of group by age was examined with the HL-ASD group dummy coded to be the reference group.

Next, cross-sectional analyses were completed at 12- and 24-months using the general linear model. Cross-sectional analyses were not completed at 15-months due to small sample sizes (Table 1). The main effect of group was evaluated at 12- and 24-months. Estimated marginal means were computed for follow-up group comparisons using the emmeans package (Lenth, 2016). Tukey adjustments were applied for post-hoc group comparisons.

General linear models were used to explore the effects of 12-month CSBS scores on 24-month language scores (MSEL receptive and expressive language raw scores; and Vineland receptive and expressive language raw scores); and 24-month CSBS scores on 36-month language scores (MSEL receptive and expressive language raw scores). NVDQ at 24 months was included as a covariate to explore the effects of social communication on later language while controlling for non-verbal cognitive skills. A decision was made *a priori* to explore associations between CSBS scores and later language scores in each group using the general linear model. The lm package (Pinheiro et al., 2021) in R was used for analyses using the general linear model.

Logistic regressions models were used to analyze the relationship between 12-month CSBS scores and 24-month diagnostic (HL-ASD vs. HL-Neg) and language outcomes (HL-Language Delay vs. HL-No Delay); as well as 24-month CSBS scores and 36-month language outcomes across HL infants. The LL-Neg group was not included for the language outcomes analysis due to the low occurrence of language delays in this group. Of the 141 LL-Neg infants who had language outcome data at 24-months, 9 met criteria for language delays.

For all the analyses, data collection site, maternal education, and sex of the infant were included as control variables. These control variables were selected *a priori* to account for differences in data collected across sites, associations between maternal education and child language skills (Hart and Risley, 1995), and sex differences in language acquisition (Eriksson et al., 2012). A false discovery rate (FDR) procedure was used to correct for multiple comparisons where multiple tests were done using the same outcome variable. Benjamini and Hochberg (1995) one-step model was used, and adjusted *p*-values are presented as *q*-values for significant associations.

Results

Sample characteristics

Table 2 contains demographic information and developmental characteristics for the HL-ASD, HL-Neg, and LL-Neg groups. Most of the participants were Caucasian (85%). The percentage of male participants was 60% for the entire sample, as compared to 78% for the HL-ASD group. Groups did not significantly differ on chronological age at the 12, 15, and 24-month time points. There was a significant positive association between group and age at 36-months ($F(2, 240) = 10.88$, $p < 0.01$), such that the LL-Neg group fell significantly below the other two groups. For all analyses involving 36-month

TABLE 2 Descriptive data for study sample by Group.

Variable	HL-ASD ($N = 81$)	HL-Neg($N = 277$)	LL-Neg ($N = 158$)	Test statistic
% Male	78	55	59	$\chi^2 = 14.10, p < 0.01$
Maternal education				$\chi^2 = 34.22, p < 0.01$
High school diploma (%)	42	31	16	
College degree (%)	31	45	39	
Graduate degree (%)	27	24	45	
Paternal education				$\chi^2 = 11.76, p = 0.02$
High school diploma (%)	43	32	23	
College degree (%)	33	38	41	
Graduate degree (%)	22	30	36	
Race				$\chi^2 = 12.91, p = 0.23$
White (%)	88	91	87	
African American (%)	1	3	6	
Asian (%)	1	1	1	
Multiracial (%)	15	9	11	
N at 12-months	62	223	115	
N at 15-months	9	15	10	
N at 24-months	65	241	129	
Age at 12-months	12.8 (0.73)	12.6 (0.66)	12.7 (0.83)	$F = 1.14, p = 0.32$
Age at 15-months	15.8 (0.81)	15.4 (0.45)	15.5 (0.40)	$F = 1.72, p = 0.19$
Age at 24-months	24.8 (1.30)	24.7 (0.91)	24.8 (0.87)	$F = 13.56, p = 0.87$
Age at 36-months	39.7 (4.58)	39.5 (4.90)	43.8 (8.53)	$F = 10.88, p < 0.01$
MSEL-ELC				
12-months	92.36 (14.97)	101.18 (2.34)	106.39 (11.48)	$F = 31.83, p < 0.01$
24-months	80.18 (17.13)	102.04 (15.60)	110.27 (15.22)	$F = 82.80, p < 0.01$
36-months	83.24 (21.29)	103.97 (18.03)	111.69 (15.31)	$F = 37.13, p < 0.01$
MSEL NVDQ				
12-months	109.38 (13.15)	113.31 (12.68)	116.39 (11.30)	$F = 5.43, p < 0.01$
24-months	87.91 (13.04)	101.98 (12.92)	108.75 (13.16)	$F = 71.99, p < 0.01$
36-months	87.28 (19.83)	105.62 (16.47)	109.14 (13.21)	$F = 29.91, p < 0.01$
MSEL expressive language raw scores				
12-months	11.01 (2.65)	12.27 (2.58)	13.01 (2.49)	$F = 5.73, p < 0.01$
24-months	18.01 (5.42)	22.43 (4.17)	23.81 (4.05)	$F = 33.45, p < 0.01$
36-months	29.24 (7.08)	33.96 (4.74)	38.23 (5.16)	$F = 30.29, p < 0.01$
MSEL receptive language raw scores				
12-months	11.67 (2.39)	12.56 (2.06)	13.76 (1.86)	$F = 12.82, p < 0.01$
24-months	18.57 (6.69)	25.27 (3.35)	26.56 (3.07)	$F = 87.20, p < 0.01$
36-months Vineland ABC	29.30 (7.30)	33.56 (4.80)	38.14 (6.26)	$F = 26.31, p < 0.01$
12-months	89.47 (14.39)	96.30 (13.98)	100.48 (9.54)	$F = 20.06, p < 0.01$
24-months	88.93 (9.75)	101.51 (10.90)	103.96 (11.43)	$F = 41.63, p < 0.01$
vineland expressive language raw scores				
12-months	16.14 (4.53)	19.00 (5.26)	20.40 (3.94)	$F = 16.71, p < 0.01$
24-months	33.44 (12.76)	45.93 (12.63)	50.01 (11.54)	$F = 14.78, p < 0.01$
Vineland receptive language raw scores				
12-months	9.55 (3.43)	11.31 (3.25)	12.46 (3.31)	$F = 38.56, p < 0.01$
24-months	18.02 (6.24)	24.07 (3.65)	25.25 (3.44)	$F = 77.16, p < 0.01$

TABLE 3 Longitudinal mixed linear model with group by age interaction effects.

CSBS scores	Group*age interaction			
	df 1	*df* 2	*F*	*q*
Emotion and eye gaze ($N = 494$)	2	321	16.26	<0.01*
Communication ($N = 511$)	2	347	8.56	<0.01*
Gestures ($N = 511$)	2	347	6.59	<0.01*
Sounds ($N = 511$)	2	347	8.23	<0.01*
Words ($N = 511$)	2	347	29.08	<0.01*
Understanding ($N = 409$)	2	159	36.85	<0.01*
Object use ($N = 485$)	2	289	6.07	0.01*

*Significant interaction that survived adaptive FDR procedure.

language scores, candidate age at 36-months was included as a control variable. Maternal education was significantly positively associated with all receptive and expressive language measures at 24-months ($p < 0.05$, $f^2 = 0.02–0.06$) and included as a control variable in all models.

Social communication development across groups

Mixed linear models revealed significant group by age interaction effects ($q < 0.01$, Table 3) for all CSBS scores. This interaction effect is represented by the widening gap over time between the HL-ASD group and the HL-Neg and LL-Neg groups (Figure 1).

Cross-sectional group differences at 12-months and 24-months

At 12-months, the main group effect was significant for all CSBS scores ($q < 0.01$, Table 4), except understanding and words. The HL-ASD group scored significantly below the LL-Neg group and HL-Neg groups on emotion and eye gaze, communication, gestures, and sounds ($p < 0.01$; Supplementary Figure 1, Supplementary results). The HL-ASD group scored significantly below the LL-Neg group on object use ($p < 0.05$; Supplementary Figure 1, Supplementary results). The HL-ASD group did not significantly differ from the LL-Neg group on words, understanding, and object use.

The HL-Neg group scored significantly below the LL-Neg group on emotion and eye gaze, and gestures at 12-months ($p < 0.05$; Supplementary Figure 1, Supplementary results). The HL-Neg infants did not differ significantly from the LL-Neg infants on communication, sounds, and object use.

At 24-months, for all CSBS scores, the main effect of group was significant ($q < 0.01$, Table 4). HL-ASD infants scored significantly below the other two groups on all CSBS scores ($p < 0.01$; Supplementary Figure 2, Supplementary results). The HL-Neg infants scored significantly below the LL-Neg group on understanding at 24-months. The HL-Neg and LL-Neg groups did not differ significantly from each other on any of the CSBS scores.

Association between 12-month social communication 24-month language skills

The interaction effects of CSBS scores by group were not significantly associated with expressive language (Supplementary Table 5, Supplementary results), and were removed from subsequent models. Once the interaction term was removed, all CSBS scores were significantly positively associated with MSEL EL measures and Vineland EL measures, with effect sizes ranging from small to medium ($q < 0.05$, Table 5). All CSBS scores except for object use and emotion and eye gaze remained significantly positively associated with MSEL EL measures ($q < 0.05$, $f^2 = 0.05–0.17$) after controlling for NVDQ. Similarly, all CSBS scores remained significantly positively associated with Vineland EL scores after controlling for NVDQ ($q < 0.05$, $f^2 = 0.06–0.14$).

The interaction effects of CSBS scores by group were not significantly associated with receptive language, except for words and Vineland RL [($F(2, 370) = 4.65$, $p < 0.05$, $f^2 0 = 0.03$), Supplementary Table 5, Supplementary results]. Follow-up analyses indicated that in the HL-ASD & LL-Neg groups, words and Vineland RL were not significantly associated with each other. However, in the HL-Neg group there was a significant positive association between words and Vineland RL ($\beta = 0.37$, $t\ (208) = 3.13$, $p = 0.01$).

For the remaining models with non-significant interaction effects of CSBS scores by group on receptive language, the interaction term was removed from subsequent models. Once the interaction effect was removed, CSBS scores that were significantly positively associated with 24-month MSEL RL, included emotion and eye gaze and gestures ($q < 0.05$, $f^2 = 0.02–0.11$). CSBS scores that were significantly positively associated with 24-month Vineland RL included emotion and eye gaze and words ($q < 0.01$, Table 5), with small effect sizes. However, after controlling for NVDQ, none of the 12-month CSBS scores were significantly associated with MSEL RL scores. Emotion and eye gaze and words scores were significantly positively associated with Vineland RL after controlling for NVDQ ($q < 0.01$, $f^2 = 0.06–0.12$).

We also explored associations between social communication skills and later language in each group of participants. In the HL-ASD group, none of the 12-month CSBS scores were associated with MSEL and Vineland language

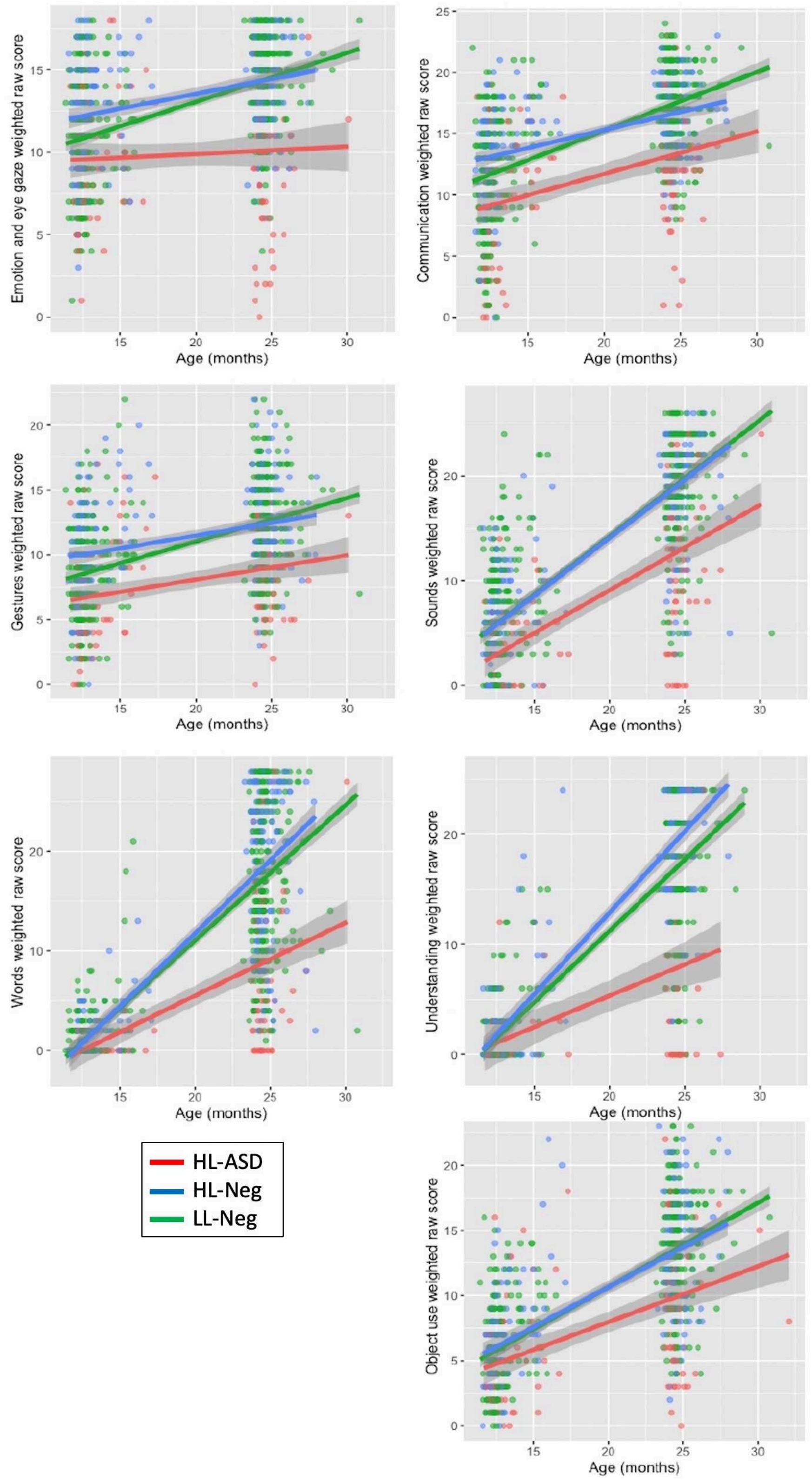

FIGURE 1
Developmental trajectory of CSBS scores by group.

TABLE 4 *Post-hoc* cross-sectional analysis exploring the main effect for group.

CSBS scores	HL-ASD (a)		HL-Neg (b)		LL-Neg (c)		Overall group comparison					*Post-Hoc* comparisons
	EMM	SE	EMM	SE	EMM	SE	*df*	SS	*F*	*q*	f^2	
12-months												
Emotion eye gaze ($N = 378$)	9.63	0.44	10.95	0.23	12.06	0.33	2	213.90	9.88	<0.01*	0.07	a < b < c
Communication ($N = 396$)	9.07	0.56	11.80	0.29	12.89	0.42	2	549.60	15.17	<0.01*	0.10	a < b, c
Gestures ($N = 396$)	6.57	0.47	8.53	0.24	9.92	0.35	2	415.30	16.45	<0.01*	0.10	a < b < c
Sounds ($N = 396$)	3.52	0.59	6.34	0.31	5.73	0.44	2	375.50	9.32	<0.01*	0.06	a < b, c
Words ($N = 396$)	0.40	0.24	1.00	0.13	0.95	0.18	2	17.86	2.62	0.25	0.02	a, b, c
Understanding ($N = 262$)	1.04	0.54	1.62	0.29	2.07	0.41	2	25.26	1.67	0.42	0.01	a, b, c
Object use ($N = 334$)	4.84	0.47	5.91	0.26	6.29	0.36	2	69.90	3.14	0.05	0.02	a, b, c
24-months												
Emotion and eye gaze ($N = 410$)	9.96	0.38	14.49	0.19	14.34	0.27	2	1,016.00	60.18	<0.01*	0.30	a < b, c
Communication ($N = 432$)	13.40	0.43	17.50	0.22	16.70	0.31	2	853.30	37.03	<0.01*	0.19	a < b, c
Gestures ($N = 432$)	9.08	0.45	12.59	0.23	12.53	0.32	2	651.90	26.63	<0.01*	0.14	a < b, c
Sounds ($N = 432$)	13.00	0.69	19.60	0.36	19.20	0.50	2	2,246.90	37.73	<0.01*	0.19	a < b, c
Words ($N = 432$)	9.55	0.99	17.75	0.51	18.78	0.71	2	4,008.20	33.43	<0.01*	0.18	a < b, c
Understanding ($N = 289$)	8.27	1.03	17.44	0.54	19.70	0.66	2	3,745.00	45.39	<0.01*	0.36	a < b < c
Object use ($N = 413$)	10.20	0.56	13.60	0.28	13.70	0.41	2	608.00	16.72	<0.01*	0.10	a < b, c

*Significant interaction that survived adaptive FDR procedure.

scores (Supplementary Table 3, Supplementary results). In contrast, the HL-Neg group demonstrated significant positive associations between all CSBS scores and MSEL EL except for object use ($q < 0.05, f^20 = 0.04–0.22$). The association of words, sounds, understanding, and gestures remained significant after controlling for NVDQ. The HL-Neg group also demonstrated significant positive associations between Vineland EL and the following CSBS scores: emotion and eye gaze, communication, gestures, sounds, words, object use and Vineland EL ($q < 0.05$, $f^2 = 0.02–0.12$). The association between Vineland EL and emotion and eye gaze, gestures, and object use and did not remain significant after controlling for NVDQ. The LL-Neg group demonstrated significant positive associations between communication, gestures, sounds, words, understanding, object use and MSEL EL ($q < 0.05, f^20 = 0.04–0.22$), which remained significant after controlling for NVDQ. The LL-Neg group also demonstrated significant positive associations between understanding and Vineland EL, and this association remained significant after controlling for NVDQ.

For MSEL-RL, the HL-Neg group demonstrated significant positive associations with emotion eye gaze, and gestures ($q < 0.05$, f^2 =0.04–0.06). Vineland RL in the HL-Neg group was significantly positively associated with words and communication ($q < 0.05$, f^2 =0.02–0.08). However, none of these associations in the HL-Neg group remained significant after controlling for NVDQ, except for words and Vineland RL. The LL-Neg group did not demonstrate any significant associations between 12-month CSBS scores and 24-month Mullen and Vineland RL scores.

Association between 24-month social communication 36-month language skills

The CSBS scores by group interaction effects were significant for MSEL EL and emotion and eye gaze ($q < 0.05$; Supplementary Table 6, Supplementary results). Follow-up groupwise analyses indicated that the HL-ASD group demonstrated a significant positive association ($\beta = 0.81$, $t(36) = 3.51$, $p < 0.01$, $f^2 = 0.42$), whereas the associations were not significant for the HL-Neg group ($\beta = -0.07$, $t(102) = -0.40$, $q = 0.69$) and LL-Neg groups ($\beta = -0.22$, $t(36) = -0.75$, $q = 0.46$). The interaction effects of CSBS scores by group on MSEL EL were not significant (Supplementary Table 6, Supplementary results) for all other models and were removed from subsequent models. Once the interaction term was removed, emotion and eye gaze, communication, gestures, sounds, words, and understanding were significantly positively associated with MSEL EL measures, with small to large effect sizes ($q < 0.05$, Table 6). All positive associations remained significant after controlling for NVDQ ($q < 0.01, f^2 = 0.08–0.69$, Table 6).

For MSEL RL, the CSBS scores by group interaction effects were significant for emotion and eye gaze and understanding ($q < 0.05$; Supplementary Table 6, Supplementary results). Follow-up groupwise analyses indicated that the HL-ASD group demonstrated a significant positive associations between emotion and eye gaze and MSEL RL ($\beta = 0.61$, $t\ (38) = 2.76$, $p < 0.01, f^2 = 0.30$), whereas the associations were not significant for the HL-Neg group ($\beta = -0.10$, $t\ (102) = -0.54$, $q =$

10.3389/fcomm.2022.977724

TABLE 5 General liner model exploring main effects of social communication skills measured at 12-months on language abilities measured at 24-months.

CSBS scores	Main effect of CSBS scores				Main effect of CSBS scores with NVDQ			
	SS	*F*	*q*-value	f^2	SS	*F*	*q*-value	f^2
MSEL EL								
Emotion eye gaze (*N* = 376)	88.2	5.13	0.02*	0.06	12.3	0.86	0.35	0.07
Communication (*N* = 393)	151	8.87	<0.01*	0.10	72.5	5.06	0.03*	0.11
Gestures (*N* = 393)	172.7	10.18	<0.01*	0.11	88.6	6.20	0.02*	0.13
Sounds (*N* = 393)	312.8	18.85	<0.01*	0.14	198.2	14.17	<0.01*	0.16
Words (*N* = 393)	136.8	8.02	<0.01*	0.04	126.6	8.93	<0.01*	0.05
Understanding (*N* = 260)	311.8	19.34	<0.01*	0.10	247.8	17.93	<0.01*	0.12
Object use (*N* = 332)	190.9	11.70	<0.01*	0.06	48.3	3.44	0.07	0.07
Vineland EL								
Emotion eye gaze (*N* = 368)	1,624	11.52	<0.01*	0.09	745	5.72	0.02*	0.10
Communication (*N* = 384)	2,497	17.95	<0.01*	0.13	1,786	13.98	<0.01*	0.14
Gestures (*N* = 384)	1,676	11.86	<0.01*	0.11	1,101	8.49	<0.01*	0.12
Sounds (*N* = 384)	2,751	19.87	<0.01*	0.11	1,951	15.32	<0.01*	0.12
Words (*N* = 396)	1,641	11.61	<0.01*	0.06	1,581	12.32	<0.01*	0.06
Understanding (*N* = 253)	1,816	13.64	<0.01*	0.08	1,430.2	11.87	<0.01*	0.09
Object use (*N* = 325)	1,689	12.10	<0.01*	0.06	669	5.14	0.02*	0.06
MSEL RL								
Emotion eye gaze (*N* = 376)	112.9	7.55	0.04*	0.11	18.5	1.58	0.44	0.15
Communication (*N* = 393)	63.8	4.34	0.05	0.11	15	1.31	0.44	0.14
Gestures (*N* = 393)	91.4	6.25	0.04*	0.11	30.2	2.66	0.36	0.14
Sounds (*N* = 393)	22.5	1.52	0.25	0.06	0.60	0.05	0.82	0.08
Words (*N* = 393)	3.9	0.26	0.61	0.02	2.3	0.20	0.82	0.02
Understanding (*N* = 260)	88.9	5.31	0.05	0.06	52.3	3.85	0.35	0.07
Object use (*N* = 332)	67.5	4.41	0.05	0.05	1.5	0.11	0.82	0.06
Vineland RL								
Emotion eye gaze (*N* = 368)	223.6	13.85	0.01*	0.11	136.4	8.76	0.01*	0.12
Communication (*N* = 384)	55.0	3.41	0.11	0.07	28.2	1.83	0.24	0.08
Gestures (*N* = 383)	39.0	2.40	0.13	0.06	17.6	1.14	0.33	0.07
Sounds (*N* = 384)	78.1	4.85	0.06	0.06	44.4	2.89	0.21	0.06
Words (*N* = 384)	149.5	9.40	<0.01*	0.06	143.5	9.49	0.01*	0.06
Understanding (*N* = 253)	49.3	3.00	0.12	0.03	36.6	2.28	0.23	0.03
Object use (*N* = 325)	35.9	2.21	0.14	0.02	6.9	0.43	0.51	0.02

*Significant main effect that survived adaptive FDR procedure.
The degree of freedom for all models was 1.

0.59) and LL-Neg groups ($\beta = -0.20$, t (36) $= -0.61$, $q = 0.54$). Associations between understanding and MSEL RL was significant for HL-ASD infants ($\beta = 0.59$, t (24) $= 4.07$, $q < 0.01$, $f^2 = 0.61$) and HL-Neg infants ($\beta = 0.20$, t (69) $= 3$, $q < 0.01$, $f^2 = 0.14$), but not for LL-Neg infants ($\beta = 0.53$, t (31) $= 2.75$, $q < 0.07$).

The interaction effects of remaining CSBS scores (communication, gesture, sounds, words, understanding, object use) by group on MSEL RL were not significant (Supplementary Table 6, Supplementary results) for all other models and were removed from subsequent models. Once the interaction term was excluded, CSBS scores that were significantly associated with 36-month MSEL RL included emotion and eye gaze, communication, sounds, words, and understanding ($q < 0.01$, Table 6), with effect sizes ranging from small to large. These associations remained significant after controlling for NVDQ ($q < 0.01$, Table 6). Gestures scores at 24-months were also significantly positively associated with MSEL RL at 36-months ($F(1, 379) = 4.86$, $q < 0.03$, $f^2 = 0.06$), however this association did not remain significant after controlling for NVDQ.

Follow-up groupwise analyses revealed unique patterns for 24-month social communication and 36-month language associations. The HL-ASD group demonstrated significant

TABLE 6 General liner model exploring main effects of social communication skills measured at 24-months on language abilities measured at 36-months.

CSBS scores	Main effect of CSBS scores				Main effect of CSBS scores with NVDQ			
	SS	*F*	*q*-value	f^2	SS	*F*	*q*-value	f^2
MSEL EL								
Emotion eye gaze ($N = 194$)	232.3	10.79	<0.01*	0.24	166.4	8.99	<0.01*	0.28
Communication ($N = 200$)	298.1	14.22	<0.01*	0.18	262.9	14.53	<0.01*	0.21
Gestures ($N = 200$)	100.6	4.57	0.04*	0.07	72.7	3.81	0.06	0.08
Sounds ($N = 200$)	935.9	53.17	<0.01*	0.56	715.54	45.57	<0.01*	0.63
Words ($N = 200$)	814.4	44.64	<0.01*	0.48	584.98	35.69	<0.01*	0.54
Understanding ($N = 144$)	486.87	23.54	<0.01*	0.63	288.36	15.19	<0.01*	0.69
Object use ($N = 192$)	72.9	3.34	0.06	0.08	33.5	1.74	0.19	0.10
MSEL RL								
Emotion eye gaze ($N = 194$)	205.6	9.62	<0.01*	0.22	138.4	7.83	<0.01*	0.27
Communication ($N = 200$)	169.3	7.88	<0.01*	0.13	139.7	7.89	<0.01*	0.16
Gestures ($N = 200$)	106	4.86	0.03*	0.06	73.9	4.09	0.05	0.07
Sounds ($N = 200$)	530.6	27.10	<0.01*	0.36	342.72	20.6	<0.01*	0.42
Words ($N = 200$)	531.1	27.13	<0.01*	0.31	321.9	19.23	<0.01*	0.37
Understanding ($N = 144$)	679.02	36.84	<0.01*	0.82	426.42	25.75	<0.01*	0.92
Object use ($N = 192$)	76.1	3.53	0.06	0.07	31.9	1.74	0.19	0.09

*Significant main effect that survived adaptive FDR procedure.
The degree of freedom for all models was 1.

associations between emotion-eye gaze, sounds, words, understanding and MSEL-EL ($q < 0.05$, $f^2 = 0.27–0.41$), with emotion-eye gaze, sound, and word associations continuing to remain significant after controlling for NVDQ (Supplementary Table 6, Supplementary results). HL-Neg group demonstrated significant associations between sounds, words, understanding and MSEL EL ($q < 0.05$, $f^2 = 0.09–0.25$). MSEL EL and sounds, words, and communication were significantly positively associated after controlling for NVDQ. For MSEL-RL, HL-ASD demonstrated significant associations for emotion-eye gaze, sounds, words, understanding ($q < 0.05$, $f^2 = 0.13–0.56$), however, none of these associations remained significant after controlling for NVDQ. HL-Neg group demonstrated significant associations for sounds, words, understanding and MSEL-RL ($q < 0.01$, $f^2 = 0.13–0.15$). These associations in the HL-Neg group continued to remain significant after controlling for NVDQ. In contrast, LL-Neg group did not demonstrate significant associations to 36-month MSEL language scores.

Social communication skills as predictors of autism diagnoses and language delays

Logistic regression was used to determine if CSBS scores predicted later autism diagnoses and language delay status. These analyses were conducted within the HL infants only. Of the HL infants 22% met criteria for signs of early language delay at 24 and 36 months (Table 7).

TABLE 7 Number of infants identified to have language delays among HL-infants by visit.

	24-months	36-months
HL	309	192
HL-language delay	69 (22%)	42 (22%)
HL-no delay	240 (78%)	150 (78%)

The models exploring 12-month CSBS scores as predictors of 24-month autism diagnosis and language delay in the high-likelihood infants did not reveal any significant associations (Table 8). However, the models exploring 24-month CSBS scores as predictors of 36-month language delay revealed a significant negative association between 24-month understanding scores and 36-month language delay outcome ($q < 0.01$, Table 9). The estimated odds ratio indicated that, holding all other social communication scores constant, the odds of a language delay increased by 0.85 times (95% CI [0.78, 0.91]) per one unit decrease in understanding scores. The sensitivity of the model was 45%, which meant that the model correctly classified infants who went on to have a language delay 45% of the time using understanding scores. The model had a specificity of 95%, which meant infants who did not go on to have a language delay were correctly classified 95% of the time using understanding

TABLE 8 Logistic regression analysis exploring main effects of social communication skills measured at 12-months on diagnostic and language outcomes measured at 24-months in HL-infants.

CSBS scores	Main effect of CSBS scores				
	Estimate	SE	*z*-value	*q*-value	OR
Autism diagnostic outcome (*N* = 163)					
Emotion eye gaze	−0.02	0.07	−0.37	0.93	0.97
Communication	−0.12	0.06	−1.99	0.33	0.89
Gestures	−0.04	0.07	−0.57	0.93	0.96
Sounds	−0.02	0.06	−0.26	0.93	0.98
Words	−0.04	0.20	−0.22	0.93	0.95
Understanding	−0.01	0.08	−0.09	0.93	0.99
Object use	−0.02	0.07	−0.24	0.93	0.98
Language delay diagnostic outcome (*N* = 139)					
Emotion eye gaze	−0.08	0.09	−0.93	0.49	0.92
Communication	−0.09	0.07	−1.30	0.48	0.91
Gestures	−0.05	0.08	−0.56	0.67	0.95
Sounds	0.09	0.08	1.09	0.48	1.09
Words	−0.39	0.29	−1.35	0.48	0.68
Understanding	−0.15	0.13	−1.14	0.48	0.86
Object use	−0.03	0.09	−0.37	0.71	0.97

TABLE 9 Logistic regression analysis exploring main effects of social communication skills measured at 24-months on language outcomes measured at 36-months in HL-infants.

CSBS scores	Main effect of CSBS scores				
	Estimate	SE	*z* value	*q value*	OR
Language delay diagnostic outcome (*N* = 103)					
Emotion eye gaze	0.00	0.10	0.03	0.98	1.00
Communication	0.29	0.16	1.79	0.17	1.34
Gestures	−0.13	0.13	−1.00	0.45	0.88
Sounds	−0.21	0.09	−2.32	0.07	0.81
Words	−0.01	0.07	−0.21	0.97	0.99
Understanding	−0.16	0.05	−3.21	<0.01*	0.85
Object use	−0.14	0.10	−1.42	0.27	0.87

scores. None of the other 24-month social communication scores predicted autism diagnosis or language delays.

Discussion

This prospective study explored social communication skills and their associations to language in infants later diagnosed autistic. Social communication skills were evaluated at 12, 15, and 24-months in a large sample of infants that were either: (a) typically developing infants with no family history of autism (LL-Neg), (b) infants with a family history of autism who were later diagnosed autistic (HL-ASD), or (c) infants with a family history of autism who were not later diagnosed autistic (HL-Neg). The clinical implications for early identification and intervention for autism are discussed below.

HL-ASD infants demonstrated lower scores on social communication assessments at 12-months-of-age across widespread domains, and these early differences became more pronounced in the second year of life. These findings add to existing research reporting that social communication difficulties are detectable using standardized assessments as early as 9- to 12-months-of-age (Bradshaw et al., 2021). HL-ASD infants demonstrated unique social communication developmental trajectories in the second year of life. As a group, they did not make significant gains on the emotion and eye gaze cluster. On the gesture cluster, HL-ASD made parallel gains when compared to the LL-Neg infants in the second year of life and remained significantly below LL-Neg and HL-Neg infants throughout. Overall, HL-ASD infants remained consistently low on emotion and eye gaze and gestures, a finding that corroborates previous reports in infants between 9- and 12-months-of-age (Rozga et al., 2011; Iverson et al., 2018; Stallworthy et al., 2021). In contrast, HL-ASD infants demonstrated a growing gap on the sounds and object use clusters. They scored significantly below the other two groups at 12-months-of-age, but this gap widened over time, with the HL-ASD group making fewer gains when compared to the other two groups. Group level differences in the areas of words and understanding were late to emerge. While HL-ASD infants did not score significantly below the other two groups at 12-months-of-age, they scored significantly below at 24-months-of-age. Similar patterns of social communication development (i.e., consistently low, growing gap, and late emerging) were reported by Bradshaw et al. (2021) between 9- and 12-months-of-age. Overall, these developmental trajectories suggest that social communication skills are an ideal target for preemptive interventions (i.e., interventions that are provided prior to formal autism diagnosis) as the goal through these interventions is to reshape symptom trajectories in the prodromal period.

Across all infants, better scores on all CSBS clusters at 12-months-of-age were related to better expressive language scores at 24-months of age. Better emotion and eye gaze and gesture scores were associated with better 24-month receptive language scores. Further, better scores on the sounds, words, and understanding clusters at 24-months-of-age were associated with better 36-month receptive and expressive language scores. In addition, infants who had better emotion and eye gaze and communication skills at 24-months also had higher 36-month expressive language scores.

This current study is the first to explore associations between a wide range social communication skills and language in infants with a high likelihood for autism as early as 12-months-of-age. Extending this research to younger ages has revealed shifts in associations between early social communication and later language, unique to infants later diagnosed autistic.

In the HL-ASD group, none of the CSBS scores at 12-months were associated with language at 24-months-of-age. In contrast, CSBS scores as early as 12-months-of age were significantly positively associated with 24-month language in the HL-Neg and LL-Neg groups. This functional association did not emerge in the HL-ASD group until 24-months-of-age, at which point emotion-eye gaze, sounds, words, and understanding were significantly positively associated with 36-month receptive and expressive language. This finding is consistent with previous research that has reported that social communication skills measured 20-months-of-age and beyond are associated with downstream language skills in autistic toddlers (Yoder et al., 2015; Delehanty et al., 2018). It is likely that HL-ASD infants did not demonstrate this association at 12-months-of-age because they had not yet acquired the relevant skills that are developmentally "upstream" from language. Once these children made advancements in the areas of emotion and eye gaze, sounds, words, and understanding, associations to later language emerged. An additional possibility is that at 12-months, skills other than social communication support later language development in the HL-ASD group. While the current study focused on social communication skills, from a "developmental cascades" perspective, it is possible that development in other domains (such as motor skills, temperament, sleep, visual attention) may impact downstream language abilities (Bradshaw et al., 2022). Future research should explore these cascading effects on language in order to elucidate alternative pathways to language in HL-ASD infants.

Pecukonis et al. (2022) recently reported surprising group differences in the association between caregiver-reported gestures at 12-months and MSEL language skills at 36 months. Significant positive associations were reported between 12-month gesture use and language in HL-Neg infants, which corroborated findings from the current study. However, they also reported negative associations between 12-month gesture and language in HL-ASD infants. The current study did not find significant negative associations between gesture and language in HL-ASD infants. These differences could be attributed to differences in measures used for gestures (parent-report vs. direct assessment) or differences in the timepoint for language skills (36 vs. 24 months). Future efforts should aim to replicate findings from the current study and Pecukonis et al. (2022) to provide clarity on the developmental relationship between gestures and language for infants who develop autism.

Although 12-month social communication skills were not associated with later language in the HL-ASD group, 24-month social communication skills were associated with later language, and this suggests that supporting early social communication development may relate to better downstream language skills. Further, HL-Neg infants are more likely to demonstrate language delays than LL-Neg infants (Miller et al., 2016; Marrus et al., 2018); and our findings revealed that supporting social communication skills in this group may relate to better language skills as well.

This study is also the first to explore the utility of social communication profiles in predicting autism diagnoses and language delays in infants with a high likelihood for autism. At 24-months, understanding scores significantly predicted 36-months language delay outcomes. While understanding scores accurately predicted high-likelihood infants who met criteria for language delays only 45% of the time; they accurately predicted infants who did not meet criteria for language delays 95% of the time. The understanding cluster measures receptive vocabulary. Our findings suggest that early receptive vocabulary can be used to screen out infants within the high likelihood group who do not need additional language interventions. Further, these findings also suggest that providing interventions in response to early receptive vocabulary developmental trajectories needs to be a key focus of preemptive interventions.

Overall, these findings suggest that preemptive interventions should follow a personalized, developmental approach and capitalize on existing social communication skills to increase bouts of shared engagement, use of intentional communication, and gesture use (i.e., skills measured in the social composite). During instances of intentional communication, the use of sounds and words can be encouraged using parental scaffolding. While infants later diagnosed autistic did not demonstrate significant differences at the group level on symbolic skills (i.e., understanding and object use) at 12-months-of-age, it is important to monitor these skills and provide interventions tailored to support symbolic skills as challenges in this area may emerge later in development.

While preemptive interventions for autism are yet to demonstrate conclusive evidence of efficacy, preliminary research reports are promising (Hampton and Rodriguez, 2021). Preemptive interventions have been successfully implemented for infants recruited based on a higher familial likelihood for autism (Green et al., 2015, 2017; Yoder et al., 2021) as well as community-based screening approaches (Watson et al., 2017; Whitehouse et al., 2021). Recent research has demonstrated that parent-implemented preemptive interventions have significant effects on child autism symptom severity and parent-reported receptive and expressive vocabulary (Whitehouse et al., 2021). Although further research is required to gather conclusive evidence of improved child outcomes, this research suggests that parents are able to implement intervention strategies with fidelity and this in turn may relate to better social communication outcomes (Watson et al., 2017; Hampton and Rodriguez, 2021; Yoder et al., 2021). Our findings provide evidence that improving social communication outcomes could confer better downstream language skills. However, in order to establish causal relationships, preemptive intervention research should explore downstream developmental effects of proximal social communication outcomes (Chawarska et al., 2009)

A limitation of this study is that the sample did not include infants with other developmental delays. Hence, the profile of social communication skills reported in this study may not be specific to autism. It should be noted, however, that approximately 30% of high likelihood infants who do not meet criteria for autism demonstrate other clinical concerns at school age, such as broader autism phenotype (BAP), Attention-Deficit/Hyperactivity Disorder (ADHD), and speech and language problems (Miller et al., 2016). Including a comparison group of infants with other developmental disabilities (such as Down's syndrome) may reveal developmental trajectories and language associations that are unique to autism. This study also had a limited sample size at 15-months. Future research should include multiple sampling points across developments to accurately characterize social communication developmental trajectories. Another important future research goal is extended follow-up of these infants. This will elucidate long-term (i.e., school-age and beyond) effects of early social communication delays. It is also important to consider heterogeneity in the early development of autism, which could impact diagnostic classification over time. While 24-month diagnostic classification is stable (Chawarska et al., 2009; Corsello et al., 2013; Guthrie et al., 2013; Barbaro and Dissanayake, 2017), extended follow-up can account for changes in diagnostic classification across groups. An additional limitation of this study is the lack of racial diversity in the sample. Future research should aim to recruit a racially diverse sample as research has reported racial disparities in the timing of autism diagnoses and access to interventions, particularly in African American autistic children (Constantino et al., 2020).

Conclusions

Infants later diagnosed autistic demonstrate widespread challenges with social communication skills as early as 12-months-of-age and this gap in social communication skills was more pronounced in the second year of life. Better social communication skills at 12-months-of-age were not associated with better with downstream receptive and expressive language skills in this group. Contrastingly, infants who did not meet criteria for autism demonstrated significant positive associations between early social communication skills and later language. This functional association only emerged in autistic infants in the second year of life, at which point 24-month social communication skills were positively associated with 36-month language skills. Further, understanding (i.e., receptive vocabulary) scores at 24-months-of-age significantly predicted language delays in infants with an older autistic sibling. Taken together, these findings support the need for preemptive interventions that are designed to respond to early developmental trajectories that consolidate into an autism diagnosis. Social communication skills, particularly sharing attention, goal-directed communication (using gestures, sounds, and words), and understanding are ideal preemptive intervention targets as they related to better downstream language abilities.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board, The University of Texas at Dallas. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

SR, MS, and JP-M contributed to conception and design of the study. AB, SM, CC, and LY organized the database. HA, SR, and MS contributed to the statistical analysis plan. SR wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted Version.

Funding

This work was supported by grants through the National Institutes of Health (R00-MH108700 PI Swanson, R01-HD055741 PI Piven, R01-HD055741-S1 PI Piven, P30-HD003110 PI Piven, U54 EB005149 PI Kikinis) and the Simons Foundation (SFARI Grant 140209). The funders had no role in study design, data collection, analysis, data interpretation, or the writing of the report.

Acknowledgments

The authors thank the IBIS children and their families for their ongoing participation in this longitudinal study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor LA declared a past co-authorship with the authors AE and JP.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2022.977724/full#supplementary-material

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, 4th Edn.* Washington, DC: Author.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®).* Philadelphia, PA: American Psychiatric Association. doi: 10.1176/appi.books.9780890425596

Barbaro, J., and Dissanayake, C. (2017). Diagnostic stability of autism spectrum disorder in toddlers prospectively identified in a community-based setting: Behavioural characteristics and predictors of change over time. *Autism* 21, 830–840. doi: 10.1177/1362361316654084

Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Ser. B Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bottema-Beutel, K. (2016). Associations between joint attention and language in autism spectrum disorder and typical development: a systematic review and meta-regression analysis: associations between joint attention. *Autism Res.* 9, 1021–1035. doi: 10.1002/aur.1624

Bradshaw, J., Klin, A., Evans, L., Klaiman, C., Saulnier, C., and McCracken, C. (2020). Development of attention from birth to 5 months in infants at risk for autism spectrum disorder. *Dev. Psychopathol.* 32, 491–501. doi: 10.1017/S0954579419000233

Bradshaw, J., McCracken, C., Pileggi, M., Brane, N., Delehanty, A., Day, T., et al. (2021). Early social communication development in infants with autism spectrum disorder. *Child Dev.* 92, 2224–2234. doi: 10.1111/cdev.13683

Bradshaw, J., Schwichtenberg, A. J., and Iverson, J. M. (2022). Capturing the complexity of autism: Applying a developmental cascades framework. *Child Develop. Perspect. cdep.* 12439. doi: 10.1111/cdep.12439

Chawarska, K., Klin, A., Paul, R., Macari, S., and Volkmar, F. (2009). A prospective study of toddlers with ASD: short-term diagnostic and cognitive outcomes. *J. Child Psychol. Psychiatry* 50, 1235–45. doi: 10.1111/j.1469-7610.2009.02101.x

Choi, B., Shah, P., Rowe, M. L., Nelson, C. A., and Tager-Flusberg, H. (2020). Gesture development, caregiver responsiveness, and language and diagnostic outcomes in infants at high and low risk for autism. *J. Autism Dev. Disord.* 50, 2556–2572. doi: 10.1007/s10803-019-03980-8

Constantino, J. N., Abbacchi, A. M., Saulnier, C., Klaiman, C., Mandell, D. S., et al. (2020). Timing of the Diagnosis of Autism in African American Children. *Pediatrics.* 146, e20193629. doi: 10.1542/peds.2019-3629

Corsello, C. M., Akshoomoff, N., and Stahmer, A. C. (2013). Diagnosis of autism spectrum disorders in 2-year-olds: a study of community practice. *J. Child Psychol. Psychiatry* 54, 178–185. doi: 10.1111/j.1469-7610.2012.02607.x

Delehanty, A. D., Stronach, S., Guthrie, W., Slate, E., and Wetherby, A. M. (2018). Verbal and nonverbal outcomes of toddlers with and without autism spectrum disorder, language delay, and global developmental delay. *Autism Dev. Lang. Impair.* 3, 239694151876476. doi: 10.1177/2396941518764764

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., et al. (2012). Differences between girls and boys in emerging language skills: evidence from 10 language communities. *Br. J. Dev. Psychol.* 30, 326–343. doi: 10.1111/j.2044-835X.2011.02042.x

Estes, A., Zwaigenbaum, L., Gu, H., St John, T., Paterson, S., and Piven, J. (2015). Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *J. Neurodev. Disord.* 7, 24. doi: 10.1186/s11689-015-9117-6

Green, J. (2020). "Intervention during the prodromal stages of autism spectrum disorder," in *Autism spectrum disorder in the first years of life: Research, assessment, and treatment*, eds K. Chawarska, and F. R. Volkmar (New York, NY: The Guilford Press).

Green, J., Charman, T., Pickles, A., Wan, M. W., Elsabbagh, M., Slonims, V., et al. (2015). Parent-mediated intervention versus no intervention for infants at high risk of autism: a parallel, single-blind, randomised trial. *Lancet Psychiatry*. 2, 133–140. doi: 10.1016/S2215-0366(14)00091-1

Green, J., Leadbitter, K., Ainsworth, J., and Bucci, S. (2022). An integrated early care pathway for autism. *Lancet Child Adolesc. Health*. 6, S2352464222000372. doi: 10.1016/S2352-4642(22)00037-2

Green, J., Pickles, A., Pasco, G., Bedford, R., Wan, M. W., Elsabbagh, M., et al. (2017). Randomised trial of a parent-mediated intervention for infants at high risk for autism: Longitudinal outcomes to age 3 years. *J. Child Psychol. Psychiatry* 58, 1330–1340. doi: 10.1111/jcpp.12728

Green, J., Wan, M. W., Guiraud, J., Holsgrove, S., McNally, J., Slonims, V., and The BASIS Team (2013). Intervention for infants at risk of developing autism: a case series. *J. Autism Dev. Disord.* 43, 2502–2514. doi: 10.1007/s10803-013-1797-8

Grzadzinski, R., Amso, D., Landa, R., Watson, L., Guralnick, M., Zwaigenbaum, L., et al. (2021). Pre-symptomatic intervention for autism spectrum disorder (ASD): defining a research agenda. *J. Neurodev. Disord.* 13, 49. doi: 10.1186/s11689-021-09393-y

Guthrie, W., Swineford, L. B., Nottke, C., and Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: stability and change in clinical diagnosis and symptom presentation: stability of early diagnosis and symptoms in ASD. *J. Child Psychol. Psychiatry* 54, 582–590. doi: 10.1111/jcpp.12008

Hampton, L. H., and Rodriguez, E. M. (2021). Preemptive interventions for infants and toddlers with a high likelihood for autism: A systematic review and meta-analysis. *Autism*. 6, 136236132110504. doi: 10.1177/13623613211050433

Hart, B., and Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: P. H. Brookes.

Iverson, J. M., Northrup, J. B., Leezenbaum, N. B., Parlade, M. V., Koterba, E. A., and West, K. L. (2018). Early gesture and vocabulary development in infant siblings of children with autism spectrum disorder. *J. Autism Dev. Disord.* 48, 55–71. doi: 10.1007/s10803-017-3297-8

Kapp, S. K. (2020). *Autistic Community and the Neurodiversity Movement: Stories from the Frontline*, 1st Edn. Singapore: Springer Singapore. doi: 10.1007/978-981-13-8437-0

Landa, R. J., Holman, K. C., and Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Arch. Gener. Psychiatry* 64, 853. doi: 10.1001/archpsyc.64.7.853

Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *J. Stat. Softw.* 69, 1–33. doi: 10.18637/jss.v069.i01

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *J. Child Psychol. Psychiatry* 36, 1365–1382. doi: 10.1111/j.1469-7610.1995.tb01669.x

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., and Pickles, A. (2006). Autism from 2 to 9 years of age. *Arch. Gener. Psychiatry* 63, 694–701. doi: 10.1001/archpsyc.63.6.694

Lord, C., Rutter, M., DiLavore, P. C, Risi, S. (2000). *Autism Diagnostic Observation Scale*. Los Angeles, CA: Western Psychological Services. doi: 10.1037/t17256-000

Marrus, N., Hall, L. P., Paterson, S. J., Elison, J. T., Wolff, J. J., Swanson, M. R., et al. (2018). Language delay aggregates in toddler siblings of children with autism spectrum disorder. *J. Neurodev. Disord.* 10, 29. doi: 10.1186/s11689-018-9247-8

Maxwell, M. (1992). *Manual for the Family Interview for Genetic Studies (FIGS)*. Bethesda, MD: National Institute of Mental Health.

Miller, M., Iosif, A.-M., Young, G. S., Hill, M., Hanzel, E. P., Hutman, T., et al. (2016). School-age outcomes of infants at risk for autism spectrum disorder. *Autism Res. Off. J. Int. Soc. Autism Res.* 9, 632–642. doi: 10.1002/aur.1572

Mullen, E. M. (1995). *Mullen scales of Early Learning,* AGS Edn. Circle Pines, MN: American Guidance Service Inc.

Ozonoff, S., Iosif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., et al. (2010). A prospective study of the emergence of early behavioral signs of autism. *J. Am. Acad. Child Adolesc. Psychiatry* 49, 256–266. doi: 10.1016/j.jaac.2009.11.009

Pecukonis, M., Young, G. S., Brian, J., Charman, T., Chawarska, K., Elsabbagh, M., et al. (2022). Early predictors of language skills at 3 years of age vary based on diagnostic outcome: a baby siblings research consortium study. *Autism Res.* 15, 1324–1335. doi: 10.1002/aur.2760

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2021). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3. 1–153, https://CRAN.R-project.org/package=nlme

Piven, J., Elison, J. T., and Zylka, M. J. (2017). Toward a conceptual framework for early brain and behavior development in autism. *Mole. Psychiatry* 22, 1385–1394. doi: 10.1038/mp.2017.131

Rozga, A., Hutman, T., Young, G. S., Rogers, S. J., Ozonoff, S., Dapretto, M., et al. (2011). Behavioral profiles of affected and unaffected siblings of children with autism: contribution of measures of mother-infant interaction and nonverbal communication. *J. Autism Dev. Disord.* 41, 287–302. doi: 10.1007/s10803-010-1051-6

Sandbank, M., Bottema-Beutel, K., Crowley, S., Cassidy, M., Dunham, C., Feldman, J. I., et al. (2020). Project AIM: autism intervention meta analysis for studies of young children. *Psychol. Bull.* 146, i–i. doi: 10.1037/bul0000215

Shen, M. D., Nordahl, C. W., Young, G. S., Wootton-Gorges, S. L., Lee, A., Liston, S. E., et al. (2013). Early brain enlargement and elevated extra-axial fluid in infants who develop autism spectrum disorder. *Brain* 136, 2825–2835. doi: 10.1093/brain/awt166

Sparrow, S. S., Cicchetti, D., and Balla, D. A. (2005). *Vineland Adaptive Behavior Scales,* 2nd Edition *(Vineland-II)* [Database record]. APA PsycTests. doi: 10.1037/t15164-000

Stallworthy, I. C., Lasch, C., Berry, D., Wolff, J. J., Pruett, J. R., Marrus, N., et al. (2021). Variability in responding to joint attention cues in the first year is associated with autism outcome. *J. Am. Acad. Child Adolesc. Psychiatry* 61, S089085672100294X. doi: 10.1016/j.jaac.2021.03.023

Swanson, M. R., Shen, M. D., Wolff, J. J., Elison, J. T., Emerson, R., Styner, M., et al. (2017). Subcortical brain and behavior phenotypes differentiate infants with autism versus language delay. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 2, 664–672. doi: 10.1016/j.bpsc.2017.07.007

Toth, K., Munson, J. N., Meltzoff, A., and Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: joint attention, imitation, and toy play. *J. Autism Dev. Disord.* 36, 993–1005. doi: 10.1007/s10803-006-0137-7

Watson, L. R., Crais, E. R., Baranek, G. T., Turner-Brown, L., Sideris, J., Wakeford, L., et al. (2017). Parent-mediated intervention for one-year-olds screened as at-risk for autism spectrum disorder: a randomized controlled trial. *J. Autism Dev. Disord.* 47, 3520–3540. doi: 10.1007/s10803-017-3268-0

Wetherby, A. M., and Prizant, B. M. (2002). *Communication and Symbolic Behavior Scales: Developmental Profile*. Paul H Brookes Publishing.

Wetherby, A. M., Watt, N., Morgan, L., and Shumway, S. (2007). Social communication profiles of children with autism spectrum disorders late in the second year of life. *J. Autism Dev. Disord.* 37, 960–975. doi: 10.1007/s10803-006-0237-4

Whitehouse, A. J. O., Varcin, K. J., Pillar, S., Billingham, W., Alvares, G. A., Barbaro, J., et al. (2021). Effect of preemptive intervention on developmental outcomes among infants showing early signs of autism: A randomized clinical trial of outcomes to diagnosis. *JAMA Pediatr*. 11, e213298. doi: 10.1001/jamapediatrics.2021.3298

Yoder, P., Stone, W. L., and Edmunds, S. R. (2021). Parent utilization of ImPACT intervention strategies is a mediator of proximal then distal social communication outcomes in younger siblings of children with ASD. *Autism* 25, 44–57. doi: 10.1177/1362361320946883

Yoder, P., Watson, L. R., and Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *J. Autism Dev. Disord.* 45, 1254–1270. doi: 10.1007/s10803-014-2286-4

Zwaigenbaum, L., Bryson, S. E., Rogers, T., Roberts, W., Brian, J. A., and Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life Lonnie. *Int. J. Dev. Neurosci.* 23, 143–152. doi: 10.1016/j.ijdevneu.2004.05.001

frontiers | Frontiers in Communication

TYPE Correction
PUBLISHED 01 February 2023
DOI 10.3389/fcomm.2023.1120237

Check for updates

OPEN ACCESS

APPROVED BY
Frontiers Editorial Office,
Frontiers Media SA, Switzerland

*CORRESPONDENCE
Shruthi Ravi
Shruthi.Ravi1@utdallas.edu

†These authors share senior authorship

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

RECEIVED 09 December 2022
ACCEPTED 02 January 2023
PUBLISHED 01 February 2023

CITATION
Ravi S, Bradshaw A, Abdi H, Meera SS, Parish-Morris J, Yankowitz L, Paterson S, Dager SR, Burrows CA, Chappell C, St.John T, Estes AM, Piven J, Swanson MR and the IBIS Network (2023) Corrigendum: Are early social communication skills a harbinger for language development in infants later diagnosed autistic?—A longitudinal study using a standardized social communication assessment. *Front. Commun.* 8:1120237. doi: 10.3389/fcomm.2023.1120237

Corrigendum: Are early social communication skills a harbinger for language development in infants later diagnosed autistic?—A longitudinal study using a standardized social communication assessment

Shruthi Ravi[1]*, Allison Bradshaw[1], Hervé Abdi[1], Shoba Sreenath Meera[2], Julia Parish-Morris[3], Lisa Yankowitz[4], Sarah Paterson[5], Stephen R. Dager[6], Catherine A. Burrows[7], Chad Chappell[8], Tanya St.John[6], Annette M. Estes[6], Joseph Piven[8†], Meghan R. Swanson[1†] and the IBIS Network

[1]Department of Psychology, The University of Texas at Dallas, Richardson, TX, United States, [2]Department of Speech Pathology & Audiology, National Institute of Mental Health and Neurosciences, Bangalore, India, [3]Center for Autism Research, Children's Hospital of Philadelphia, Philadelphia, PA, United States, [4]Boston Children's Hospital, Harvard Medical School, Boston, MA, United States, [5]The James S. McDonnell Foundation, St. Louis, MO, United States, [6]Department of Radiology, University of Washington, Seattle, WA, United States, [7]Department of Pediatrics, University of Minnesota, Minneapolis, MN, United States, [8]Department of Psychiatry, University of North Carolina, Chapel Hill, NC, United States

KEYWORDS

autism, language, social communication, longitudinal, infancy

A corrigendum on
Are early social communication skills a harbinger for language development in infants later diagnosed autistic?—A longitudinal study using a standardized social communication assessment

by Ravi, S., Bradshaw, A., Abdi, H., Meera, S. S., Parish-Morris, J., Yankowitz, L., Paterson, S., Dager, S. R., Burrows, C. A., Chappell, C., St.John, T., Estes, A. M., Piven, J., Swanson, M. R., and the IBIS Network. (2022). *Front. Commun*. 7:977724. doi: 10.3389/fcomm.2022.977724

In the published article, there were two errors in Table 4, *Post-hoc* cross-sectional analysis exploring the main effect for group, as published. There is a missing "<" symbol for understanding scores at 24-months, and an incorrect "<" symbol for object use scores at 24 months under the *post-Hoc* comparisons column. The corrected Table 4, *Post-hoc* cross-sectional analysis exploring the main effect for group, appears below.

The authors apologize for this error and state that this does not change the scientific conclusions made in the article in any way. The original article has been updated.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

TABLE 4 *Post-hoc* cross-sectional analysis exploring the main effect for group.

CSBS scores	HL-ASD (a)		HL-Neg (b)		LL-Neg (c)		Overall group comparison					*Post-Hoc* comparisons
	EMM	SE	EMM	SE	EMM	SE	*df*	SS	*F*	*q*	f^2	
12-months												
Emotion eye gaze (*N* = 378)	9.63	0.44	10.95	0.23	12.06	0.33	2	213.90	9.88	<0.01*	0.07	a < b < c
Communication (*N* = 396)	9.07	0.56	11.80	0.29	12.89	0.42	2	549.60	15.17	<0.01*	0.10	a < b, c
Gestures (*N* = 396)	6.57	0.47	8.53	0.24	9.92	0.35	2	415.30	16.45	<0.01*	0.10	a < b < c
Sounds (*N* = 396)	3.52	0.59	6.34	0.31	5.73	0.44	2	375.50	9.32	<0.01*	0.06	a < b, c
Words (*N* = 396)	0.40	0.24	1.00	0.13	0.95	0.18	2	17.86	2.62	0.25	0.02	a, b, c
Understanding (*N* = 262)	1.04	0.54	1.62	0.29	2.07	0.41	2	25.26	1.67	0.42	0.01	a, b, c
Object use (*N* = 334)	4.84	0.47	5.91	0.26	6.29	0.36	2	69.90	3.14	0.05	0.02	a, b, c
24-months												
Emotion and eye gaze (*N* = 410)	9.96	0.38	14.49	0.19	14.34	0.27	2	1,016.00	60.18	<0.01*	0.30	a < b, c
Communication (*N* = 432)	13.40	0.43	17.50	0.22	16.70	0.31	2	853.30	37.03	<0.01*	0.19	a < b, c
Gestures (*N* = 432)	9.08	0.45	12.59	0.23	12.53	0.32	2	651.90	26.63	<0.01*	0.14	a < b, c
Sounds (*N* = 432)	13.00	0.69	19.60	0.36	19.20	0.50	2	2,246.90	37.73	<0.01*	0.19	a < b, c
Words (*N* = 432)	9.55	0.99	17.75	0.51	18.78	0.71	2	4,008.20	33.43	<0.01*	0.18	a < b, c
Understanding (*N* = 289)	8.27	1.03	17.44	0.54	19.70	0.66	2	3,745.00	45.39	<0.01*	0.36	a < b < c
Object use (*N* = 413)	10.20	0.56	13.60	0.28	13.70	0.41	2	608.00	16.72	<0.01*	0.10	a < b, c

*Significant interaction that survived adaptive FDR procedure.

10.3389/fcomm.2023.1120237

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 21 October 2022
DOI 10.3389/fpsyg.2022.999396

Check for updates

OPEN ACCESS

EDITED BY
Leonard Abbeduto,
University of California,
Davis, United States

REVIEWED BY
Marie Thérèse Le Normand,
Institut National de la Santé et de la
Recherche Médicale (INSERM), France
Audra Sterling,
University of Wisconsin-Madison,
United States
Anne Hoffmann,
Rush University,
United States

*CORRESPONDENCE
Amanda Mankovich
amanda.mankovich@uconn.edu

SPECIALTY SECTION
This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

RECEIVED 21 July 2022
ACCEPTED 23 September 2022
PUBLISHED 21 October 2022

CITATION
Mankovich A, Blume J, Wittke K,
Mastergeorge AM, Paxton A and
Naigles LR (2022) Say that again:
Quantifying patterns of production for
children with autism using recurrence
analysis.
Front. Psychol. 13:999396.
doi: 10.3389/fpsyg.2022.999396

Say that again: Quantifying patterns of production for children with autism using recurrence analysis

Amanda Mankovich[1]*, Jessica Blume[2], Kacie Wittke[3], Ann M. Mastergeorge[2], Alexandra Paxton[4,5] and Letitia R. Naigles[1]

[1]Department of Psychological Sciences, UConn Child Language Lab, University of Connecticut, Storrs, CT, United States, [2]Human Development and Family Sciences, Research in Early Developmental Studies Lab, Texas Tech University, Lubbock, TX, United States, [3]Department of Speech, Language, and Hearing Sciences, University of Connecticut, Storrs, CT, United States, [4]Department of Psychological Sciences, University of Connecticut, Storrs, CT, United States, [5]Center for the Ecological Study of Perception and Action, University of Connecticut, Storrs, CT, United States

The current research study characterized syntactic productivity across a range of 5-year-old children with autism and explored the degree to which this productivity was associated with standardized measures of language and autism symptomatology. Natural language samples were transcribed from play-based interactions between a clinician and participants with an autism diagnosis. Speech samples were parsed for grammatical morphemes and were used to generate measures of MLU and total number of utterances. We applied categorical recurrence quantification analysis, a technique used to quantify patterns of repetition in behaviors, to the children's noun-related and verb-related speech. Recurrence metrics captured the degree to which children repeated specific lexical/grammatical units (i.e., recurrence rate) and the degree to which children repeated combinations of lexical/grammatical units (i.e., percent determinism). Findings indicated that beyond capturing patterns shown in traditional linguistic analysis, recurrence can reveal differences in the speech productions of children with autism spectrum disorder at the lexical and grammatical levels. We also found that the degree of repeating noun-related units and grammatical units was related to MLU and ADOS Severity Score, while the degree of repeating unit combinations (e.g., saying "the big fluffy dog" or the determiner-adjective-adjective-noun construction multiple times), in general, was only related to MLU.

KEYWORDS

autism, grammar, productivity, language development, recurrence

Introduction

In this paper, we use "children with autism" or "children with ASD" as part of a choice to focus on person-first language with this specific developmental sample. We recognize that adults in the autistic community have increasingly advocated for identity-first language (Vivanti, 2020), but this preference has not yet been investigated or established in children. Albeit outside the scope of the current work, we encourage future researchers to investigate preferences for identity- versus person-first language in children so that scholars and others in the field can honor the needs of this community.

The development of grammar marks a shift from the ability to construct relatively simple sentences (e.g., "want ball") to the ability to express more complex ideas (e.g., "I want the large green ball"). Interestingly, compared to typically developing peers, many studies have reported that children with autism spectrum disorder (ASD) exhibit a much wider range of spoken language abilities, including their acquisition and use of grammar (see Eigsti et al., 2011, for review). Variation in production across the spectrum has been demonstrated through measurements of utterance length (e.g., mean length of utterance, or MLU), utterance complexity (e.g., grammatical morphemes and clauses), and amount of word-/utterance-level repetition of a social partner (e.g., echolalia; Kjelgaard and Tager-Flusberg, 2001).

Researchers have proposed that variations in language production are based on *why* children with ASD communicate (Chevallier et al., 2012; Yoder et al., 2015; Mundy et al., 2019; Su et al., 2021). Children with ASD may communicate selectively because of their varying levels of social motivations, such as whether or not they are intentional in their communication and if they are, how varied their pragmatic functions are (e.g., requesting, seeking information, liking, social maintaining, and social orienting). Consistent with this idea, the *elicited bootstrapping hypothesis*, an extension of the transactional model of language development, has suggested that these differences in social motivations may activate a chain reaction with consequences for language production (Camarata and Yoder, 2002; Sameroff, 2009; Su et al., 2021). That is, reduced motivations within social contexts may suppress interest in and production of communication bids. Fewer attempts to communicate thereby provide fewer opportunities to elicit and absorb communicative responses, limiting children's access to functional language models, which may also reduce how much the child speaks.

This variability in social interest to communicate likely contributes to a broad range of language production profiles observed among children with ASD. For instance, if a child is unmotivated to talk within a social interaction, they may say very little to their communication partners, or they may only communicate for a restricted range of pragmatic functions, such as to request (e.g., "*I want bear*"). Additionally, they may use a frozen phrase such as "I want ______," rarely using that same pronoun "I" with other verbs. Such restricted and repetitive production profiles make it challenging to assess whether the child's language knowledge is abstract (e.g., manifesting subject-verb-object structure), and whether their language use is productive or creative. Producing additional utterances within the turn, such as "We bought the toys yesterday" or "I like cuddly animals at the zoo," points to both abstract and productive usage, but requires more talk and hence more motivation to talk.

Linguists have commonly referred to the ability to creatively combine units of meaning (morphology) into complex structures (syntax) as *productivity* (see Baker, 1979; Pinker, 1989; Tomasello, 2000; Hoff, 2012). It is not immediately clear what the wide range of spoken language levels across only a few contexts implies for productivity in ASD. Understanding productivity is critical: Productivity can have trickle-down effects on other components of language, impacting communicative competence (Yorio, 1980; Pinker, 1989; Tomasello, 2000). For instance, children with more frequent and varied productions may later develop a broader vocabulary, which enables them to talk about a wider range of topics. A better grasp of how early grammar manifests productivity among children with autism may help therapists select the most effective targets in clinical sessions.

The objective of the current study is to quantify indicators of productivity across a range of verbal children with ASD and to characterize how these children might vary in their productivity. We introduce a new method of characterizing productivity—namely, recurrence analysis, a nonlinear time series analysis technique used across several disciplines to capture underlying structural patterns of the system (Leonardi, 2012; Webber and Marwan, 2015). Because recurrence analysis involves continuous measurements, it may be particularly well-suited in order to precisely and accurately capture the variability in children's language productivity across the autism spectrum.

Variability in the grammar of children with ASD

Recent work has focused on exploring the nature of structural language production in autism, specifically syntax and morphology (e.g., Park et al., 2012; Zhou et al., 2015; see Boucher, 2012, for review). Of particular interest has been whether children with ASD have typically developing morphological and syntactic language use. Compared with typically developing children (either age-matched or language-matched), the development of syntax and morphology in speech is frequently protracted for children with ASD (Bartolucci et al., 1980; Howlin, 1984; Eigsti et al., 2007; Park et al., 2012; Zhou et al., 2015; Brynskov et al., 2017; Chin et al., 2018; see Boucher, 2012, for review). This line of work suggests that children with ASD produce less complex speech than matched TD children, often measured by *mean length of utterance* (MLU), which counts the morphemes a child uses in their utterances.

In one study, Eigsti et al. (2007) recorded language samples during free play from 5-year-olds with ASD and from TD children matched on vocabulary, talkativeness, and non-verbal mental age.

Compared to TD children, children with ASD produced utterances that were less syntactically complex (i.e., containing fewer verb phrases, noun phrases, and sentence structures), and shorter (i.e., smaller MLU). Thus, these children with ASD appeared to experience syntactic delays separate from lexical achievements.

A longitudinal study by Tek et al. (2014) found both similar and different patterns to Eigsti et al. (2007) cross-sectional data. Across 24 months of development, one ASD subgroup (32 months old at study onset) showed slower growth in MLU and total number of utterances compared to a TD group matched on expressive language skills (20 months old at study onset). This ASD group also lagged on the production of several specific grammatical elements, including a range of verb types and markers plus noun plurals. In contrast, another ASD subgroup developed grammar at similar rates to the TD group (see also Bartolucci et al., 1980; Howlin, 1984; Park et al., 2012).

Thus, more recent work suggests that not all children with ASD follow the same language acquisition trajectories (Kjelgaard and Tager-Flusberg, 2001; Modyanova et al., 2017; Wittke et al., 2017; see Naigles and Chin, 2015, for review). For instance, when using standardized assessments, Kjelgaard and Tager-Flusberg (2001) found at least three language-related subgroups of children with ASD, including those with language impairment, who exhibited language difficulties across all tested syntactic and semantic domains, those with borderline language deficits, who exhibited fewer language difficulties across tested syntactic and semantic domains, and those with neurotypical language. This work marked a call to characterize the entire spectrum of language abilities in ASD, particularly beyond just vocabulary size.

More recent work has continued to compare grammar use in subgroups of children with ASD. For example, Modyanova et al. (2017) examined subject-verb agreement in the elicited productions of 3- to 16-year-old children with ASD possessing normal language (ALN) and those with language impairment (ALI). Those in the ALI group performed more poorly on their elicited production of the present, past regular, and past irregular tenses compared to the ALN group. However, some children with ASD in both ALI and ALN groups performed similarly to TD children, providing further evidence of variability across the spectrum. Moreover, Wittke et al. (2017) characterized sub-phenotypes for grammatical abilities in the speech of 5-year-olds with ASD who engaged in semi-structured play activities. Their analysis focused on children's usage of Brown's (1973) 14 grammatical morphemes, and described three subgroups for the verbal children in their sample: One whose children were highly talkative and virtually error-free in grammatical usage, one whose children were highly talkative but produced numerous grammatical errors, and one whose children produced both fewer and shorter utterances, but whose utterances were relatively error-free.

Taken together, these studies on the heterogeneity of language production in ASD suggest that traditional language sample descriptors like MLU and total utterances do not capture language heterogeneity in describing patterns of typical versus slow and/or grammatically impaired language trajectories, thus warranting more dynamic grammatical analysis strategies. Moreover, in order to understand *productivity* in this population, we will argue that it is important to think about the degree to which children combine new grammatical structures independently from the degree to which they combine words, and keep in mind that MLU conflates word and grammatical unit combinations. Furthermore, the context and topics of the samples contribute to variability in grammatical usage (Kover et al., 2014). As we describe below, studies investigating productivity in children with ASD have yielded mixed results, in part because the measures of productivity have not clearly distinguished word combinations from grammatical combinations.

Assessments of productivity in ASD compared to TD[1]

Among TD individuals, productivity is usually demonstrated when a person uses a grammatical construction (a) with five or more lexical items (Rispoli et al., 2009), (b) with novel lexical items (Pinker, 1989; Akhtar and Tomasello, 1997), (c) with different morphological endings (Akhtar and Tomasello, 1997; Tomasello et al., 1997), and/or (d) consistently across obligatory contexts (Brown, 1973). In contrast to studies of TD children, which have yielded estimates of consistent productivity by the age of 2 years, examinations of productivity in speech among preschool-aged children with ASD are very limited and have yielded mixed results. That is, some children are found to be consistently productive across grammatical constructions, whereas others show productivity with some constructions but not others (see Roberts et al., 2004; Eigsti et al., 2007; Park et al., 2012; Chin et al., 2018; Le Normand et al., 2018). For instance, Roberts et al. (2004) found no distinguishable differences in the degree to which 5- to 15-year-olds with ASD and language-matched TD children produced past and present tense markers for familiar verbs across obligatory contexts. Similarly, Le Normand et al. (2018) recorded child productions during a narrative-elicitation task and found that the ASD group consisting of 5-year-old French speakers did not differ from the age-matched TD group in their production of verbs, pronouns, the imperfect tense, past participle, and case markers across obligatory contexts. However, their ASD group did produce significantly fewer nouns,

1 It is straightforward to assess productivity within comprehension/novel word studies, such as those that investigate whether children can use sentence structures to figure out the meaning of unknown words. Indeed, children can use sentence structures to identify words (e.g., for TD research, see Naigles et al., 2005,2009; Shulman and Guberman, 2007; for ASD research see Naigles et al., 2011). However, it is also important to demonstrate productivity in actual production and that is the aim of the current paper.

adjectives, determiners, and prepositions, meaning that they appeared less productive on these measures. Le Normand and colleagues suggested that nominal morphology may be more difficult for children with ASD to master than verbal morphology.

Furthermore, Eigsti et al. (2007) found that their 5-year-old autism group used significantly fewer subject-verb-object sequences/sentences with three or more different verbs, showing less advanced productivity than their TD group, whereas Park et al. (2012) reported less productivity in preschool-aged children with autism's spontaneous usage of plurals, "ing," and 3rd person singular "-s," but were as productive as the TD group in the usage of articles, auxiliary verbs, and copula verbs. Interestingly, Park et al. (2012) also assessed productivity *via* elicited production of the past tense and plural and found that children whose elicited production of the past tense was not productive nonetheless used the past tense productively in their spontaneous speech.

Additional mixed findings come from a data-rich case study by Chin et al. (2018). Using a Speechome Recorder to collect longitudinal home-based language samples, a 3-year-old child who was later diagnosed with autism was found to produce language comparable to a 2-year-old TD child (matched on language complexity across all the visits) in the number of different verbs they used with each tense/aspect, indicating more advanced productivity. However, compared to the TD child the child with autism produced conventional past, present, and future tenses with fewer verbs and less consistently across obligatory contexts (i.e., less advanced productivity). In other words, the child with autism showed the ability to use grammatical morphemes related to verb tense/aspect but did not do so as flexibly as the TD peer.

Taken together, these findings highlight that establishing the level of productivity manifested by children with ASD in their speech is difficult. Previous studies have primarily examined two types of measures to assess productivity: elicited production scores, from semi-structured procedures meant to elicit specific morphemes, and measures of spontaneous speech from naturalistic language samples. However, elicited production tasks may not be ideal for revealing productivity in children with ASD, because these tasks rely on good participation and social attention. For example, many elicitation tasks provide children with 1–2 stimulus images and prompt children to produce a one-word response using open-ended questions (e.g., "Tell me what he did to the leaves?") or cloze procedure scaffolding (e.g., "What happened? The boy….[raked]"). Children may also be prompted to produce contrasting morpheme markers using learned non-words that correspond to paired stimulus images (e.g., "How many are there? [one/two wug/wugz]"). Lack of productivity within these tasks, then, could arise because the children do not understand the tasks and so do not provide the correct words, or sometimes even any words, for productivity to be assessed (Boucher, 2012).

Beyond these specific procedures in a research context, we know that measuring language in autism comes with challenges (Tager-Flusberg, 2000). Children with autism often present with differences in social behaviors (e.g., differences in levels of attention in structured tasks) and atypical language behaviors like delayed or immediate echolalia (i.e., the delayed or immediate repetition of a social partner's utterances; Tager-Flusberg and Calkins, 1990; American Psychiatric Association, 2013). Differences in attention, motivations to communicate, and test-taking skills may make it challenging to elicit long, rich productions in structured contexts, such as during standardized language testing or even semi-structured language interactions (Scarborough et al., 1991; Koegel et al., 1997; Condouris et al., 2003; Su et al., 2021; see Boucher, 2012, for review). For instance, if a clinician tries to elicit a narrative language sample where a child shares a personal story or retells a story from a book, but that child is not interested in the topic, they may produce less language than they might with another topic. And, even if they did produce some language, we might not expect it to be as productive in length, content, and grammatical structures as they would be in the context involving the topic that interested them. In other words, language samples derived from a less engaging context may not be as representative of linguistic skill. As indicated by Kover et al. (2014), the ADOS may offer a more appropriate language sampling context since it comprises several activities, varying across modules and sessions (e.g., Module 2 includes a birthday party task and a snack, whereas Module 3 does not). However, Kover et al. (2014) also point out that the context and speech partners also contribute to children's proclivity to use a wide range of grammatical devices. Park et al. (2012) suggested that differences in procedures (i.e., semi-structured play versus free play versus elicitation tasks) could account for discrepancies in results between their research and other research. Thus, an approach to production data across a range of activities that potentially taps into varied interests would therefore be critical if we want to characterize children with ASD's full range of abilities.

Another limitation of productivity studies lies in their statistical approaches. Although they report a large degree of variability in performance during productivity assessments (e.g., Kjelgaard and Tager-Flusberg, 2001; Park et al., 2012; Tek et al., 2014), results have been based on aggregate mean scores (i.e., counts of morphemes and words). Mean scores likely mask interesting patterns of behavior by eliding important variability, and measures of production beyond frequency may provide insights into differences in productivity for these children (Hoff, 2006; see also Müller-Frommeyer et al., 2020).

Gaps in the literature

While language differences within ASD have been broadly characterized within the literature, several key open questions still exist. First, language development studies of children with ASD have largely focused on group-level differences between children with ASD and age-matched TD peers. However, ASD exists on a spectrum of language abilities that range from minor to severe. The vast range of possible language production outcomes for ASD has not yet been thoroughly investigated.

Second, although we know that lexical and grammatical production abilities range from average to highly impaired, what these differences in language abilities mean for the productivity of syntax—or the degree to which specific lexical and grammatical items are used with different items—remains unclear. For example, we might expect a child who is not productive to only use the word “the” with the word “cat,” whereas a child who is productive would use “the” with all sorts of nouns. Degrees of productivity could be indicated by different recurrence measures, in that children who may be less recurrent in their individual lexical and grammatical productions may also be more recurrent in their *patterns* of productions. For instance, if a child just produces noun phrases (e.g., “the cat,” “a big bear,” and “the bank”), they are highly recurrent in individual grammatical productions (e.g., repeating determiner-noun or determiner-adjective-noun) but less recurrent across a range of grammatical phrases. Having a more advanced syntax means that the child is moving beyond noun phrases; that is, a productive speaker would link noun phrases using verb phrases (e.g., “*would love to* play with the cat”) and prepositional phrases (e.g., “I would love to play with the cat *in the* morning”). This type of analysis is considerably more sensitive than a gross language measure like MLU, which captures the length of utterances but not the grammatical complexity or novelty of word combinations.

Finally, approaches to these group-level differences have been based on composite scores from either standardized tests, lab-based paradigms, or spontaneous speech measures. These measures have been compared using traditional methods of analysis (e.g., means and ranges). However, these methods of analysis make key assumptions about the degree to which different activities elicit the same types of talk. For instance, traditional analyses would suggest that a child who produces rich talk in one task but less advanced talk across several other tasks is relatively unproductive. These analyses are unable to capture data that seem complex or irregular (i.e., children alter speech by task) but may actually involve predictable underlying structures. These analyses are thus problematic given differences in social motivations to talk in autistic individuals (see Chevallier et al., 2012) and the context-sensitivity of language production even among TD individuals (Müller-Frommeyer et al., 2020). These traditional methods also make assumptions about the nature of syntactic abilities within ASD and how components of a linguistic system interact. For instance, earlier analyses of grammatical abilities and productivity have not captured the relative sequential occurrence of recurrent words and grammatical units. That is, currently, it is unclear how individual items (i.e., words and grammatical units) unfold relative to one another across a whole language transcript. This is problematic since the ordering of particular words and grammatical units is essential to understanding the nature of the productivity of syntax.

A nonlinear approach to studying productivity would allow for the representation of linear interactions within child language as well as a broad range of other special component interactions informative to syntax that often get masked by summative analyses. Furthermore, this approach does not make assumptions about the distribution of data points across a sequence or even their stationarity (i.e., how the mean state changes across a sequence of behaviors); this is meaningful for small data sets, as well as data sets that contain outliers. This is true of many language studies containing heterogeneous groups of children with ASD. Thus, one potentially valuable tool to characterize the unfolding of grammatical abilities in ASD into a fruitful syntax typology is RQA, a technique to understand how units of speech repeat across stretches of transcriptions.

Microlevel assessment of language production

Many studies have focused on standardized testing and language production scores to characterize children’s early language abilities. Furthermore, most assessments of linguistic repetition are not measured quantitatively so degrees of repetition are not really known. An informative alternative to characterizing language abilities would be a more microlevel assessment of children’s productions with a fine-grained analysis of their actual linguistic and grammatical structures—and more specifically, how frequently and in what ways these structures are being repeated. Understanding the nature of repetitions of words and grammar is important because it may provide insights into the degree to which children combine meanings of units in a creative way (i.e., productivity). For example, children who are repetitive in their word combinations (i.e., saying the same words in the same order), perhaps due to delayed echolalia, are likely less productive than children who utter repetitions of grammatical combinations.

Recurrence quantification analysis (RQA) is a nonlinear approach that quantifies change in a system over time (see Marwan et al., 2007; Webber and Marwan, 2015). RQA allows researchers to quantify how a time series repeats values or patterns of values across a period of observation to provide insights into the relative deterministic properties and flow of changes of the target phenomenon (e.g., types of words and grammatical units). While a comprehensive description of RQA is beyond the scope of the current work, we provide a conceptual overview of its principles and procedures; further methodological details and empirical applications can be found in Riley and Van Orden (2005), Orsucci et al. (2006), Coco and Dale (2014), and Leonardi (2012).

Categorical RQA is a variant of RQA that specifically examines the structures and patterns within discrete data, such as language (e.g., Dale and Spivey, 2006). In general, recurrence—or *repetition*—between adult interlocutors has been considered “good” at the pragmatic level because it indicates that the interlocutors are aligned in semantic interests and thereby engaged in the same conversation (i.e., semantic alignment; Dale and Spivey, 2006; Fusaroli et al., 2020, Unpublished manuscript[2]).

2 Fusaroli, R., Weed, E., Fein, D., Naigles, L. (2020). *Caregiver linguistic alignment to autistic and typically developing children: a natural language processing approach illuminates the interactive components of language* Development. Unpublished manuscript.

However, completely verbatim repetition of the addressee's previous speech would be considered less ideal since it would not further the dialog, or could reflect the echolalia that may be a reflection of restricted behavior/interests. Thus, recurrence could be inflated by echolalia or perseveration. Recurrence of *specific* patterns, though, could reflect the rehearsal of newly acquired structures with the implied goal of morpheme mastery in functional social communication contexts.

To date, the research comparing grammar and word recurrence has been limited (see Leonardi, 2012, for review). Previous researchers using RQA have focused on (1) lexical mirroring of two TD interlocutors (Dale and Spivey, 2006) and (2) the changes in language styles (i.e., broad function word category items such as pronouns, articles, prepositions, auxiliary verbs, adverbs, conjunctions, and negations) of a single TD interlocutor (Müller-Frommeyer et al., 2020). For instance, in an analysis of recurrence in language styles, Müller-Frommeyer and colleagues found that recurrence rate (i.e., the degree of repetition; RR) was perfectly correlated with the proportion of function words, indicating that our RQA-based approach is meaningful when compared against more traditional metrics. However, compared to monologues, conversations elicited a higher determinism of function words (i.e., a measure of how structured repetitions are across speech). Findings indicate that metrics such as determinism can shed light on the patterning of language which cannot be captured by counts and proportions (i.e., recurrence rate).

We suggest that at the *grammatical* and *word* levels, high recurrence of individual items is indicative of less advanced grammatical and lexical production abilities because the child would simply be repeating themselves. For example, a highly lexically recurrent child might say "the ball" three times without adding in any further details about its size, shape, and capabilities (i.e., to be thrown and bounced). Such lexical repetitions may signify echolalic speech. A highly grammatically recurrent child might reuse the same parts of speech over and over again (i.e., determiner-noun; "the cat," "a bag," "my toy"). This latter child might be expected to be less productive as well, as they are not trying out a variety of grammatical units. However, other recurrence parameters focusing on the patterning of words (e.g., percent determinism; %DET) capture something more than simply word count or proportions, including features of the communicative context (e.g., having a conversational partner changed the structure of how function words were used in Müller-Frommeyer et al., 2020). Regular structure in how these items pattern (i.e., %DET) could be indicative of adapting language style to another person across the course of a conversation. This adaptiveness might therefore provide evidence of more advanced grammatical and lexical abilities because the child is practicing new ways to combine units.

Applying RQA to understanding language heterogeneity in autism would address three important gaps in various literatures. From a measurement perspective, assessments of language abilities do not currently respect the continuous nature of the phenomena: Most productivity and repetitive speech measures are currently all-or-none, despite our understanding that autistic language exists on a spectrum. From a methodological perspective, although scholars have claimed that RQA can uncover some structural differences in language, studies have not yet directly compared the grammatical and word levels of analysis. From a language development perspective, researchers have yet to explore the sequential structures that make up noun and verb phrases at both the lexical and grammatical levels. Understanding how repetitive language patterns are structured within these types of phrases has implications for how spoken language production is assessed and described in this population. In summary, RQA has been used to assess diversity and alignment of semantic and lexical productions primarily within typically developing populations. Thus, tackling these topics *via* RQA will add valuable information to understanding the nature of early productivity in ASD.

Current study

The primary goal of this research is to more subtly characterize the language production of a heterogeneous sample of children with ASD.

We do this first by focusing on the degree to which lexical and grammatical units repeat within the language data from 5-year-old children with ASD. To answer this question, we reanalyze the dataset from Wittke et al. (2017) due to the heterogeneity of syntactic ability within its sample (including, e.g., children who were highly talkative or minimally talkative, children who produced many or few grammatical errors; see Wittke et al., 2017, for additional information about participants and tasks). These data provide an excellent opportunity to apply RQA to capture this variability because the summative analyses used in the initial study may have masked meaningful language information in the sample. Because learning the structure of grammar involves learning how to combine both words and grammatical elements (e.g., nouns, verbs, morphemes) in rule-governed ways, we quantify the degree to which children repeat *specific lexical items* (and *the grammatical units that make up these lexical items*) with items they have never heard in combination before, what we call "syntactic recurrence."

Second, although Wittke et al. (2017) previously assigned the children to three subgroups based on their NVIQ and percent of grammatical errors, the present analyses do not focus on these subgroups. Instead, we focus on individual differences in the production of phrasal constructions across this sample. At the micro (individual) level, we explore whether repetitions are indicative of language measures that Wittke et al. (2017) calculated from the language samples (e.g., mean length of utterance and total number of utterances).

Third, we investigate these questions by using nonlinear methods (i.e., RQA) to quantify patterns of repetition across an individual child's speech. Within this type of analysis, each word in the child's transcript is a sequential datum. Each lexical item is isolated in the transcript and is then divided into morphological and syntactic units. We specifically focus on noun phrases since this grammatical form class develops the earliest (Gentner, 1982; Goldfield and Renick, 1990; Fenson et al., 1994). We also focus on verb phrases since they are crucial pieces for children to start building their very first sentences (Gleitman, 1990; Bloom, 1993).

The current study involved several hypotheses about the mappings between RQA and linguistic structure, not necessarily specific to ASD. Broadly speaking, we test whether more advanced syntax, measured *via* traditional linguistic measures and then *via* RQA, could be an indicator that a child is more productive (i.e., less recurrent). In particular, we hypothesized that producing more utterances overall would be associated with a lower RR, but also with longer sequences (i.e., higher %DET), of repeated units. We also predicted that more complex language (i.e., higher MLU) would be associated with less repetition (lower RR), and with longer sequences at the lexical and grammatical levels of noun and verb phrases (higher %DET).

Materials and methods

Corpus

The participant dataset for the current study started with the 189 children with ASD from the Autism Phenome Project (APP). The APP is a longitudinal project conducted at the Medical Investigation of Neurodevelopmental Disorders (MIND) Institute (University of California, Davis), and it examines the neurobiological, genetic, and behavioral features of autism. Children were recruited within northern California with exclusionary criteria based on diagnosis, age, and language exposure (i.e., children were only exposed to English or to both English and Spanish). The first time the children participated in the APP was at age 3 years (Wave 1), often following the child's initial diagnosis of ASD. However, almost 100 children returned for additional assessments through the APP around 5 years of age (Wave 3; $n = 98$).

Child participants of the APP at Wave 3 engaged in extensive behavioral testing, including standardized language assessments. The comprehensive assessment battery included the *Autism Diagnostic Observation Schedule* (ADOS; Lord et al., 2000), for confirmation of autism diagnostic status; the *Differential Ability Scale*, *Second Edition* (DAS-II; Elliott, 2007), to obtain a non-verbal IQ score; and the *Peabody Picture Vocabulary Test*, *Third Edition* (PPVT-3; Dunn and Dunn, 1997) and *Expressive One-Word Picture Vocabulary Test*, *Third Edition* (EOWPVT-3; Brownell, 2000), to assess both receptive and expressive vocabulary abilities. Previously, Wittke et al. (2017) classified the children based on their language and non-verbal IQ scores (see Table 1). Classifications included: (1) *High Verbal* children, scoring in the typical range (standard scores of 85 and above) for both non-verbal and vocabulary language testing; (2) *Low Verbal* children, whose non-verbal IQ standard scores ranged from 71 to 85 and with standardized testing commensurate with their non-verbal IQ; and (3) *Minimally Verbal* children, whose non-verbal IQ and vocabulary performance was significantly below average (i.e., standard scores of 70 or less). Here, we treat ASD symptomatology and language as continuous variables in order to take advantage of increased variance in the data and to identify patterns with further nuance in the dataset, but a previous analysis of this dataset grouped participants into discrete categories. We present their descriptive statistics in Table 1 to provide an overview of the dataset.

TABLE 1 Means for original groups based on standardized test scores.

Measure	High verbal ($n = 38$)	Low verbal ($n = 11$)	Minimally verbal ($n = 33$)
	M (SD)	M (SD)	M (SD)
Age (in months)	68.63 (12.21)	66.20 (7.60)	68.88 (12.52)
NVIQ	102.95 (11.68)	78.70 (2.54)	56.29[a] (10.22)
ADOS	11.89 (5.13)	17.80 (4.92)	22.25 (2.76)
DAS verbal	48.44 (8.80)	32.90 (9.35)	13.81[b] (6.52)
PPVT-3	98.47 (14.33)	75.63 (17.77)	44.67 (10.56)
EOWPVT-3	101.03 (16.52)	76.00 (12.63)	60.94 (7.86)

NVIQ, standard score on differential ability scale, second edition (DAS-II); ADOS, Autism Diagnostic Observation Schedule; DAS verbal = T-score on DAS-II; PPVT-3, standard score on Peabody Picture Vocabulary Test, third edition; EOWPVT-3, standard score on Expressive One-Word Picture Vocabulary Test, third edition. This table is drawn from Wittke et al. (2017).

[a]Only 14 participants in the *Minimally Verbal* group were able to participate in the *DAS-II* testing. The remainder of this group completed the *Mullen Scale of Early Learning* (MSEL; Mullen, 1995) at Wave 3, and their mean group *T*-scores on this measure was 20 ($SD = 0$), indicating floor-level performance for those children who completed the MSEL.

[b]This reflects the group mean for only the 14 participants in this group who participated in the DAS-II.

All children were autistic and were diagnosed based on the DSM-IV American Psychiatric Association (2000). Additional exclusion criteria were applied for the current study after screening assessment performance within the available data sample. One child was excluded because autism diagnostic criteria were not met based on ADOS cutoff scores at Wave 3. Another child was excluded because performance on expressive language and speech production measures were affected by intelligibility difficulties exacerbated by suspected childhood apraxia of speech. Furthermore, because the focal research question in the current study concerned language production, an additional 29 children were excluded because they did not produce enough language (i.e., at least 20 utterances) during the ADOS, which was used for retrospectively transcribing spontaneous language samples. This is perhaps unsurprising given that all of these children had also been classified as Minimally Verbal, although two participants from the Minimally Verbal group were included in the sample since they did produce spontaneous language, $N(\text{utterances}) = 33$ and 124. Video recordings were not available for an additional 16 children due to recording errors (i.e., session not taped or file corrupted), and so they were also excluded from this analysis.

The final sample comprised 51 of the original 98 children (Wave 3 of the study; $M_{age} = 68.84$, $SD = 12.77$), all of whom had language transcriptions collected from ADOS recordings. The sample included 36 males, 13 females, and 2 children whose sex was not reported. The sample is predominantly male, consistent with evidence that the rate of diagnosis is higher in males and consistent with the growing consensus that females are likely under-diagnosed due to differences in ASD symptomatology that are not well-captured by current assessment tools (Kanner, 1943; Asperger, 1944; Fombonne, 2009; Kreiser and White, 2014). Descriptive statistics for this broader sample can be found in Table 2.

Transcriptions

As stated, recordings of previous behavioral testing were used for collecting language transcripts for this sample. Children engaged with investigators, administrators, and parents in semi-structured tasks from the ADOS that afforded high levels of spontaneous and unprompted language production (Tager-Flusberg et al., 2009). ADOS tasks were generally administered in the standardized order for each Module, although the clinicians occasionally administered tasks out of order when the child's participation required a change in task type to increase motivation and engagement. Whether the tasks were administered in the standardized sequence or out of the order, all the assigned tasks for these language samples were still transcribed. Of the children in our sample, 25 completed ADOS Module 2, 25 completed ADOS Module 3, and only one completed ADOS Module 1. Language production samples were derived from these tasks and used to construct participants' grammatical profiles. Language-transcribed tasks varied slightly by ADOS Module administered but generally included: Free Play, Birthday Party, Bubble Play, Snack, Make-Believe Play, Conversation, Description of a Picture, Telling a Story from a Book, Cartoons, and Creating a Story. Although a previous study found that the ADOS yielded less complex and productive language from children with ASD than a parent–child play sample (Kover et al., 2014), those researchers included only the first 15 min of the ADOS for their language sample. We aimed to maximize the potential for language output by including selective tasks that encourage language rather than press for social responses only. All audiotapes were transcribed word-for-word by the third author and an undergraduate research assistant. Audiotapes were listened to multiple times and transcribed verbatim. If an utterance or its parts could not be identified after three passes, it was marked as unintelligible. Transcription reliability was reached *via* a consensus process where transcribers watched video recordings together and checked for differences in codes or errors (Shriberg et al., 1984). All discrepancies were discussed by the transcription team until at least 90% inter-rater agreement (range of 92–98%) was achieved; if line agreement was unable to be achieved, such utterances were consequently coded as unintelligible.

Each utterance was then assigned to a speaker—the child, the parent, or the administrator—but only children's utterances are included in the current analysis to focus on their individual language use. Given that we were not interested in how much children were repeating others (i.e., echolalia), rather our focus was on how much children were repeating themselves, we included all speech that was produced in our analyses. All transcripts were analyzed using the Computerized Language Analysis (CLAN) software in CHAT format (MacWhinney, 2008). This software takes the words from a text file and categorizes them according to their free and bound morphemes for a categorical analysis at the morpheme level.

TABLE 2 Sample means for standardized testing and spontaneous speech measures.

Measure	M (SD)	Min	Max
Age (ADOS)	68.84 (12.77)	54	112
NVIQ	95.45 (17.17)	48	146
ADOS (SA)	9.15 (2.10)	3	17
ADOS (RRB)	4.41 (2.26)	0	8
ADOS (Total)	13.57 (5.77)	4	24
ADOS severity score	6.69 (2.10)	2	10
Length of transcript (in minutes)	17.67 (4.22)	10.5	28.5
Total utterances	126.47 (55.39)	29	263
MLU	4.30 (1.62	1.9	9.14
% Ungrammatical utterances	0.08 (0.06)	0	0.29
% Echolalic utterances	0.03 (0.06)	0	0.29
TTR	0.37 (0.11)	0.2	0.67
Verb token	91.27 (53.33)	10	213
Verb type	31.25 (15.53)	5	65
Verb TTR	0.39 (0.11)	0.25	0.69
Noun token	83.59 (47.42)	14	178
Noun type	45.51 (23.43)	9	88
Noun TTR	0.58 (0.11)	0.35	0.85

NVIQ, standard score on differential ability scale, second edition (DAS-II); ADOS, Autism Diagnostic Observation Schedule. MLU, Mean Length of Utterance. TTR, Type-Token Ratio.

Coding

Our purpose was to analyze both the lexical and grammatical levels of children's speech production using RQA. As stated earlier, because nouns are one of the first lexical items that children produce, we analyzed the elements of noun phrases. Furthermore, because verbs are necessary to form meaningful sentences, we also analyzed many elements of verb phrases. Thus, language transcriptions were specifically annotated for noun phrase or verb phrase, lexical and grammatical, components (see Tables 3, 4). In addition, to further distinguish noun and verb coding, we did not include any of the noun phrase structures in the verb phrase-related lexical and grammatical coding (see lines 1, 3, 4, and 5 in Table 4). Hence, verb coding is more properly called verb-related rather than verb phrase. CLAN conventions were used to mark morphological aspects of speech transcriptions and syntactic errors.

TABLE 3 Noun-related category codes.

Sub-level of analysis	Grammatical category	Lexical examples
Syntax	1. Common Noun	dog, spoon, book
	2. Proper Noun	Tuesday, Polar Express, Dora
	3. Pronoun	it, that, those, her, I, him, itself, yourself
	4. Determiner	the, a, an, that, those
	5. Adjective	pretty, old, nice, funny
	6. Adverb	over, next, once, about, today, just, all
	7. Gerund	flying, fishing, jumping, swimming
	8. Wh-question	who, what, when, where, why, how
	9. Number	six, seven, eight
	10. Preposition	(go) in (the house), (look) at (the dinosaur)
Morphology	11. Plural	-s, -es, children
	12. Possessive	-'s, -his, her, my, your, their, our
	13. Superlative	worst, best
	14. Comparative	better, older

The coded items within children's noun -related syntactic and morphology constructions. All morphology items were double coded as nouns (e.g., plural and possessive) or adjectives (e.g., superlative and comparative). The order of these items was preserved across transcriptions. All non-noun-related syntactic and morphology items (including adjectives and plurals outside of the noun phrase-related syntactic and morphology constructions) were coded as random number sequences.

TABLE 4 Verb-related category codes.

Sub-level of analysis	Grammatical category	Lexical examples
Syntax	1. Verb	go, see, play, want
	2. Adverb	(what's gonna happen) now
	3. Preposition	(make) up (a story)
	4. Negative	not
	5. Infinitive	to
Morphology	6. 1st/3rd person singular	was
	7. 1st person singular	are, am
	8. 3rd person singular	is, does, wants
	9. Present participle	gonna, doing, destroying
	10. Past participle	stuck, broke, seen
	11. Present tense	are
	12. Past tense	got, dropped, went, did
	13. Modal/Conditional	would, does, can
	14. Progressive	(what else is happen)-ing
	15. Copula	(here it) is
	16. Auxiliary	(what) is (he doing)

The coded items within children's verb-related syntactic and morphology constructions. All morphology items were double coded as verbs. The order of these items was preserved across transcriptions. All non-verb-related syntactic and morphology items (including adjectives and plurals outside of the verb-related syntactic and morphology construction) were coded as random number sequences and therefore not included in the current analyses. The nouns within the verb-related constructions were also coded as random number sequences (e.g., the word "them" in "go get them" would be coded as a random number), and again not included in the current analyses.

To provide a richer picture of the dataset, we provide an example of the coding below. In this example, a child is responding to a prompt about make-believe play with action figures and tools. Of particular interest is the child's raw speech: Blue text represents all noun phrase-related components, orange text represents all verb-related components, and black text represents components not involved in noun phrases or verb-related constructions. Again, notice that when the child says, "*you knock it you get more power that way*," the words "*you*," "*it*," and "*more power that way*" are marked as parts of noun phrases for the noun-related coding (in blue). By contrast, the verb-related construction coding (in orange) is largely based on morphology and ignores the nouns entirely.

Example 1

Lexical:	**He can jump** super high.
Grammatical:	**Pronoun-modal-verb**-1-2.
Lexical:	Higher **than you can fly**.
Grammatical:	3-**preposition-pronoun-modal-verb**.
Lexical:	**Pretend that's just a baseball** and **you can find it to get more power**.
Grammatical:	**Verb-pronoun-3rd person-adverb-determiner-noun**-4-**pronoun-modal-verb-pronoun-infinitive-verb-noun-noun**.
Lexical:	**Watch you can hold this up** and **how** long **you knock it you get more power that way**.
Grammatical:	**Verb-pronoun-modal-verb-pronoun-adverb**-5-wh**pronoun**-6-**pronoun-verb-pronoun-pronoun-verb-adverb-noun-determiner-noun**.

All raw text was then converted to numerically identified categories (e.g., all nouns coded as ""1," all pronouns coded as "2"). This coding was critical for RQA to reveal how children reuse noun- and verb-related lexical and grammatical structures. To prevent RQA from capturing repeating patterns of non-target grammatical structures, items identified as not being part of noun-related or verb-related lexicon/grammar were coded as unique (i.e., non-repeating) values; this ensured that RQA could only "see" the patterns of language that we were interested in studying here. The coded words within each sentence were strung together in a way that maintained the temporal order of the speech.

Categorical recurrence quantification analysis

In the current work, we apply RQA to the coded transcripts of child language to examine how patterns of children's noun- and verb-related phrases change over time. Thus, this new application involves characterizing the lexical and grammatical constructions of the noun and verb-related phrases within a child's "series" of speech, in which each word in the child's transcript is a sequential measurement (*cf.* Dale and Spivey, 2006). This is the focus we apply here. That is, we characterize the degree to which an *individual* child repeats specific lexical/grammatical items alone and in combination with items they have never repeated in combination before. For instance, in one example of lexical repetition across noun phrases, a child said:

Example 2

Lexical: The frog starts going like (unintelligible word).
Grammatical: Determiner-noun-verb-thirdpersonsingular-verb-presentparticiple-preposition.
Lexical: The frogs are gonna invade the city.
Grammatical: Determiner-noun-plural-auxiliary-present-verb-presentparticiple-verb-determiner-noun.
Lexical: The frogs are leaving trying to invade the city.
Grammatical: Determiner-noun-plural-auxiliary-present-verb-presentparticiple-verb-presentparticiple-1-verb-determiner-noun.

Notice the repetitions of "the frog" and "the city" across the utterances. Furthermore, the child is consistently using determiners with their nouns to form noun phrases. In contrast, another child said:

Example 3

Lexical: Look a frog mom!
Grammatical: Verb-determiner-noun-noun.
Lexical: They're flying.
Grammatical: Pronoun-present-verb-presentparticiple.
Lexical: Hey mom look at frogs are doing.
Grammatical: 1-noun-verb-preposition-noun-plural-auxiliary-present-verb-presentparticiple.

Notice that this second child produces noun phrases with much less repetition both in their lexical items (i.e., "a frog mom," "they," and "frogs" referring to the same concepts) and grammatical items (i.e., determiner, adjective, pronoun, and noun).

As shown in these examples, we identify repetitions in individual categories and across sequences of categories—here, words and grammatical units. By comparing these data, we can characterize how the trajectories of word sequences and grammatical constructions might be more vs. less consistent (i.e., frequently vs. infrequently repeated) within a single speech sample.

A strength of using RQA to quantify patterns across an individual child's speech transcript is that it can be used to examine very short or very long time-series data without assuming a normal distribution of the data (Carello and Moreno, 2005). Although transcriptions varied in the amount of time the children participated in each activity and the number of utterances produced, we decided not to cut longer transcriptions short because these differences in language production are interesting for understanding the wide range of language abilities of children with ASD.

We conducted RQA on the lexical and grammatical data for each child's transcription using the "crqa" package (*version 1.0.9*; Coco and Dale, 2014) from R in RStudio (*version 1.1.423*; R Core Team, 2021). First, we constructed a recurrence matrix that indicates when a time series returned to a given state (e.g., word repetitions across a transcription). Given that we conducted categorical RQA based on the type of data available, this recurrence matrix included only *exact* repetitions of the categorical state under consideration (e.g., each specific lexical item) across the *entire* time series, even lagged across time (similar conceptually to autocorrelation). A separate recurrence matrix was created for each noun- and verb-related lexical and grammatical time series for each child, resulting in four matrices per child.

As a technical point, calculating recurrence matrices from categorical data requires the researcher to provide a unique categorical identifier for each item of interest so that the recurrence matrix will identify any repetition of the same values in the time series. However, if a researcher wishes to remove data from consideration—say, if items in a specific class are not of interest to the given research question—the researcher must be sure to code the data accordingly: If all items outside of the class of interest are given the same categorical identifier, those not-of-interest items will appear as repetitions in the recurrence matrix, skewing the later steps. In the present study, we were exclusively interested in noun-related and verb-related lexical and grammatical items, so all other items in other classes were given random categorical identifiers (i.e., non-repeating negative numbers) to be sure they were not considered as moments of recurrence in the analysis.

Visualization

Each recurrence matrix was plotted to create a recurrence plot (RP; Marwan, 2008), which allows a qualitative inspection of how key features of sequential data change across time (see Figures 1, 2). Each point on the plot represents a single repeated item in the child's production at different points across the transcript. In the present study, RP markings specifically indicate all points within a transcription in which the child repeats either a noun-related or verb-related lexical or grammatical item. For example, an RP for the lexical items in noun-related sequences with the text from Example 2 would pull out repetitions (represented as filled-in points) with the words *the*, *frogs*, and *city*.

If we had analyzed the lexical items in verb-related sequences, an RP for the same text would identify no recurrent sequences since no verb-related repetitions exist (e.g., exact repetitions of "*are gonna invade*" as a verb trigram). That is, there would be no recurring dots for these verb-related lexical items. However, RPs would pull out repetitions in the individual noun-related grammatical units (e.g., repeating determiners five times across Example 2) and verb-related grammatical units (e.g., repeating verbs seven times across Example 2). Thus, each diagonal line represents repetitions in sequences that the child produced at different times throughout the transcript. For instance, the noun phrase "the frogs" would be represented as a diagonal line on the RP since it is repeated twice verbatim.

Metrics

In addition to visual inspection, we can quantify the patterns and sequences of points on RPs to yield a variety of

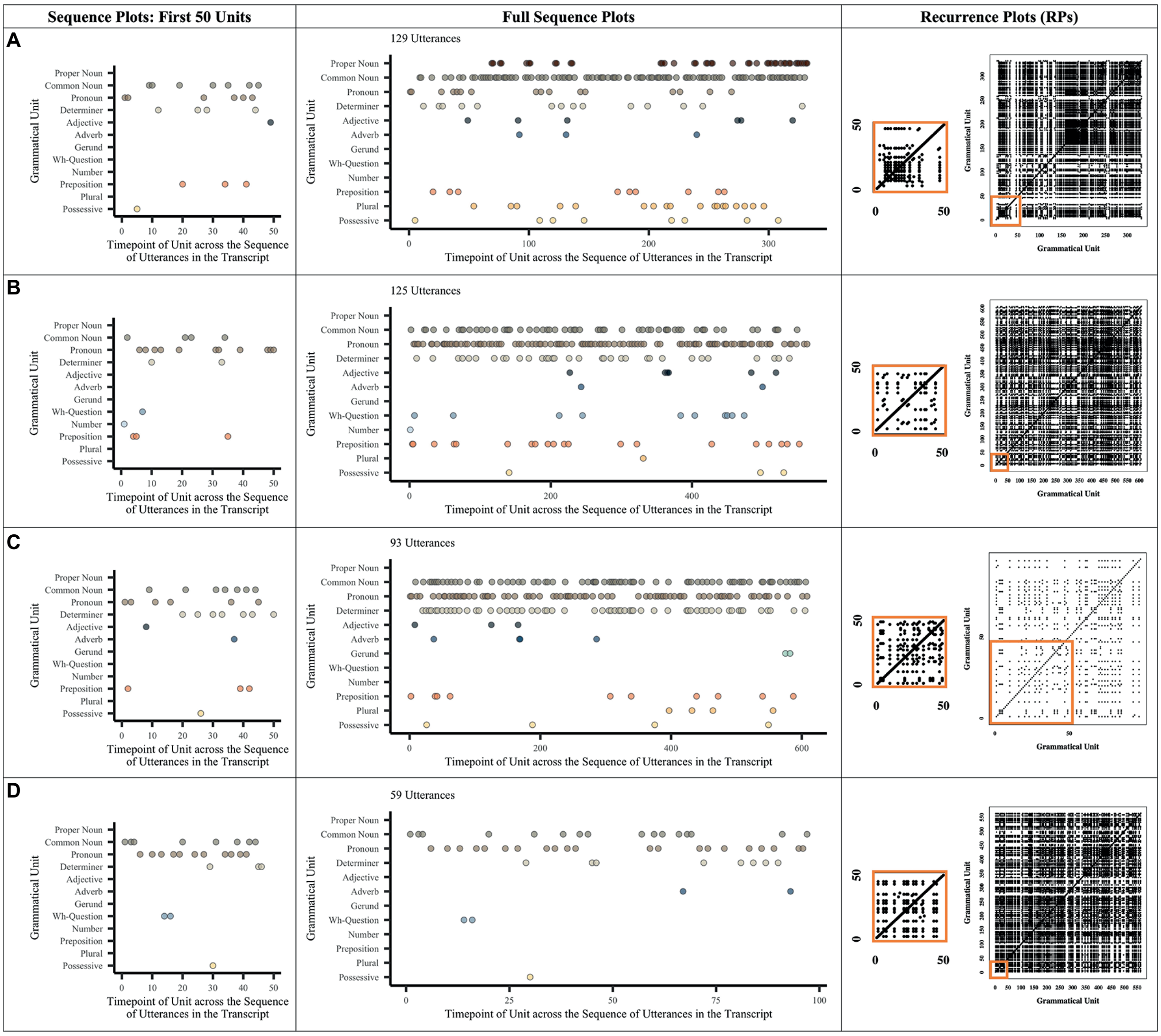

FIGURE 1
Example Plots for the Noun-related Grammar Production of Four Children. The space not covered by dots in the sequence graph represents instances when a child did not either use one of the noun-related grammatical units listed or produced other units not in the noun phrase (e.g., verb-related units, coordinators, and adjectives). Recall that RR is based on a percentage, not on counts. Child **(A)** produced noun-related speech high in %DET and high in RR. Child **(B)** produced high %DET but low RR. Child **(C)** produced low %DET but higher RR. Child **(D)** produced low %DET and low RR. Looking at children **(A)** and **(D)**, each who produced a similar number of utterances, we see that A has a denser RP and more lines than **(D)**.

metrics. Here, we specifically focus on *recurrence rate* and *determinism*. Recurrence rate (RR) captures the percentage of the RP containing filled-in points (relative to all possible points); high RR indicates frequent reuse of lexical or grammatical units. For example, we could track "ice cream" in a single child's transcript: "You got me *ice_cream*. Big ice_creams. You'll have vanilla and I'll have white *ice_cream*." In this example, note that—since only *exact* repetitions would count as recurrent for noun-related lexical items—the word "ice cream" is only counted as repeating twice; the plural "ice creams" is not included. Low RR indicates infrequent reuse of the lexical or grammatical units (e.g., the word "big" in the previous example was only produced once).

When recurrent points occur in succession to create line structures, we can visualize a repeating trajectory. The percent of recurrent points on the RP that involve these diagonal line structures (i.e., two or more consecutive points) is known as percent determinism (also simply called determinism; DET). Determinism can reveal whether strings of repeated structures occur across the same contexts. Note that these repetitions themselves need not be sequential: That is, the repeated strings can occur across the entire transcript as well and are treated the same way. High %DET indicates that children frequently repeat the same lexical or grammatical combinations. For example, consider "ball" in this excerpt of a single child's transcript:

Example 4

Lexical:	Can I play with **the ball**?
Grammatical:	Modal-pronoun-verb-preposition-**determiner-noun**
Lexical:	Where's **the ball**?
Grammatical:	**Pronoun-thirdpersonsingular-determiner-noun**
Lexical:	He likes balls.
Grammatical:	Pronoun-verb-thirdpersonsingular-**noun-plural**
Lexical:	Of dogs that like to play ball.
Grammatical:	**Preposition-noun-plural-pronoun**-verb-1-verb-noun
Lexical:	He likes to play with all **the balls**.
Grammatical:	Pronoun-verb-thirdpersonsingular-2-verb-**preposition-noun-determiner-noun-plural**
Lexical:	And mine too but I **do not let** him have **the balls** but I **do not let** him have **the balls** because.
Grammatical:	3-pronoun-adverb-4-pronoun-**auxiliary-verb**-pronoun-verb-**determiner-noun-plural**-5-pronoun-**auxiliary-verb**-pronoun-verb-**determiner-noun-plural**-6
Lexical:	There's some balls that can that he can choke on it.
Grammatical:	**Pronoun-thirdpersonsingular**-noun-**noun-plural-pronoun**-pronoun-modal-verb-preposition-pronoun

Noun-related units referred to in the text explanation are in blue while verb-related units are in orange. The bolded darker blue (versus the non-bolded lighter blue) indicates that the noun-related lexical or grammatical units are a part of a deterministic sequence; the bolded darker orange (versus the non-bolded lighter orange) indicates that that verb-related lexical or grammatical units are a part of a deterministic sequence. Unlike in the prior examples, the black font in this example indicates that the words/grammatical units are not being counted as part of a deterministic structure. In this example, the child repeats "the ball" twice and "the balls" three times. A closer look at these phrases reveals that the child frequently combines grammatical units in the same way (e.g., preposition-noun; determiner-noun; determiner-noun-plural; noun-plural-pronoun). Lower determinism indicates that children are testing out many different unit combinations (e.g., only repeating the verb-related words "do not let" in this example).

The center line of each RP—the *line of identity* (LOI)—indicates lag-zero. By *lag-zero* (as it is called in autocorrelation), we mean all instances when that moment in the time series is compared to itself; this means that RR is always equal to 1 for the LOI. These self-comparison values do not vary across the children and are therefore ignored in RQA.

Statistical analyses

All analyses were completed using R in RStudio (*version 1.1.423*; R Core Team, 2021). Current best practices for RQA were applied to the data (see Carello and Moreno, 2005; Riley and Van Orden, 2005). Our primary analytic approach was to use linear models to predict changes in RR and %DET, respectively, by Type (Noun-related vs. Verb-related) and Analysis Level (Grammar vs. Words). By also including more macro spontaneous speech metrics in the model (i.e., MLU and Total Number of Utterances,), we can account for variance directly from the structure of the children's language, and we can directly compare the dynamical approach to the traditional approach. Autism Severity Score was included in the model to explore the degree to which repetitiveness was a facet of language development versus a characteristic of being autistic. Supplementary analyses controlling for NVIQ did not improve model fits when predicting either RR or %DET, and so NVIQ was not included in the models.

In interpreting RQA results, it is important to note that many metrics are not inherently meaningful. That is, they are often more useful as relative metrics compared across conditions (e.g., between experimental conditions, between two interlocutors) *via* inferential statistics. However, this could be potentially problematic in the case of understanding whether the observed values differ from those values that might be expected simply by chance. We address this concern using *approximate permutation tests*, which allow a researcher to create and test surrogate time series (i.e., use itself as a baseline; see Chiovaro et al., 2021, and Paxton and Dale, 2017). Permutation tests go beyond the raw frequencies of categories to test the degree to which the *structure* of the categories across the transcript can be found together more often than would be expected by chance (i.e., the baseline). Through these permutation tests, we can evaluate whether categories of words and grammatical units are organized in meaningful ways.

Here, we conducted tests for significance with confidence intervals at the upper and lower bounds of the 95th percentile (comparable to alpha criteria of 0.05). We then created 100 permutations of each participant's transcript (i.e., removing category dependencies across the transcript but maintaining raw frequencies) and conducted RQA on each of these permutations. We compared this output to what we might expect to see by chance, again preserving the participant-level variability (i.e., comparing the observed values from a given participant to the permutation values created from that same participant's data). The proportion of times that the real-time series' values exceed the baseline time series' values is used as the alpha criterion for significance. However, because we maintain the frequencies of the original time series, it is critical to note that permutation tests can *only* be used to establish baselines for RQA metrics that rely on sequences—here, meaning that we can only examine %DET and not RR. Of the permutation tests run for the %DET of noun-related and verb-related words and grammar, respectively (i.e., four measures), we find that 80.39% of noun-related grammatical unit data ($n = 51$; $p_{median} < 0.001$, $p_{sd} = 0.22$), 89.36% of the noun-related lexical data ($n = 47$; $p_{median} < 0.001$; $p_{sd} = 0.13$), 90% of the verb-related grammatical data ($n = 50$; $p_{median} < 0.001$; $p_{sd} = 0.04$), and 97.92% of the verb-related lexical data ($n = 48$; $p_{median} < 0.001$; $p_{sd} = 0.02$) are above the criterion. This means that, in general, the observed structures within the data tend to appear together more than what would be expected by chance.

Results

Characterizing the sequences of grammatical units

Since there are so many possible lexical units within any given noun- or verb-related construction and many fewer possible grammatical units, we only show visualizations of the grammatical unit data. The left-hand side of Figures 1, 2 show sequence figures to characterize children's production of syntactic and morphological units across the span of a single transcript. Note that since each child may vary in the number of grammatical units that they might produce, their x-axes can vary. Each point represents a single syntactic or morphological unit and the order in which they occur (and reoccur) over the course of a transcript. Each sequence figure shows the sheer quantity of units that a single child produces.

Figures 1, 2 also highlight differences in the degree to which children use certain grammatical items within the same ADOS protocol. For instance, children A and D from Figure 1 produce speech that is similar in quantity (i.e., number of utterances and number of noun-related grammatical units); however, D produced a wider range of grammatical units overall (see full sequence plots). For instance, D produced many more pronouns, determiners, wh-questions, and prepositions overall, while A produced many more nouns and number units. In contrast, B produced fewer utterances than A and D but still produced a wide range of grammatical units. Child C produced the fewest utterances and the fewest grammatical units (i.e., did not produce proper nouns, adjectives, gerunds, numbers, prepositions, or plurals).

Figure 2's full sequence plots show that even children who are more similar in utterance quantity (i.e., children F and G) may produce similar numbers of verb-related grammatical units. For instance, while child F produced prepositions, the negative, the 1st- or 3rd-person singular, the past participle, the past tense, and the auxiliary, child G did not. Thus, child F produced more instances and a greater variety of verb-related grammatical units. Children E and G highlight the opposite pattern: Child E produced far fewer utterances than both F and G; however, although E and G produced a different number of utterances, they both similarly produced a small number of grammatical units, especially relative to F. Sequence graphs thus show that RQA is a good measure to capture the differences in how children produce their noun-related and verb-related sequences.

To the right of the sequence plots in Figures 1, 2 are the example corresponding recurrence plots (RPs). RPs also highlight

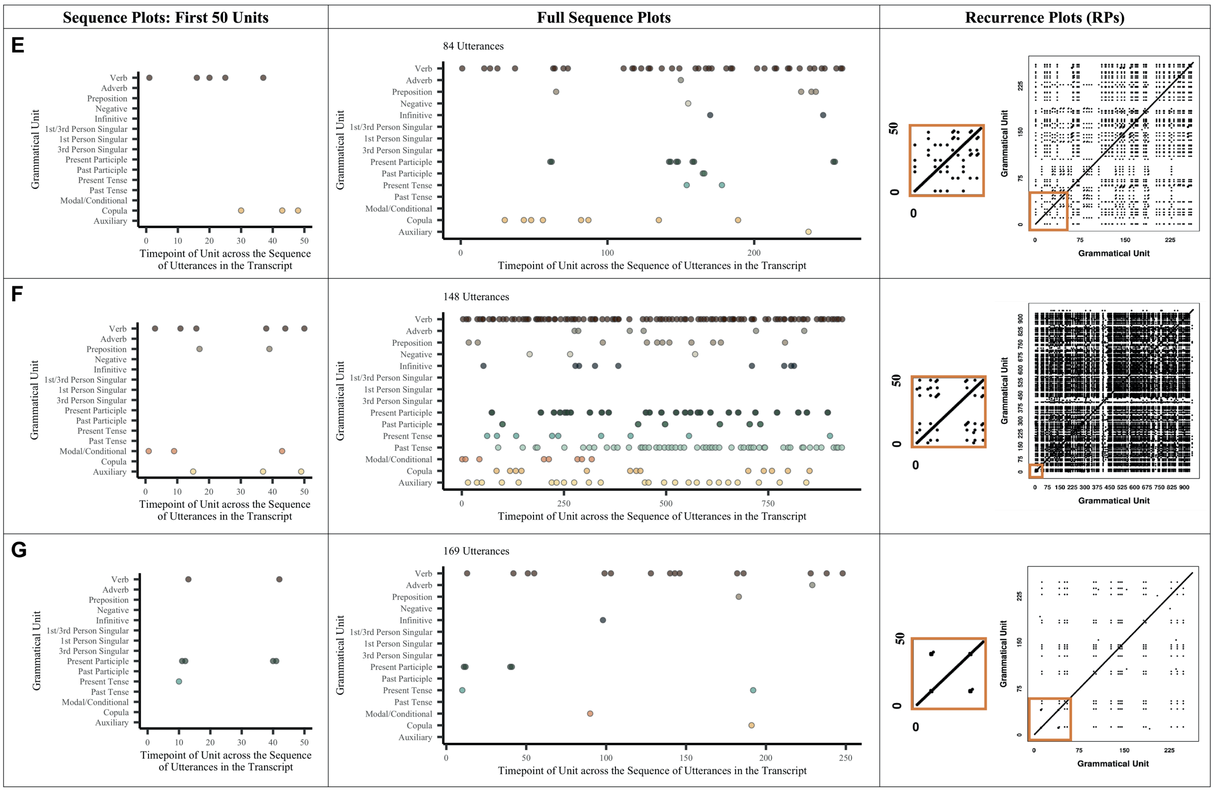

FIGURE 2
Example Plots for the Verb-related Grammar Production of Three Children. The space not covered by dots in the sequence graph represents instances when a child did not either use one of the verb-related grammatical units listed or produced other units not in the verb phrase (e.g., noun-related units, coordinators, and prepositions). Recall that RR is based on a percentage, not on counts. No child produced verb-related speech high in %DET and high in RR. Child **(E)** produced high %DET but low RR. Child **(F)** produced low %DET but higher RR. Child **(G)** produced low %DET and low RR.

the variability of units within our sample. Again, note that the *x*- and *y*-axes represent the categories (word or grammatical units) that children produce across the sequence of a transcript. Because each child may vary in the number of different grammatical units that they produce, their *x*- and *y*-axes differ on the RPs; accordingly, this means that the size of a single point will be larger or smaller on the graph, depending on the total number of possible points. A filled-in space indicates that the child is revisiting a previously used category (i.e., contributing to RR), while the line structures indicate that the child is revisiting a previously used sequence of categories (i.e., contributing to %DET).

The RPs in Figure 1 show that Child A repeats many unit combinations verbatim (possibly indicating less advanced production), whereas B repeats less but often repeats combinations of "preposition determiner noun" and "determiner noun plural" (indicating more advanced production; see sequence plot of first 50 units). C repeats words often but produces few new words across many new sequences (indicating moderately advanced production; e.g., "determiner noun" and "noun"), whereas D keeps using new units in new combinations without revisiting prior ones (indicating moderately advanced production; e.g., "determiner noun" versus "preposition pronoun"). This contrast between C and D is particularly striking in the sequence plots for their first 50 grammatical units. That is, these plots show that C repeats a few units (i.e., common noun and pronoun) quite frequently; D only repeats pronouns frequently. The RPs for verb-related words and grammar were calculated in the same manner.

Recurrence rate (RR)

Our analyses examined the degree to which children tend to reuse the same lexical or grammatical units (i.e., RR) by Type (Noun-related vs. Verb-related) and whether these RR values correlated with the Total Number of Utterances, MLU, and ADOS Severity Score at that visit. Descriptive data for RR by Type (Noun-related vs. Verb-related) and Analysis Level (Lexicon vs. Grammar) are provided in Table 5. These data are visualized in Figure 3. Note that the predictor variable Total Number of Utterances was moderately correlated with both MLU and ADOS Severity Score, whereas MLU and ADOS Severity Score were strongly correlated with one another (see Table 6).

Linear modeling was carried out to investigate whether these variables could significantly predict RR.[3] Results indicated that the model explained 87.32% of the variance in RR and that the model was a significant predictor of RR, $F(15,188) = 94.17$, $p < 0.001$. For clarity and flow, model results—including unstandardized betas and confidence intervals—can be found in Table 7 rather than in the text.

The analysis revealed a main effect of Type, such that noun-related speech involved a higher RR ($M = 3.71$, $SD = 3.00$) than verb-related speech ($M = 1.42$, $SD = 1.26$). We found a main effect of Analysis Level, in which the RR of grammatical units ($M = 4.46$, $SD = 2.40$) was higher than the RR of lexical units ($M = 0.68$, $SD = 0.51$). The two-way interaction between Type and Analysis Level was not significant; however, based on visual inspection of Figure 3, Panel A, we conducted follow-up analyses Tukey's *post-hoc t*-tests comparing the RR of noun-related and verb-related grammatical and lexical units. Our analyses revealed that RR was significantly higher for noun-related words than verb-related words [$B = 0.70$, $t(176) = 3.91$, $p < 0.01$] and for noun-related grammar than verb-related grammar [$B = 3.89$, $t(176) = 21.81$, $p < 0.001$]. Results also revealed a higher RR for noun-related grammar than noun-related words, $B = 5.38$, $t(176) = 30.15$, $p < 0.001$. Similarly, RR was higher for verb-related grammar than verb-related words, $B = 2.19$, $t(176) = 12.24$, $p < 0.001$.

Generally speaking, the analysis further revealed that the Total Number of Utterances was not related to RR in any way. More specifically, we found a main effect of MLU, with MLU increasing as RR decreases. Results revealed a Type-by-MLU interaction. Interactions are visualized in Figure 4. While RR does not vary for verb-related items by MLU, it does for noun-related items, with RR lower for noun-related items when MLU is short (see Figure 4, Panel A). Furthermore, we found a significant Analysis-Level-by-MLU interaction: Although RR and MLU do not change according to lexical units, RR of grammatical units is positively correlated with MLU when the MLU is short but plateaus when MLU is longer (see Figure 4, Panel B). Results also revealed that ADOS Severity Scores positively predicted overall RR.

TABLE 5 Means and standard deviations for RQA metrics.

Phrase type	Measure	M (SD)	Lower 95% CI	Upper 95% CI
Noun	Lexicon RR	1.02 (0.49)	0.74	1.30
	Grammar RR	6.40 (1.78)	6.12	6.68
	Lexicon %DET	12.94 (15.05)	10.20	15.70
	Grammar %DET	21.44 (7.67)	18.70	24.20
Verb	Lexicon RR	0.33 (0.20)	0.05	0.61
	Grammar RR	2.51 (0.86)	2.23	2.79
	Lexicon %DET	14.63 (8.44)	11.90	17.40
	Grammar %DET	18.31 (6.87)	15.5	21.10

Percent determinism (%DET)

We examined characteristics of the degree to which units tend to fall on repeated sequences of the same grouping of units (i.e., %DET) by Type, Analysis Level, MLU, Total Number of

3 We first attempted to use random effects to account for the variance from the participants in a linear mixed-effects modeling approach. However, these random effects led to overfitting in the model in which the estimate of variance became extremely small (i.e., e-16). Therefore, the random effects by participant were dropped from the model, resulting in a linear modeling approach. The same issue emerged for all other models, so random effects are not included in any models.

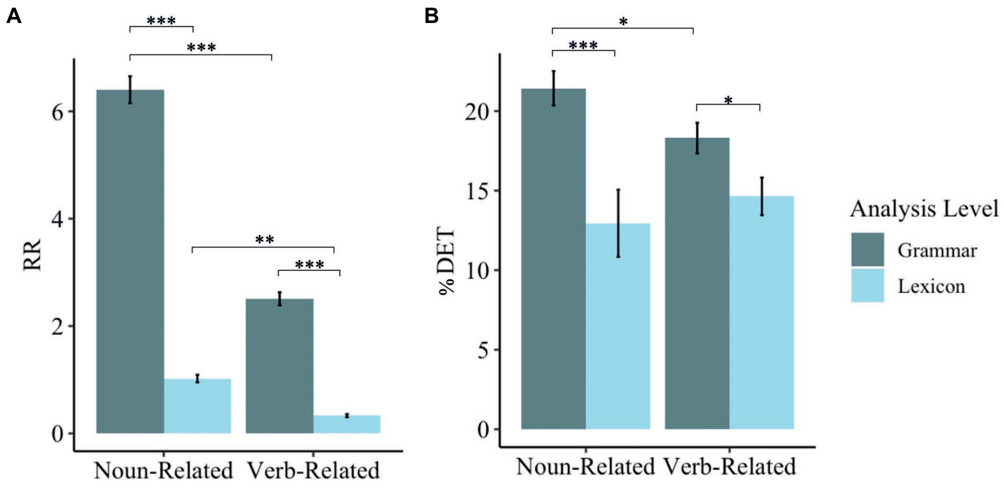

FIGURE 3
RQA metrics for the components of noun-related and verb-related phrases. ***=$p<0.001$; **=$p<0.01$; *=$p<0.05$. Error bars represent standard errors. **(A)** Shows the mean recurrence rate across the sample. **(B)** Shows the mean percent determinism across the sample.

TABLE 6 Correlations between predictor variables.

Measure	1	2	3	4
1. Type				
2. Level of analysis	0			
3. MLU	0	0		
4. Total utterances	0	0	0.32***	
5. ADOS severity score	0	0	−0.59***	−0.21***

Each value represents a correlation coefficient (*r*). ***indicates $p < 0.001$.

Utterances, and ADOS Severity Score. The lexicon %DET values capture how much the children are combining the same words that are in noun-related and verb-related sequences, respectively, in the same way (i.e., productivity). Figure 5 shows the variability in the %DET of the 51 children.

Visual inspection of the figure reveals that children seem to be more productive in their verb-related lexicon than in their noun-related lexicon, as demonstrated by the fewer children repeating the same lexical sequences in their verb-related constructions. Example 4 (above) shows an excerpt from one child and highlights this difference in the %DET of noun-related and verb-related lexicons. In blue bold ink are the repeating noun-related word combinations and grammar combinations. In contrast, the orange bold ink highlights the repeating verb-related word combinations and grammar combinations. While some children reuse word combinations frequently (e.g., "the balls" in Example 4; see Figure 4), in general, it is to a much smaller extent than the degree to which they repeat grammatical combinations (e.g., determiner-noun-plural and auxiliary-verb). Note that while several noun-related combinations are repeated in Example 4, many combinations are new. Descriptive data for %DET by Type and Analysis Level are provided in Table 5.

We used linear regression to test whether main effects of and interactions between Type, Analysis Level, MLU, Total Number of Utterances, and ADOS Severity Score predicted variance in %DET. The model was statistically significant, $F(15,188) = 3.40$, $p < 0.001$, $R^2 = 0.15$. As with our analyses of RR, we present all model results—including unstandardized betas and confidence intervals corresponding to main effects and interactions—in Table 8.

Type did not significantly predict %DET ($p = 0.79$), suggesting that children did not vary in the degree to which they combined units for noun-related and verb-related words and grammar. Children also did not alter their deterministic productions by analysis level ($p = 0.61$). However, based on visual inspection of Figure 3B, we conducted follow-up *t*-tests comparing %DET for words and grammar of noun-related and verb-related units. Results revealed a higher %DET for grammatical units compared to lexical units for both noun-related units [$t(50) = 3.63$, $p < 0.001$] and verb-related units [$t(50) = 2.47$, $p < 0.05$]. We also found a higher %DET for noun-related grammar than verb-related grammar ($p < 0.05$, $d = 0.36$), but this did not hold for words ($p = 0.38$). In general, MLU was positively associated with %DET. This association only emerged once MLU reached approximately 4. Moreover, we found a trending interaction between Level of Analysis and MLU ($p = 0.06$) (see Figure. 4, Panel C). Although the %DET of lexical units does not vary by MLU, it does for grammatical units, with %DET higher for grammatical units when MLU is larger. Finally, we also found a trending positive association between %DET and ADOS Severity Score ($p = 0.067$). No other main effects or interactions were statistically significant.

Discussion

The current study presented an innovative technique (i.e., recurrence quantification analysis) for measuring the productivity

10.3389/fpsyg.2022.999396

TABLE 7 Regression results for the model predicting RR.

Predictor	B	lower 95% CI	upper 95% CI	SE	*t*
(Intercept)	6.93***	5.22	8.62	0.87	7.97
Type	−3.88**	−6.31	−1.45	1.23	−3.16
Analysis level	−6.23***	−8.66	−3.81	1.23	−5.07
MLU	−0.40***	−0.60	−0.20	0.10	−3.91
Total utterances	0	−0.01	0	0.002	−0.47
ADOS severity score	0.20**	0.05	0.35	0.08	2.63
Type: Analysis level	3.55*	0.12	6.98	1.74	2.04
Type: MLU	0.35*	0.06	0.63	0.14	2.14
Type: Total utterances	0	−0.01	0	0.003	−0.69
Type: ADOS severity score	−0.18	−0.39	0.03	0.11	−1.68
Analysis level: MLU	0.40**	0.11	0.68	0.14	2.74
Analysis level: Total utterances	0	−0.01	0.01	0.003	0.11
Analysis level: ADOS severity score	−0.13	−0.35	0.08	0.11	−1.25
Type: Analysis level: MLU	−0.35	−0.75	0.06	0.01	−1.69
Type: Analysis level: Total utterances	0	−0.01	0.01	0.005	0.35
Type: Analysis level: ADOS severity score	0.14	−0.16	0.44	0.15	0.91

$R^2 = 0.873$***. *indicates $p < 0.05$. **indicates $p < 0.01$. ***indicates $p < 0.001$.

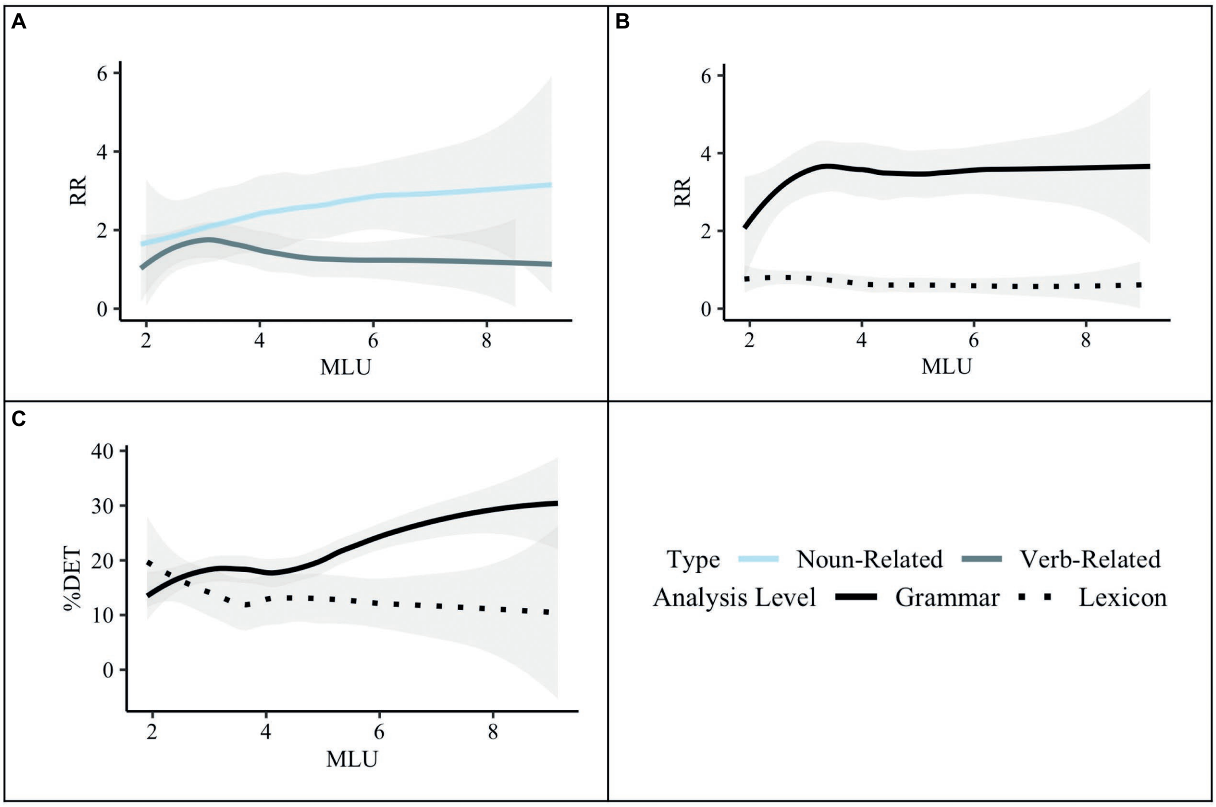

FIGURE 4
Predictors for RR and %DET from mixed effects modeling. Panel **(A)**, shows predictors for RR by type and MLU. Panel **(B)**, shows predictors for RR by the level of analysis and MLU. Panel **(C)**, shows predictors for %DET by Level-of-Analysis and MLU. Panels **(B)** and **(C)**, show the RR and %DET, respectively, for grammar and lexicon collapsing across noun-related and verb-related items, whereas Panel **(A)**, shows the RR for the noun-related and verb-related items collapsing across grammar and lexicon.

of syntax; this technique can consider the dynamic nature of syntax and the variability in how productivity unfolds in running conversations. RQA provided a way to capture gradations of repetitions (e.g., quantity, diversity, and sequences) to shed light on a wide spectrum of language use in children with ASD. For instance, using this technique, we explored individual differences in the productivity of noun-related and verb-related speech. Our first hypothesis was not supported since we found that degree of talk and recurrence metrics were unrelated.

In contrast, our findings were consistent with our second hypothesis, that recurrence measures would be associated with MLU. Our major finding here was that RR was related to MLU, as children with higher MLUs repeated noun-related grammatical units *less* across the entire MLU range, while children who repeated verb-related grammatical units *more* produced longer utterances but only up to MLUs of 3–4. Notably, determinism provided an even more detailed look into the structures that make up productivity than are made possible by traditional composite linguistic measures. For example, while determinism was not related to total number of utterances, it was related to MLU, thus lending even more credibility to our second hypothesis. This finding highlights how challenging it is to establish productivity in children who consistently produce short utterances. Children with ASD who produced longer utterances manifested more productivity; thus, they were not just repeating the same utterances over and over. Furthermore, the determinism of grammatical units was what seemed to drive this relationship with MLU. That is, children who repeated grammatical combinations also produced more complex language, signifying the importance of creating varied grammatical constructions for early productivity.

We also investigated how RQA measures compared to well-established linguistic analyses in a sample of 5-year-olds with autism from the Autism Phenome Project dataset (Wittke et al., 2017). Our analyses revealed that the recurrence rate of grammatical and lexical units within noun-related and verb-related speech mapped onto traditional linguistic analyses; for example, grammatical units were repeated more than lexical units. Measures of determinism further illuminated gradations in the productivity of grammatical language use for children with ASD. As expected, grammar was more productive (i.e., higher %DET) than words in both noun-related and verb-related speech sequences. Noun-related grammar usage was more productive than verb-related grammar usage, but no significant noun-verb differences were found for words. Thus, RQA and traditional linguistic analyses—at least to some extent—identify similar signals.

In broad brush, our findings are consistent with the elicited bootstrapping hypothesis. Although we did not directly measure

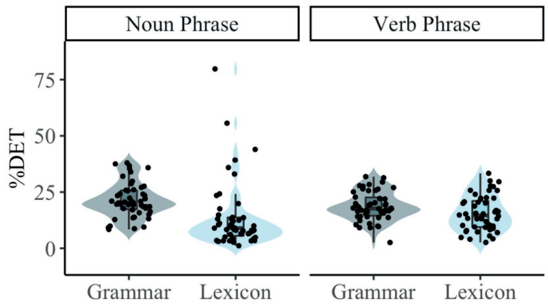

FIGURE 5
Variability in children's combinations of phrasal units by analysis level. One child is combining the same noun-related words over and over (e.g., "Happy Birthday," "Scooby Doo snack," "upside down"). However, this child only produced 35 utterances, most of which were very short (MLU of 2.67) and so, although they produced the highest word determinism, even adding one new grammatical unit to their noun-related lexical repetitions would make them less grammatically deterministic (e.g., the child said "pronoun-adverb-adverb and noun-adverb-adverb"). This same child also produced the highest percent of echolalic utterances in the sample. Similarly, five other children repeated the same words in noun phrases. Most of the other children generated a lot of possible word combinations for noun phrases in each utterance.

TABLE 8 Regression results predicting %DET.

Predictor	B	Lower 95% CI	Upper 95% CI	SE	*t*
(Intercept)	1.80	−16.36	19.97	9.21	0.20
Type	3.49	−22.20	29.18	13.02	0.27
Analysis level	6.57	−19.12	32.26	13.02	0.50
MLU	3.10**	0.97	5.24	1.08	2.87
Total utterances	−0.03	−0.08	0.02	0.03	−1.10
ADOS severity score	1.48†	−0.11	3.07	0.81	1.84
Type: Analysis level	9.73	−26.60	46.06	18.42	0.53
Type: MLU	0.12	−2.90	3.14	1.53	0.08
Type: Total utterances	0.01	−0.07	0.08	0.04	0.18
Type: ADOS severity score	−1.19	−3.44	1.05	1.14	−1.05
Analysis level: MLU	−2.89†	−5.91	0.13	1.53	−1.89
Analysis level: Total utterances	−0.01	−0.08	0.07	0.04	−0.15
Analysis level: ADOS severity score	−0.29	−2.53	1.96	1.14	−0.25
Type: Analysis level: MLU	−1.50	−5.77	2.77	2.16	−0.70
Type: Analysis level: Total utterances	0.02	−0.08	0.13	0.05	0.43
Type: Analysis level: ADOS severity score	−0.19	−3.37	2.99	1.61	−0.12

$R^2 = 0.150$***. †indicated $p < 0.07$. *indicates $p < 0.05$. ***indicates $p < 0.001$.

social motivation nor attempted to test the complete theoretical model, we consider autism severity scores a parallel to social motivation (see also Naigles and Chin, 2015; Thomas et al., 2022), and our results showed that children with less social motivation were more repetitive. That is, they repeated sequences of words more as well as individual words more, either immediately or further along in the conversation. In what follows, we explore possible explanations for why specific patterns of repetition emerged across the different types of speech and levels of analysis, and consider possible explanations for the reported associations between recurrence measures and traditional linguistic measures.

Recurrence metrics relate to productivity

Recurrence does not equate to being more (or less) talkative

In general, we found no association between the number of utterances and either recurrence metrics (i.e., RR and %DET). This is plausible given that producing fewer utterances does not mean that the children are not producing rich utterances when they do talk. For instance, two children in our sample produced only 29 utterances but varied in the complexity of those utterances. One child repeated noun-related grammatical units moderately (e.g., "Ah, I do not pop *bubbles*. *Bubbles* go. Ah, *bubbles* pop pop,"), while the other repeated noun-related grammatical units frequently but had more complex language (e.g., "*I* want to play balloon. *I* want *the mommy*'s phone. Clean up *the toys*").

Recurrence captures individual differences in productivity

We did find that children who repeated grammatical units more frequently (e.g., more Determiner-Noun or Verb-ing sequences) produced *longer* utterances overall. In a way, this is necessary, as the repetition of grammatical units means that there are indeed sequences of units, hence longer utterances. This was particularly evident when the children's speech was in the early phases of becoming more complex.

However, it seems that RR only matters for the onset of grammatical speech and then the relationship plateaus, with a lot of variation in repetitions for high MLU (see Figure 4B). Possibly there is a plateau because RR does not differentiate between the child who just says Determiner-Noun all the time versus the child who says Determiner-Noun and Verb-ing, which would be captured by %DET. Thus, this shift in patterning likely reflects the shift over to multiword speech. There are a few reasons this might occur. First, it could be that there are just fewer children with larger MLUs driving this effect. However, the distribution of scores in grammatical unit repetition (see Figure 4) indicates that this is not the case.

Second, perhaps at first children produce lots of pronouns, which keeps their MLU short. For instance, a child who produces less complex speech might be more likely to frequently say "get *it*." For these children, relying on these specific grammatical units may hamper or delay their production of longer utterances. Thus, children's longer utterances do not just involve saying the same items over and over. Rather, longer utterances involve—in ASD as in TD-—fairly morphologically or syntactically rich sentences (e.g., sentences with multiple clauses that contain adjectives, prepositions, adverbs, and verbs).

Finally, it could be that since there are only so many repetitions in noun-related speech that one can do in the span of English grammar. After a certain point, there is only a set number of ways that English can support noun-related grammatical recurrence. For instance, one could produce sentences with determiner-noun or determiner-adjective-noun to form a legal noun phrase; however, adjective-determiner-preposition-noun would not be an appropriate noun phrase construction in the English language. So, if the child is repeating lots of noun-related grammatical units then they are probably producing shorter utterances and if they are producing shorter utterances then they are probably repeating a lot of noun-related grammatical units. Repeating noun-related grammatical units (e.g., "A baby. A firetruck. A boy. With the pants.") does not enable the child to produce longer utterances, because it is the verbs that extend the utterance length (e.g., "They will not stand up. Yeah they were eating. And then they come in. And they took the food away."). Either way, the data suggest that these children may benefit from therapy to address verb-related speech.

Perhaps the closest analog to our own study is Lieven et al. (2009), who focused on the productions of four TD children. Consistent with our findings, they reported that noun-related (so-called REF) repetitions are more frequent than verb-related (so-called PROCESS) repetitions. In contrast to the current findings, they also report generally less repetitiveness (i.e., fewer repeated multi-word utterances) in children with higher MLUs (and across development for one child, with their MLU changing from 1.6 to 2.2). Slight differences between our results and Lieven et al.'s (2009) research may be due to their decision to confine analyses to multi-word utterances that have been repeated (which means they did not count repetitions of single words that might appear across utterances), their use of a traceback method (which means they had to more arbitrarily decide what was versus was not repeated), and their four-child sample size (which resulted in a much smaller MLU range of 1.6–2.2). Upon analyzing the data further by type (i.e., noun-related speech versus verb-related speech) they revealed that PROCESS-related/verb-related multi-word repetitions increase with MLU, which potentially matches our increase from an MLU of 2–4. It could be that this association between verb-related grammatical units and MLU reflects a shift from learning how to combine basic words to learning how to combine grammar in more complex ways. This would indicate that the value of RR may lie in its potential to capture emerging complexity in grammatical constructions, but beyond this shift, RR is less informative.

This interpretation of the data is partially consistent with the hypothesis of predictive impairment in ASD (Sinha et al., 2014); one component of this hypothesis suggests that challenges in prediction lead to overly repetitive behavior as compensation. Similarly, our analyses found that higher RRs for noun-related structures were associated with more advanced language; however, high RRs for verb-related structures did not yield the same expected association. Note, though, that all children repeated themselves at least somewhat; we conjecture that it was the children who found *a few structures* to consistently refer back to while testing new structures who were the ones with better language abilities. Therefore, all of the differences in methods considered, the data broadly suggest that our children with ASD are not markedly different from the TD children in their RR; variations in findings are likely based on the language level (MLU), not ASD presence (see also Weismer and Saffran 2022).

Interestingly, the only significant relationship that emerged with %DET was MLU, as children producing longer utterances combined the same grammatical units more frequently, showing more advanced productivity since they are practicing the same sentence structures. For example, a child with an MLU of 3.04 repeated the sequence "started to took off" frequently while saying "Then it started to take off. The. To took off. It started to took off already." Notice that this child is building on each of the repeated sequences in different ways across each sentence. In contrast, higher MLU children combined the same grammatical units more frequently (i.e., showing more advanced productivity). For instance, the child from Example 4 had an MLU of 5.29 and repeated determiner-noun-plural sequences several times within the brief excerpt. Our results suggest that determinism goes beyond frequency counts, providing more detail on the structure of productivity. Not only do children with ASD vary in their usage of words and grammatical units, but—depending on their language skills—children with ASD exhibit different dynamics in their speech patterns, too.

Recurrence measures mirror and extend traditional linguistic analyses

Across the children in this study, a large degree of variability was evident in the repetitions of words and grammatical units in noun-related and verb-related speech, in amount of talk, in types of words and grammatical units, and in combinations of these units. This variability is consistent with previous work documenting a vast heterogeneity in the language skills of children with ASD and this variability spans their lexicon, syntax, and morphology (Kjelgaard and Tager-Flusberg, 2001; Modyanova et al., 2017). This variability is perhaps not surprising given existing work claiming that motivations to communicate may actually alter the degree to which children on the autism spectrum exhibit repetitive speech (i.e., elicited production theory; Camarata and Yoder, 2002; Sameroff, 2009). Thus, in our language sample, we may possibly be capturing these differences in motivation to communicate across the different activities. However, these data cannot parse out whether repetitions occur because the child is problem-solving their social partner's intent, affirming their preference by imitating or producing a self-regulatory behavior (i.e., stimming).

Increased recurrence of nouns and grammatical structures

We found more repetitions in noun-related speech than verb-related speech, of noun-related words than verb-related words, and of grammatical units than lexical units. Such findings are consistent with the structure of the English language, of our choice of lexical and grammatical units, and of the ADOS protocol. A closer look at type and token distributions of units in noun-related and verb-related speech can help explain why these patterns might emerge.

The data revealed that differences in repetitions by speech type may emerge because children tended to produce many different noun-related words (average number of noun-related word tokens = 267, range = 27–636 words), but only a few of these units were repeated frequently. By contrast, children produced fewer verb-related words overall (average number of verb-related word tokens = 145, range = 12–338 words) but repeated a greater variety of them. These differences in variety and volubility in noun- and verb-related production are consistent with other research on TD children. For instance, researchers have found that of the earliest words that TD children produce, over half are nouns, while less than 25% are verbs (Stern, 1924; Nelson, 1973; Fenson et al., 1994). Further, TD children produce many more noun types (see Sandhofer et al., 2000) and more noun tokens (Tardif et al., 1997).

Another possible explanation for these findings is that the noun-related units are largely syntactic (10 possible syntactic items versus 4 possible morphological items), whereas the verb-related units are mostly morphological, not fully syntactic (5 possible syntactic items versus 11 possible morphological items). That is, fewer grammatical items comprised noun-related speech (i.e., 14 possible items) than verb-related speech (i.e., 16 possible items; see Tables 3, 4), leading to more repetitions in noun-related speech. Finally, because there are fewer grammatical items than lexical items (in both noun and verb phrases), it is unsurprising that RR is lower for lexicon than grammar (see Naigles et al., 2009, for documentation of productivity in verbs). These analyses, therefore, show that RR is capable of capturing the difference between noun-related and verb-related speech and grammar and lexicon and so analyses are consistent with traditional linguistic analyses.

Our finding that children more frequently combined grammatical units in the same ways compared to word units, for both noun-related and verb-related speech, likely emerged because there are simply many more words that children could choose to combine compared to grammatical units (i.e., "a cute dog" would be flagged as a different combination than "a fluffy dog"). We also found that children combined noun-related grammatical units more so than verb-related grammatical units, but this difference in speech type did not hold for words. This is likely a facet of our coding, in that we coded for more ways to appropriately combine grammatical units of noun-related speech than verb-related speech, given our choice to not code for verb argument structure.

Recurrence captures autism symptomatology

Our analyses exploring the relationships between RR, %DET, and autism diagnosis-related metrics revealed some interesting nuances to help explain extant research. Primarily, we found that children who were generally more repetitive tended to present with more autism traits; this matches the broader ASD literature, which suggests that repetitive behaviors are common in ASD (Tager-Flusberg and Calkins, 1990; American Psychiatric Association, 2013). Thus, calculating and comparing RRs of speech for both autistic and non-autistic individuals could further help refine the prediction impairment hypothesis (Weismer and Saffran, 2022).

We also build on the existing literature about language profiles in 4- to 8-year-old children with autism (see Van Santen et al., 2013; Thomas et al., 2022). For instance, Van Santen et al. (2013) reported no difference in intra-turn self-repeats of words by autism diagnosis. However, we found that children who were generally more repetitive (across lexical and grammatical units) tended to demonstrate more autistic traits, suggesting that Van Santen et al. (2013) may have not captured the relevant metric of repetition. Thus, we add that like repetitions at the lexical level, repetitions at a finer granularity of measurement (i.e., grammatical units and different parts of speech) may also provide informative data points to understand differences across the spectrum. We extend previous findings by including ASD participants with a wider range of IQ scores and participants ranging from talkative to minimally talkative (whereas Van Santen et al. exclusively focused on low verbal children), making the current findings more representative of the ASD population.

Limitations and future directions

While these results are intriguing, there are several limitations within the present study. First, the current data did not include any comparison groups for the ASD group, making it difficult to assess the degree to which variability in recurrence is unique to autism or characteristic of broader language heterogeneity in all children. To better describe the productivity of syntax in autism, it would be important to conduct studies that involved a TD group, a Developmental Language Disorder group, more age groups, a language-matched group, and an age-matched group.

Second, these data are drawn just from interactions during the ADOS, with the child engaging with a clinician. However, child speech and more importantly, the degree to which that speech is repetitive, can vary by interactional context. For instance, Gladfelter and VanZuiden (2020) found that school-aged children with ASD repeated themselves less frequently (i.e., self-repeating) during storytelling compared to during play-based contexts. These findings suggest that context can shape the degree to which children repeat: more unstructured contexts, as in the current study, may involve more lexical repetitions, which could be an indicator of less productive speech (see also Kover et al., 2014, for differences in the number of unique words by context). This raises the possibility that the language samples collected in the ADOS can underestimate linguistic complexity. Relatedly, Naigles et al. (2009) found more productivity in the verb use of TD children within parent diaries, presumably because this format required all verb use to be written down across the children's daily activities. Work across a variety of contexts, therefore, suggests a broader need to study language in autism within more naturalistic and a wider variety of settings. Perhaps this issue could be tackled *via* the LENA system, which can capture many settings of talk at home. At present, LENA recordings are not as well analyzed as traditional free-play interactions, as LENA outputs the presence of speech and auto-generates word counts but not types of words, syntactic complexity, or transcriptions of the speech itself. Furthermore, research has suggested that it is not yet useful for detecting speech vocalizations in ASD (Wang et al., 2017; Jones et al., 2019; Sulek et al., 2022). Although LENA's raw data are not yet amenable to lexical or grammatical RQA, LENA transcripts could reveal what the child is saying over the entire day. Coding these transcripts for RQA would be an important next step in this avenue of research.

Third, given our focus on child language production, we did not assess the role of the social partner in prompting repetitive or productive speech. However, across our sample, there was large variability in the degree to which parents were present and involved for ADOS administrations. The degree to which this social partner, and even the clinicians and the experimenters, contribute to the reported patterns is unclear. Further characterization of recurrence in speech should involve more conversations with parents (see Fusaroli et al., 2020, Unpublished manuscript), cross-recurrence with different conversational partners (e.g., parents, clinicians, and strangers), and a comparison to intra-child recurrence for TD groups (see Dale and Spivey, 2006; Müller-Frommeyer et al., 2019). This type of work could be applied to analyze coherence in content within speakers (e.g., auto-scoring essays; Angus et al., 2012). It could also be helpful for assessing the degree to which speakers are on the same page (i.e., semantic alignment; Dale and Spivey, 2006; Fusaroli et al., 2020, Unpublished manuscript). Relatedly, we also have not considered how self-repetition, studied here, relates to echolalia, or the repetition of the speech of others. In our sample, only a few children produced a substantial number of echolalic utterances, and their RRs varied hugely, so drawing conclusions about this relationship was unwarranted. However, with a bigger sample of children producing more echolalic utterances, the relationship with self-repetitions could be studied in more depth.

Fourth, we have not included analyses that might map RQA metrics onto the subgroups that Wittke et al. (2017) first identified. Since the proportion of ungrammatical utterances (which was a key grouping variable for Wittke et al., 2017) correlated with RQA metrics, we might expect that RQA could pull out additional things from the subgroups to characterize these children in even greater detail.

Finally, our particular interests in understanding repetitions in noun-related and verb-related speech led us to remove all other data from consideration in our analyses and consider only two kinds of RQA metrics. This coding made it impossible to tell whether other parts of speech generate unique recurrence patterns. However, clearly, there are many other parts of speech (e.g., prepositional phrases and adverbial phrases). Other researchers have started to look at recurrence in grammar (Dale and Spivey, 2006; Müller-Frommeyer et al., 2019) but have not yet assessed all parts of speech. One approach to examining categories of speech might be to use recurrence block representation analyses created by Xu and Yu (2016). This approach would generate recurrence plots that showed where in time certain categories were chunked. Future work should also examine other RQA metrics not analyzed here; for example, in taking a dynamical systems approach, it may be valuable to explore attractor strength (or the relative "pull" of different kinds of behaviors) through the RQA metric known as maximum line length (or maxline; e.g., Pellecchia et al., 2005). While outside of the scope of the current article, future exploratory or confirmatory analyses of RQA metrics may provide valuable insights into these and similar phenomena.

Conclusion

Autistic individuals comprise a diverse population with a diverse set of skills. This study is a first step in understanding the real-time syntactic structures that characterize the diverse range of language abilities in young children with ASD. While the current study did not attempt to model the entire elicited bootstrapping theory framework, we affirm that differences in early social motivation prompt a series of shifts in children with ASD's language production and reciprocal language input. Based on the recurring patterns of grammar and lexicon observed within a rich, naturalistic, spontaneous language sampling opportunity, we emphasize that these productions were still characterized by complex and adaptive content not restricted to repetitive speech or echolalia. Results suggest that we should perhaps refocus from aggregate measures to consider many of the nuanced patterns that emerge across the span of a conversation.

The primary contribution of the current study is a technique for quantifying patterns of repetition in language automatically. This type of technique could help guide assessments and interventions in capturing and tapping into underlying mechanisms of repetitive language use in autism. That is, findings from this work, if replicated, may assist clinicians design more powerfully targeted therapies for developing early language use. Our RQA analyses showed that both grammatical productivity and lexical productivity were related to language competence in different ways to this heterogeneous sample of children with ASD. Beyond more traditional measures like MLU, it appears that *less* repetition in noun-related grammar leads to longer utterances, whereas *more* repetition of verb-related grammar leads to longer utterances (up to MLUs of 3–4 morphemes). This could benefit clinicians to more strategically structure their language interventions, working on increasing the diversity of lexical items while emphasizing the importance of grammatical repetition, particularly for verb-related units. A parallel in this treatment philosophy is seen in harnessing statistical learning for children with specific language impairment (SLI), now more commonly known as developmental language disorder (DLD; see Plante et al., 2014 and Plante and Gomez, 2018 for more information). We also suggest that—although it is important to capture simple single unit repetitions (i.e., repetitive speech and RR)—measures of how children combine these units (i.e., %DET) can shed light on how children are building their sentences (i.e., testing out new structures versus relying on the same structures over and over again). Findings ultimately suggest that fine-grained measures such as RQA metrics may have the power to illuminate this continuum of productivity in children with ASD.

Data availability statement

Datasets on which analyses were performed are available at https://github.com/amandamankovich/ASD-Recurrence-Analysis. The raw data supporting the conclusions of this article will be made available by the authors, after consultation for privacy reasons with the APP team at the MIND Institute.

Author contributions

AMM designed the original data collection. AM, KW, AP, and LRN worked together on the design, coding, and analyses of the current study. AP developed the RQA code in R and RStudio. AM, JB, and LRN worked together on the primary write-up of this study. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by NSF-IGERT DGE-1144399 to the University of Connecticut in addition to NIH R01 MH089626-1 and NIMH to the MIND Institute, Autism Phenome Project. AM was awarded a UConn Science of Learning Training Fellowship.

Acknowledgments

We gratefully acknowledge the contributions of Dr. Sally Ozonoff and Dr. Sally J. Rogers, and the entire APP team, to the design and data collection of T3 of the original APP project.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Akhtar, N., and Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Dev. Psychol.* 33, 952–965. doi: 10.1037/0012-1649.33.6.952

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders. 5th Edn*. Arlington, VA: American Psychiatric Publishing.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders. 4th Edn*. Washington, DC: American Psychiatric Publishing. Text Revision.

Angus, D., Smith, A. E., and Wiles, J. (2012). Conceptual recurrence plots: revealing patterns in human discourse. *IEEE Trans. Vis. Comput. Graph.* 18, 988–997. doi: 10.1109/TVCG.2011.100

Asperger, H. (1944). Die "Autistischen psychopathen," im kindesalter. *Arch. Psychiatr. Nervenkr.* 117, 76–136. doi: 10.1007/BF01837709

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguist. Inq.* 10, 533–581.

Bartolucci, G., Pierce, S. J., and Streiner, D. L. (1980). Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. *J. Autism Dev. Disord.* 10, 39–50. doi: 10.1007/BF02408431

Bloom, L. (1993). *Language Development from Two to Three*. Cambridge, UK: Cambridge University Press.

Boucher, J. (2012). Research review: structural language in autistic spectrum disorder–characteristics and causes. *J. Child Psychol. Psychiatry* 53, 219–233. doi: 10.1111/j.1469-7610.2011.02508.x

Brown, R. (1973). *A First Language, The Early Stages*. Cambridge, MA: Harvard University Press. doi: 10.4159/harvard.9780674732469

Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test. 3rd Edn*. Novato, CA: Academic Therapy Publications. doi: 10.1007/978-0-387-79948-3_1544

Brynskov, C., Eigsti, I. M., Jørgensen, M., Lemcke, S., Bohn, O. S., and Krøjgaard, P. (2017). Syntax and morphology in Danish-speaking children with autism spectrum disorder. *J. Autism Dev. Disord.* 47, 373–383. doi: 10.1007/s10803-016-2962-7

Camarata, S., and Yoder, P. (2002). Language transactions during development and intervention: theoretical implications for developmental neuroscience. *Int. J. Dev. Neurosci.* 20, 459–465. doi: 10.1016/s0736-5748(02)00044-8

Carello, C., and Moreno, M. (2005). "Why nonlinear methods," in *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*. eds. M. A. Riley and G. C. Van Orden (Arlington, VA: National Science Foundation), 1–25.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., and Schultz, R. T. (2012). The social motivation theory of autism. *Trends Cogn. Sci.* 16, 231–239. doi: 10.1016/j.tics.2012.02.007

Chin, I., Goodwin, M. S., Vosoughi, S., Roy, D., and Naigles, L. R. (2018). Dense home-based recordings reveal typical and atypical development of tense/aspect in a child with delayed language development. *J. Child Lang.* 45, 1–34. doi: 10.1017/S0305000916000696

Chiovaro, M., Windsor, L. C., Windsor, A., and Paxton, A. (2021). Online social cohesion reflects real-world group action in Syria during the Arab spring. *PLoS One* 16:e0254087. doi: 10.1371/journal.pone.0254087

Coco, M. I., and Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: an R package. *Front. Psychol.* 5, 1–14. doi: 10.3389/fpsyg.2014.00510

Condouris, K., Meyer, E., and Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measures of spontaneous speech in children with autism. *Am. J. Speech Lang. Pathol.* 12, 349–358. doi: 10.1044/1058-0360(2003/080)

Dale, R., and Spivey, M. J. (2006). Unraveling the dyad: using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Lang. Learn.* 56, 391–430. doi: 10.1111/j.1467-9922.2006.00372.x

Dunn, L. M., and Dunn, D. M. (1997). *Peabody Picture Vocabulary Test. 3rd Edn*. Circle Pines, MN: American Guidance Service. doi: 10.1007/978-1-4419-1698-3_531

Eigsti, I. M., Bennetto, L., and Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *J. Autism Dev. Disord.* 37, 1007–1023. doi: 10.1007/s10803-006-0239-2

Eigsti, I. M., de Marchena, A. B., Schuh, J. M., and Kelley, E. (2011). Language acquisition in autism spectrum disorders: a developmental review. *Res. Autism Spectr. Disord.* 5, 681–691. doi: 10.1016/j.rasd.2010.09.001

Elliott, C. D. (2007). *Differential Ability Scale. 2nd Edn. (DAS-II)*. San Antonio, TX: Harcourt Assessment. doi: 10.1177/0829573507302967

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., et al. (1994). Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* 59:i-185. doi: 10.2307/1166093

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatr. Res.* 65, 591–598. doi: 10.1203/PDR.0b013e31819e7203

Gentner, D. (1982). "Why nouns are learned before verbs: linguistic relativity versus natural partitioning," in *Language Development*. ed. S. A. Kuczaj, Language, Thought, And Culture, vol. *2* (Hillsdale, NJ: Lawrence Erlbaum Associates), 301–334.

Gladfelter, A., and VanZuiden, C. (2020). The influence of language context on repetitive speech use in children with autism spectrum disorder. *Am. J. Speech Lang. Pathol.* 29, 327–334. doi: 10.1044/2019_AJSLP-19-00003

Gleitman, L. (1990). The structural sources of verb meanings. *Lang. Acquis.* 1, 3–55. doi: 10.1207/s15327817la0101_2

Goldfield, B. A., and Reznick, J. S. (1990). Early lexical acquisition: rate, content, and the vocabulary spurt. *J. Child Lang.* 17, 171–183. doi: 10.1017/s0305000900013167

Hoff, E. (2006). The uses of longitudinal data and person-centered analyses in the study of cognitive and language development. *Merrill-Palmer Q.* 52, 633–644. doi: 10.1353/mpq.2006.0028

Hoff, E. (ed.) (2012). *Research Methods in Child Language: A Practical Guide*. Oxford: Blackwell.

Howlin, P. (1984). The acquisition of grammatical morphemes in autistic children: a critique and replication of the findings of Bartolucci, Pierce, and Streiner, 1980. *J. Autism Dev. Disord.* 14, 127–136. doi: 10.1007/BF02409656

Jones, R. M., Plesa Skwerer, D., Pawar, R., Hamo, A., Carberry, C., Ajodan, E. L., et al. (2019). How effective is LENA in detecting speech vocalizations and language produced by children and adolescents with ASD in different contexts? *Autism Res.* 12, 628–635. doi: 10.1002/aur.2071

Kanner, L. (1943). Autistic disturbances of affective contact. *Nerv. Child* 2, 217–250.

Kjelgaard, M. M., and Tager-Flusberg, H. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Lang. Cogn. Process.* 16, 287–308. doi: 10.1080/01690960042000058

Koegel, L. K., Koegel, R. L., and Smith, A. (1997). Variables related to differences in standardized test outcomes for children with autism. *J. Autism Dev. Disord.* 27, 233–243. doi: 10.1023/a:1025894213424

Kover, S. T., Davidson, M. M., Sindberg, H. A., and Weismer, S. E. (2014). Use of the ADOS for assessing spontaneous expressive language in young children with ASD: a comparison of sampling contexts. *J. Speech Lang. Hear. Res.* 57, 2221–2233. doi: 10.1044/2014_JSLHR-L-13-0330

Kreiser, N. L., and White, S. W. (2014). ASD in females: are we overstating the gender difference in diagnosis? *Clin. Child. Fam. Psychol. Rev.* 17, 67–84. doi: 10.1007/s10567-013-0148-9

Le Normand, M. T., Blanc, R., Caldani, S., and Bonnet-Brilhault, F. (2018). Disrupted behaviour in grammatical morphology in French speakers with autism spectrum disorders. *Clin. Linguist. Phon.* 32, 706–720. doi: 10.1080/02699206.2018.1425917

Leonardi, G. (2012). The study of language and conversation with recurrence analysis methods. *Psychol. Lang. Commun.* 16, 165–183. doi: 10.2478/v10057-012-0012-x

Lieven, E., Salomo, D., and Tomasello, M. (2009). Two-year-old children's production of multiword utterances: a usage-based analysis. *Cogn. Linguist.* 20, 481–507. doi: 10.1515/COGL.2009.022

Lord, C., Risi, S., Lambrecht, L., Cook, E. H. Jr., Leventhal, B. L., et al. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.* 30, 205–223. doi: 10.1023/A:1005592401947

MacWhinney, B. (2008). "Enriching CHILDES for morphosyntactic analysis," in *Trends in Corpus Research: Finding Structure in Data*. ed. H. Behrens (Amsterdam: Benjamins) doi: 10.1075/tilar.6.09mac

Marwan, N. (2008). A historical review of recurrence plots. *Eur. Phys. J. Spec. Top.* 164, 3–12. doi: 10.1140/epjst/e2008-00829-1

Marwan, N., Romano, M. C., Thiel, M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Phys. Rep.* 438, 237–329. doi: 10.1016/j.physrep.2006.11.001

Modyanova, N., Perovic, A., and Wexler, K. (2017). Grammar is differentially impaired in subgroups of autism spectrum disorders: evidence from an investigation of tense marking and morphosyntax. *Front. Psychol.* 8, 1–23. doi: 10.3389/fpsyg.2017.00320

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Service Inc.

Müller-Frommeyer, L. C., Frommeyer, N. A. M., and Kauffeld, S. (2019). Introducing rLSM: an integrated metric assessing temporal reciprocity in language style matching. *Behav. Res. Methods* 51, 1343–1359. doi: 10.3758/s13428-018-1078-8

Müller-Frommeyer, L. C., Kauffeld, S., and Paxton, A. (2020). Beyond consistency: contextual dependency of language style in monolog and conversation. *Cogn. Sci.* 44:e12834. doi: 10.1111/cogs.12834

Mundy, P. C. (2019). Individual differences, social attention, and the history of the social motivation hypotheses of autism. *Behav. Brain Sci.* 42:E103. doi: 10.1017/S0140525X18002509

Naigles, L. R., Bavin, E. L., and Smith, M. A. (2005). Toddlers recognize verbs in novel situations and sentences. *Dev. Sci.* 8, 424–431. doi: 10.1111/j.1467-7687.2005.00431.x

Naigles, L. R., and Chin, I. (2015). "Language in children with autism spectrum disorders," in *The Cambridge Handbook of Child Language*. eds. E. L. Bavin and L. R. Naigles (Cambridge: Cambridge University Press), 637–658. doi: 10.1017/CBO9781316095829.029

Naigles, L. R., Hoff, E., Vear, D., Tomasello, M., Brandt, S., Waxman, S. R., et al. (2009). Flexibility in early verb use: evidence from a multiple-N dairy study: III. A general description of early verb growth and use. *Monogr. Soc. Res. Child Dev.* 74, 32–39. doi: 10.1111/j.1540-5834.2009.00516.x

Naigles, L. R., Kelty, E., Jaffery, R., and Fein, D. (2011). Abstractness and continuity in the syntactic development of young children with autism. *Autism Res.* 4, 422–437. doi: 10.1002/aur.223

Nelson, K. (1973). Structure and strategy in learning to talk. *Monogr. Soc. Res. Child Dev.* 38, 1–2. doi: 10.2307/1165788

Orsucci, F., Giuliani, A., Webber, C. Jr., Zbilut, J., Fonagy, P., and Mazza, M. (2006). Combinatorics and synchronization in natural semiotics. *Phys. A Stat. Mech. Appl.* 361, 665–676. doi: 10.1016/j.physa.2005.06.044

Park, C. J., Yelland, G. W., Taffe, J. R., and Gray, K. M. (2012). Morphological and syntactic skills in language samples of preschool aged children with autism: atypical development? *Int. J. Speech Lang. Pathol.* 14, 95–108. doi: 10.3109/17549507.2011.645555

Paxton, A., and Dale, R. (2017). Interpersonal movement synchrony responds to high-and low-level conversational constraints. *Front. Psychol.* 8, 1–16. doi: 10.3389/fpsyg.2017.01135

Pellecchia, G. L., Shockley, K., and Turvey, M. T. (2005). Concurrent cognitive task modulates coordination dynamics. *Cogn. Sci.* 29, 531–557. doi: 10.1207/s15516709cog0000_12

Pinker, S. (1989). "Resolving a Learnability Paradox in the Acquisition of the Verb Lexicon," in *The teachability of language*. eds. M. Rice and R. L. Scbiefelbusch Baltimore: Paul H. Brookes Publishing.

Plante, E., and Gómez, R. L. (2018). Learning without trying: the clinical relevance of statistical learning. *Lang. Speech Hear. Serv. Sch.* 49, 710–722. doi: 10.1044/2018_LSHSS-STLT1-17-0131

Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., et al. (2014). Variability in the language input to children enhances learning in a treatment context. *Am. J. Speech Lang. Pathol.* 23, 530–545. doi: 10.1044/2014_AJSLP-13-0038

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Riley, E. B. M. A., and Van Orden, G. U. Y. C. (2005). *Tutorials in Contemporary Nonlinear Methods*. Arlington, VA: National Science Foundation.

Rispoli, M., Hadley, P. A., and Holt, J. K. (2009). The growth of tense productivity. *J. Speech Lang. Hear. Res.* 52, 930–944. doi: 10.1044/1092-4388(2009/08-0079)

Roberts, J. A., Rice, M. L., and Tager Flusberg, H. (2004). Tense marking in children with autism. *Appl. Psycholinguist.* 25, 429–448. doi: 10.1017/S0142716404001201

Sameroff, A. (2009). "The transactional model" in *The Transactional Model of Development: How Children and Contexts Shape Each Other*. ed. A. Sameroff (Washington, DC: American Psychological Association), 3–21. doi: 10.1037/11877-001

Sandhofer, C. M., Smith, L. B., and Luo, J. (2000). Counting nouns and verbs in the input: differential frequencies, different kinds of learning? *J. Child Lang.* 27, 561–585. doi: 10.1017/S0305000900004256

Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. E., and Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language-disordered groups. *Appl. Psycholinguist.* 12, 23–46. doi: 10.1017/S014271640000936X

Shriberg, L. D., Kwiatkowski, J., and Hoffman, K. (1984). A procedure for phonetic transcriptions by consensus. *J. Speech Hear. Res.* 27, 456–465. doi: 10.1044/jshr.2703.456

Shulman, C., and Guberman, A. (2007). Acquisition of verb meaning through syntactic cues: a comparison of children with autism, children with specific language impairment (SLI) and children with typical language development (TLD). *J. Child Lang.* 34, 411–423. doi: 10.1017/s0305000906007963

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., et al. (2014). Autism as a disorder of prediction. *Proc. Natl. Acad. Sci.* 111, 15220–15225. doi: 10.1073/pnas.1416797111

Stern, W. (1924). *Psychology of Early Childhood up to the Sixth Year of Age*. New York: Holt.

Su, P. L., Rogers, S. J., Estes, A., and Yoder, P. (2021). The role of early social motivation in explaining variability in functional language in toddlers with autism spectrum disorder. *Autism* 25, 244–257. doi: 10.1177/1362361320953260

Sulek, R., Smith, J., Bent, C. A., Hudry, K., Trembath, D., Vivanti, G., et al. (2022). The utility of LENA as an indicator of developmental outcomes for young children with autism. *Int. J. Lang. Commun. Disord.* 57, 103–111. doi: 10.1111/1460-6984.12678

Tager-Flusberg, H. (2000). "The challenge of studying language development in children with autism," in *Methods for Studying Language Production*. eds. L. Menn and N. Bernstein Ratner (Mahwah, NJ: Erlbaum), 313–331.

Tager-Flusberg, H., and Calkins, S. (1990). Does imitation facilitate the acquisition of grammar? Evidence from a study of autistic, Down's syndrome and normal children. *J. Child Lang.* 17, 591–606. doi: 10.1017/s0305000900010898

Tager-Flusberg, T., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., et al. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *J. Speech Lang. Hear. Res.* 52, 643–652. doi: 10.1044/1092-4388(2009/08-0136)

Tardif, T., Shatz, M., and Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian, and Mandarin. *J. Child Lang.* 24, 535–565. doi: 10.1017/S030500099700319X

Tek, S., Mesite, L., Fein, D., and Naigles, L. (2014). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorders. *J. Autism Dev. Disord.* 44, 75–89. doi: 10.1007/s10803-013-1853-4

Thomas, R. P., Wittke, K., Blume, J., Mastergeorge, A. M., and Naigles, L. (2022). Predicting language in children with ASD using spontaneous language samples and standardized measures. Department of Psychological Sciences, University of Connecticut. Manuscript submitted for publication. doi: 10.1007/s10803-022-05691-z

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition* 74, 209–253. doi: 10.1016/S0010-0277(99)00069-4

Tomasello, M., Akhtar, N., Dodson, K., and Rekau, L. (1997). Differential productivity in young children's use of nouns and verbs. *J. Child Lang.* 24, 373–387. doi: 10.1017/S0305000997003085

Van Santen, J. P., Sproat, R. W., and Hill, A. P. (2013). Quantifying repetitive speech in autism spectrum disorders and language impairment. *Autism Res.* 6, 372–383. doi: 10.1002/aur.1301

Vivanti, G. (2020). Ask the editor: what is the most appropriate way to talk about individuals with a diagnosis of autism? *J. Autism Dev. Disord.* 50, 691–693. doi: 10.1007/s10803-019-04280-x

Wang, Y., Hartman, M., Abdul Aziz, N. A., Arora, S., Shi, L., and Tunison, E. (2017). A systematic review of the use of LENA technology. *Am. Ann. Deaf* 162, 295–311. doi: 10.1353/aad.2017.0028

Webber, C. L., and Marwan, N. (2015). Recurrence quantification analysis. *Theory Best Pract.* doi: 10.1007/978-3-319-07155-8

Weismer, S. E., and Saffran, J. R. (2022). Differences in prediction may underlie language disorder in autism. *Front. Psychol.* 13:897187. doi: 10.3389/fpsyg.2022.897187

Wittke, K., Mastergeorge, A. M., Ozonoff, S., Rogers, S. J., and Naigles, L. R. (2017). Grammatical language impairment in autism spectrum disorder: exploring language phenotypes beyond standardized testing. *Front. Psychol.* 8, 1–12. doi: 10.3389/fpsyg.2017.0053

Xu, T., and Yu, C. (2016). "Quantifying Joint Activities using Cross-Recurrence Block Representation." in *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. eds. A. Papafragou, D. Grodner, D. Mirman and J. C. Trueswell, 1997-2002.

Yoder, P., Watson, L. R., and Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *J. Autism Dev. Disord.* 45, 1254–1270. doi: 10.1007/s10803-014-2286-4

Yorio, C. A. (1980). Conventionalized language forms and the development of communicative competence. *TESOL Q.* 14, 433–442. doi: 10.2307/3586232

Zhou, P., Crain, S., Gao, L., Tang, Y., and Jia, M. (2015). The use of grammatical morphemes by mandarin-speaking children with high functioning autism. *J. Autism Dev. Disord.* 45, 1428–1436. doi: 10.1007/s10803-014-2304-6

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 22 December 2022
DOI 10.3389/fpsyg.2022.929379

Check for updates

OPEN ACCESS

EDITED BY
Marisa Filipe,
Faculdade de Letras da Universidade de Lisboa, Portugal

REVIEWED BY
Eliseo Diez-Itza,
University of Oviedo, Spain
Ricardo Scott,
University of Alicante,
Spain

*CORRESPONDENCE
Anne Hoffmann,
✉ anne_hoffmann@rush.edu

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 26 April 2022
ACCEPTED 05 December 2022
PUBLISHED 22 December 2022

CITATION
Hoffmann A (2022) Communication in fragile X syndrome: Patterns and implications for assessment and intervention.
Front. Psychol. 13:929379.
doi: 10.3389/fpsyg.2022.929379

Communication in fragile X syndrome: Patterns and implications for assessment and intervention

Anne Hoffmann[1,2]*

[1]Department of Communication Disorders and Sciences, Rush University Medical Center, Chicago, IL, United States, [2]Department of Pediatrics, Rush University Medical Center, Chicago, IL, United States

Fragile X syndrome (FXS) is the most common cause of inherited intellectual disability and is associated with a high rate of autism diagnosis. Language delays have been noted in the areas of overall communication and the specific areas of receptive, expressive, and pragmatic language, as well as in development of speech sounds and literacy. It has been widely noted that those individuals with a diagnosis of both FXS and autism tend to have more significant intellectual disability and language disorder. In this study, the research exploring the FXS language phenotype is presented, and the roles of cognition, autistic symptomatology, and gender are highlighted as possible. Implications for assessment and intervention approaches based on the strengths and weaknesses of the FXS language phenotype are provided.

KEYWORDS

fragile X syndrome, language, communication, intellectual disability, assessment, intervention

1. Introduction

1.1. Fragile X syndrome

Fragile X syndrome (FXS) is the most common inherited form of intellectual disability with an estimated prevalence of 1/4,000–1/6,000 in males and 1/8,000 in females in the Western world (Coffee et al., 2009). While research in other areas of the world has been historically limited, there is some evidence that prevalence is lower in some Eastern countries, such as China, and higher in some Middle-Eastern countries, such as Egypt (Meguid et al., 2007; Niu et al., 2017). This single-gene disorder stems from the expansion of a trinucleotide sequence (CGG) on the X-chromosome (Willemsen et al., 2011). Once the expansion reaches >200 repeats, it is termed a full mutation and typically the gene becomes methylated, which results in the gene being turned off and production of fragile X messenger ribonucleic protein (FMRP) is reduced or ceased (Kaufmann and Reiss, 1999). FMRP is critical for overall development, and its reduction or absence is the underlying factor in the phenotypic expression of FXS (Casingal et al., 2020). The variance in prevalence between males and females is secondary to the x-linked nature of FXS, as females carry a "protective-X" which may mitigate the effects of the methylated gene (Loesch et al., 2004). The effects of this altered level of FMRP are pervasive, with clinically significant developmental delay, learning disabilities, social and behavioral challenges, anxiety, and executive function deficits being commonly reported (Gallagher and Hallahan, 2012).

Two additional areas frequently associated with FXS are intellectual disability (ID) and autistic characteristics, with increased language delay noted with increased ID and severity of autistic features (Oakes et al., 2013). Studies indicate that the majority of males with FXS will have a moderate to severe ID (Hessl et al., 2009) and 25% of females will have some form of ID (Hagerman et al., 1992). The rate of autism diagnosis is much higher in FXS than in typical development (TD), with approximately 50%–67% of males and 20% of females meeting criteria for autism spectrum disorder (ASD; Wang et al., 2010). This range likely stems from multiple sources, including variance in how ASD is diagnosed (e.g., parent report vs. direct measure of current behavior; standardized assessment vs. clinical judgment). Further, the question of whether the ASD present in FXS is the same ASD found in non-syndromic cases has been the topic of substantial debate (see Abbeduto et al., 2014 for review). The debate has primarily hinged on the observation that those individuals with FXS who also meet criteria for ASD (hereafter referred to as FXS + ASD) have lower intelligence quotient (IQ) on average than those who do not meet criteria (referred to as FXS-O; Bailey et al., 2001; Kaufmann et al., 2004; Lewis et al., 2006), which asks the question of whether FXS + ASD simply represents the more affected end of the spectrum of FXS phenotypic presentation. While that question is beyond the scope of this paper, in an effort to clarify research findings, we will highlight those studies that have compared FXS-O and FXS + ASD when such distinctions are possible.

For this review, we consider several areas in communication and language development. Communication refers to the broad concept of how an individual relays and receives messages with others, including the prelinguistic communication associated with very early development. This is frequently included in measures of adaptive behavior and the mode of communication can vary (e.g., gestures, use of a speech-generating device, spoken messages). As multiple studies have used communication in its broadest sense to assess if individuals possessed this capacity, we have included it as a separate category, in addition to language. Language is a form of communication that utilizes a specific set of symbols mutually understood by the creator and receiver of the messages (Gumperz, 1967) and for this review, we use this to refer to spoken language. Within language, we discuss receptive language (what is understood), expressive language (how an individual communicates), and pragmatic language (how communication is used in social contexts). Within receptive and expressive language, we examine overall patterns as well as the separate areas of morphology and syntax (i.e., morphosyntax/grammar) and vocabulary as permitted by the research that has been done in these areas. We also review current findings for speech sound and literacy development. Comparisons with other groups, most commonly Down syndrome (DS) and idiopathic ASD, will be highlighted to demonstrate phenotypic-specific tendencies in communication. The roles of gender, cognition, and autistic symptomatology in the communication profile are considered as possible. For interpretation of findings, infants refer to children 1 year of age and younger, very young children refer to those individuals ages 1 to 3 years, children (i.e., boys and girls) to those individuals aged 4–11 years, adolescents to those individuals aged 12–17, and adults (i.e., men and women) to those individuals 18 years and older. For overarching trends across the lifespan, the terms males and females are used. We also use the terms boys/men/males and girls/women/females to refer to biological sex as determined at birth. Finally, implications of the FXS language phenotype for clinical assessment and intervention are considered.

2. Materials and methods

For the current study, a comprehensive literature search was developed and run by an experienced medical librarian in October 2022 in the following databases: PubMed/MEDLINE, Scopus, CINAHL, PsycINFO, ComDisDom, the Cochrane Database of Controlled Trials, and the Cochrane Database of Systematic Reviews. Google Scholar was searched as well. Both controlled vocabularies (e.g., MeSH terms) and keywords in the title or abstract fields were searched. There were no restrictions on geography or age of participants. Animal studies were excluded. Additionally, a hand search was conducted of the reference lists of selected articles. A reproducible search strategy is attached, see Appendix 1. This initial search resulted in 2319 studies being imported for screening, of which 1,132 duplicates were removed, leaving 1,187 studies to be screened using title and abstract. These were screened and 990 were excluded secondary to one of the following criteria (1) no language outcomes; (2) participants did not include individuals with FXS; (3) was not a peer-reviewed study (e.g., book chapter, dissertation); and (4) article was not available in English. This resulted in 197 studies being assessed *via* full-text review. Twenty-seven of these studies were excluded for the following reasons: outcomes (17, outcomes did not include separate communication measures); study design (5, only case studies were provided); patient population (2, full mutation FXS was not included or details regarding the FXS performance were not provided); and Article was unavailable (3). Of the remaining 170 articles, 5 were review articles and 55 were published prior to 2009, which was the date of the latest comprehensive review. As such, the focus of this paper will be on research found within the remaining 110 studies, with comparisons drawn to the findings of previous reviews. The PRISMA diagram can be seen in Figure 1.

3. Results and discussion

3.1. Communication

As noted in previous reviews (Abbeduto and Hagerman, 1997; Abbeduto et al., 2007; Finestack et al., 2009), individuals with FXS

FIGURE 1
PRISMA diagram.

evidence communication delays early in development. Measurable delays have been noted as young as 6 months for males with FXS (Wheeler et al., 2021). This results in many individuals with FXS remaining as prelinguistic communicators far later than what is seen in typical development (Brady et al., 2006). These delays extend beyond the milestones of spoken language; the areas of communicative gestures, eye gaze, vocalizations, and communicative functions have all been shown as delayed relative to typical development (Flenthrope and Brady, 2010; Hinton et al., 2013; Marschik et al., 2014; Kover et al., 2015; Hahn et al., 2017; Rague et al., 2018; Hughes et al., 2019; Mattie and Hamrick, 2022).

Despite these delays, there is clear evidence that individuals with FXS do progress in their communicative ability (Bailey et al., 2009; Wheeler et al., 2021). Bailey et al. (2009) performed a study in which a survey was distributed to a large sample of individuals with FXS and their caregivers to assess adaptive behavior, including communication. Participants ranged in age from 1 to 62 years. Results indicated that the majority of adult males and females with FXS had reached functional communication levels by adulthood (i.e., single words or signs) and most females had reached advanced communication levels (i.e., complex sentences and conversations). Of note, there were increased percentages of individuals in each age group demonstrating the various communication skills (e.g., single words, signs, complex speech), indicating that skills continued to develop, albeit at a slower pace than TD. Because growth in FXS is slower than in TD, standard scores will sometimes show a decline (Klaiman et al., 2014). However, it is important to note that the decline in standard scores does not necessarily mean a loss of skill. Rather, as has been demonstrated in performance on cognitive assessments, the rate of skill acquisition in FXS often does not show the rapid acceleration of growth found in typical development, which increases the gap between FXS and neurotypical individuals (Hall et al., 2008).

3.1.1. Related factors

The role of gender in communicative development has demonstrated the expected strengths in females relative to males with FXS although females with FXS may still show delays relative to TD (Caravella and Roberts, 2017; Wheeler et al., 2021). Females with FXS often manifest delays by the age of 12 months, with a rate of growth that while faster than males with FXS is still slower than TD (Caravella and Roberts, 2017; Wheeler et al., 2021). However, studies have noted significant variability in communicative performance in females, so higher or lower performance is possible (Bailey et al., 2009; Flenthrope and Brady, 2010; Klaiman et al., 2014; Wheeler et al., 2021).

Some research has focused on cognitive processes that might underlie the communication delay found in FXS. There is evidence that very young children with FXS had atypical face-scanning patterns, suggesting differences in visual attention (D'Souza et al., 2015, 2020). In general, nonverbal cognition has strong relationships with communication ability across the lifespan, with lower nonverbal ability being correlated to decreased communication (Reisinger et al., 2019). However, in general, it appears that early communication is roughly commensurate with other developmental areas (Reisinger et al., 2019). Other studies have suggested that variance in parental input could impact communication development, as increased maternal responsivity is associated with steeper trajectories of growth (Warren et al., 2017).

Recent research has found that increased autistic symptomatology is generally associated with greater communication delay in FXS, and this is in agreement with previous research (Abbeduto and Hagerman, 1997; Abbeduto et al., 2007; Finestack et al., 2009; Fielding-Gebhardt and Warren, 2019; Mattie and Hamrick, 2022). The impact of autistic symptomatology has been demonstrated in gestures, gaze shift, and initiation of joint attention (Flenthrope and Brady, 2010; Hahn et al., 2016, 2017; Brewe et al., 2018; Rague et al., 2018; Hughes et al., 2019). However, the majority of these studies also found strong correlations between these same areas and nonverbal cognition, which speaks toward the difficulty in separating these two characteristics in the FXS phenotype, a challenge that has been discussed at length (Abbeduto et al., 2014).

3.2. Receptive language

As noted in previous reviews, receptive language is delayed with impairments evidence in comprehension of vocabulary and morphosyntactic structures (Abbeduto and Hagerman, 1997; Abbeduto et al., 2007; Finestack et al., 2009). The review by Finestack and Abbeduto (2010) presented mixed study results when comparing receptive language in FXS to children with TD matched on nonverbal cognitive development. Some studies have found that receptive language in FXS was on par with controls matched on nonverbal cognition (Abbeduto et al., 2003; Roberts et al., 2007) while others show the FXS group falling below (Roberts et al., 2001; Price et al., 2007). Recent studies have examined specific receptive language domains, although as highlighted below, there is still debate.

3.2.1. Receptive vocabulary

When examining specific areas of receptive language in individuals with FXS, vocabulary has appeared as a relative strength, with skills in this area outpacing syntax and sometimes nonverbal cognition in adolescents and adults (Thurman et al., 2017b; Hoffmann et al., 2019). Receptive vocabulary increases with age (Brady et al., 2020), and its position as a relative strength has been shown across development (Thurman et al., 2017b). When comparing receptive vocabulary in FXS to what is found in other neurodevelopmental diagnoses, there have been mixed findings. Some studies have found that children and adolescents with FXS have stronger receptive vocabulary skills than individuals with DS or ASD matched on nonverbal cognition (Thurman et al., 2017b; Del Hoyo Soriano et al., 2018; Thurman and Hoyos Alvarez, 2020). Others find no difference between the groups (Finestack et al., 2013).

3.2.2. Receptive morphosyntax

Comprehension of grammar has been shown as commensurate with nonverbal cognition in some studies (Thurman and Hoyos Alvarez, 2020) and below nonverbal cognition in others (Oakes et al., 2013). It is possible that there are certain contexts which impact receptive morphosyntax. Oakes et al. (2013) propose that comprehension of sentences with a high demand of auditory sequencing or ones that lack lexical supports might be more problematic for individuals with FXS. When compared to peers with TD matched on nonverbal cognition, male children and adolescents with FXS still tend to fall below on receptive morphosyntax measures (Finestack et al., 2013; Oakes et al., 2013; Martin et al., 2013b), but there are studies that show similar performance (Finestack and Abbeduto, 2010). Comparisons with other groups have found equivalent skills between FXS and both ASD and DS when individuals are matched on nonverbal IQ (NVIQ; Finestack et al., 2013; Thurman et al., 2017b).

While the mixed results make a summary difficult, there are clearly delays relative to chronological age. The variance in study results is likely to stem from methodological differences. For example, the age of participants, the inclusion/exclusion of females, how language and cognition were assessed, whether age equivalent scores were used, these could all impact how groups compare against each other.

3.2.3. Related factors

Studies specifically examining receptive language in females with FXS have found the expected trend of more preserved abilities as compared to males, although many of the participants still fall below chronological age expectations (Roberts et al., 2007; Sterling and Abbeduto, 2012; Joga-Elvira et al., 2021). Sterling and Abbeduto (2012) found that similar to males with FXS, females also had receptive vocabulary skills that were generally higher than their nonverbal cognitive ability, although there was considerable variation across participants. Receptive syntax is generally weaker than receptive vocabulary, just as was described in males with FXS (Oakes et al., 2013). While studies specifically examining receptive language in females with FXS are limited, several have included females within the participant group. Many of these studies also found that while the females had stronger language skills overall, they had similar relationships between receptive language and other traits (e.g., autistic symptomatology, nonverbal cognition) as males with FXS (Finestack et al., 2013; Oakes et al., 2013; Hoffmann et al., 2019). However, Brady et al. (2020) and Pierpont et al. (2011) both found a steeper trajectory for some receptive language skills in female children and adolescents with FXS as compared to males.

Across studies, nonverbal cognition has been demonstrated as an important factor for receptive language. Brady et al. (2020) found that NVIQ, as well as parenting style, was related to growth in receptive and expressive vocabulary over time. Pierpont et al. (2011) examined specific cognitive areas, with phonological memory and working memory being strongly correlated to receptive vocabulary and syntax in boys with FXS, while in girls, overall cognition was strongly correlated but not those specific subdomains.

The role of autism in receptive language is closely related to cognition. Thurman and Hoyos Alvarez (2020) and Thurman et al. (2017b) found that autistic symptomatology and nonverbal cognition predicted receptive vocabulary in boys with FXS regardless of ASD status. Interestingly, the type of autistic symptomatology was important, with severity in restricted and repetitive behaviors having strong correlations to delays in receptive vocabulary and other language areas for children (Thurman and Hoyos Alvarez, 2020). In adolescents and adult with FXS, recent studies have not shown a difference in receptive language based on autism status once analyses are adjusted for nonverbal cognition (McDuffie et al., 2012; Hoffmann et al., 2019). However, when autistic symptomatology was assessed as a continuous metric, it was a significant predictor of receptive vocabulary and grammar (McDuffie et al., 2012). This suggests that the relationship between autistic behaviors in FXS may benefit from a more nuanced assessment than a simple categorical approach.

3.3. Expressive language

As in receptive language, there is general consensus that expressive language in FXS is significantly delayed relative to chronological age expectations (Abbeduto and Hagerman, 1997; Abbeduto et al., 2007; Finestack et al., 2009). These delays have been found in previous studies in both expressive vocabulary and expressive morphosyntax, when assessed through traditional standardized assessment as well as language sampling. Previous reviews have described the expressive language ability of males with FXS as falling below that of children with TD matched on cognition (Finestack et al., 2009), but more mixed findings are reported in vocabulary and morphosyntax (Abbeduto and Hagerman, 1997; Abbeduto et al., 2007).

3.3.1. Expressive vocabulary

Recent studies have shown expressive vocabulary in boys with FXS as impaired relative to TD children matched on nonverbal mental age (Kover et al., 2012; Martin et al., 2013b). Longitudinal studies using standardized measures have found increases in vocabulary in childhood and adolescence, although there may be a decrease in rate of growth during late childhood (Martin et al., 2013b; Brady et al., 2020). When lexical diversity—a measure of expressive vocabulary—has been calculated from language samples, there seems to be a decrease in number of different words used by adolescent males in conversation, despite an increase in the talkativeness (Del Hoyo Soriano et al., 2020).

Comparisons have been made between individuals with FXS and those with ASD and DS, matched on either nonverbal cognition, mean length utterance (MLU), and or autistic symptomatology. Individuals with FXS have generally had stronger performance on expressive vocabulary measures than individuals with DS matched on nonverbal cognition (Finestack et al., 2013; Martin et al., 2013b). FXS as compared to ASD has not evidenced differences in lexical diversity when NVIQ was used for matching (Kover et al., 2012), but when MLU and autistic symptomatology were used the FXS group had fewer different words than the ASD group (Hilvert et al., 2020). However, the participants with FXS in Hilvert et al.'s study had much lower scores on a standardized assessment of vocabulary and NVIQ than the group with ASD which could impact their performance.

3.3.2. Expressive morphosyntax

Expressive morphosyntax is also below what is seen in TD when nonverbal cognition is controlled (Estigarribia et al., 2012), and there is also evidence that boys with FXS have more impairment in expressive grammar as compared to expressive vocabulary (Martin et al., 2013b). When specific grammatical forms are examined, individuals with FXS seem to follow an atypical developmental pattern. While children with FXS fall below children with TD matched on nonverbal mental age in general measures of expressive grammar and MLU, they acquire some later developing forms (e.g., third-person singular markers) earlier than would be predicted by MLU (Estigarribia et al., 2011; Sterling et al., 2012; Komesidou et al., 2017).

Compared to groups with developmental language disorder (DLD) and TD matched on MLU, boys with FXS performed better on certain morphological structures such as finiteness marking than the group with DLD and even out-performed the group with TD on third-person singular forms (Haebig et al., 2016). This could indicate that in FXS, MLU does not have the same relationship to specific morphological forms that is seen in TD (Rice and Wexler, 2001; Rice et al., 2010; Haebig et al., 2016). When compared to individuals with DS, frequently noted as having relative weakness in expressive language skills, individuals with FXS have mostly been found as having stronger expressive morphosyntax (Martin et al., 2009; Finestack and Abbeduto, 2010; Finestack et al., 2013; Martin et al., 2013a).

Longitudinally, Martin et al. found that the boys with FXS did make gains over time on standardized assessments of expressive morphosyntax, but the rate of growth was slower than what is seen in TD, similar to what was seen in the group with DS. This slower growth has been replicated in other studies, and the possibility of a plateau in skill development has been noted (Komesidou et al., 2017). When adolescents with FXS were assessed over time using language samples, there was a decrease in syntactic complexity despite an increase in the overall amount of utterances (Del Hoyo Soriano et al., 2020). These could reflect a discrepancy between growth in standardized assessment as compared to functional use of structural language, as was seen in expressive vocabulary.

3.3.3. Related factors

Early expressive language delays occur in both males and females with FXS, although as in other areas, females tend to be more mildly affected (Brady et al., 2006). Research specifically comparing males and females with FXS in expressive language has found the expected trends of stronger performance and growth in females, with considerable individual variability (Finestack and Abbeduto, 2010; Komesidou et al., 2017). Some research indicates that female children and adolescents with FXS have MLU within the age expectations, and that NVIQ is not predictive of this ability (Sterling and Abbeduto, 2012). Others have found that NVIQ is predictive of either MLU (Komesidou et al., 2017) or complex syntax (Kover and Abbeduto, 2019). Given the tendency of males with FXS to have complex syntax above what their MLU would predict, this is an area that merits further research.

Several studies have found that expressive language ability and growth is predicted by nonverbal cognitive skills (Price et al., 2008; Pierpont et al., 2011; Estigarribia et al., 2012; Martin et al., 2013b; Komesidou et al., 2017). As in receptive language, phonological and working memory appear correlated with expressive vocabulary and syntax (Pierpont et al., 2011; Estigarribia et al., 2012; Kover and Abbeduto, 2019).

There is evidence that autistic symptomatology is linked to increased expressive language deficits across development. A study examining parent-reported early milestones found an average delay in first words of 3 months for very young boys with FXS-O and 13 months for FXS + ASD (Hinton et al., 2013). However, a study that examined early gesture usage did not find a correlation between autistic symptomatology and gestural delay once nonverbal ability was added as covariate (Rague et al., 2018). Haebig and Sterling (2017) compared receptive-expressive vocabulary profiles in adolescents with FXS + ASD and ASD. They found that despite having similar profiles of autistic symptomatology, the groups differed significantly in their vocabulary profiles, with the participants with ASD having a high rate of gaps in receptive-expressive vocabulary skills that favored expressive vocabulary and participants with FXS + ASD having a much lower rate (Haebig and Sterling, 2017). In addition, there is some evidence that boys with FXS + ASD show atypical acquisition of grammatical morphemes in a manner more similar to what is seen in ASD, although this has not included a comparison to boys with FXS-O (Sterling, 2018). Studies examining syntax in boys with FXS-O and FXS + ASD have not consistently found differences between the two once NVIQ is considered (Roberts et al., 2007; Kover and Abbeduto, 2010; Estigarribia et al., 2012.

3.4. Pragmatic language

Pragmatic language refers to the use of communication in social contexts, including communicative exchanges, production of contingent and appropriate messages, understanding varying points of view, etc. (American Speech-Language-Hearing Association, 2022). This is a core deficit in ASD and given the high rate of ASD diagnosis in FXS, it is unsurprising that this is a frequent area of weakness. In previous reviews, FXS has been noted as having difficulty in initiating and maintaining discourse, repairing communication breakdowns, and creating narratives. Increased rates of pragmatic deficits are also noted in populations with intellectual disability, language disorder, attention deficits, and other neurodevelopmental disorders (Tager-Flusberg, 2004; Towbin et al., 2005; Hoffmann et al., 2013; Smith et al., 2017; Diez-Itza et al., 2022). Pragmatic expectations are derived from cultural expectations (Hyter, 2007), creating some level of variance in terms of what is considered typical, but there are common patterns that emerge in FXS regardless of culture.

Aside from the linguistic characteristics that will be discussed, there are non-spoken elements to pragmatic language that are atypical in the FXS phenotype. Eye gaze aversion has been extensively noted as occurring regardless of the presence of other autistic symptomatology (see Hagerman and Hagerman, 2002 for review) and in both males and females, although females do continue to show increased variability in presentation (Hessl et al., 2006; Hall et al., 2009). Other nonverbal areas that are reported as being atypical in boys with FXS are intonation, gesture use, and facial expression (Klusek et al., 2014). When comparing FXS to ASD, there is mixed evidence. Some studies found that boys with FXS + ASD perform similarly to boys with ASD matched on chronological age and language ability (Losh et al., 2012; Klusek et al., 2014). Other research has found that individuals with FXS + ASD have some key differences in core autistic traits when compared to ASD (Wolff et al., 2012; McDuffie et al., 2015; Lee et al., 2016; Thurman et al., 2017b). A study by McDuffie et al. (2015) found that boys with FXS + ASD matched to a group of boys with ASD on both chronological age and autistic symptomatology had different patterns of symptoms. The group with FXS + ASD manifested less impairment in social smiling, range of facial expressions, gesture use, and restricted interests than the group with ASD. There is also evidence that social responsivity is less impaired in FXS + ASD than ASD (Wolff et al., 2012; Thurman et al., 2017b; Hong et al., 2019).

Assessments of meta-pragmatics (i.e., the understanding of what *should* occur) have found that males with FXS perform similarly to individuals with other forms of ID (e.g., DS; Losh et al., 2012; Klusek et al., 2014), and higher than individuals with ASD (Losh et al., 2012). However, functional use of those same skills, as measured by caregiver report, reveals similar performance between boys with FXS and ASD (Losh et al., 2012) and weaker performance than boys with DS (Del Hoyo Soriano et al., 2018). This suggests that the manifestation of pragmatic deficits during interactions is not reflective solely of intellectual disability.

Narrative ability (i.e., story-telling) is a key element of social interaction. In FXS, there is demonstrated impairment in narrative processing and creation (Estigarribia et al., 2011). However, in some specific areas (e.g., inferential language and providing introductory details), children and adolescents with FXS perform at similar or higher levels as TD children matched on nonverbal cognition (Finestack et al., 2012; Hogan-Brown et al., 2013). Comparisons to other groups have shown no difference in narrative macrostructure for boys with FXS and individuals with DS, ASD, and TD matched on either language or cognition (Finestack et al., 2012; Hogan-Brown et al., 2013).

Conversational analyses have revealed that males with FXS produce significantly more non-contingent remarks (i.e., responses that are tangential to the preceding remark) than males with TD who are matched on language ability (Wolf-Schein et al., 1987; Sudhalter and Belser, 2001; Martin et al., 2013b) as well as reduced usage of conversational repair strategies (Abbeduto et al., 2008; Barstein et al., 2018). Another key finding noted consistently in language analyses of FXS is excessive self-repetition of certain phrases and topics, also termed perseveration (Losh et al., 2012; Martin et al., 2012, 2013b, 2018; Del Hoyo Soriano et al., 2018; Friedman et al., 2018; Diez-Itza et al., 2022). This repetition is found in several forms, including immediate repetition of a specific word or phrase (e.g., "She's gonna be a statue, gonna be a statue"), repetition of a specific conversational device that does not add information to the conversation (e.g., "Right on"), or repeatedly returning to a specific topic of conversation (Murphy and Abbeduto, 2007). There is evidence that this is a key phenotypic element to FXS, as it is found regardless of non-verbal cognitive or language ability and in both males and females with FXS (Martin et al., 2018; Hoffmann et al., 2022). Interestingly, levels of self-repetition have distinguished groups with FXS and ASD, with FXS showing higher levels of topic and phrase repetition and ASD showing higher rates of conversational device repetition (Hilvert et al., 2020; Hoffmann et al., 2022).

3.4.1. Related factors

As in other areas, females with FXS frequently demonstrate less severe pragmatic impairment than males, although there is considerable variability (Abbeduto et al., 2008; Thurman et al., 2017a; Martin et al., 2020; Neal et al., 2022). Girls show deficits in signaling of non-comprehension as compared to TD peers matched on cognition, and there has been some research showing decreased responsivity in girls with FXS as they reach adolescence when asked to repair a communication breakdown (Thurman et al., 2017a; Martin et al., 2020). Females with FXS who also meet criteria for ASD have been shown to be less likely to signal non-comprehension, initiate conversation, or make contingent remarks in conversation than those with FXS-O or individuals with DS and TD matched on nonverbal cognition.

In infants with FXS, lower NVIQ has been shown as related to reduced initiation of joint attention (Brewe et al., 2018). Nonverbal cognition was correlated to overall ASD severity and predictive of the level of restricted and repetitive behaviors (RRBs; Abbeduto et al., 2020). However, in Haebig et al. (2020), NVIQ did not account for different performance on measures of autistic symptomatology. Similarly, a study examining question usage in boys with FXS + ASD did not find NVIQ correlated to the rate of inappropriate questions, personal questions, or requests for clarification (Friedman et al., 2020).

Some studies have found evidence of group differences based on ASD diagnosis, with boys with FXS + ASD demonstrating more impairment in pragmatic understanding and skills than FXS-O even after controlling for nonverbal cognition (Losh et al., 2012; Martin et al., 2013b; Klusek et al., 2014). The pattern of autistic symptomatology seems to vary with age. McDuffie et al. (2015) found that increased rates of RRBs were the determining factor for a comorbid diagnosis of ASD for children and adolescent males with FXS. However, when male adolescents and young adults were assessed for autistic traits, there were few RRBs with the exception of stereotyped and idiosyncratic language and more impairment in the social affective domain (Abbeduto et al., 2019).

3.5. Speech

A review of speech sound development by Barnes et al. (2006) describes a pattern of reduced intelligibility in FXS as compared to TD. Formal assessments of articulation found that boys with FXS have error patterns similar to nonverbal mental-age-matched boys with TD on single-word tasks (Paul et al., 1984; Roberts et al., 2005) and that there are increased errors on multisyllabic words as compared to single syllable words with significant effects for both nonverbal cognition and chronological age (Barnes, 2006).

Recent studies have reflected these same findings, with on-going evidence of articulation deficits as well as atypical rate of speech (Madison et al., 1986; Sudhalter et al., 1990; Ferrier et al., 1991; Belser and Sudhalter, 2001). Intelligibility in connected speech is lower than what would be predicted by performance on single words for males (Barnes, 2006; Barnes et al., 2006). This is evidenced by similar performance to boys with TD matched on nonverbal cognition on single-word tasks, but significantly lower performance on measures assessing intelligibility in connected speech (Barnes et al., 2006). Boys with FXS have also shown lower intelligibility in connected speech than boys with ASD matched on autism severity (Hilvert et al., 2020). Compared to boys with DS matched on nonverbal cognition, boys with FXS typically have better performance on all speech-sound and intelligibility tasks (Barnes et al., 2009; Kover et al., 2012; Martin et al., 2018). Acoustical analyses of speech samples have also revealed that the perceived rapid rate of speaking may stem from fewer pauses between words instead of faster rate of articulation (Zajac et al., 2006). There is also evidence that up to 50% of young adult males with FXS meet criteria for cluttering, a fluency disorder that is associated with irregular rate of speech and decreased intelligibility, with the unexpected finding that nonverbal cognition was positively correlated with increased risk of cluttering (Bangert et al., 2022).

3.5.1. Related factors

At this time, we are unable to find any published studies examining speech sound patterns in females with FXS.

Nonverbal cognition has shown strong relations to intelligibility, with lower NVIQ being associated with lower intelligibility (Barnes et al., 2006; Shaffer et al., 2020). Similarly, individuals with FXS+ASD have shown a tendency to have decreased intelligibility compared to FXS-O (Barnes et al., 2009; Estigarribia et al., 2011; Kover et al., 2012; Klusek et al., 2014; Shaffer et al., 2020), but there have been exceptions (Barnes et al., 2009; Estigarribia et al., 2011).

3.6. Literacy

Limited research exists regarding literacy development in FXS, as such this discussion will not separate out related factors. A large national survey of families living with individuals with FXS revealed that only 44% of adult males were able to read basic picture books and just 59% knew letter sounds (Bailey et al., 2009). A study comparing boys with FXS to boys with TD matched on nonverbal cognition found that boys with FXS had similar or superior performance on word reading and passage comprehension (Klusek et al., 2014). However, this same study found that phonological awareness was lower in the boys with FXS as compared to boys with TD, and that this skill was significantly correlated with autistic symptomatology (Klusek et al., 2014). A follow-up study for these same participants demonstrated that the boys with FXS acquired phonological awareness at a similar rate to the boys with TD once nonverbal cognition was controlled, although both this study and others have found a plateau in phonological awareness growth for boys with FXS at around the age of 10 years (Roberts et al., 2005; Bailey et al., 2009; Adlof et al., 2015).

Despite the relative strength found in early word recognition, there is general consensus that phonological awareness is an important predictor of reading ability, just as in typical development (Roberts et al., 2005; Bailey et al., 2009; Adlof et al., 2015). Research with adolescent boys with FXS has strengthened that understanding as phonological awareness skills had a strong positive relationship with oral word reading ability (Adlof et al., 2018).

4. Clinical implications

4.1. Assessment

Standardized language and educational assessments of individuals with FXS are central to the creation of an appropriate intervention plan (Salvia et al., 2016). Unfortunately, given their global language delays, there are frequently limited options for norm-referenced standardized assessments that have items for the appropriate skills (Hoffmann et al., 2020). As an older individual with FXS may still be at an early developmental language level, e.g., an adult who is at the two-word phrase level, an assessment that expects fluent, multi-word utterances would be inappropriate. This is especially true for the areas of syntax and morphology, which as discussed above can be specific areas of weakness. Clinicians are often faced with the choice of using an assessment that is appropriate for an individual's chronological age or using one that is appropriate for their language level. Hoffmann et al. (2020) found that the majority of individuals with FXS across a wide-age range were able to complete a standardized assessment meant for their chronological age, but that a significant percentage did not achieve a valid score (i.e., they received a score at the floor of the assessment, which does not reflect language variability).

This lack of appropriate measures often forces the use of instruments outside of their intended age range, which creates the difficulty of what scores to report. While age-equivalency scores are still frequently seen in both research and clinical reports, they are concerning psychometrically as they do not represent an equal interval scale (Salvia et al., 2007). This lack of appropriate measures has been cited as a leading cause of the failure of several clinical trials in FXS (Berry-Kravis et al., 2013b; Budimirovic et al., 2017), besides limiting the ability of clinicians to accurately assess their clients.

One option that can be considered is caregiver report, these are frequently used as they can provide information about behaviors across contexts as well as skills that are difficult to elicit in clinical or educational settings. Three commonly used caregiver report measures have been adapted for the specific profiles found in FXS, and are used to assess maladaptive behavior (Kerr et al., 2014) and social-communication/responsivity (Kidd et al., 2014). However, caregiver reports need to be combined with objective measures to gain an accurate picture of functioning (Bishop and McDonald, 2009).

Another choice that allows for an objective measure of expressive language across a wide range of language abilities is communication or language sampling. For individuals relying on primarily non-speaking means of communication (e.g., triadic eye gaze, gestures), communication sampling can allow for assessment of those often subtle behaviors (Brady et al., 2012; Hahn et al., 2017). These have been shown as effective in a wide range of populations and ages, including FXS (Brady et al., 2012; Hahn et al., 2017). For individuals regularly using two-to-three-word phrases, an expressive language sampling (ELS) protocol has been developed and shows strong psychometrics in its use in FXS (Berry-Kravis et al., 2013a; Abbeduto et al., 2020; Shaffer et al., 2020). It has been shown to differentiate between diagnoses, and to be able to characterize syntax, vocabulary, and pragmatics in FXS and other populations with varying levels of language ability (Abbeduto et al., 2020; Shaffer et al., 2020; Hoffmann et al., 2022).

Given these findings, clinicians will need to rely on a combination of clinical reasoning and research-based recommendations. What is clear is that assessment of individuals with FXS will likely require a clinician to think outside of the traditional norm-referenced standardized assessments. In order to gain an accurate understanding of ability, it is likely that multiple types of assessment will need to be used.

4.2. Intervention

Most individuals with FXS will receive services early in life, with declining rates of service utilization as they age (Martin et al., 2013a). There is growing research indicating that increased caregiver responsivity with young children is highly predictive of later language ability in FXS (Brady et al., 2014, 2020). There has also been some research as to how a parent-mediated intervention can increase social responsivity in children with FXS (Alfieri et al., 2021). This means that caregivers should be actively involved in treatment and clinicians should pay particular attention to fostering more responsive interactions. This includes supporting the use of augmentative and alternative communication (AAC) in the home, which caregivers report as being a useful tool for addressing complex communication needs in FXS (Schladant and Dowling, 2020).

As children become older, caregivers are still an important tool for improving language as there is evidence that caregiver responsivity practices can remain effective later in development with some adjustments (e.g., more commenting and fewer questions; Brady et al., 2020). Shared book-reading has also been shown as an effective tool for increasing the likelihood of sustained verbal interactions between school-aged children with FXS and their caregivers (McDuffie et al., 2016a, 2018; Nelson et al., 2018). The caregivers increased their use of language facilitation strategies (e.g., intonation prompts, modeling of story-related grammar and vocabulary) and the children showed gains in vocabulary and inferential language. The benefit to incorporating a book into this intervention is that it also continues to build on the print awareness and narrative structure needed for literacy (Justice et al., 2009). These practices that have focused on educating caregivers in communication techniques have also been proven effective when delivered *via* telehealth, opening up additional possibilities for families who may have trouble finding a provider familiar with them nearby (McDuffie et al., 2016a,b, 2018; Bullard et al., 2017; Abbeduto et al., 2020; Shaffer et al., 2020; Hoffmann et al., 2022).

Given the growing evidence that reading skills in FXS follow the same path as in TD, i.e., phonological awareness leading to increased oral word reading ability, clinicians should consider how to effectively target this area. Whereas earlier recommendations focused only on whole word recognition (Braden, 2002) secondary to concerns about weaknesses in sequential processing (Hodapp et al., 1992), there is now evidence that individuals with FXS may benefit from the traditional phonics-based approach (Adlof et al., 2018). Adlof et al. (2018) examined whether a widely available computer-based phonics program would be appropriate for a group of adolescent and young adults with FXS. They found that most of the participants (which included both males and females) were able to access and use the intervention which had been developed for use in general education.

These findings provide guidance to clinicians, although future studies examining how to support higher level language skills and school-based practices are still needed. Currently, it appears that embedding language learning opportunities in interactions that happen frequently and consistently are key elements to early language development, similar to what is recommended for other populations with language delay (Snyder et al., 2015). Similarly, growing research indicates that the key elements needed for literacy in the general populations are the same ones needed for individuals with FXS, and they can be supported by already available techniques. While it is likely that clinicians will need to modify to accommodate the FXS phenotype (e.g., providing increased repetition, structuring activities around breaks to decrease anxiety), it is also important to note that it appears that commonly used and recommended approaches to intervention are effective.

5. Conclusion

Language in FXS has benefitted from extensive research, highlighting its unique pattern of strengths and weaknesses. In general, individuals with FXS have stronger receptive than expressive language skills, and this tendency begins early in development. In both receptive and expressive language, vocabulary is often an area of strength, as compared to morphology and syntax, and at times exceeds what is expected given nonverbal cognitive abilities. Pragmatics are an area of weakness, although the role that autism comorbidity plays is still a question. Repetitive language appears to be a key component of the FXS phenotype, and its presence is independent of both IQ and autism status. The importance of considering cognition when analyzing language trends is clear, a common theme throughout the research is that when NVIQ is considered, many of the differences between FXS-O and FXS + ASD do not remain. Speech intelligibility is also an area of concern, with correlations to nonverbal cognition. Finally, literacy is an area that has received little attention, despite reports that individuals with FXS have extremely limited literacy skills.

Despite the well-established understanding of language abilities in this population, it is vital that future studies continue to extend assessment and intervention approaches to this population. While the benefits of caregiver responsivity have been made clear, there is scant research on other methods of supporting communication in individuals with FXS, especially once they reach school-age or above, despite clear evidence that they have significant needs. These areas must be addressed if we are to provide the necessary tools for best outcomes over the long-term, and likely includes how to afford caregivers with the required supports over the lifespan.

Author contributions

AH contributed to the writing of initial draft, subsequent edits, and final preparation of the manuscript.

Funding

Rush University, College of Health Sciences, provides funds for open-access publication fees.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abbeduto, L., Berry-Kravis, E., Sterling, A., Sherman, S., Edgin, J. O., McDuffie, A., et al. (2020). Expressive language sampling as a source of outcome measures for treatment studies in fragile X syndrome: feasibility, practice effects, test-retest reliability, and construct validity. *J. Neurodev. Disord.* 12:10. doi: 10.1186/s11689-020-09313-6

Abbeduto, L., Brady, N., and Kover, S. T. (2007). Language development and fragile X syndrome: profiles, syndrome-specificity, and within-syndrome differences. *Ment. Retard. Dev. Disabil. Res. Rev.* 13, 36–46. doi: 10.1002/mrdd.20142

Abbeduto, L., and Hagerman, R. J. (1997). Language and communication in fragile X syndrome. *Ment. Retard. Dev. Disabil. Res. Rev.* 3, 313–322. doi: 10.1002/(SICI)1098-2779(1997)3:4<313::AID-MRDD6>3.0.CO;2-O

Abbeduto, L., McDuffie, A., and Thurman, A. J. (2014). The fragile X syndrome-autism comorbidity: what do we really know? *Front. Genet.* 5:355. doi: 10.3389/fgene.2014.00355

Abbeduto, L., Murphy, M. M., Cawthon, S. W., Richmond, E. K., Weissman, M. D., Karadottir, S., et al. (2003). Receptive language skills of adolescents and young adults with down syndrome or fragile X syndrome. *Am. J. Ment. Retard.* 108, 149–160. doi: 10.1352/0895-8017(2003)108<0149:RLSOAA>2.0.CO;2

Abbeduto, L., Murphy, M. M., Kover, S. T., Giles, N. D., Karadottir, S., Amman, A., et al. (2008). Signaling noncomprehension of language: a comparison of fragile X syndrome and down syndrome. *Am. J. Ment. Retard.* 113, 214–230. doi: 10.1352/0895-8017(2008)113[214:SNOLAC]2.0.CO;2

Abbeduto, L., Thurman, A. J., McDuffie, A., Klusek, J., Feigles, R. T., Ted Brown, W., et al. (2019). ASD comorbidity in fragile X syndrome: symptom profile and predictors of symptom severity in adolescent and young adult males. *J. Autism Dev. Disord.* 49, 960–977. doi: 10.1007/s10803-018-3796-2

Adlof, S. M., Klusek, J., Hoffmann, A., Chitwood, K. L., Brazendale, A., Riley, K., et al. (2018). Reading in children with fragile X syndrome: phonological awareness and feasibility of intervention. *Am. J. Intellect. Dev. Disabil.* 123, 193–211. doi: 10.1352/1944-7558-123.3.193

Adlof, S. M., Klusek, J., Shinkareva, S. V., Robinson, M. L., and Roberts, J. E. (2015). Phonological awareness and reading in boys with fragile X syndrome. *J. Child Psychol. Psychiatry* 56, 30–39. doi: 10.1111/jcpp.12267

Alfieri, P., Scibelli, F., Casula, L., Piga, S., Napoli, E., Valeri, G., et al. (2021). Cooperative parent-mediated therapy in children with fragile X syndrome and Williams Beuren syndrome: a pilot RCT study of a transdiagnostic intervention-preliminary data. *Brain Sci.* 12. doi: 10.3390/brainsci12010008

American Speech-Language-Hearing Association. (2022). *Language in brief.*

Bailey, D. B., Hatton, D. D., Skinner, M., and Mesibov, G. (2001). Autistic behavior, FMR1 protein, and developmental trajectories in young males with fragile X syndrome. *J. Autism Dev. Disord.* 31, 165–174. doi: 10.1023/A:1010747131386

Bailey, D. B., Raspa, M., Holiday, D., Bishop, E., and Olmsted, M. (2009). Functional skills of individuals with fragile x syndrome: a lifespan cross-sectional analysis. *Am. J. Intellect. Dev. Disabil.* 114, 289–303. doi: 10.1352/1944-7558-114.4.289-303

Bangert, K., Scott, K. S., Adams, C., Kisenwether, J. S., Giuffre, L., Reed, J., et al. (2022). Cluttering in the speech of young men with fragile X syndrome. *J. Speech Lang. Hear. Res.* 65, 954–969. doi: 10.1044/2021_JSLHR-21-00446

Barnes, E. F. (2006). Motor speech function in young males with fragile X syndrome. *Persp. Neurophysiol. Neurogen. Speech Lang. Disord.* 16, 3–7. doi: 10.1044/nnsld16.2.3

Barnes, E., Roberts, J., Long, S. H., Martin, G. E., Berni, M. C., Mandulak, K. C., et al. (2009). Phonological accuracy and intelligibility in connected speech of boys with fragile X syndrome or down syndrome. *J. Speech Lang. Hear. Res.* 52, 1048–1061. doi: 10.1044/1092-4388(2009/08-0001)

Barnes, E. F., Roberts, J., Mirrett, P., Sideris, J., and Misenheimer, J. (2006). A comparison of oral structure and oral-motor function in young males with fragile X syndrome and down syndrome. *J. Speech Lang. Hear. Res.* 49, 903–917. doi: 10.1044/1092-4388(2006/065)

Barstein, J., Martin, G. E., Lee, M., and Losh, M. (2018). A duck wearing boots?! Pragmatic language strategies for repairing communication breakdowns across genetically based neurodevelopmental disabilities. *J. Speech Lang. Hear. Res.* 61, 1440–1454. doi: 10.1044/2018_JSLHR-L-17-0064

Belser, R. C., and Sudhalter, V. (2001). Conversational characteristics of children with fragile X syndrome: repetitive speech. *Am. J. Ment. Retard.* 106, 28–38. doi: 10.1352/0895-8017(2001)106<0028:CCOCWF>2.0.CO;2

Berry-Kravis, E., Doll, E., Sterling, A., Kover, S. T., Schroeder, S. M., Mathur, S., et al. (2013a). Development of an expressive language sampling procedure in fragile X syndrome: a pilot study. *J. Dev. Behav. Pediatr.* 34, 245–251. doi: 10.1097/DBP.0b013e31828742fc

Berry-Kravis, E., Hessl, D., Abbeduto, L., Reiss, A. L., Beckel-Mitchener, A., Urv, T. K., et al. (2013b). Outcome measures for clinical trials in fragile X syndrome. *J. Dev. Behav. Pediatr.* 34, 508–522. doi: 10.1097/DBP.0b013e31829d1f20

Bishop, D. V., and McDonald, D. (2009). Identifying language impairment in children: combining language test scores with parental report. *Int. J. Lang. Commun. Disord.* 44, 600–615. doi: 10.1080/13682820802259662

Braden, M. L. (2002). "Academic interventions," in *Fragile X syndrome: Diagnosis, treatment and research*. 3rd Edn. eds. R. J. Hagerman and P. J. Hagerman (Baltimore, MD: JHU Press), 428–464.

Brady, N. C., Fleming, K., Bredin-Oja, S. L., Fielding-Gebhardt, H., and Warren, S. F. (2020). Language development from early childhood to adolescence in youths with fragile X syndrome. *J. Speech Lang. Hear. Res.* 63, 3727–3742. doi: 10.1044/2020_JSLHR-20-00198

Brady, N. C., Fleming, K., Thiemann-Bourque, K., Olswang, L., Dowden, P., Saunders, M. D., et al. (2012). Development of the communication complexity scale. *Am. J. Speech Lang. Pathol.* 21, 16–28. doi: 10.1044/1058-0360(2011/10-0099)

Brady, N., Skinner, D., Roberts, J., and Hennon, E. (2006). Communication in young children with fragile X syndrome: a qualitative study of mothers' perspectives. *Am. J. Speech Lang. Pathol.* 15, 353–364. doi: 10.1044/1058-0360(2006/033)

Brady, N., Warren, S. F., Fleming, K., Keller, J., and Sterling, A. (2014). Effect of sustained maternal responsivity on later vocabulary development in children with fragile X syndrome. *J. Speech Lang. Hear. Res.* 57, 212–226. doi: 10.1044/1092-4388(2013/12-0341)

Brewe, A. M., Reisinger, D. L., Adlof, S. M., and Roberts, J. E. (2018). Initiating joint attention use in infants at high-risk for autism spectrum disorder. *J. Intellect. Disabil. Res.* 62, 842–853. doi: 10.1111/jir.12539

Budimirovic, D. B., Berry-Kravis, E., Erickson, C. A., Hall, S. S., Hessl, D., Reiss, A. L., et al. (2017). Updated report on tools to measure outcomes of clinical trials in fragile X syndrome. *J. Neurodev. Disord.* 9, 1–36. doi: 10.1186/s11689-017-9193-x

Bullard, L., McDuffie, A., and Abbeduto, L. (2017). Distance delivery of a parent-implemented language intervention for young boys with fragile X syndrome. *Autism Dev. Lang. Impairment* 2:239694151772869. doi: 10.1177/2396941517728690

Caravella, K. E., and Roberts, J. E. (2017). Adaptive skill trajectories in infants with fragile X syndrome contrasted to typical controls and infants at high risk for autism. *Res. Autism Spectr. Disord.* 40, 1–12. doi: 10.1016/j.rasd.2017.05.002

Casingal, C. R., Kikkawa, T., Inada, H., Sasaki, Y., and Osumi, N. (2020). Identification of FMRP target mRNAs in the developmental brain: FMRP might coordinate Ras/MAPK, Wnt/β-catenin, and mTOR signaling during corticogenesis. *Mol. Brain* 13, 1–13. doi: 10.1186/s13041-020-00706-1

Coffee, B., Keith, K., Albizua, I., Malone, T., Mowrey, J., Sherman, S. L., et al. (2009). Incidence of fragile X syndrome by newborn screening for methylated FMR1 DNA. *Am. J. Hum. Genet.* 85, 503–514. doi: 10.1016/j.ajhg.2009.09.007

Del Hoyo Soriano, L., Thurman, A. J., and Abbeduto, L. (2018). Specificity: a phenotypic comparison of communication-relevant domains between youth with down syndrome and fragile X syndrome. *Front. Genet.* 9:424. doi: 10.3389/fgene.2018.00424

Del Hoyo Soriano, L., Thurman, A. J., Harvey, D., Kover, S. T., and Abbeduto, L. (2020). Expressive language development in adolescents with down syndrome and fragile X syndrome: change over time and the role of family-related factors. *J. Neurodev. Disord.* 12:18. doi: 10.1186/s11689-020-09320-7

Diez-Itza, E., Viejo, A., and Fernández-Urquiza, M. (2022). Pragmatic profiles of adults with fragile X syndrome and Williams syndrome. *Brain Sci.* 12. doi: 10.3390/brainsci12030385

D'Souza, D., D'Souza, H., Johnson, M. H., and Karmiloff-Smith, A. (2015). Concurrent relations between face scanning and language: a cross-syndrome infant study. *PLoS One* 10:e0139319. doi: 10.1371/journal.pone.0139319

D'Souza, D., D'Souza, H., Jones, E. J. H., and Karmiloff-Smith, A. (2020). Attentional abilities constrain language development: a cross-syndrome infant/toddler study. *Dev. Sci.* 23:e12961. doi: 10.1111/desc.12961

Estigarribia, B., Martin, G. E., and Roberts, J. E. (2012). Cognitive, environmental, and linguistic predictors of syntax in fragile X syndrome and down syndrome. *J. Speech Lang. Hear. Res.* 55, 1600–1612. doi: 10.1044/1092-4388(2012/10-0153)

Estigarribia, B., Roberts, J. E., Sideris, J., and Price, J. (2011). Expressive morphosyntax in boys with fragile X syndrome with and without autism spectrum disorder. *Int. J. Lang. Commun. Disord.* 46, 100824014249025–100824014249230. doi: 10.3109/13682822.2010.487885

Ferrier, L. J., Bashir, A. S., Meryash, D. L., Johnston, J., and Wolff, P. (1991). Conversational skills of individuals with fragile-X syndrome: a comparison with autism and down syndrome. *Dev. Med. Child Neurol.* 33, 776–788. doi: 10.1111/j.1469-8749.1991.tb14961.x

Fielding-Gebhardt, H., and Warren, S. F. (2019). Early predictors of later expressive language in boys with fragile X syndrome. *Am. J. Intellect. Dev. Disabil.* 124, 11–24. doi: 10.1352/1944-7558-124.1.11

Finestack, L. H., and Abbeduto, L. (2010). Expressive language profiles of verbally expressive adolescents and young adults with down syndrome or fragile X syndrome. *J. Speech Lang. Hear. Res.* 53, 1334–1348. doi: 10.1044/1092-4388(2010/09-0125)

Finestack, L. H., Palmer, M., and Abbeduto, L. (2012). Macrostructural narrative language of adolescents and young adults with down syndrome or fragile X syndrome. *Am. J. Speech Lang. Pathol.* 21, 29–46. doi: 10.1044/1058-0360(2011/10-0095)

Finestack, L. H., Richmond, E. K., and Abbeduto, L. (2009). Language development in individuals with fragile X syndrome. *Top. Lang. Disord.* 29, 133–148. doi: 10.1097/tld.0b013e3181a72016

Finestack, L. H., Sterling, A. M., and Abbeduto, L. (2013). Discriminating down syndrome and fragile X syndrome based on language ability. *J. Child Lang.* 40, 244–265. doi: 10.1017/S0305000912000207

Flenthrope, J. L., and Brady, N. C. (2010). Relationships between early gestures and later language in children with fragile X syndrome. *Am. J. Speech Lang. Pathol.* 19, 135–142. doi: 10.1044/1058-0360(2009/09-0018)

Friedman, L., Lorang, E., Hilvert, E., and Sterling, A. (2020). "Are we done yet?" question asking in boys with fragile X syndrome and idiopathic autism spectrum disorder. *J. Speech Lang. Hear. Res.* 63, 1822–1834. doi: 10.1044/2020_JSLHR-19-00139

Friedman, L., Sterling, A., and Barton-Hulsey, A. (2018). Gaze avoidance and perseverative language in fragile X syndrome and autism spectrum disorder: brief report. *Dev. Neurorehabil.* 21, 137–140. doi: 10.1080/17518423.2018.1424264

Gallagher, A., and Hallahan, B. (2012). Fragile X-associated disorders: a clinical overview. *J. Neurol.* 259, 401–413. doi: 10.1007/s00415-011-6161-3

Gumperz, J. (1967). Language and communication. *Am. Acad. Polit. Soc. Sci.* 373, 219–231. doi: 10.1177/000271626737300112

Haebig, E., and Sterling, A. (2017). Investigating the receptive-expressive vocabulary profile in children with idiopathic asd and comorbid asd and fragile X syndrome. *J. Autism Dev. Disord.* 47, 260–274. doi: 10.1007/s10803-016-2921-3

Haebig, E., Sterling, A., Barton-Hulsey, A., and Friedman, L. (2020). Rates and predictors of co-occurring autism spectrum disorder in boys with fragile X syndrome. *Autism Dev. Lang. Impair.* 5:2396941520905328.

Haebig, E., Sterling, A., and Hoover, J. (2016). Examining the language phenotype in children with typical development, specific language impairment, and fragile X syndrome. *J. Speech Lang. Hear. Res.* 59, 1046–1058. doi: 10.1044/2016_JSLHR-L-15-0185

Hagerman, R. J., and Hagerman, P. J. (2002). *Fragile X syndrome: Diagnosis, treatment, and research*. Baltimore, MD: JHU Press..

Hagerman, R. J., Jackson, C., Amiri, K., O'Connor, R., Sobesky, W., and Silverman, A. C. (1992). Girls with fragile X syndrome: physical and neurocognitive status and outcome. *Pediatrics* 89, 395–400. doi: 10.1542/peds.89.3.395

Hahn, L. J., Brady, N. C., Fleming, K. K., and Warren, S. F. (2016). Joint engagement and early language in young children with fragile X syndrome. *J. Speech Lang. Hear. Res.* 59, 1087–1098. doi: 10.1044/2016_JSLHR-L-15-0005

Hahn, L. J., Brady, N. C., McCary, L., Rague, L., and Roberts, J. E. (2017). Early social communication in infants with fragile X syndrome and infant siblings of children with autism spectrum disorder. *Res. Dev. Disabil.* 71, 169–180. doi: 10.1016/j.ridd.2017.10.004

Hall, S. S., Burns, D. D., Lightbody, A. A., and Reiss, A. L. (2008). Longitudinal changes in intellectual development in children with fragile X syndrome. *J. Abnorm. Child Psychol.* 36, 927–939. doi: 10.1007/s10802-008-9223-y

Hall, S. S., Lightbody, A. A., Huffman, L. C., Lazzeroni, L. C., and Reiss, A. L. (2009). Physiological correlates of social avoidance behavior in children and adolescents with fragile X syndrome. *J. Am. Acad. Child Adolesc. Psychiatry* 48, 320–329. doi: 10.1097/CHI.0b013e318195bd15

Hessl, D., Glaser, B., Dyer-Friedman, J., and Reiss, A. L. (2006). Social behavior and cortisol reactivity in children with fragile X syndrome. *J. Child Psychol. Psychiatry* 47, 602–610. doi: 10.1111/j.1469-7610.2005.01556.x

Hessl, D., Nguyen, D. V., Green, C., Chavez, A., Tassone, F., Hagerman, R. J., et al. (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: the case of fragile X syndrome. *J. Neurodev. Disord.* 1, 33–45. doi: 10.1007/s11689-008-9001-8

Hilvert, E., Sterling, A., Haebig, E., and Friedman, L. (2020). Expressive language abilities of boys with idiopathic autism spectrum disorder and boys with fragile X syndrome + autism spectrum disorder: cross-context comparisons. *Autism Dev. Lang. Impairment* 5:239694152091211. doi: 10.1177/2396941520912118

Hinton, R., Budimirovic, D. B., Marschik, P. B., Talisa, V. B., Einspieler, C., Gipson, T., et al. (2013). Parental reports on early language and motor milestones in fragile X syndrome with and without autism spectrum disorders. *Dev. Neurorehabil.* 16, 58–66. doi: 10.3109/17518423.2012.704414

Hodapp, R. M., Leckman, J. F., Dykens, E. M., Sparrow, S. S., Zelinsky, D., and Ort, S. (1992). K-ABC profiles in children with fragile X syndrome, down syndrome, and nonspecific mental retardation. *Am. J. Ment. Retard.* 97, 39–46. PMID: 1386742

Hoffmann, A., Krause, S. E., Wuu, J., Leurgans, S., Guter, S. J. Jr., Block, S. S., et al. (2019). Vocabulary comprehension in adults with fragile X syndrome (FXS). *J. Neurodev. Disord.* 11:25. doi: 10.1186/s11689-019-9285-x

Hoffmann, A., Martens, M. A., Fox, R., Rabidoux, P., and Andridge, R. (2013). Pragmatic language assessment in Williams syndrome: a comparison of the test of pragmatic language—2 and the Children's communication checklist—2. *Am. J. Speech Lang. Pathol.* 22, 198–204. doi: 10.1044/1058-0360(2012/11-0131)

Hoffmann, A., Thurman, A. J., Sterling, A., Kover, S. T., Finestack, L., Berry-Kravis, E., et al. (2022). Analysis of a repetitive language coding system: comparisons between fragile X syndrome, autism, and down syndrome. *Brain Sci.* 12. doi: 10.3390/brainsci12050575

Hoffmann, A., Wang, A., Berger, N., Cordeiro, L., Shaffer, R., Tartaglia, N., et al. (2020). Language across the lifespan in fragile X syndrome: characteristics and considerations for assessment. *Brain Sci.* 10. doi: 10.3390/brainsci10040212

Hogan-Brown, A. L., Losh, M., Martin, G. E., and Mueffelmann, D. J. (2013). An investigation of narrative ability in boys with autism and fragile X syndrome. *Am. J. Intellect. Dev. Disabil.* 118, 77–94. doi: 10.1352/1944-7558-118.2.77

Hong, M. P., Eckert, E. M., Pedapati, E. V., Shaffer, R. C., Dominick, K. C., Wink, L. K., et al. (2019). Differentiating social preference and social anxiety phenotypes in fragile X syndrome using an eye gaze analysis: a pilot study. *J. Neurodev. Disord.* 11, 1–10. doi: 10.1186/s11689-019-9262-4

Hughes, K. R., Hogan, A. L., Roberts, J. E., and Klusek, J. (2019). Gesture frequency and function in infants with fragile X syndrome and infant siblings of children with autism spectrum disorder. *J. Speech Lang. Hear. Res.* 62, 2386–2399. doi: 10.1044/2019_JSLHR-L-17-0491

Hyter, Y. D. (2007). Pragmatic language assessment: a pragmatics-as-social practice model. *Top. Lang. Disord.* 27, 128–145. doi: 10.1097/01.TLD.0000269929.41751.6b

Joga-Elvira, L., Roche-Martínez, A., Joga, M. L., Jacas-Escarcelle, C., and Brun-Gasca, C. (2021). Language in young females with fragile X syndrome: influence on the neurocognitive profile and adaptive behavior. *Am. J. Med. Genet.* 185, 1448–1460. doi: 10.1002/ajmg.a.62130

Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., and Hunt, A. (2009). Accelerating preschoolers' early literacy development through classroom-based teacher–child storybook reading and explicit print referencing. *Lang. Speech Hear. Serv. Sch.* 40, 67–85. doi: 10.1044/0161-1461(2008/07-0098)

Kaufmann, W. E., Cortell, R., Kau, A. S. M., Bukelis, I., Tierney, E., Gray, R. M., et al. (2004). Autism spectrum disorder in fragile X syndrome: communication, social interaction, and specific behaviors. *Am. J. Med. Genet.* 129A, 225–234. doi: 10.1002/ajmg.a.30229

Kaufmann, W. E., and Reiss, A. L. (1999). Molecular and cellular genetics of fragile X syndrome. *Am. J. Med. Genet.* 88, 11–24. doi: 10.1002/(SICI)1096-8628(19990205)88:1<11::AID-AJMG3>3.0.CO;2-O

Kerr, C., Breheny, K., Lloyd, A., Brazier, J., Bailey Jr, D. B., Berry-Kravis, E., et al. (2014). Developing a utility index for the aberrant behavior checklist (ABC-C) for fragile X syndrome. *Qual. Life Res.* 24, 1–10. doi: 10.1007/s11136-014-0759-8

Kidd, S., Lachiewicz, A., and Barbouth, D. (2014). Fragile X syndrome: a review of associated medical problems. *Pediatrics* 134, 995–1005. doi: 10.1542/peds.2013-4301

Klaiman, C., Quintin, E. M., Jo, B., Lightbody, A. A., Hazlett, H. C., Piven, J., et al. (2014). Longitudinal profiles of adaptive behavior in fragile X syndrome. *Pediatrics* 134, 315–324. doi: 10.1542/peds.2013-3990

Klusek, J., Hunt, A. W., Mirrett, P. L., Hatton, D. D., Hooper, S. R., Roberts, J. E., et al. (2014). Reading and phonological skills in boys with fragile X syndrome. *J. Autism Dev. Disord.* 45, 1699–1711. doi: 10.1007/s10803-014-2328-y

Klusek, J., Martin, G. E., and Losh, M. (2014). A comparison of pragmatic language in boys with autism and fragile X syndrome. *J. Speech Lang. Hear. Res.* 57, 1692–1707. doi: 10.1044/2014_JSLHR-L-13-0064

Komesidou, R., Brady, N. C., Fleming, K., Esplund, A., and Warren, S. F. (2017). Growth of expressive syntax in children with fragile X syndrome. *J. Speech Lang. Hear. Res.* 60, 422–434. doi: 10.1044/2016_JSLHR-L-15-0360

Kover, S. T., and Abbeduto, L. (2010). Expressive language in male adolescents with fragile X syndrome with and without comorbid autism. *J. Intellect. Disabil. Res.* 54, 246–265. doi: 10.1111/j.1365-2788.2010.01255.x

Kover, S. T., and Abbeduto, L. (2019). Syntactic ability of girls with fragile X syndrome: phonological memory and discourse demands on complex sentence use. *Am. J. Intellect. Dev. Disabil.* 124, 511–534. doi: 10.1352/1944-7558-124.6.511

Kover, S. T., McCary, L. M., Ingram, A. M., Hatton, D. D., and Roberts, J. E. (2015). Language development in infants and toddlers with fragile X syndrome: change over time and the role of attention. *Am. J. Intellect. Dev. Disabil.* 120, 125–144. doi: 10.1352/1944-7558-120.2.125

Kover, S. T., McDuffie, A., Abbeduto, L., and Brown, W. T. (2012). Effects of sampling context on spontaneous expressive language in males with fragile X syndrome or down syndrome. *J. Speech Lang. Hear. Res.* 55, 1022–1038. doi: 10.1044/1092-4388(2011/11-0075)

Lee, M., Martin, G. E., Berry-Kravis, E., and Losh, M. (2016). A developmental, longitudinal investigation of autism phenotypic profiles in fragile X syndrome. *J. Neurodev. Disord.* 8:47. doi: 10.1186/s11689-016-9179-0

Lewis, P., Abbeduto, L., Murphy, M., Richmond, E., Giles, N., Bruno, L., et al. (2006). Cognitive, language and social-cognitive skills of individuals with fragile X syndrome with and without autism. *J. Intellect. Disabil. Res.* 50, 532–545. doi: 10.1111/j.1365-2788.2006.00803.x

Loesch, D. Z., Huggins, R. M., and Hagerman, R. J. (2004). Phenotypic variation and FMRP levels in fragile X. *Ment. Retard. Dev. Disabil. Res. Rev.* 10, 31–41. doi: 10.1002/mrdd.20006

Losh, M., Martin, G. E., Klusek, J., Hogan-Brown, A. L., and Sideris, J. (2012). Social communication and theory of mind in boys with autism and fragile x syndrome. *Front. Psychol.* 3:266. doi: 10.3389/fpsyg.2012.00266

Madison, L. S., George, C., and Moeschler, J. (1986). Cognitive functioning in the fragile-X syndrome: a study of intellectual, memory and communication skills. *J. Ment. Defic. Res.* 30, 129–148. doi: 10.1111/j.1365-2788.1986.tb01306.x

Marschik, P. B., Bartl-Pokorny, K. D., Sigafoos, J., Urlesberger, L., Pokorny, F., Didden, R., et al. (2014). Development of socio-communicative skills in 9- to 12-month-old individuals with fragile X syndrome. *Res. Dev. Disabil.* 35, 597–602. doi: 10.1016/j.ridd.2014.01.004

Martin, G. E., Ausderau, K. K., Raspa, M., Bishop, E., Mallya, U., and Bailey, D. B. Jr. (2013a). Therapy service use among individuals with fragile X syndrome: findings from a US parent survey. *J. Intellect. Disabil. Res.* 57, 837–849. doi: 10.1111/j.1365-2788.2012.01608.x

Martin, G. E., Barstein, J., Patel, S., Lee, M., Henry, L., and Losh, M. (2020). Longitudinal analysis of communication repair skills across three neurodevelopmental disabilities. *Int. J. Lang. Commun. Disord.* 55, 26–42. doi: 10.1111/1460-6984.12500

Martin, G. E., Bush, L., Klusek, J., Patel, S., and Losh, M. (2018). A multimethod analysis of pragmatic skills in children and adolescents with fragile X syndrome, autism spectrum disorder, and down syndrome. *J. Speech Lang. Hear. Res.* 61, 3023–3037. doi: 10.1044/2018_JSLHR-L-18-0008

Martin, G. E., Klusek, J., Estigarribia, B., and Roberts, J. E. (2009). Language characteristics of individuals with down syndrome. *Top. Lang. Disord.* 29, 112–132. doi: 10.1097/TLD.0b013e3181a71fe1

Martin, G. E., Losh, M., Estigarribia, B., Sideris, J., and Roberts, J. (2013b). Longitudinal profiles of expressive vocabulary, syntax and pragmatic language in boys with fragile X syndrome or down syndrome. *Int. J. Lang. Commun. Disord.* 48, 432–443. doi: 10.1111/1460-6984.12019

Martin, G. E., Roberts, J. E., Helm-Estabrooks, N., Sideris, J., Vanderbilt, J., and Moskowitz, L. (2012). Perseveration in the connected speech of boys with fragile X syndrome with and without autism spectrum disorder. *Am. J. Intellect. Dev. Disabil.* 117, 384–399. doi: 10.1352/1944-7558-117.5.384

Mattie, L. J., and Hamrick, L. R. (2022). Early communication development in infants and toddlers with fragile X syndrome. *Autism Dev. Lang. Impaimentr.* 7:239694152210994. doi: 10.1177/23969415221099403

McDuffie, A., Banasik, A., Bullard, L., Nelson, S., Feigles, R. T., Hagerman, R., et al. (2018). Distance delivery of a spoken language intervention for school-aged and adolescent boys with fragile X syndrome. *Dev. Neurorehabil.* 21, 48–63. doi: 10.1080/17518423.2017.1369189

McDuffie, A., Kover, S., Abbeduto, L., Lewis, P., and Brown, T. (2012). Profiles of receptive and expressive language abilities in boys with comorbid fragile X syndrome and autism. *Am. J. Intellect. Dev. Disabil.* 117, 18–32. doi: 10.1352/1944-7558-117.1.18

McDuffie, A., Machalicek, W., Bullard, L., Nelson, S., Mello, M., Tempero-Feigles, R., et al. (2016a). A spoken-language intervention for school-aged boys with fragile X syndrome. *Am. J. Intellect. Dev. Disabil.* 121, 236–265. doi: 10.1352/1944-7558-121.3.236

McDuffie, A., Oakes, A., Machalicek, W., Ma, M., Bullard, L., Nelson, S., et al. (2016b). Early language intervention using distance video-teleconferencing: a pilot study of young boys with fragile X syndrome and their mothers. *Am. J. Speech Lang. Pathol.* 25, 46–66. doi: 10.1044/2015_AJSLP-14-0137

McDuffie, A., Thurman, A. J., Hagerman, R. J., and Abbeduto, L. (2015). Symptoms of autism in males with fragile X syndrome: a comparison to nonsyndromic ASD using current ADI-R scores. *J. Autism Dev. Disord.* 45, 1925–1937. doi: 10.1007/s10803-013-2013-6

Meguid, N. A., Abdel-Raouf, E. R., Dardir, A. A., and El Awady, M. K. (2007). Prevalence of fragile X syndrome amove school-age Egyptian males. *World J. Pediatr.* 3, 271–275.

Murphy, M., and Abbeduto, L. (2007). Gender differences in repetitive language in fragile X syndrome. *J. Intellect. Disabil. Res.* 51, 387–400. doi: 10.1111/j.1365-2788.2006.00888.x

Neal, C. N., Brady, N. C., and Fleming, K. K. (2022). Narrative analysis in adolescents with fragile X syndrome. *Am. J. Intellect. Dev. Disabil.* 127, 11–28. doi: 10.1352/1944-7558-127.1.11

Nelson, S., McDuffie, A., Banasik, A., Feigles, R. T., Thurman, A. J., and Abbeduto, L. (2018). Inferential language use by school-aged boys with fragile X syndrome: effects of a parent-implemented spoken language intervention. *J. Commun. Disord.* 72, 64–76. doi: 10.1016/j.jcomdis.2018.02.007

Niu, M., Han, Y., Dy, A. B. C., Du, J., Jin, H., Qin, J., et al. (2017). Autism symptoms in fragile X syndrome. *J. Child Neurol.* 32, 903–949.

Oakes, A., Kover, S. T., and Abbeduto, L. (2013). Language comprehension profiles of young adolescents with fragile X syndrome. *Am. J. Speech Lang. Pathol.* 22, 615–626. doi: 10.1044/1058-0360(2013/12-0109)

Paul, R., Cohen, D., Breg, R., Watson, M., and Hermal, S. (1984). Fragile X syndrome: its relations to speech and language disorders. *J. Speech Hear. Disord.* 49, 328–332. doi: 10.1044/jshd.4903.328

Pierpont, E. I., Richmond, E. K., Abbeduto, L., Kover, S. T., and Brown, W. T. (2011). Contributions of phonological and verbal working memory to language development in adolescents with fragile X syndrome. *J. Neurodev. Disord.* 3, 335–347. doi: 10.1007/s11689-011-9095-2

Price, J. R., Roberts, J. E., Hennon, E. A., Berni, M. C., Anderson, K. L., and Sideris, J. (2008). Syntactic complexity during conversation of boys with fragile X syndrome and down syndrome. *J. Speech Lang. Hear. Res.* 51, 3–15. doi: 10.1044/1092-4388(2008/001)

Price, J., Roberts, J., Vandergrift, N., and Martin, G. (2007). Language comprehension in boys with fragile X syndrome and boys with down syndrome. *J. Intellect. Disabil. Res.* 51, 318–326. doi: 10.1111/j.1365-2788.2006.00881.x

Rague, L., Caravella, K., Tonnsen, B., Klusek, J., and Roberts, J. (2018). Early gesture use in fragile X syndrome. *J. Intellect. Disabil. Res.* 62, 625–636. doi: 10.1111/jir.12498

Reisinger, D. L., Shaffer, R. C., Pedapati, E. V., Dominick, K. C., and Erickson, C. A. (2019). A pilot quantitative evaluation of early life language development in fragile X syndrome. *Brain Sci.* 9. doi: 10.3390/brainsci9020027

Rice, M. L., Smolik, F., Perpich, D., Thompson, T., Rytting, N., and Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *J. Speech Lang. Hear. Res.* 53, 333–349. doi: 10.1044/1092-4388(2009/08-0183)

Rice, M. L., and Wexler, K. (2001). *Test of early grammatical impairment* (TEGI): Pearson Publishing.

Roberts, J. E., Mirrett, P., and Burchinal, M. (2001). Receptive and expressive communication development of young males with fragile X syndrome. *Am. J. Ment. Retard.* 106, 216–230. doi: 10.1352/0895-8017(2001)106<0216:RAECDO>2.0.CO;2

Roberts, J., Price, J., Barnes, E., Nelson, L., Burchinal, M., Hennon, E. A., et al. (2007). Receptive vocabulary, expressive vocabulary, and speech production of boys with fragile X syndrome in comparison to boys with down syndrome. *Am. J. Ment. Retard.* 112, 177–193. doi: 10.1352/0895-8017(2007)112[177:RVEVAS]2.0.CO;2

Roberts, J. E., Schaaf, J. M., Skinner, M., Wheeler, A., Hooper, S., Hatton, D. D., et al. (2005). Academic skills of boys with fragile X syndrome: profiles and predictors. *Am. J. Ment. Retard.* 110, 107–120. doi: 10.1352/0895-8017(2005)110<107:ASOBWF>2.0.CO;2

Salvia, J., Ysseldyke, J. E., and Bolt, S. (2007). *Assessment in special and inclusive education*. Belmont, CA: Wadsworth.

Salvia, J., Ysseldyke, J., and Witmer, S. (2016). *Assessment in special and inclusive education*. Boston, MA: Cengage Learning.

Schladant, M., and Dowling, M. (2020). Parent perspectives on augmentative and alternative communication integration for children with fragile X syndrome: it starts in the home. *Intellect. Dev. Disabil.* 58, 409–421. doi: 10.1352/1934-9556-58.5.409

Shaffer, R. C., Schmitt, L., John Thurman, A., Abbeduto, L., Hong, M., Pedapati, E., et al. (2020). The relationship between expressive language sampling and clinical measures in fragile X syndrome and typical development. *Brain Sci.* 10. doi: 10.3390/brainsci10020066

Smith, E., Næss, K.-A. B., and Jarrold, C. (2017). Assessing pragmatic communication in children with down syndrome. *J. Commun. Disord.* 68, 10–23. doi: 10.1016/j.jcomdis.2017.06.003

Snyder, P. A., Rakap, S., Hemmeter, M. L., McLaughlin, T. W., Sandall, S., and McLean, M. E. (2015). Naturalistic instructional approaches in early learning: a systematic review. *J. Early Interv.* 37, 69–97. doi: 10.1177/1053815115595461

Sterling, A. (2018). Grammar in boys with idiopathic autism spectrum disorder and boys with fragile X syndrome plus autism spectrum disorder. *J. Speech Lang. Hear. Res.* 61, 857–869. doi: 10.1044/2017_JSLHR-L-17-0248

Sterling, A., and Abbeduto, L. (2012). Language development in school-age girls with fragile X syndrome. *J. Intellect. Disabil. Res.* 56, 974–983. doi: 10.1111/j.1365-2788.2012.01578.x

Sterling, A. M., Rice, M. L., and Warren, S. F. (2012). Finiteness marking in boys with fragile X syndrome. *J. Speech Lang. Hear. Res.* 55, 1704–1715. doi: 10.1044/1092-4388(2012/10-0106)

Sudhalter, V., and Belser, R. C. (2001). Conversational characteristics of children with fragile X syndrome: tangential language. *Am. J. Ment. Retard.* 106, 389–400. doi: 10.1352/0895-8017(2001)106<0389:CCOCWF>2.0.CO;2

Sudhalter, V., Cohen, I. L., Silverman, W., and Wolf-Schein, E. G. (1990). Conversational analyses of males with fragile X, down syndrome, and autism: comparison of the emergence of deviant language. *Am. J. Ment. Retard.* 94, 431–441. PMID: 2137003

Tager-Flusberg, H. (2004). "Do autism and specific language impairment represent overlapping language disorders?" in *Developmental Language Disorders: From Phenotypes to Etiologies*. eds. M. Rice and S. Warren (Mahwah, NJ: Lawrence Erlbaum Associates), 42–63.

Thurman, A. J., and Hoyos Alvarez, C. (2020). Language performance in preschool-aged boys with nonsyndromic autism spectrum disorder or fragile X syndrome. *J. Autism Dev. Disord.* 50, 1621–1638. doi: 10.1007/s10803-019-03919-z

Thurman, A. J., Kover, S. T., Ted Brown, W., Harvey, D. J., and Abbeduto, L. (2017a). Noncomprehension signaling in males and females with fragile X syndrome. *J. Speech Lang. Hear. Res.* 60, 1606–1621. doi: 10.1044/2016_JSLHR-L-15-0358

Thurman, A. J., McDuffie, A., Hagerman, R. J., Josol, C. K., and Abbeduto, L. (2017b). Language skills of males with fragile X syndrome or nonsyndromic autism spectrum disorder. *J. Autism Dev. Disord.* 47, 728–743. doi: 10.1007/s10803-016-3003-2

Towbin, K. E., Pradella, A., Gorrindo, T., Pine, D. S., and Leibenluft, E. (2005). Autism spectrum traits in children with mood and anxiety disorders. *J. Child Adolesc. Psychopharmacol.* 15, 452–464. doi: 10.1089/cap.2005.15.452

Wang, L. W., Berry-Kravis, E., and Hagerman, R. J. (2010). Fragile X: leading the way for targeted treatments in autism. *Neurotherapeutics* 7, 264–274. doi: 10.1016/j.nurt.2010.05.005

Warren, S. F., Brady, N., Fleming, K. K., and Hahn, L. J. (2017). The longitudinal effects of parenting on adaptive behavior in children with fragile X syndrome. *J. Autism Dev. Disord.* 47, 768–784. doi: 10.1007/s10803-016-2999-7

Wheeler, A. C., Gwaltney, A., Raspa, M., Okoniewski, K. C., Berry-Kravis, E., Botteron, K. N., et al. (2021). Emergence of developmental delay in infants and toddlers with an FMR1 mutation. *Pediatrics* 147. doi: 10.1542/peds.2020-011528

Willemsen, R., Levenga, J., and Oostra, B. A. (2011). CGG repeat in the FMR1 gene: size matters. *Clin. Genet.* 80, 214–225. doi: 10.1111/j.1399-0004.2011.01723.x

Wolff, J. J., Bodfish, J. W., Hazlett, H. C., Lightbody, A. A., Reiss, A. L., and Piven, J. (2012). Evidence of a distinct behavioral phenotype in young boys with fragile X syndrome and autism. *J. Am. Acad. Child Adolesc. Psychiatry* 51, 1324–1332. doi: 10.1016/j.jaac.2012.09.001

Wolf-Schein, E., Sudhalter, V., Cohen, I., Fisch, G., Hanson, D., Pfadt, A., et al. (1987). Speech-language and the fragile X syndrome: initial findings. *ASHA* 29, 35–38.

Zajac, D. J., Roberts, J. E., Hennon, E. A., Harris, A. A., Barnes, E. F., and Misenheimer, J. (2006). Articulation rate and vowel space characteristics of young males with fragile X syndrome: preliminary acoustic findings. *J. Speech Lang. Hear. Res.* 49, 1147–1155. doi: 10.1044/1092-4388(2006/082)

Appendix 1

Search terms for systematic review.

PubMed: 743.

("fragile X"[tiab] OR "Fragile X Syndrome"[Mesh])

AND (Language[tiab] OR communicat*[tiab] OR conversation[tiab] OR language[tiab] OR linguistic*[tiab] OR literacy[tiab] OR literate[tiab] OR narration[tiab] OR non-verbal[tiab] OR speak[tiab] OR speech[tiab] OR talk*[tiab] OR verbal*[tiab] OR "Language"[Mesh] OR "Literacy"[Mesh] OR "Speech"[Mesh] OR "Narration"[Mesh] OR "Nonverbal Communication"[Mesh]).

NOT (mice OR (animals[mesh] NOT humans[mesh])).

Scopus: 623.

(TITLE-ABS ("fragile X"))

AND (TITLE-ABS (language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*)).

AND NOT (mice OR mouse).

ComDisDom: 46.

title((fragile X).

AND (Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*)).

OR abstract((fragile X).

AND (Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*)).

CINAHL: 282.

((MH "Fragile X Syndrome")

OR TI "fragile x."

OR AB "fragile x").

AND (((MH "Communication") OR (MH "Language") OR (MH "Nonverbal Communication") OR (MH "Verbal Behavior")).

OR TI ((Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*)).

OR AB ((Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*))).

NOT ((MH "Mice") OR (mice OR mouse)).

PsycINFO: 583.

(DE "Fragile X Syndrome."

OR TI "Fragile X."

OR AB "Fragile X").

AND ((DE "Communication" OR DE "Nonverbal Communication" OR DE "Verbal Communication" OR DE "Language").

OR TI ((Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*)).

OR AB ((Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*))).

NOT (mice OR mouse).

Cochrane Database of Systematic Reviews: 5.

Cochrane Central Register of Controlled Trials: 44.

("fragile X") AND (Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*) in Record Title.

OR ("fragile X") AND (Language OR communicat* OR conversation OR language OR linguistic* OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk* OR verbal*) in Abstract.

- (Word variations have been searched).

Google Scholar: top 35, sorted by relevance, citations and patents removed.

("fragile X") AND (Language OR communication OR conversation OR language OR linguistic OR literacy OR literate OR narration OR non-verbal OR speak OR speech OR talk OR verbal) -mice.