frontiers | Research Topics

Speech perception and language development in individuals with special educational needs

Edited by
Yuan Chen, Mingshuang Li and Xinyue Qiu

Published in
Frontiers in Education

ISSN 1664-8714
ISBN 978-2-8325-7515-4
DOI 10.3389/978-2-8325-7515-4

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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

Speech perception and language development in individuals with special educational needs

Topic editors

Yuan Chen — The Education University of Hong Kong, Hong Kong, SAR China

Mingshuang Li — California State University, Northridge, United States

Xinyue Qiu — Beijing Union University, China

Citation

Chen, Y., Li, M., Qiu, X., eds. (2026). *Speech perception and language development in individuals with special educational needs*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-7515-4

Table of contents

frontiers | Frontiers in Education

TYPE Editorial
PUBLISHED 30 January 2026
DOI 10.3389/feduc.2026.1790429

OPEN ACCESS

EDITED AND REVIEWED BY
Geoff Lindsay,
University of Warwick, United Kingdom

*CORRESPONDENCE
Mingshuang Li
mingshuang.li@csun.edu

RECEIVED 18 January 2026
REVISED 19 January 2026
ACCEPTED 20 January 2026
PUBLISHED 30 January 2026

CITATION
Li M, Chen Y and Qiu X (2026) Editorial: Speech perception and language development in individuals with special educational needs. *Front. Educ.* 11:1790429. doi: 10.3389/feduc.2026.1790429

Editorial: Speech perception and language development in individuals with special educational needs

Mingshuang Li[1]*, Yuan Chen[2] and Xinyue Qiu[3,4]

[1]Department of Communication Disorders and Sciences, California State University, Northridge, Northridge, CA, United States, [2]Department of Special Education and Counselling, Faculty of Education and Human Development, The Education University of Hong Kong, Hong Kong, Hong Kong SAR, China, [3]Department of Special Education, College of Special Education, Beijing Union University, Beijing, China, [4]Center for Educational Research on Developmental Disabilities, Beijing Union University, Beijing, China

KEYWORDS

augmentative and alternative communication (AAC), autism spectrum disorder (ASD), hearing impairment, language development, special educational needs (SEN), speech perception, telehealth

Editorial on the Research Topic
Speech perception and language development in individuals with special educational needs

Introduction

Individuals with special educational needs (SEN), including hearing impairment, autism spectrum disorder (ASD), intellectual disabilities, specific learning difficulties, and developmental language disorder (DLD), often face distinct challenges in speech processing and language development. These challenges may reflect reduced access to clear sensory input, differences in attention and learning, and difficulties integrating linguistic and social cues in real time. When such barriers recur across school, clinic, and home settings, everyday communication becomes less frequent and less effective, limiting opportunities for practice and learning. Consequently, communication difficulties in SEN can impede academic progress, social participation, and emotional wellbeing, particularly when demands are high and support is inconsistent.

This Research Topic, *Speech Perception and Language Development in Individuals with Special Educational Needs*, was conceived to provide a collaborative forum for researchers, educators, and clinicians to share advances spanning theoretical frameworks, assessment and diagnosis, intervention and educational supports, technology-enabled practice, and cultural and linguistic considerations. Across the nine accepted articles, a consistent message emerges: outcomes in SEN are shaped by interactions among sensory access, cognitive-linguistic processes, learning environments, and the quality and availability of supports across home, school, and clinical settings. Importantly, these contributions highlight that progress depends not only on individual capacities, but also on partner behaviors, instructional design, organizational capacity, and the accessibility of resources that enable participation. Collectively, these articles move beyond deficit-only accounts

toward practical, context-sensitive approaches to assessment and support that promote inclusion, communication, and sustained engagement.

Communication environments and augmentative and alternative communication (AAC): participation depends on partners and settings

Egeland-Eriksen et al. highlight that effective AAC implementation is enabled by communication-partner knowledge and skills, consistent modeling, accessible materials and aids, supportive attitudes, and organizational conditions that sustain practice. Their findings reinforce that AAC success is not solely determined by an individual's profile; rather, it is co-constructed through everyday opportunities, routines, and shared responsibility across educational teams.

Telehealth delivery in speech-language services: assessment validity, rapport, and parent-mediated implementation

As telehealth expands, Du et al. foreground tele-assessment validity, showing that remote administration can shift parent-child interaction dynamics with implications for fidelity and performance. Building from measurement to service delivery, Hao et al. emphasize rapport as a key facilitator of engagement and learning in tele-practice, while noting barriers such as technology constraints, sensory/attention challenges, and the need for effective caregiver coordination. Together, these studies argue that telehealth effectiveness and equity depend not only on platforms, but on rigorous attention to validity, teachable interactional strategies, and clear caregiver guidance, especially when remote modalities are used to broaden access for diverse and bilingual families.

Structural language and narrative discourse: refining educational planning for ASD and DLD

Two articles address language profiles using complementary levels of analysis. Andreou et al. compare core structural domains (e.g., phonology, morphosyntax and vocabulary), contributing to differential profiling relevant to assessment and intervention planning. Andreou and Lemoni focuses on narrative macrostructure (story structure, complexity, and internal state terms), emphasizing discourse-level skills that are highly relevant to classroom learning and social communication. Together, the two studies highlight the value of pairing standardized assessments with functional language tasks to provide a more complete picture of individuals' strengths, needs, and real-world communication demands.

Pragmatics in syndromic populations: toward ecologically valid profiling

Moraleda Sepulveda et al. compare pragmatic competence between individuals with 22q11.2 deletion syndrome (22q11.2DS) and 22q11.2 duplication syndrome (22q11.2DupS) using naturalistic conversational sampling combined with a pragmatic profiling approach. By foregrounding real-world interaction, this work supports assessment strategies that are clinically meaningful and sensitive to within-syndrome heterogeneity, which is an important step for individualized supports that target functional communication and participation.

Family wellbeing in hearing-loss comorbidity: broadening outcome targets

Alkahtani et al. examined maternal quality of life in families of children with Down syndrome, including those with and without hearing loss. By centering caregiver wellbeing within the context of hearing-loss comorbidity, the study reframes SEN outcomes as fundamentally family-centered, not child-only. From an editorial perspective, it highlights a broader message: addressing hearing-loss comorbidities should be viewed not as an optional add-on, but as a pathway to strengthening participation, sustaining engagement with services, and supporting the relational and social ecology in which development unfolds.

Educational underachievement and contextual determinants: integrating learning conditions with communication

Assogba et al. highlight multi-level predictors of academic difficulties, including cognitive measures, nutritional diversity, household/structural resources, distance to school, and absenteeism. This contribution is a timely reminder that language and learning trajectories are embedded within broader ecological conditions. Designing interventions for SEN populations requires coordination across educational, health, and community systems, particularly in resource-variable contexts.

Early literacy and dyslexia: addressing heterogeneity beyond single-path approaches

Finally, McMurray et al. challenge strictly phoneme-to-grapheme-only approaches for learners showing severe phonological and orthographic difficulties, and highlight the need for broader strategy repertoires. This article reinforces a cross-cutting theme of the Research Topic: SEN populations are heterogeneous, and effective educational responses often require

flexible, multi-component approaches tailored to learner profiles and contexts.

Conclusions

Collectively, these nine articles advance understanding of SEN by connecting mechanisms (e.g., speech, language, and related processing) with functional outcomes (e.g., participation and learning) and real-world implementation (e.g., schools, families, telehealth, and community contexts). Across the Topic, several shared priorities emerge: assessment approaches that are ecologically valid and sensitive to within-group heterogeneity; interventions that explicitly incorporate partner training, environmental design, and sustained organizational support; and technology-enabled practice evaluated not only for effectiveness but also through an equity and implementation lens, with careful attention to validity when service delivery shifts to remote modalities. Future work is expected to broaden cross-linguistic and cross-cultural perspectives. It may also adopt integrated measures that link structural language, pragmatics, and participation. In addition, interventions need to be tested over time to understand how supports operate and build across home, school, and clinical settings. With rigorous research and clear pathways to practice, the field may strengthen learning environments and support improved long-term outcomes for individuals with SEN.

Author contributions

ML: Writing – review & editing, Writing – original draft. YC: Writing – review & editing. XQ: Writing – review & editing.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. ChatGPT (OpenAI; model: GPT-5.2) was used exclusively for language editing (grammar, clarity, and style). The tool did not generate any new scientific content, data, results, interpretations, or references. The author(s) reviewed and verified all AI-assisted text and take full responsibility for the final manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 19 February 2025
DOI 10.3389/feduc.2025.1535819

OPEN ACCESS

EDITED BY
Yuan Chen,
The Education University of Hong Kong, Hong Kong SAR, China

REVIEWED BY
Weifeng Han,
Flinders University, Australia
Grace Cristina Ferreira-Donati,
Adastra Desenvolvimento e Comportamento Humano, Brazil

*CORRESPONDENCE
Sobh Chahboun
✉ sch@dmmh.no

RECEIVED 27 November 2024
ACCEPTED 21 January 2025
PUBLISHED 19 February 2025

CITATION
Egeland-Eriksen S, Bjørkøy I, Pettersen M and Chahboun S (2025) Breaking the silence: Norwegian teachers' perspectives on adapting language environments for AAC users.
Front. Educ. 10:1535819.
doi: 10.3389/feduc.2025.1535819

Breaking the silence: Norwegian teachers' perspectives on adapting language environments for AAC users

Siv Egeland-Eriksen[1], Ingrid Bjørkøy[2], Marit Pettersen[2] and Sobh Chahboun[2]*

[1]Krossen School, Kristiansand, Norway, [2]Queen Maud University College, Trondheim, Norway

Introduction: Many children experience challenges and limitations in spoken language, necessitating the use of alternative and augmentative communication (AAC) to communicate effectively. Learning to communicate with AAC involves a systematic training process, and it is essential to adapt the language environment to meet the specific needs of AAC users. The study aims to enhance knowledge on establishing a positive language environment for children using AAC.

Methods: A qualitative approach was employed, involving three semi-structured interviews with teachers from three distinct units within the same county in Norway. The interviews focused on elements such as teachers' knowledge and skills in AAC, communication partner and language model, use of communication materials and aids, and attitudes toward AAC.

Results: Key findings highlight that teachers' knowledge and skills in AAC, the role of communication partners and language models, the availability and use of communication materials and aids, and positive attitudes significantly contribute to creating a positive language environment for children using AAC. These elements enable effective communication and support language development.

Discussion: A combination of knowledge and effective organization is essential for prioritizing the creation of a positive language environment. This, in turn, enables children who use AAC to develop and acquire practical language skills that they can use throughout their lives. The study underscores the importance of systematic training, supportive attitudes, and tailored communication resources in enhancing the language environment for AAC users.

KEYWORDS

special education, children, augmentative and alternative communication, language environment, Norwegian teachers' perspectives, inclusive communication strategies

Introduction

Language acquisition is crucial for children's development (Høigård, 2004, p.14). Differentiating between language and speech is essential. Various types of language exist regardless of the sensory mode used. Some people require the use of alternative and augmentative communication (AAC) if verbal language is inadequate (Tetzchner, 2019, p. 276).

AAC encompasses all forms of communication assistance, including gestures made independently or with the aid of a communication device. Learning to communicate with AAC involves a systematic training process (Østvik, 2008a). The language environment of the AAC user must be adapted to suit their specific needs.

Many individuals experience growing up in a language environment where they have limitations in spoken language, either partially or completely (Blackstone and Berg, 2003). Studies

indicate that between 0.4 and 1.2 per cent of the population needs augmentative and alternative communication (AAC) (Tetzchner, 2019, cited in Karlsen and Næss, 2015, p. 5). In a report to the parliament on learning and cohesion, it is pointed out that approximately 6,500 children aged 1–18 in Norway cannot use speech as their way of communicating (Meld. Stanza 18, 2011, line. 83). It is challenging to determine the exact number of children who use AAC in Norway.

When individuals using AAC need to engage in society at the same level as others, they face a bigger obstacle. They require the surrounding environment to be organized for their benefit. AAC users require a specialized language environment to mature and acquire a practical language for lifelong use. This language setting needs to include multiple individuals who communicate in the same way as the person in question. In the context of communication, there should be language models who are more skilled communicators than the AAC user. For AAC users to be fully included in a school setting, it is crucial for other students to also utilize the same communication method (Tetzchner and Martinsen, 2014, p.310). AAC users require a supportive language environment to feel a sense of belonging, communicate effectively, establish an identity, and engage with the community.

The main goal of this research is to examine teachers' perspectives on the language setting for individuals using AAC in special education. Inquiries for research: *What experiences do teachers go through when establishing a positive language environment for children with special needs who utilize alternative and augmentative communication?*

A lot of AAC users attend their nearby school and are, to different extents, involved in a classroom environment. Multiple specialized departments are associated with nearby schools, which house numerous AAC users within their group. The language setting in a community school with an AAC student will vary compared to a specialized department where all students use AAC. This research examines the linguistic atmosphere within specialized departments that cater to groups consisting solely of AAC users.

Communication

The word *communication* comes from the Latin word communicare and means "to make common." Communicating is an interaction that takes place between people (Næss, 2015, p. 16). It has always been important for people to communicate with other people. We can communicate with language or signs and it acts as a link between people. The United Nations claims that communicating with the outside world is a basic need and a human right to communicate (Von Bernstorff, 2008). By communicating with others, we can express feelings, thoughts, intentions, attitudes, answer questions or comment on others' statements (Postholm, 2005, p. 68). We use speech, sounds, gestures, body language and facial expressions during communication, which means that communication is multimodal. One can also communicate with aids that are non-electronic or electronic (Blackstone and Berg, 2003, p. 12).

People who, for various reasons, have problems with communication, have a communication difficulty. The term communication difficulties are a collective term because the cause of the difficulty and the degree of difficulty vary in severity from person to person. Difficulties with communication can be acquired or developmental (Næss, 2015, p. 25). There are two primary categories of language challenges. The initial category consists of individuals with particular language challenges that are primarily caused by their linguistic issues. The language challenges in the second group stem from a developmental disability that is not the primary cause (Rygvold, 2004, p. 202). In this research, the participants collaborated with children who use AAC and experience varying degrees of communication challenges.

People who cannot communicate with spoken language must communicate in another ways use augmentative and alternative communication (AAC). People may be born with an illness, condition or an injury that makes them unable to use verbal speech in communication. The way all people express themselves is language (Statped, 2024). There are great differences between AAC users, but the common denominator is that they need an alternative form of communication that must supplement or replace spoken language (Tetzchner and Martinsen, 2014, p. 2). When finding the right form of communication for an AAC user, there are several aspects that must be considered. When choosing a communication aid, it is important to carry out a survey to get some answers on mobility, communication skills and the AAC user's expected development (Næss, 2015, p. 29).

Tetzchner and Martinsen (2014), leading theorists in Norway, categorize AAC users into three main groups. The group of expressive means comprises AAC users who comprehend spoken language well but cannot communicate using it. The AAC users in this group will rely on their AAC device indefinitely. The language support group utilizes AAC to aid in the development of spoken language. The AAC form of communication is temporary, not permanent. The language alternative group consists of AAC users who communicate using AAC as their primary language. Communication partners must utilize the appropriate form of communication when interacting with the AAC user (p. 66).

If an AAC user needs a talking AAC device, touchpad or computer AAC, it is called aided communication. When the way in which the AAC user communicates is a physical form separate from himself and the communicative expression is the picture or drawing, then pointing to the AAC aid is aided communication. Unaided communication is when the AAC user is able to perform the linguistic expressions themselves. Hand signs are the main form in this category, but eye blinking can also be unaided communication. There is dependent communication if the communication partner must put together or interpret what is being communicated. With independent communication, the AAC user can formulate what is communicated themselves, independently of the communication partner (Tetzchner and Martinsen, 2014, p. 8, 9).

An important element for communicative development in people is a language environment where you get linguistic stimulation as a result of access to other people's use of the language (Kunnskapsbanken, 2024). AAC users need to be exposed to other uses of language and given the opportunity to communicate themselves. Østvik has looked at the concept of language environment and looked at what different components a language environment consists of. He has broken down the language environment into eight categories which are illustrated below:

Eight components of language environment (Østvik, 2008b):

- Conceptual understanding
- Communication partners
- Physical environment
- Communication materials and aids
- Meaningful and motivating activities, theme, situations and environment
- Language models
- Barriers
- Attitudes

Conceptual understanding

The understanding the people in the language environment have of the AAC user's language. Sign language, spoken language and other forms of non-verbal communication are languages. How one understands the concept of language will have an impact on how we facilitate communication for AAC users (Østvik, 2008b) (Figure 1).

Terminological considerations in defining AAC

The terminology surrounding Augmentative and Alternative Communication (AAC) reflects varied theoretical perspectives and disciplinary conventions. In this article, AAC is described as a form of non-verbal communication, emphasizing its distinction from spoken language. However, it is important to acknowledge an alternative perspective where AAC, due to its symbolic nature, may be categorized as non-vocal (or non-oral) verbal communication.

Verbal communication is often defined as the use of structured symbolic systems to convey meaning (Von Tetzchner and Martinsen, 2002). AAC systems, such as communication boards, speech-generating devices, or manual signing, rely on structured symbols and syntax, aligning them with this definition. Thus, AAC could be described as verbal communication that substitutes or supplements oral expression. This interpretation emphasizes the linguistic and cognitive dimensions of AAC, supporting its role in fostering language acquisition and interaction.

In this study, however, the term "non-verbal" is used to highlight the non-oral characteristics of AAC and to maintain consistency with its practical focus on educational and communication strategies for AAC users. By adopting this terminology, we align with frameworks commonly employed in special education (e.g., Blackstone and Berg, 2003; Østvik, 2008b) while recognizing the validity of alternative terminological choices.

Including this clarification provides readers with a nuanced understanding of AAC's dual classification as both verbal (symbolic) and non-verbal (non-oral) communication. This distinction is critical for framing educational practices and policies that accommodate diverse conceptualizations of AAC.

Communication partners

The person who communicates with AAC users is called a communication partner, and is the most important resource for achieving good communication for AAC users (Østvik, 2008b). The Directorate of Education claims in the AAC guide that those who work in schools should have AAC knowledge and skills. Employees must recognize that AAC is as valuable and important as spoken language (Utdanningsdirektoratet, 2023). There are several AAC users who do not feel that their form of communication is used in their language environment (Karlsen, 2020, p. 85). Many people in the language environment of AAC users have no or little knowledge and experience with AAC; this also applies to teachers (Beukelman and Light, 2020, p. 136).

The person who has the greatest responsibility for facilitating communication and for the communication process itself is the communication partner (Næss, 2015, p. 32). Typical patterns of communication between AAC user and communication partner are that the communication partner is often dominant. They set a lot of guidelines for communication, ask yes/no questions, can interrupt often, have the focus on the technology, rarely confirm the content of what is being communicated and do not always give the AAC user the opportunity to answer (Blackstone and Berg, 2003, p. 13). One must have patience and dare to wait, then the probability that the AAC user will show initiative in communication will be greater (Tetzchner and Martinsen, 2014, p. 165). AAC users are more likely to develop learned helplessness and stop responding, which occurs when the AAC user does not believe that their response has an impact on the surrounding

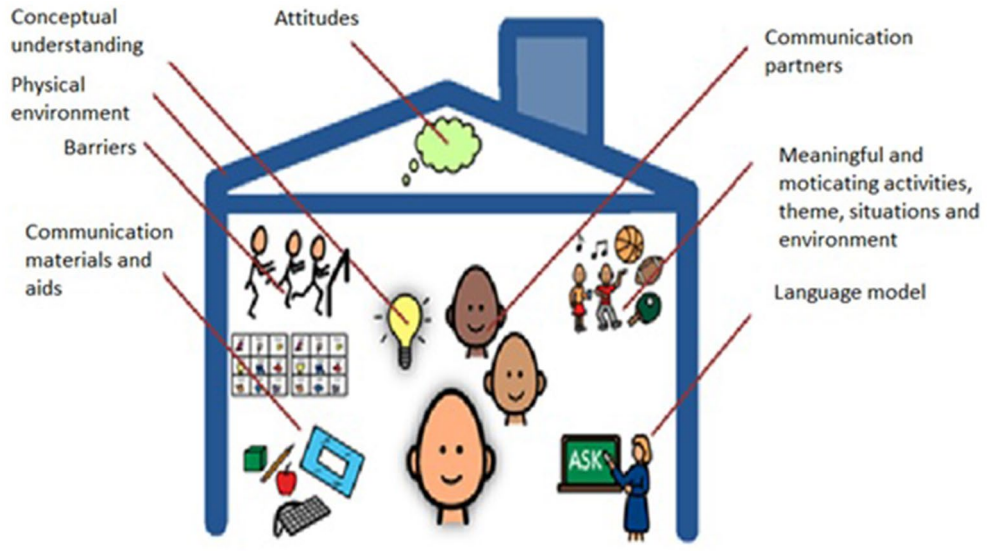


FIGURE 1
Language environment (Kunnskapsbanken) (Østvik, 2008b).

environment (Feeley and Jones, 2011, p. 280). Communication partners must tolerate silence, confirm and interpret, be motivated, patient and interested in the AAC user (Næss, 2015, p. 32).

Physical environment

The physical framework in which the AAC users are located constitutes the physical environment. Access to communication materials and aids, and communication partners is important for communication. The organized activities must be made possible for participation and adapted to AAC users. One must work to ensure that the physical environment does not hinder communication (Østvik, 2008b, p. 19). In article § 9–3, it is specified that pupils have the right to learning materials, inventory and necessary equipment (Directorate of Education, 2021).

Communication materials and aids

The aid with which the AAC user communicates must be available at all times (Thunberg, 2015, p. 120). The people in the surrounding language environment must see the importance of the communication aid being easily available, at all times and with all communication partners (Tetzchner and Stadskleiv, 2016, p. 20). In the UN Convention on the Rights of Persons with Disabilities (CRPD - Convention on the Rights of Persons with Disabilities) in chapter nine, which deals with accessibility, it is specified that people with disabilities must have access to communication on an equal basis with other people (UN, 2008).

Meaningful and motivating activities, theme, situations and environment

We get to know the outside world by participating in activities and situations that engage us. By unfolding ourselves in the world, we learn what possibilities we have. We feel that we can influence and take control. Great demands are placed on the people in the environment around AAC users to make communication, participation and learning possible. Employees who work in schools must facilitate learning and have special knowledge about this for AAC users (Slåtta, 2021, p. 98). It is important that AAC users are in a language environment that facilitates communication, interaction and learning. AAC users must be allowed to use their form of communication in motivating and meaningful activities (Østvik, 2008b, p. 20).

Language models

There must be language models in the language environment so that the AAC users receive linguistic stimulation. There must be people in the language environment who model and are more linguistically competent than the AAC user. In this way, the AAC user can imitate, learn and experience their form of communication in use (Østvik, 2008b, p. 20). The AAC user is exposed to communication based on their level of development, and can in this way develop vocabulary, understanding of concepts, language content, form and its use in practice (Karlsen, 2020, p. 93).

Many school staff lack knowledge of communication aids and experience in using them, which makes them good language models (Karlsen, 2020, p. 85). The staff in the school must receive training and the opportunity to use the different forms of communication. That could be a good use of money and time to provide good training to the people in the language environment of AAC users (Tetzchner and Martinsen, 2014, p. 322).

Barriers

Barriers in the language environment are factors that limit the possibility of communication for AAC users. The UN Convention points out barriers that prevent AAC users from participating in society effectively, and uses the terms "environmental barriers" and "attitudinal barriers." The people in the language environment should work toward creating change to minimize or remove barriers (Østvik, 2008b, p. 21). If one manages to uncover barriers, these might transform into opportunities (Skogdal, 2015).

Attitudes

The employees within the language environment of AAC users can promote or inhibit development with their attitudes. The best way to combat barriers is the attitudes, and they are reflected in our expectations of communication and what we believe are important aspects in a language environment. How inclined the employees are to change their own practice in order to promote a good language environment for AAC users is shown in their attitudes (Østvik, 2008b). In CRPD chapter eight it is stated that one must "promote positive attitudes toward and have greater societal awareness of people with disabilities" (UN, 2008).

Methods

This study has a qualitative approach, as it is based on human experiences. The informants are three teachers who work with AAC users and create a positive language environment for them. It is their experiences and reflections that are highlighted. A qualitative research interview has been conducted where the purpose is to obtain information from the informants about their life world, and then to interpret the empirical evidence (Kvale and Brinkmann, 2017, p. 22). As an attempt to include most relevant topics, an interview guide was prepared. The interview was semi-structured, as it gave more flexibility in the interview situation. The selection was criteria-based, as it was of crucial importance that the informants were teachers who worked with AAC users in a specialized department. After the interview, complete transcriptions were carried out, and Østvik (2018) eight components for the language environment were chosen as categories for the interpretation and analysis process.

Informants

All three participants are women, and to preserve their anonymity, fictive names have been assigned (Table 1).

TABLE 1 Participants' profiles and backgrounds.

Informant (Fictive names)	Education	Number of years worked as a teacher	Number of years working with AAC
Anne	Child Protection Pedagogy (bachelor) Special pedagogy (master) AAC	5 years	10 years
Bente	Social worker (bachelor) Sports Pedagogy Guidance AAC	8 years	8 years
Celine	Teacher AAC	16 years	16 years

Semi-structured interviews

The study includes a semi-structured interview guide (see Appendix 1) designed to explore teachers' experiences in promoting effective language environments for students utilizing augmentative and alternative communication (AAC). The interviews were conducted at the teachers' workplaces, specifically within competence departments in the same county in Norway. The duration of the interviews ranged from 30 min to an hour, as is common for semi-structured interviews. All interviews were conducted by the same researcher, ensuring consistency in data collection. The study employed a qualitative methodological approach, grounded in phenomenology and hermeneutics, to investigate the subjective experiences of participants and provide an interpretive analysis of the data. The three participating teachers represented diverse educational levels, including primary, secondary, and high school settings. Key themes of the interview guide included teachers' experiences with facilitating language environments, challenges encountered in supporting AAC users, strategies for overcoming barriers, and reflections on the roles of communication partners and physical environments.

Ethical considerations

Ethical guidelines and guidelines have been followed according to Postholm (2005) to ensure that the research process (Postholm, 2005, p. 145). The national research ethics committee (NESH) has ethical guidelines that apply to research. Three important considerations to take are that those who participate in the research should not be harmed by participating, the informants' right to self-determination and that researchers respect the informants (Johannessen et al., 2021, p. 45). In the study in question, approval has been sought from SIKT - the knowledge sector's service provider. The participants received all the necessary information concerning the current study.

Results and discussion

The current research focuses on what experiences teachers go through when establishing a positive language environment for children with special needs who utilize alternative and augmentative communication.

All informants work daily at a school with students who need augmentative and alternative communication (AAC) and possess great expertise when it comes to creating a language environment that is adapted to AAC users. Their experiences and reflections have been highlighted to support their view of the language environment for AAC users. The following elements were identified as essential for this study:

Knowledge and skills in AAC

The Norwegian Directorate of Education highlighted the importance of school staff having skills and knowledge of AAC (Utdanningsdirektoratet, 2023). Tetzchner and Martinsen (2014) highlight that staff must be given the opportunity to use AAC and receive training in AAC. The training is necessary to get competent staff to use AAC. Bente says that working to increase the competence at AAC must be fixed in the timetable, otherwise the staff will not be able to do it. Management must prioritize training in AAC, in order to increase knowledge and skills. The informants point out that there should be an AAC manager in each specialized department, which all the informants confirmed existed at their place of work. The informants pointed out reasons why there may be little AAC skills and knowledge among the employees could be the management's attitudes toward the importance of training and further education, and how much time the management sets aside to increase AAC knowledge and skills during working hours. The informants want more time to practice with the employees and share experiences.

The informants also pointed out that the employees themselves must want to increase their knowledge and skills. Bente says that the staff can also be a barrier, if you do not bother or find it stressful to learn the communication aids. It is not only knowledge and skills in relation to the practical use of communication aids that are important, but one must also have knowledge of the application process, be familiar with relevant material and available symbols.

It is crucial to have knowledge and skills about AAC if one is to manage to create a good language environment for AAC users. Staff must be given time for practical tasks around AAC, share experiences and practice with other staff. The staff need further education and courses in AAC to increase their competence.

Communication partner and language model

There must be people in the language environment who model the form of communication that the AAC user uses. In this way, the AAC user can learn, imitate and experience (Østvik, 2008b).

The informants shared many good reflections and experiences around this topic. Bente says that it is the employees' job to use the communication aids to the greatest extent possible. We must model, model and model. Celine agreed and added that if the AAC users are to learn to use communication aids, there must be AAC competence among the staff. Tetzchner and Martinsen (2014) say that more AAC users become communicative underachievers. This is because there is a lack of skills and knowledge about communication aids in the language environment, which leads to a lack of communication opportunities. Bente claims that the employees must know where to find the words in the communication aid. «*You cannot sit and browse and not find out yourself. In advance, you have to practice, try yourself out and really get to know the communication aid.* ».

Anne also said that the adults must point, point, point and point, in the same way as we talk, talk and talk to a small child. We must model and be good language models. Celine goes further and says that in an ideal world, the employees would use AAC when speaking to each other too, so that the AAC users are bathed in language, just as children are bathed in verbal speech. AAC users need to see that their form of communication is being used. The staff must also be realistic, patient and celebrate the small advances in the language development of AAC users.

In order to be good language models, knowledge and skills in AAC must be the basis, the findings show. Previous studies also points out the importance of the communication partner having to take greater responsibility and facilitate communication with the AAC user. Bente says that you must have the expectation that you will get a response in communication with AAC users. She claims that «*you have to give the AAC user enough time to respond, be patient and not talk too much yourself, as you often tend to do in communication with AAC users.* » If a communication aid is used in communication, it takes extra time to communicate what you want. The informants were aware of this by giving AAC users time in communication, but were also honest that you do not always have that time at your disposal.

The informants and the literature emphasized that one must be interested, patient and motivated to be a good communication partner. You have to tolerate the silence that may occur and be honest if there is something you do not understand what the AAC user was trying to communicate.

Communication materials and aids

Communication materials and aids must always be available (Thunberg, 2015).

Anne says that she tries to encourage her colleagues to carry communication aids everywhere, dare to remind each other of it, and that they can thus point and model in all situations. She goes on to say that «*we have to bathe the AAC users in language.* » If the employees do not bring the communication aids, they violate what is written in CRPD chapter 12 that AAC material must be available. The informants were very aware of this, and they strive to always have the communication aid with them, but that it can also be challenging to do at all times.

The informants had different ways of having communication material available at their workplaces. Two of the schools have large AAC boards with symbols outside in the schoolyard so that students from the competence department and the local school can use them. This creates a common focus and attention on AAC in the school. All the informants had symbols on them, either in the form of aprons with Velcro with symbols or key rings around the neck with symbols. All classrooms and other living spaces have blackboards and boards with accessible symbols.

Both the literature and the informants claim that it is important to communicate about what you want and with whom you want. When choosing a communication aid, there are several factors that must be considered, and the various aids have advantages and disadvantages. Celine says that not all communication aids can withstand all kinds of weather. Anne explained that it can be easier to use sign-to-speech if you have to give a message to something far away, instead of a symbol that you have to be close to in order to communicate with it. The advantage of symbols is that many people who do not know AAC can understand symbols, while sign-to-speech is something you have to know. The different communication materials and aids cover different needs.

Availability is also affected by materials disappearing, a lack of licenses, aids running out of power, technical problems arising or being destroyed, say the informants. The communication aids must also be always updated, so that the AAC user has the opportunity to develop his language. Anne says that it is important to show the AAC users that they keep the communication aids up to date and show that they are important and use them. Both the literature and the informants point out that there must be one person who has primary responsibility for the communication aid.

Attitudes

A person's attitudes reflect what one considers important in a language environment. This once more highlights our anticipation for communication and the readiness to modify our own methods to enhance the language setting for AAC users (Østvik, 2008b). Being aware of one's own attitudes is necessary when working on one's mindset. Celine openly admits that attitudes are the most difficult thing to address and the most challenging to deal with. She also emphasizes the importance of comprehending why you need to do certain tasks, rather than simply following instructions, in order to perform more effectively. Employees need to be knowledgeable of the goal, understand their choices to reach the goal, and grasp the significance of their work. Those who assist AAC users should consider their own attitudes toward AAC. This should be carried out independently, within the team one belongs to, or with the support of the administration.

What view one has of communication affects how we communicate and work with communication. The employees must, according to the AAC guide (2023), recognize that AAC is just as valuable and just as important as communication with verbal speech. Bente believes that we must think of AAC as the student's language. The attitudes we have toward AAC and toward the communication aid are important. Celine points out that the employees who have the attitude that this is a job they want to do, are

TABLE 2 Summary of the main findings of the study.

Key findings	Summary	Recommendations	Congruence/discordance
Knowledge and skills in AAC	Staff need consistent training and knowledge of AAC, practical use, application processes, and material knowledge. Challenges include management's attitude and time constraints.	Management should prioritize AAC training, allocate time for practice, and appoint AAC managers in departments.	*Congruent:* All informants agree on the need for consistent training and noted similar challenges.
Communication partner and language model	Staff must model AAC communication extensively to help users learn. Effective communication requires patience, practice, and realistic expectations.	Staff should practice AAC tools, model communication consistently, and remain patient to support AAC users' learning.	*Congruent:* All informants emphasize the importance of modeling and agree on the need for staff competence.
Communication materials and aids	Communication aids must be accessible, maintained, and updated. Examples include large AAC boards and portable materials like Velcro symbols.	Ensure consistent availability of aids, assign responsibility for maintenance, and encourage use in diverse settings.	*Mostly congruent:* All informants stressed availability, though specific approaches to implementation varied slightly.
Attitudes	Positive attitudes toward AAC are essential. Staff must value AAC as equally important as verbal speech and remain motivated and reflective.	Foster positive attitudes through reflection, team discussions, and administrative support to enhance AAC integration.	*Partially congruent:* Informants agreed on the importance of attitudes but highlighted varying levels of staff motivation and challenges in fostering reflection.

the people who need to work with AAC users. Lack of understanding of why it is important can hinder a good language environment. Our attitudes can be directly reflected in the access AAC users have to their language. Anne says that if someone thinks that this child cannot communicate, then there is little point in communicating with that child. That AAC users can understand and make themselves understood is extremely important. AAC users must be given the opportunity to express their thoughts, needs and feelings. Bente points out that the student does not have the opportunity to communicate if the adults do not bother to take communication aid with them. If you have an attitude that AAC is not that important, then AAC users will not be able to communicate what they want. The attitudes one has toward AAC aids are decisive for what we put into it.

AAC users are raised in a society where most individuals communicate verbally, so educators must ensure that language is easily understandable for AAC users in school settings. According to the CRPD, it is important to encourage positive views of individuals with disabilities. Working on our attitudes presents a wonderful chance to reduce or eliminate barriers in the language environment of AAC users. Employees need to be conscious of the topic of attitudes, despite the challenge of discussing it. The participants concurred that the motivation to assist AAC users, along with the mindset toward the task, impacts the quality of work done, ultimately shaping the language atmosphere for AAC users.

Summary of findings and recommendations for AAC implementation

The table below summarizes the key findings from the study, highlighting the identified themes and corresponding recommendations for enhancing the language environment for AAC users. Additionally, the table includes an assessment of congruence or discordance among the three informants, illustrating the degree of agreement on each theme based on their reflections and experiences. This presentation provides a clear overview of the results while contextualizing the perspectives of the participants (Table 2).

Conclusion

Teachers and other staff members at the school are obligated to create a supportive and stimulating language and learning atmosphere for children with special needs overall, especially those who use AAC.

A combination of knowledge and organization is key in prioritizing the wellbeing of these children. Furthermore, it appears that serving as an effective language model and communication ally enhances language and communication acquisition for AAC users. Adults in the school, including educators, need to be proficient in AAC to effectively introduce it to the students. The AAC user's language needs to be current and accessible constantly. Furthermore, educators need to acknowledge that AAC is equally valuable and crucial compared to verbal communication, and they must maintain a positive attitude in their everyday teaching.

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 humans were approved by Norwegian Agency for Shared Services in Education and Research. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SE-E: Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft. IB: Validation, Visualization, Writing – review & editing. MP: Validation, Visualization,

Writing – review & editing. SC: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Beukelman, D. R., and Light, J. C. (2020). Augmentative & Alternative Communication. *5. utg* Edn. Baltimore, MD: Paul H Brooks Publishing Co.

Blackstone, S. W., and Berg, M. H. (2003). Social networks: en kommunikasjonsoversikt for mennesker med store kommunikasjonsvansker og deres kommunikasjonspartnere. Levanger: Trøndelag kompetansesenter.

Feeley, K. M., and Jones, E. A. (2011). Obtaining and maintaining communicative interactions. I AAC strategies for individuals with moderate to severe disabilities. New York, NY: Brookes publishing company, 279–310.

Høigård, A. (2004). Barns språkutvikling. Oslo: Universitetsforlaget.

Johannessen, A., Tufte, P. A., and Christoffersen, L. (2021). Introduksjon til samfunnsvitenskapelig metode (6. utg.). Oslo: Abstrakt forlag.

Karlsen, A. V. (2020). Flere med i klassefellesskapet: digitale læringsressurser som støtte. Oslo: Universitetsforlaget.

Karlsen, A. V., and Næss, K. A. B. (2015). God kommunikasjon med AAC-brukere. Bergen: Fagbokforlaget.

Kunnskapsbanken. (2024). Språkmiljø. Available at: https://www.kunnskapsbanken.net/kommunikasjon/sprakmiljo/ (Accessed July 4, 2024).

Kvale, S., and Brinkmann, S. (2017). Det kvalitative forskningsintervju (3. utg.). Oslo: Gyldendal Akademisk.

Meld. Stanza 18 (2011). Læring og fellesskap. Available at: https://www.regjeringen.no/contentassets/baeeee60df7c4637a72fec2a18273d8b/no/pdfs/stm201020110018000dddpdfs.pdf (Accessed October 5, 2024).

Næss, K. A. B. (2015). "God kommunikasjon med AAC-brukere" in God kommunikasjon med AAC-brukere. eds. I. K. A. B. Næss and A. V. Karlsen (Bergen: Fagbokforlaget), 15–46.

Østvik, J. (2008a). Hva er Alternativ og supplerende kommunikasjon (AAC): og hvem har behov for det? Dialog: medlemsblad for ISAAC Norge, 2008 (Ekstra nummer), 4–7.

Østvik, J. (2008b). Språkmiljø - tanker om prinsipielle sider og praktiske sider. *Dialog: medlemsblad for ISAAC Norge* 2008, 18–22.

Østvik, J. (2018). Språkmiljø - tanker om prinsipielle og praktiske sider [illustrasjon]. Available at: https://www.kunnskapsbanken.net/sprakmiljo-tanker-om-prinsipielle-og-praktiske-sider/

Postholm, M. B. (2005). Kvalitativ metode. Oslo: Universitetsforlaget.

Rygvold, A. L. (2004) Språkvansker hos barn. Ibefring, E. & R Tangen. (Red.), Spesialpedagogikk (s. 197–216). Oslo: Cappelen Akademisk Forlag.

Skogdal, S. (2015) Mulighetsbetingelser for deltagelse og kommunikasjon i skolen. I. K. A. B. Næss and A. V. Karlsen (Red.), God kommunikasjon med AAC-brukere (s. 217–238). Bergen: Fagbokforlaget.

Slåtta, K. (2021) Deltagelse og engasjement. I T. Horgen, K. Slåtta and Gjermestad (Red.), Multifunksjonshemming, livsutfoldelse og læring (2. utg.). (s. 83–99). Oslo: Universitetsforlaget.

Statped. (2024). Alternativ og supplerende kommunikasjon. Available at: https://www.statped.no/AAC/ (Accessed April 7, 2024).

Tetzchner, S. V. (2019). Barne- og ungdomspsykologi. (1. utg.). Copenhagen: Gyldendal.

Tetzchner, S. V., and Martinsen, H. (2014). Alternativ og supplerende kommunikasjon. En innføring i tegnspråkopplæring og bruk av kommunikasjonshjelpemidler for mennesker med språk og kommunikasjonsvansker. Copenhagen: Gyldendal.

Tetzchner, S. V., and Stadskleiv, K. (2016). Constructing a language in alternative forms. The silent partner? Language, interaction and aided communication. London: Guildford J&R Press Ltd., 27–34.

Thunberg, G. (2015). Kommunikationshjälpmedel. I K. A. B. Næss and A. V. Karlsen (Red.), God kommunikasjon med AAC-brukere (s. 119–158). Bergen: Fagbokforlaget.

Directorate of Education. (2021). Veilederen spesialundervisning. Available at: https://www.udir.no/laring-og-trivsel/spesialpedagogikk/spesialundervisning/Spesialundervisning/Retten/2.6/ (Accessed October 5, 2024).

Utdanningsdirektoratet. (2023). Alternativ og supplerende kommunikasjon (AAC). Available at: https://www.udir.no/laring-og-trivsel/spesialpedagogikk/alternativ-og-supplerende-kommunikasjon-AAC/

Von Tetzchner, S., and Martinsen, H. (2002). Alternativ og supplerende kommunikasjon: en innføring i tegnspråksopplæring og bruk av kommunikasjonshjelpemidler for mennesker med språk-og kommunikasjonsvansker. Oslo: Gyldendal akademisk.

Von Bernstorff, J. (2008). The changing fortunes of the Universal Declaration of Human Rights: Genesis and symbolic dimensions of the turn to rights in international law. *Eur. J. Int. Law*. 19, 903–924.

Appendix

Interview guide
Introduction

- Brief introduction of the researcher and the purpose of the study.
- Explanation of the participant's rights, including confidentiality and voluntary participation.
- Request for permission to record the interview.

Main topics and questions

1. **Experiences with Language Environments**
   - Can you describe your experiences with facilitating a good language environment for students using AAC?
   - What do you consider to be the key factors in creating an effective language environment?

2. Challenges
   - What challenges have you encountered when working to support students who use AAC?
   - How do these challenges impact your ability to facilitate a good language environment?

3. Strategies and solutions
   - What strategies have you employed to overcome these challenges?
   - Are there specific methods or tools that you find particularly effective in supporting students with AAC needs?

4. Communication partners
   - How do you perceive the role of communication partners in the language environment?
   - What is your approach to engaging peers, teachers, or other staff as communication partners?

5. Physical environment
   - How does the physical environment influence the language opportunities for students using AAC?
   - Can you provide examples of adjustments or changes made to the environment to support AAC users?

6. Reflections and recommendations
   - Based on your experiences, what recommendations would you give to others working with AAC students to improve language environments?
   - Are there specific resources or support systems you feel are lacking or needed to enhance your work?

Closing

- Is there anything else you would like to add or highlight about your experiences with AAC students and language environments?
- Thank the participant for their time and contributions.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 23 July 2025
DOI 10.3389/feduc.2025.1626433

OPEN ACCESS

EDITED BY
Xinyue Qiu,
Beijing Union University, China

REVIEWED BY
Megan Israelsen-Augenstein,
West Virginia University, United States
Feifan Wang,
Shanghai Normal University, China

*CORRESPONDENCE
Georgia Andreou
✉ andreou@uth.gr;
✉ galemoni@uth.gr

RECEIVED 10 May 2025
ACCEPTED 04 July 2025
PUBLISHED 23 July 2025

CITATION
Andreou G and Lemoni G (2025) Narrative skills of children with developmental language disorder: retelling in macrostructure.
Front. Educ. 10:1626433.
doi: 10.3389/feduc.2025.1626433

Narrative skills of children with developmental language disorder: retelling in macrostructure

Georgia Andreou* and Garyfallia Lemoni

Department of Special Education, University of Thessaly, Volos, Greece

Introduction: Developmental Language Disorder is a common developmental disorder that affects 7% of both preschool and school-aged children. Children with DLD typically demonstrate simpler syntax, higher rates of grammatical errors and greater difficulty acquiring new vocabulary in comparison to typically developing (TD) children. Research has shown that children with DLD have significant language difficulties that result in poor narrative performance.

Methods: In this paper, the narrative skills of monolingual children with DLD and typically developing (TD) children are examined at the macrostructural level, using one of the most common methods of assessing narrative skills, picture retelling. The sample consists of 100 preschool and school-aged children (50 with DLD and 50 TD), aged 5 to 11 years, who were matched according to chronological age, socioeconomic status and language input received at home. The parameters measured are story structure, structural complexity and Internal State Terms (IST). The research hypotheses of the study were: (a) Children with DLD will present lower performance than TD children in the parameter of story structure, (b) Children with DLD are expected to present lower performance than TD children in the parameter of structural complexity, (c) Children with DLD will present lower performance than TD children in the parameter of expressing Internal State Terms and (d) There will be a significant dependence between the groups (children with DLD and TD children) and the subcategories of the story structure.

Results: The results confirmed all the above hypotheses except for the hypothesis that there is a significant dependence between the groups of children in the subcategories of the story structure.

Discussion: The findings of our study revealed that the narratives skills of children with DLD are more affected at the level of macrostructure than those of children with Typical Development (TD). However, a significant dependence between the groups (children with DLD and TD children) and the subcategories of the story structure was found only in five out of sixteen components in the three episodes of the story, a finding which is discussed on the basis of the limitations included in the present study.

KEYWORDS

developmental language disorder, narrative skills, macrostructure, retelling, typically developed children

1 Introduction

1.1 The development of narrative skills in children

The development of narrative skills in children starts from birth, when they begin to participate in interactions with older language users and develop gradually as they grow up. From infancy, children begin to acquire an understanding of the structure, use, and meaning of language and then they use one- or two-word phrases to communicate (Ralli and Sidiropoulou, 2012). Around the age of two, children begin to combine a group of ideas using words like "then" and "and" to connect sentences. After that, they are able to sequence story elements together without causal or time links. As children develop, they tell primitive stories with basic elements such as setting, main characters, and topic. After that stage, the stories of children begin to follow a predictable timeline. Finally, around the age of 5–7 years, children can tell stories with a true plot and well-developed storyline. At this age there is character development, sequencing of events, a problem, and a solution in children's narrations (Hutson-Nechkash, 2001).

Previous studies have shown that children's narrative abilities develop extensively during preschool and early school years (Pearson, 2002; Schneider et al., 2006). Narrative ability is an important skill for both children's school performance as well as their daily communication and constitutes a strong predictor for their later language skills. Thus, children with poor performance in narratives tend to show persistent language problems in lexical and syntactic skills (Botting et al., 2001; Mazlan et al., 2024). Narrative data may provide information about various aspects of children's language skills, such as their ability to structure complex discourse (Fiestas and Peña, 2004) or to narrate how story characters think and feel (Burris and Brown, 2014). For these reasons, studies of children's narratives have become increasingly popular in recent years, with both monolingual and bilingual children speaking a number of different languages being investigated on the development of their narrative ability (Lindgren, 2022).

1.2 Developmental language disorder (DLD)

Developmental Language Disorder (DLD) is a neurodevelopmental disorder characterized by persistent language difficulties in comprehension and/or production. It first emerges in early childhood in the absence of sensory, intellectual, or neurological problems and affects roughly 7% of the general population (Norbury et al., 2016). DLD was previously known as Specific Language Impairment until the year 2017, when the 2017 Delphi consensus took place and Developmental Language Disorder (DLD) was established instead, following Bishop et al. (2017).

DLD is a disorder with great heterogeneity and a wide range of communication difficulties that, although manifested in childhood, still occur in adulthood (Botting, 2010). Children diagnosed with DLD as preschoolers often present difficulties in their social–emotional development later on and they also demonstrate lower levels of school performance (Vissers and Koolen, 2016). Also, DLD is characterized by perceptual and expressive linguistic deficits which include extensive use of immature phonological processes (Aguilar-Mediavilla et al., 2002), slow word retrieval, naming errors (McGregor et al., 2002) and shorter mean length of utterance (Redmond, 2004) than typical development. Thus, the ability of children with DLD to compose and transmit oral narratives appropriate for their age is affected. It has also been found that poor expressive abilities of children with DLD in early childhood are the best predictor of reading problems and dyslexia in school-aged children (Lyytinen et al., 2015), thus placing DLD children at a further disadvantage compared to their peers (Tomas and Vissers, 2019).

1.3 Narrative skills in children with DLD

Oral narration is important to children, and the skill remains decisive across the lifespan. Narratives comprise several linguistic elements (e.g., syntax, morphology, semantics, pragmatics) and children with Developmental Language Disorder (DLD) are particularly vulnerable to experiencing difficulties with storytelling not only in language comprehension but in language production as well (Pauls and Archibald, 2021).

The narratives of children with DLD are shorter, show problems in complexity, grammaticality, coherence and fluency and contain a lower information or plot value compared to the narratives of typically developing (TD) peers (Christensen, 2019). Children with DLD demonstrate difficulty with many aspects of narration, such as making logical connections between story events, establishing a sense of continuity or describing characters' feelings or intentions (Reilly et al., 2004).

The narrative skills of children with DLD may develop slowly and their stories may not contain the most advanced elements that TD children include in their narratives such as the setting, the characters and the plot. Research has shown that compared to their peers, children with DLD produce fewer complex stories (Fey et al., 2004), more confused or deficient (Jones, 2015) and with more grammatical errors (Guo et al., 2008).

In particular, Fey et al. (2004), who studied the oral and written story skills of second and fourth grade children with DLD, found that in both grades children with DLD produced shorter and poorer stories and made more grammatical errors than typically developing children. Moreover, research data coming from a long-term study of storytelling skills in preschool children in Sweden have shown that children develop their storytelling skills over time, but not at the level of children of typical development at the age of ten (Reuterskiöld et al., 2011). In another study, Vandewalle et al. (2012) found that children with DLD at the age between 5–8 years, although they have good literacy development, they continue to show reduced narrative skills, with problems in vocabulary and morphology (Wellman et al., 2011).

Furthermore, it has been found that children with DLD seem to produce less complete and immature narratives in relation to size, lexical diversity, phrase complexity, and content (Gillam and Pearson, 2004). In addition, their narratives seem to contain less details that make the story more complete, such as links, fewer causal connections between events (Hayward et al., 2007) and fewer elements of story grammar (Leonard, 2014). In addition, Cleave et al. (2010) found reduced productivity, limited literary language, and several syntactic errors in children's with DLD storytelling at the age of four, while lexical-grammatical problems were identified in children of five and eight years old (Thomson, 2005).

Other studies reveal difficulties for DLD children in morphology, such as in the production of clitical objects and the use of pronouns (Leonard, 2014), in semantics of words (Befi-Lopes et al., 2008) and lexical deficits (Leonard and Deevy, 2004). According to Tribushinina et al. (2015) these children fail to incorporate effectively syntactic and semantic elements into word processing.

In Greece, a small number of studies have been conducted on the exploration of children's narrative skills at school and early school age. Research by Tsimpli et al. (2016) in monolingual and bilingual children with Specific Language Impairment (SLI) showed differences between typically development children and children with DLD in microstructure. Other studies by Theodorou and Grohmann (2010) and Theodorou et al. (2012) found that preschool and early school children with DLD who speak Cypriot-Greek presented a significantly lower performance in storytelling than their TD peers. In addition, the results of a research study conducted by Mpaka et al. (2012) in Greek students, showed that children with DLD compared to children of the corresponding linguistic but not chronological age, presented significantly lower performance in their narrations.

1.4 Narrative tasks: macro- and microstructure

The narratives children produce are generally analyzed at two different levels, namely macrostructure and microstructure. The term macrostructure refers to the overall content and organization of the story (Govindarajan and Paradis, 2022). The two most widely used models of macrostructure are story grammar and high point analysis. Such approaches focus on recognizing the key components of a story, the sequence of events, and the episodic structure of a story (Justice et al., 2006). According to the Story Grammar model, a story has (1) a *Setting* that introduces the time, place, and characters in the story, (2) an *Initiating Event* that sets up the problem or dilemma in the story, (3) an *Internal Response* or the character's response to the Initiating Event, (4) an *Attempt* of the character to solve the problem, (5) the *Outcome or the result* of the previous action, and (6) a *Response* or how a story character responds to the outcome (Govindarajan and Paradis, 2022).

On the other hand, the term microstructure refers to the language content of the discourse. It is an analysis of the linguistic structures used to produce stories. It includes measures of productivity and measures of complexity (Justice et al., 2006). Microstructure refers to the word- and sentence-level components of a story, such as the variety of vocabulary, clarity of cohesion or pronominal references or complexity of syntax and the use of referential, temporal and causal linking devices (Lindgren, 2022).

The types of narratives used in language acquisition research can be either fictional or personal (Fioretti et al., 2019). Personal narratives come from the narrator's life experiences and they are the first and most important type of narration that small children acquire. On the other hand, fictional narratives describe imaginary events, characters, and settings. Unlike personal narratives, which are based on real-life experiences, fictional narratives are created from the imagination and can include elements of fantasy, adventure and mystery. Fictional narratives often follow a specific structure, including components such as characters, settings, initiating events, internal responses, plans, actions, consequences, and conclusions (Gillam and Pearson, 2004).

Fictional narratives are usually evaluated through story retelling and through story telling. In the first case, children repeat a story they have just heard or through story generation tasks, in which children may produce a story while looking at a wordless picture book. In the story retelling children listen to stories told by the researcher and are asked to tell the stories back to the researcher whereas in story telling or generation task children tell a story while looking at a wordless picture book (Vandewalle et al., 2012).

1.5 Narrative macrostructure: DLD vs TD children

There are a lot of studies comparing the narrative skills of DLD children to those of children with TD. Yet, the results seem to be conflicting for narrative macrostructure. Some studies have found children with TD to obtain higher story grammar scores or include more narrative content, that is, more story grammar components, producing more coherent stories (Mäkinen et al., 2014; Norbury et al., 2016; Kunnari et al., 2016; Mazlan et al., 2024), whereas, other studies have not found macrostructure to differentiate TD from DLD groups (Tsimpli et al., 2016). The conflicting findings are attributed, in part, to methodological differences, and more specifically whether a story retell or a story generation task was used, with story generation being a more difficult task (Schneider et al., 2005).

Several studies found significantly higher scores of TD children compared to children with DLD at the macrostructure level in the retelling task or story telling task. More specifically, in a study with Croatian-speaking monolinguals it was found that children with TD outperformed those with DLD at the macrostructure level in both conditions of story retelling and storytelling (Kraljević et al., 2020). The stories produced by children with DLD were shorter and they were generally assessed as more modest in that they lacked important structural components, such as the problem of the story. The study by Sheng et al. (2020) in Mandarin- speaking children with TD and those at risk for DLD found a difference between the two groups on story structure in narratives elicited in the retelling mode, but greater difference between the groups in the story telling mode. Also, the grammaticality and productivity of DLD children were relatively preserved but story macrostructure, lexical diversity, and sentence complexity were vulnerable. Another study with Mandarin-speaking children (Torng and Sah, 2020) revealed that the narratives of children with DLD included significantly less story grammar components, less evaluative comments and were less coherent than those of TD controls. In addition, Xue et al. (2022) tried to capture the features of narratives for school-aged Mandarin-speaking children with SLI. The results revealed that across grades, for macrostructure, children with SLI lagged behind TD children in narrative pattern scores. Furthermore, Andreou and Lemoni (2020), in their systematic review on the narrative skills of monolingual and bilingual pre-school and primary school children with DLD, reported significant differences in the narrative performance between monolinguals with and without DLD and between bilinguals with and without DLD.

In addition, the studies by Otwinowska et al. (2020), and Wehmeier (2019) found significantly higher scores between TD and DLD children in retelling with pictures than in storytelling. In another study, Altman et al. (2024) examined the role of narrative microstructure (production of words and sentences) and narrative

macrostructure (organization of events) in the use of Internal State Terms (ISTs) in narratives of bilingual children with developmental language disorder (DLD) in their school language. The results revealed that at the macrostructure level children with DLD performed weaker in six out of the seven story grammar elements in their narratives than bilinguals with TD. For Internal State Terms (ISTs) and macrostructure, bilinguals with DLD produced fewer linguistic ISTs in the story structure component of Attempts than their peers with TD.

In their research Lin et al. (2024) tried to shed light in the relationship between Executive Functions in a daily life context and performance on two narrative tasks of Mandarin-speaking preschoolers with DLD and theirTD controls. The subjects completed a story generation and a story recall task. The results showed the TD group outperformed the DLD group on narrative macrostructure and microstructure. In another study Lin et al. (2024) the differences in narrative abilities of Malay-speaking school-age children with and without DLD are examined. TD children outperformed children with DLD in both narrative production and comprehension with TD ones constructing a higher combination of Goals, Attempts, and Outcomes components than children with DLD.

On the other hand, Tsimpli et al. (2016) in their research on narrative production in monolingual and bilingual children with Specific Language Impairment (SLI) found that bilingual children with SLI were found to attain similar levels of performance, and even to outperform monolingual children with SLI in macrostructure yet, there were differences between TD children and children with SLI in microstructure. Roch et al. (2016) found a significant, but relatively small difference in the story structure score, with higher scores in retelling. In another study, Soodla and Kikas (2010) found no consistent difference in marking all story structure components among 6- to 8-year-old children with typical and delayed language development.

Also, in another study Altman et al. (2016) investigated the macrostructure, microstructure, and Internal State Terms in the narratives of English–Hebrew bilingual preschool children with and without SLI. The macrostructure results showed similar performance in both languages for children with TLD and those diagnosed with SLI.

There were neither group nor language differences regarding Goals, Attempts, Outcomes (GAO) proportion and GAO per episode. Yet, an analysis of ISTs revealed more ISTs in children's L2, in particular, more mental verbs, especially early acquired perceptual and motivational verbs such as "see" and "want."

There is not a lot of research in the Greek language on the narrative skills of children with Developmental Language Disorder (DLD) since most research is in the English language, which is considered a language with limited morphological grammar (Haspelmath and Sims, 2010). Conducting this research in the Greek language is very important, as it is a highly declinable language with rich morphology, which comprises eleven parts of speech, out of which six are declinable. These come in a great variety of morphological forms, as the language distinguishes a large number of regular declension categories for nouns, adjectives, and verbs (Baldzis et al., 2005).

Based on the above, the aim of the present study is to study the narrative skills of children with Developmental Language Disorder (DLD) at the macrostructural level. It is expected that the narrative abilities of children with DLD will be more affected at the level of macrostructure than those of children with Typical Development (TD). More specifically, the research hypotheses of the study are the following: (a) Children with DLD will present lower performance than TD children in the parameter of story structure, (b) children with DLD are expected to present lower performance than TD children in the parameter of structural complexity, (c) children with DLD will present lower performance than TD children in the parameter of expressing Internal State Terms (ISTs) and (d) there will be a significant dependence between the groups (children with DLD and TD children) and the subcategories of the story structure.

2 Materials and methods

2.1 Participants

A total of 100 pre-school and first school age children 5 to 11 years old participated in the study, who were matched on chronological age, socioeconomic status and language input received at home, according to the answers given in the questionnaire administered to the whole of the sample. For each DLD child a TD child was selected from the same school and area. The experimental group included 50 children with Developmental Language Disorder (DLD) while the control group included 50 participants with Typical Development (TD). All participants were monolingual Greek language speakers. Testing was conducted in the cities of Volos, Athens and Thessaloniki.

The children from the experimental group were chosen based on their diagnosis from KE. D. A. S. Y, which are support centers for the Diagnosis, Assessment, and Counseling for people with special educational needs under the supervision of the Ministry of Education. Another inclusion criterion was the speech and language pathologists' diagnostic reports from six institutions in which the participants underwent language therapy. The exclusion criteria for establishing this diagnosis were the presence of a cognitive disability and/or hearing impairment. Children with DLD, according to the details given in their official diagnoses, had deficits in one or more language domains namely expressive language (e.g., vocabulary, grammar), receptive language (e.g., understanding instructions) morphosyntax, semantics and discourse.

For the selection of the children consisting the TD group, the following criteria were taken into account: (a) no language difficulties reported by parents, teachers, or clinicians, (b) no history of speech-language therapy, no cognitive, neurological, or psychiatric disorders, (c) normal hearing, (d) age-appropriate performance both at school as well as in their daily communication and (e) their performance in the language tests administered, as they described below, which was within the average range, according to the cut off percentile score given for each test. More specifically, the scores obtained for each of the tests given were: (a) Raven's Progressive Matrices test (CPM): ≥ 25o percentile, (b) Raven's Vocabulary Scales- Crichton Vocabulary Scales (CVS): ≥ 10o percentile (c) the Greek version Test of expressive vocabulary: ≥ 25o percentile and (d) the Action Pictures: informational and grammatical proficiency test: ≥ 10o percentile.

2.2 Instruments

The tests administered for the selection of the TD group were: (a) Raven's Progressive Matrices test (CPM) (Raven, 2015; Sideridis et al., 2015), (b) Raven's Vocabulary Scales- Crichton Vocabulary Scales (CVS) (Sideridis et al., 2015), (c) the Greek version of the Word Finding

Vocabulary Test (Renfrew, 1995) "Test of expressive vocabulary" (Vogindroukas et al., 2009a) and (d) the "Action Pictures: informational and grammatical proficiency test" (Vogindroukas et al., 2009b).

Raven's Progressive Matrices (CPM) measure the individual's nonverbal ability to draw inferences in a visuospatial context. The Crichton Vocabulary Scales (CVS) assess the individual's verbal ability, which is related to the familiarity that a person has with specific concepts and verbal information. The combination of the results from the use of the two scales (CPM and CVS) is indicated for the most comprehensive assessment of general cognitive ability. The "Test of expressive vocabulary" is a reliable language assessment tool, which accurately measures the lexical abilities of children aged 4–8 whereas the "Action Pictures: informational and grammatical proficiency test" is a diagnostic tool which focuses on two areas of language, morphosyntax and pragmatics, during linguistic expression, offering specific information on the child's strengths and weaknesses in those language domains.

The main instrument of the study, which was used to assess both groups in story retelling with pictures, was Multilingual Assessment Instrument for Narratives (MAIN) (Gagarina et al., 2019). This tool can be used to assess *narrative comprehension and production* in children between 3 and 11 years of age. The parameters measured by the test, which measures retelling with pictures and storytelling with pictures are *story structure, structural complexity and Internal State Terms*. In our study we measured only the narrative production of the children and, based on the protocol in the production section, the highest score for the story structure parameter is 17 points, the highest score for the structural complexity is 15 points and one point is awarded for each Internal State Term (IST). The total number of IST in tokens is calculated. The list of suggested ISTs is long and is drawn from the following categories: Perceptual state terms, e.g., see, hear, feel, smell; Physiological state terms, e.g., thirsty, hungry, tired, sore, hurt(ing); Consciousness terms, e.g., alive, awake, asleep; Emotion terms, e.g., sad, happy, glad, angry, worried, disappointed; afraid, scared, proud, brave, (feel)safe, pleased, surprised; Mental verbs, e.g., want, think, know, forget, decide, believe, wonder, have/ make a plan; Linguistic verbs/ verbs of saying/ telling, e.g., say, call, shout, warn, ask.

The highest total points in all three categories indicate better performance. An experimental design was followed and the sample of children with DLD of preschool and early school age was compared with that of TD at the macrostructural level. The subcategories for story structure are the following: A1 setting, A2 IST initiating event, A3 Goal, A4 Attempt, A5 Outcome, A6 IST reaction, A7 IST initiating event, A8 Goal, A9 Attempt, A10 Outcome, A11 IST reaction, A12 IST initiating event, A13 Goal, A14 Attempt, A15 Outcome, A16 IST reaction. Each component is awarded 0 or 1 points except for setting that is awarded 0, 1 or 2 points.

The subcategories for structural complexity are the following: Number of attempt-outcome sequences (maximum 3 points), Number of single Goals-without Attempt or Outcome (maximum 3 points), Number of Goals- Attempts/Goals-Outcomes sequences (maximum 6 points), Number of Goals-Attempts-Outcomes sequences (maximum 3 points).

2.3 Procedure

Before the beginning of the research process, parents and teachers of the children of the two groups were informed on the content of the research, the data collection tools and the method of recording the data. A statement signed by the parents of both TD and DLD children was obtained, in which it was stated clearly that participation in the research was not mandatory and that they could withdraw at any time they wished during the research process.

Regarding the administration of the MAIN, each participant was presented with one story to tell and one story to retell but in this article only the results of the story retelling are presented. The testing was conducted in quiet classrooms, or the library of the children's schools and participants were told that they had to choose among four different stories, although the stories were the same. In such a way a condition of an unshared context was created, in which the participant was convinced that the examiner does not know which story will be presented nor the content of the stories. During testing, the examiner was not allowed to give prompting questions that could affect the content and structure of the participants' performance in the story retelling process. Each participant produced two stories that were recorded and transcribed. Transcription and coding were carried out by the researcher and by a monolingual speaker of Greek who had undergone special training for coding.

Word-by-word transcription for each of the samples indicated at least 94% agreement with the corresponding original. *TurboScribe* was used which is an online tool that was used to convert audio files into accurate text in seconds. Also, *oTranscribe* was used, a free online tool that makes interview transcription easy. Transcripts were used for evaluating story structure. All stories produced by children were analyzed using the scoring protocol for analysis at a macrolevel, which was developed and provided with the test materials. The scoring sheet developed for use with MAIN contains a list of structural components for each episode, as well as examples of each component.

2.4 Data analysis

For the statistical analysis IBM SPSS Statistics 29.00.00 software was used. The normality of the distribution for the retelling with pictures variable was tested using the Kolmogorov–Smirnov test. This test showed that the variables related to story structure performance ($Z = 0.150$, $p < 0.001$), structural complexity performance ($Z = 0.185$, $p < 0.001$) and internal state terms ($Z = 0.217$, $p < 0.001$) deviate significantly from the normal distribution. Therefore, Mann–Whitney U test was used to compare the two groups (DLD vs. TD). It is a non-parametric alternative test to the independent sample t-test that compares two sample means from the same population and tests whether they are equal. Researchers usually use the Mann–Whitney U test when they have ordinal data or when they cannot meet the assumptions of the t-test (Nachar, 2008). Chi-square test of independence was used to examine if there is a significant dependence between the groups of children (DLD vs. TD) in the subcategories of the story structure. All comparisons were made at a significance level of 5%.

3 Results

The demographic characteristics of the children who participated in this study (50 children with TD and 50 children with DLD) are presented in Table 1. The majority of children in the DLD group were

boys (n = 31, 62%), whereas the TD group had an equal number of boys and girls (50% each). Both groups had comparable mean ages (7.7 years for the DLD group and 7.8 years for the TD group). Additionally, 80% (n = 20) of the participants in the DLD group were school-aged children, while 78% (n = 39) of those in the TD group were school-aged children. The majority of children with DLD 84% (n = 37) and those with typical development 78% (n = 39) resided in urban areas. Finally, 90% (n = 45) of the children in the DLD group had received some form of intervention (speech therapy), compared to only 22% (n = 11) in the TD group (Table 1).

The results of Mann–Whitney U test indicate that there is a significant difference between DLD and TD students in story structure performance (U = 537.0, p < 0.001). (Table 2). The findings show that TD children outperform those with DLD in story structure in the retelling with pictures task (DLD group: M = 9.3, Md = 9.0, SD = 1.6; TD group: M = 11.1, Md = 11.0, SD = 1.7) (Figures 1, 2).

Also, the results indicate that there is a significant difference between DLD and TD students in structural complexity performance (U = 834.5, p = 0.004). The findings show that TD students have a higher level of skills in structural complexity in the retelling with pictures task than those of students with DLD (DLD group: M = 4.0, Md = 3.0, SD = 2.6; TD group: M = 5.8, Md = 6.0, SD = 3.3) (Figure 3).

Moreover, the results of the Mann–Whitney U test indicate that there is a significant difference between DLD and TD students in Internal State Terms (U = 718.0, p = 0.001). The findings show that TD students presented a higher performance in Internal State Terms compared to that of DLD students (DLD group: M = 2.5, Md = 3.0, SD = 1.1; TD group: M = 3.5, Md = 3.0, SD = 1.2) (Figure 4).

In Table 3 the frequencies and % for story picture items in the two groups of children are presented. From the chi-square test of independence, a significant dependence was presented between the group of children (DLD vs. TD) and performance in the following categories: A4: Attempt [$\chi^2(1) = 8.306$, $p = 0.004$], A6: IST as reaction [$\chi^2(1) = 4.762$, $p = 0.029$], A9:attempt [$\chi^2(1) = 4.320$, $p = 0.038$], A11: IST as reaction [$\chi^2(1) = 5.797$, $p = 0.016$] and A15: outcome [$\chi^2(1) = 7.527$, $p = 0.006$]. In these categoriesTD children achieved a score of 1 (correct response) at a significantly greater proportion compared to DLD ones.

The results of our study showed that TD children scored higher than DLD ones in all subcategories of the story structure, however, only in the components of A4: Attempt, A6: IST as reaction, A9: attempt, A11: IST as reaction and A15: outcome the results were statistically significant.

Specifically, in the category A4, 96% of TD students achieved a score of 1while the corresponding percentage for DLD children was 76%. Similarly, in the category A6, 24% of typically developing students achieved a score of 1, compared to 8% of DLD students. In the category A9, 84% of typically developing students achieved a score of 1, while the corresponding percentage for DLD children was 66%. In the category A11, 58% of typically developing students achieved a score of 1, whereas 34% of DLD students did. Finally, in the category A15, 100% of typically developing students achieved a score of 1, compared to 86% of DLD students.

TABLE 1 Sample demographics.

Individual -level variables		Group			
		DLD		TD	
		n	%	*n*	%
Gender	Boy	31	62%	25	50%
	Girl	19	38%	25	50%
Age (in years)	M (SD)	7.7 (2.0)		7.8 (2.0)	
	Range: Min-Max	5–11		4–11	
Level of education	Preschool	10	20%	11	22%
	School	40	80%	39	78%
Residential area	Urban	37	84%	39	78%
	Semi-urban	13	16%	11	22%
	Rural	0	0%	1	2%
Intervention	Yes	45	90%	11	22%
	No	5	10%	39	78%

4 Discussion

The present study compared the narrative performance of a group of children with DLD and children with TD and the aim of this study was to determine whether these two groups differ in their ability to structure a story in the condition of a story retelling task with pictures. It was expected that the narrative abilities of children with DLD would be more affected at the level of macrostructure than those of children with Typical Development (TD).

More specifically, our first hypothesis was that children with DLD will present lower performance than TD children in the parameter of story structure. According to the results of the study, children with DLD had a weaker performance than that of children with TD in the story structure components and therefore our hypothesis is confirmed. Our findings for Greek speaking children with DLD confirm those for children with DLD, who are speakers of different languages.

More specifically, the findings of the present study agree with Blom and Boerma (2016) who found that the Language Impaired group performed weaker than the TD group in all the stages of their assessment. In particular, the two groups were assessed at wave/stage 1 in story comprehension and production and one year later they were assessed at wave 2. At wave 1, the LI group performed weaker than the TD group in both tasks and at wave 2 the groups performed similarly on story comprehensionbut on story generation, the TD group still outperformed the LI group. Also, our study is consistent with Boerma et al. (2016) who found that Dutch-speaking children with LI produced fewer story structure elements and expressed a smaller number of Internal State Terms than children with TD. It is worth mentioning that the language impaired group scored lower than the TD group on all measures (grammar, grammatical complexity (mean length of utterance), verbal short-term and working memory, and sustained attention) except expressive vocabulary.

Also, our findings accord with those of Kraljević et al. (2020) with Croatian-speaking monolinguals, which showed that the stories produced by children with DLD were shorter and were generally assessed as more modest as they lacked important structural components, such as the problem of the story. Pham et al. (2019), in his study with Vietnamese-speaking monolinguals, also found weaker

TABLE 2 Mean scores for retelling with pictures in story structure, structural complexity and internal state terms of TD and DLD and children.

Macrostructure elements	DLD (n = 50)				TD (n = 50)				U	p
	M	SD	Md	IQR	M	SD	Md	IQR		
Story structure	9.3	1.6	9.0	3.0	11.1	1.7	11.0	2.0	537.0	0.001*
Structural complexity	4.0	2.6	3.0	4.0	5.8	3.3	6.0	4.0	834.5	0.004*
Internal state terms	2.5	1.1	3.0	1.0	3.5	1.2	3.0	1.0	718.0	0.001*

M, mean; SD, standard deviation; Md, median; IQR, interquartile Range.
*Statistically significant result.

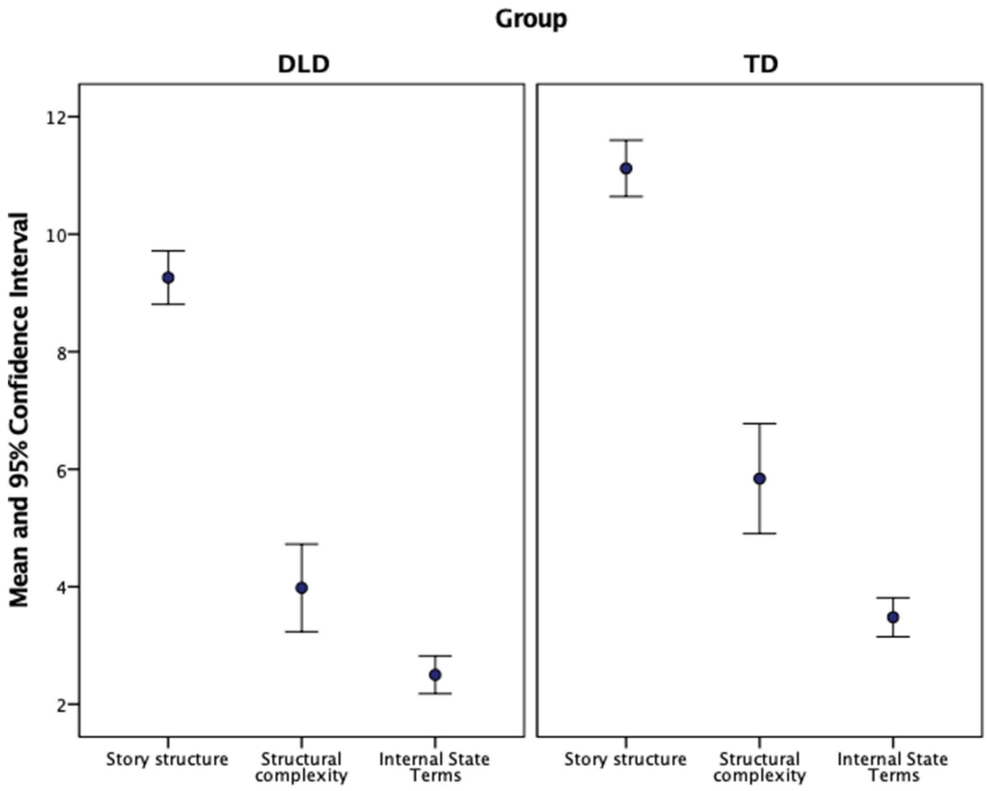


FIGURE 1
DLD and TD performance in the subcategories of story structure, structural complexity and ISTs parameters.

performance on the part of DLD children in narrative macrostructure as compared to DLD ones.

Andreou and Lemoni (2020), in their systematic review on the narrative skills of monolingual and bilingual pre-school and primary school children with DLD, reported significant differences in the narrative performance between monolinguals with and without DLD and between bilinguals with and without DLD. Also, our findings agree with those of Sheng et al. (2020) in Mandarin- speaking children with TD and those at risk for DLD who found better performance in story-retell than story-tell on measures of overall story structure and percentage of complex clauses. The grammaticality and productivity of DLD children were relatively preserved but story macrostructure, lexical diversity, and sentence complexity were vulnerable.

The second hypothesis of our study is that children with DLD will present lower performance than TD children in the parameter of structural complexity. The above hypothesis is confirmed since the findings of our study showed that children with DLD performed weaker than TD children in this parameter. Our findings agree with those of Sheng et al. (2020) who found a marginal difference in the structural complexity scores and a significant difference in overall story structure scores.

Yet, our findings do not agree with those of Tsimpli et al. (2016) who found no significant difference either between TD monolingual children and monolingual children with SLI or between TD bilingual and bilingual children with SLI. Also, they do not agree with Altman et al. (2016) who found neither group nor language differences regarding GAO proportion and GAO per episode. These studies did not use the story structure score but instead analyzed a score for story complexity for two narratives combined (Tsimpli et al., 2016) or counted only goals, attempts and outcomes in the narratives, respectively, (Altman et al., 2016). Although macrostructure results revealed similar performance in both languages for children with TD and those with SLI, microstructure analysis of verbal productivity, length of communication units, and lexical diversity distinguished

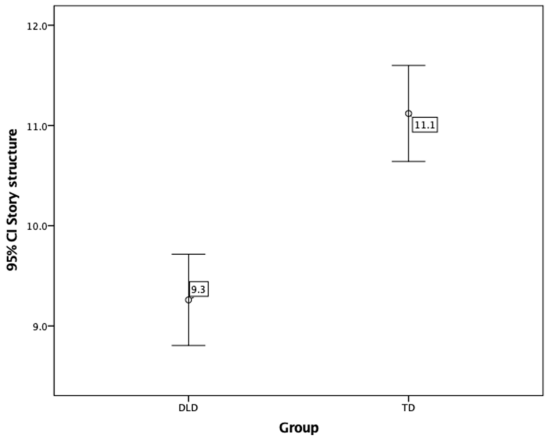


FIGURE 2
DLD and TD performance in the subcategory of story structure.

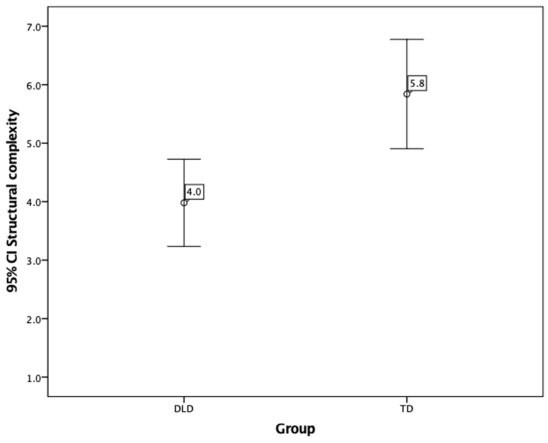


FIGURE 3
DLD and TD performance in the subcategory of structural complexity.

children with TD from those with SLI. The difference in the results

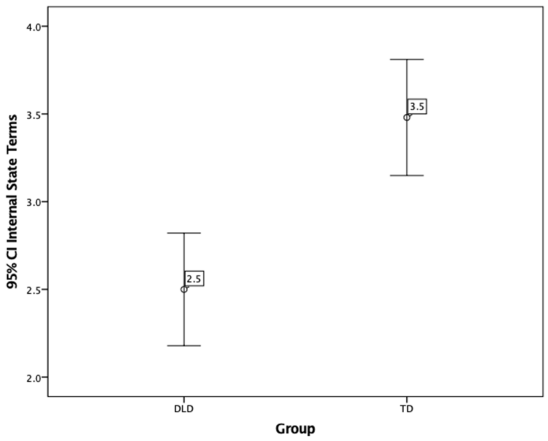


FIGURE 4
DLD and TD performance in the subcategory of ISTs.

TABLE 3 Frequencies and % for story picture items for DLD and TD children.

Story structure parameters		Group				χ^2	*P*
		DLD		TD			
		n	%	*n*	%		
A1 Setting	0	17	34.0%	11	22.0%	3.514	0.173
	1	33	66.0%	37	74.0%		
	2	0	0.0%	2	4.0%		
A2 IST initiating event	0	14	28.0%	8	16.0%	2.098	0.148
	1	36	72.0%	42	84.0%		
A3 Goal	0	32	64.0%	23	46.0%	3.273	0.070
	1	18	36.0%	27	54.0%		
A4 Attempt	0	12	24.0%	2	4.0%	8.306	0.004*
	1	38	76.0%	48	96.0%		
A5 Outcome	0	6	12.2%	3	6.0%	1.168	0.280
	1	43	87.8%	47	94.0%		
A6 IST reaction	0	46	92.0%	38	76.0%	4.762	0.029*
	1	4	8.0%	12	24.0%		
A7 IST initiating event	0	12	24.0%	9	18.0%	0.541	0.461
	1	38	76.0%	41	82.0%		
A8 Goal	0	40	80.0%	40	80.0%	0.000	1.000
	1	10	20.0%	10	20.0%		
A9 Attempt	0	17	34.0%	8	16.0%	4.320	0.038*
	1	33	66.0%	42	84.0%		
A10 Outcome	0	5	10.0%	1	2.0%	2.837	0.092
	1	45	90.0%	49	98.0%		
A11 IST reaction	0	33	66.0%	21	42.0%	5.797	0.016*
	1	17	34.0%	29	58.0%		
A12 IST initiating event	0	10	20.0%	10	20.0%	0.000	1.000
	1	40	80.0%	40	80.0%		
A13 Goal	0	31	62.0%	32	64.0%	0.043	0.836
	1	19	38.0%	18	36.0%		
A14 Attempt	0	28	56.0%	20	40.0%	2.564	0.109
	1	22	44.0%	30	60.0%		
A15 Outcome	0	7	14.0%	0	0.0%	7.527	0.006*
	1	43	86.0%	50	100.0%		
A16 IST reaction	0	27	54.0%	18	36.0%	3.273	0.070
	1	23	46.0%	32	64.0%		

*Statistically significant result.

between our study and the aforementioned ones could be attributed to the fact that in our study structural complexity is measured as a part of a whole scoring pattern including the full range of story grammar elements and Internal State Terms whereas the previous studies used only the Goal-Attempt-Outcome sequence score. In other words, this discrepancy in the findings comes as a result of the different methodology and scoring between our study and the ones

by Tsimpli et al. (2016) and Altman et al. (2016). Also, the participants of the two previous studies consisted of bilinguals while our study includes monolinguals.

The third hypothesis of this study is that children with DLD will present lower performance than TD children in the parameter of expressing Internal State Terms (ISTs). The findings of our study showed that children with DLD performed weaker than TD children in expressing Internal State Terms in the categories of perception/cognition, desires, intentions, consciousness, emotions, mentality, decisions and language, so our hypothesis is confirmed.

Our findings agree with the studies of (Boerma et al., 2016; Greenhalgh and Strong, 2001; Johnston et al., 2001) which revealed delays in understanding and producing ISTs among children with DLD. Also, our research aligns with two other studies that examined the causal relations in bilingual children with DLD showing that they use fewer causal relations in their narratives (Fichman et al., 2017; Kupersmitt and Armon-Lotem, 2019). Both reported the difficulty of children with DLD in producing causal relations, which are linked to the ability to create inferences about characters' intentions and mental states. Also, our findings are consistent with the study of Boerma et al., (2016) with bilinguals which showed that DLD children speaking Dutch as L2 used fewer ISTs than TD children. Our results also agree with those of Tsimpli et al. (2016) who collected narratives from children speaking Greek as L2 and reported that children with DLD used fewer ISTs terms than TD children. Moreover, our research aligns with the research of Altman et al. (2024), who examined the role of narrative microstructure and narrative macrostructure in the use of Internal State Terms (ISTs) in narratives of bilingual children with developmental language disorder (DLD) in their school language (SL). In terms of ISTs and macrostructure, bilinguals with DLD produced fewer linguistic ISTs in Attempts than their peers with TD. According to De Villiers, (2007) children with DLD usually demonstrate grammatical difficulties, poor vocabulary knowledge or deficient syntactic skills which may influence the production of Internal State verbs which require more complex complements.

The fourth hypothesis is that there will be significant dependence between the groups of children (DLD vs. TD) in the subcategories of the story structure. The results of our study showed that there is a relatively significant dependence between the groups of children on the subcategories of the story structure and as a result our hypothesis is partially confirmed. In particular, TD children scored higher than DLD ones in all subcategories of the story structure, however, only in the components of IST reaction, attempt and partially outcome the results were statistically significant (attempt and IST reaction of the first episode, the attempt and IST reaction of the second episode and the outcome in the third episode). In these categories, TD children achieved a score of 1 (correct response) at a significantly greater proportion compared to DLD ones. Considering some further explanation why the majority of components did not yield differences makes us skeptical of the sample size. The small size of the sample might not have enough power to detect a difference even if it exists. Also, there may have been a ceiling effect which occurs because the measure used is easy and it has an upper limit, causing many participants to achieve the highest possible score.

It is expected that TD children will be able to mark the situation, the initiating event, the goal and all attempts and consequences (i.e., outcome) in the year prior to entering school around five years old and to progress more quickly in this developmental pattern of narrative competence than children with DLD (To et al., 2010). Yet, in our study DLD children had a better performance in the individual components of the story structure than expected. A possible explanation for this result could be that having an adult model benefited both groups in sentence complexity and story macrostructure and potentially helped maintain the performance in TD children (Sheng et al., 2020).

In addition, DLD children might not have been relied exclusively on their linguistic competence but they might have been affected from the exposure to a prior audiovisual model, in other words the picture retelling task, which seems to have benefited both groups. Previous research has shown that visual elicitation decreases processing demands and facilitates the process of recalling information (Kraljević et al., 2020).

Our findings are partially in line with the study of Kraljević et al. (2020) who showed that in the retelling task children with TD more frequently marked all parts of the story (except reaction) than children with DLD. IST Reaction seems to represent the part of the episode that relates to the expression of the feelings and attitudes of the story characters. Therefore, DLD children may exhibit a lower level of empathy and emotional regulation than their peers with TD (Kraljević et al., 2020). These results suggest that TD children can not only produce a story with a well-formed structure but are also more likely to express their inner feelings about the events. On the other hand, the expression of their feelings about the events seems to be a weakness for DLD children who managed to do well in other story structure elements.

Also, our findings are partially consistent with Kraljević et al. (2020) who found a difference in their research in the component of outcome (77% compared to 51%) between TD and DLD groups. In our study there is a difference in the outcome component only in the third episode (100% compared to 86%) between TD and DLD groups. A possible explanation for this difference in the outcome component could be the influence of the presented model of the story which facilitates the process of recalling information.

Regarding the pedagogical implications of our study it is evident that this research comes to fill in the gap that exists in the Greek literature, but also to validate and expand the research data of the international literature, since in Greece few studies have been conducted on the narrative skills of children with DLD with most of them focusing on both monolinguals and bilinguals (Tsimpli et al., 2016) or ASD (Peristeri et al., 2017). Moreover, most research studies concern preschool age children, with a significant lack of research concerning school age children or combining both.

5 Limitations and directions for future research

Our research provided valuable results but is subject to some limitations. One of these is that both DLD and TD children produced shorter narratives than expected not only in the lower but in the higher grades as well. As literature supports the majority of children between the 5th-6th years of age are capable of constructing fully formed narratives and with many story-structure components as age increases from 3 to 9 years old (Khan et al., 2016). At that age, a hierarchical increase of the Mean Length of Utterance (MLU) across the different age groups is observed (Safwat et al., 2013).

A possible interpretation for the short length of their narrations could have been their anxiety or fatigue during their assessment. According to previous research, when children narrate they say more and make longer narrations if they do not see the interaction as a test (McCabe and Rollins,

1994). On the other hand, fatigue sometimes may be the reason for low performance in narrative assessment or misinterpreted as language impairment (Peña et al., 2006). This might have affected the representativeness of the narrative measures used. Therefore, the specific narrative measures need to be investigated through longer narratives in future studies in order to confirm the present findings.

Also, there is a need for further research with larger sample size in order to fully confirm previous research in the field and elucidate the specific difficulties the children with DLD face in the domain of narrative skills in the Greek language. Furthermore, another limitation derives from the fact that no results of microstructure analysis are included in this paper. Such results could have provided a more holistic profile of narrative competence in DLD and strengthened the discussion of linguistic complexity.

Moreover, concerning task effects, the results from previous studies indicate that researchers need to be cautious when using the different stories of MAIN. These stories which are intended to be parallel both in their macrostructure and in their comprehension questions, may not be completely comparable (Lindgren, 2022). Maybe it would be interesting to repeat our study one or two years later to check if the narration capability of DLD and TD children continues to develop gradually. Obviously, story-retell constitutes a valuable form of narrative assessment and should be further investigated in future larger scale studies. The empirical findings of this study aim to broaden the scope of the existing research on children with DLD indicating deficits in their narrative skills. Additionally, they can lead to the creation of educational interventions based on storytelling aiming to improve the language skills of children with DLD.

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 humans were approved by the University of Thessaly ethics committee. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

GA: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. GL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Aguilar-Mediavilla, E., Sanz-Torrent, M., and Serra-Raventós, M. (2002). A comparative study of the phonology of pre-school children with specific language impairment (SLI), language delay (LD) and Normal acquisition. *Clin. Linguist. Phon.* 16, 573–596. doi: 10.1080/02699200210148394

Altman, C., Armon-Lotem, S., Fichman, S., and Walters, J. (2016). Macrostructure, microstructure, and mental state terms in the narratives of English–Hebrew bilingual preschool children with and without specific language impairment. *Appl. Psycholinguist.* 37, 165–193. doi: 10.1017/S0142716415000466

Altman, C., Fichman, S., Perry, N., Osher, P., and Walters, J. (2024). Internal state terms in the narratives of bilingual children with developmental language disorder: the role of microstructure and macrostructure. *Lang. Speech Hear. Serv. Sch.* 55, 1039–1053. doi: 10.1044/2024_LSHSS-23-00170

Andreou, G., and Lemoni, G. (2020). Narrative skills of monolingual and bilingual pre-school and primary school children with developmental language disorder (DLD): a systematic review. *Open J. Mod. Linguist.* 10, 429–458. doi: 10.4236/ojml.2020.105026

Baldzis, S. D., Kolalas, S., and Eumeridou, E. (2005). Computational modern Greek morphological lexicon – an efficient and comprehensive system for morphological analysis and synthesis. *Literary Linguist. Comput.* 20, 153–187.

Befi-Lopes, D. M., Bento, A. C. P., and Perissinoto, J. (2008). Narração de histórias por crianças com distúrbioespecífico de linguagem. *Pró-Fono Rev. Atual. Cient.* 20, 93–98. doi: 10.1590/S0104-56872008000200004

Bishop, D. V., 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

Blom, E., and Boerma, T. (2016). Why do children with language impairment have difficulties with narrative macrostructure? *Res. Dev. Disabil.* 55, 301–311. doi: 10.1016/j.ridd.2016.05.001

Botting, N. (2010). It's not (just) what you do, but the way that you do it: factors that determine narrative ability in atypical language learners. *Dev. Med. Child Neurol.* 52, 886–887. doi: 10.1111/j.1469-8749.2010.03651.x

Botting, N., Faragher, B., Simkin, Z., Knox, E., and Conti-Ramsden, G. (2001). Predicting pathways of specific language impairment: what differentiates good and poor outcome? *J. Child Psychol. Psychiatry* 42, 1013–1020. doi: 10.1111/1469-7610.00799

Boerma, T., Leseman, P., Timmermeister, M., Wijnen, F., and Blom, E. (2016). Narrative abilities of monolingual and bilingual children with and without language

impairment: implications for clinical practice. *International Journal of Language and Communication Disorders*, 51, 626–638. doi: 10.1111/1460-6984.12234

Burris, S. E., and Brown, D. D. (2014). When all children comprehend: increasing the external validity of narrative comprehension development research. *Front. Psychol.* 5:168. doi: 10.3389/fpsyg.2014.00168

Christensen, R. V. (2019). Grammatical characteristics of narratives told by Danish children with developmental language disorder. *Int. J. Lang. Commun.* 49, 130–148.

Cleave, P. L., Girolametto, L., Chen, X., and Johnson, C. J. (2010). Narrative abilities in monolingual and dual language learning children with specific language impairment. *J. Commun. Disord.* 43, 511–522. doi: 10.1016/j.jcomdis.2010.05.005

Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. B., and Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *J. Speech Lang. Hear. Res.* 47, 1301–1318. doi: 10.1044/1092-4388(2004/098)

Fiestas, C. E., and Peña, E. D. (2004). Narrative discourse in bilingual children: language and task effects. *Lang. Speech Hear. Serv. Sch.* 35, 155–168. doi: 10.1044/0161-1461(2004/016)

Fichman, S., Altman, C., Voloskovich, A., Armon-Lotem, S., and Walters, J. (2017). Story grammar elements and causal relations in the narratives of Russian–Hebrew bilingual children with SLI and typical language development. *Journal of Communication Disorders*, 69, 72–93.

Fioretti, C., Pascuzzi, D., and Smorti, A. (2019). Narrative and narrativization of a journey: Differences between personal and fictional narratives. *The Open Psychology Journal*, 12, 205–215. Available at: https://psycnet.apa.org/doi/10.2174/1874350101912010205

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U., et al. (2019). MAIN: multilingual assessment instrument for narratives - revised. *ZAS Papers Linguistics* 63, 1–39. doi: 10.21248/zaspil.63.2019.516

Gillam, R. B., and Pearson, N. (2004). Test of narrative language. Austin, TX: Pro-Ed.

Govindarajan, K., and Paradis, J. (2022). Narrative macrostructure and microstructure profiles of bilingual children with autism spectrum disorder: differentiation from bilingual children with developmental language disorder and typical development. *Appl. Psycholinguist.* 43, 1359–1390. doi: 10.1017/S0142716422000406

Greenhalgh, K. S., and Strong, C. J. (2001). Literate Language Features in Spoken Narratives of Children With Typical Language and Children With Language Impairments. *Language, Speech, and Hearing Services in Schools*, 32, 114–125. doi: 10.1044/0161-1461(2001/010)

Guo, L. Y., Tomblin, J. B., and Samelson, V. (2008). Speech disruptions in the narratives of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 51, 722–738.

Haspelmath, M., and Sims, A. (2010). Understanding morphology. *2nd* Edn. London, UK: Routledge.

Hayward, D. V., Gillam, R. B., and Lien, P. (2007). Retelling a script-based story: do children with and without language impairments focus on script and story elements? *Am. J. Speech Lang. Pathol.* 16, 235–245. doi: 10.1044/1058-0360(2007/028)

Hutson-Nechkash, P. (2001). Narrative toolbox: Blueprints for storybuilding. Eau Claire, WI: Thinking Publications.

Jones, S. Tutt., "Story generation in five school-aged children with language impairment" (2015). All Theses and Dissertations. 6143. Available at: https://scholarsarchive.byu.edu/etd/6143

Johnston, J. R., Miller, J., and Tallal, P. (2001). Use of cognitive state predicates by language-impaired children. *International journal of language and communication disorders*, 36, 349–370. doi: 10.1080/13682820110045838

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg, S. L., and Gillam, R. B. (2006). The index of narrative microstructure: a clinical tool for analyzing school age children's narrative performance. *Am. J. Speech Lang. Pathol.* 15, 177–191. doi: 10.1044/1058-0360(2006/017)

Khan, K. S., Gugiu, M. R., Justice, L. M., Bowles, R. P., Skibbe, L. E., and Piasta, S. B. (2016). Age-related progressions in story structure in young children's narratives. *J Speech Language Hearing Research* 59, 1395–1408. doi: 10.1044/2016_JSLHR-L-15-0275

Kraljević, J. K., Hržica, G., and Gorup, I. V. (2020). A comparative macrostructural analysis of narrative discourse in children with typical language development and children with developmental language disorder. *Drustv. Istraz.* 29, 453–470. doi: 10.5559/di.29.3.06

Kunnari, S., Välimaa, T., and Laukkanen-Nevala, P. (2016). Macrostructure in the narratives of monolingual Finnish and bilingual Finnish Swedish children. *Appl. Psycholinguist.* 37, 123–144. doi: 10.1017/S0142716415000442

Kupersmitt, J. R., and Armon-Lotem, S. (2019). The linguistic expression of causal relations in picture-based narratives: A comparative study of bilingual and monolingual children with TLD and DLD. *First Language*, 39, 319–343. doi: 10.1177/0142723719831927

Leonard, L. B. (2014). Children with specific language impairment and their contribution to the study of language development. *J. Child Lang.* 41, 38–47. doi: 10.1017/S0305000914000130

Leonard, L. B., and Deevy, P. (2004). "Lexical deficits in specific language impairment" in Classification of developmental language disorders: Theoretical issues and clinical implications. eds. L. Verhoeven and H. Van Balkom (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 209–233.

Lin, Y., Sheng, L., Shi, H., Yan, W., and Zhang, Y. (2024). Narrative generation and narrative recall recruit different executive functions in preschoolers with and without developmental language disorder. *Clin. Linguist. Phon.* 39, 1–20. doi: 10.1080/02699206.2024.2441306

Lindgren, J. (2022). The development of narrative skills in monolingual Swedish-speaking children aged 4 to 9: a longitudinal study. *J. Child Lang.* 49, 1281–1294. doi: 10.1017/S030500092100057X

Lyytinen, H., Erskine, J., Hämäläinen, J., Torppa, M., and Ronimus, M. (2015). Dyslexia-early identification and prevention: highlights from the Jyväskylä longitudinal study of dyslexia. *Curr. Dev. Disord. Rep.* 2, 330–338. doi: 10.1007/s40474-015-0067-1

Mäkinen, L., Loukusa, S., Laukkanen, P., Leinonen, E., and Kunnari, S. (2014). Linguistic and pragmatic aspects of narration in Finnish typically developing children and children with specific language impairment. *Clin. Linguist. Phon.* 28, 413–427. doi: 10.3109/02699206.2013.875592

Mazlan, I. R., Hassnan, N. M., and Ahmad Rusli, Y. (2024). A comparison of narrative abilities in Malay school-age typically developing children and children with developmental language disorder. *Clin. Linguist. Phon.* 39, 233–254. doi: 10.1080/02699206.2024.2359462

McCabe, A., and Rollins, P. R. (1994). Assessment of preschool narrative skills. *Am. J. Speech-Lang. Pathol.* 3, 45–56. doi: 10.1044/1058-0360.0301.45

McGregor, K. K., Newman, R. M., Reilly, R. M., and Capone, N. (2002). Semantic representation and naming in children with specific language impairment. *J. Speech Lang. Hear. Res.* 45, 998–1014. doi: 10.1044/1092-4388(2002/081)

Mpaka, F., Vlassopoulou, M., and Drakos, G. (2012). "Μελέτη του αφηγηματικού λόγου σε παιδιά με ειδική αναπτυξιακή διαταραχή (ΕΑΔ) του λόγου μέσω του Bus Story Test" in 12ο πανελλήνιο συνέδριο Λογοπεδικών- Λογοθεραπευτών: Κλινικές Παρεμβάσεις στις Διαταραχές Λόγου και Επικοινωνίας (Αθήνα: Πανελλήνιος Σύλλογος Λογοπεδικών-Λογοθεραπευτών).

Nachar, N. (2008). The Mann-Whitney U: a test for assessing whether two independent samples come from the same distribution. *Tutorials Quant. Methods Psychol.* 4, 13–20. doi: 10.20982/tqmp.04.1.p013

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* 57, 1247–1257. doi: 10.1111/jcpp.12573

Otwinowska, A., Mieszkowska, K., Białecka-Pikul, M., Opacki, M., and Haman, E. (2020). Retelling a model story improves the narratives of polish-English bilingual children. *Int. J. Bilingual Educ. Bilingualism* 23, 1083–1107. doi: 10.1080/13670050.2018.1434124

Pauls, L. J., and Archibald, L. M. (2021). Cognitive and linguistic effects of narrative-based language intervention in children with developmental language disorder. *Autism Developmental Language Impairments* 6, 23969415211015867. doi: 10.1177/23969415211015867

Pearson, B. Z. (2002). "Narrative competence among monolingual and bilingual school children in Miami" in Language and literacy in bilingual children. eds. D. K. Oller and R. E. Eilers (Clevedon: Multilingual Matters) 135–174.

Peña, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., et al. (2006). Dynamic assessment of school-age children's narrative ability: an experimental investigation of classification accuracy. *J. Speech Language Hearing Research* 49, 1037–1057. doi: 10.1044/1092-4388(2006/074)

Peristeri, E., Andreou, M., and Tsimpli, I. F. M. (2017). Syntactic and story structure complexity in the narratives of high- and low-language ability children with autism Spectrum disorder. *Front. Psychol.* 8:2027. doi: 10.3389/fpsyg.2017.02027

Pham, G. T., Pruitt-Lord, S., Snow, C. E., Nguyen, Y. H. T., Phạm, B., Dao, T. B. T., et al. (2019). Identifying developmental language disorder in Vietnamese children. *J. Speech Lang. Hear. Res.* 62, 1452–1467. doi: 10.1044/2019_JSLHR-L-18-0305

Ralli, A. M., and Sidiropoulou, T. (2012). The importance of retelling for the development of narrative skills of preschool children (in Greek). *ΕπιστήμεςΑγωγής* 1, 13–15.

Raven, C. J., RAVEN'S Educational CPM/CVS. (2015). (Greek version) G. Sideridis, F. Antoniou, A. Mouzaki and Simos. Greek Editions Motivo, Athens.

Redmond, S. M. (2004). Conversational profiles of children with ADHD, SLI and typical development. *Clin. Linguist. Phon.* 18, 107–125. doi: 10.1080/02699200310001611612

Reilly, J., Losh, M., Bellugi, U., and Wulfeck, B. (2004). "Frog, where are you?" narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain Lang.* 88, 229–247. doi: 10.1016/S0093-934X(03)00101-9

Renfrew, C. E. (1995). *Word Finding Vocabulary Test*. Bicester: Winslow.

Reuterskiöld, C., Hansson, K., and Sahlén, B. (2011). Narrative skills in Swedish children with language impairment. *J. Commun. Disord.* 44, 733–744. doi: 10.1016/j.jcomdis.2011.04.010

Roch, M., Florit, E., and Levorato, C. (2016). Narrative competence of Italian–English bilingual children between 5 and 7 years. *Appl. Psycholinguist.* 37, 49–67. doi: 10.1017/S0142716415000417

Safwat, R. F., EL-Dessouky, H. M., Shohdi, S. S., and Hussien, I. A. (2013). Assessment of narrative skills in preschool children. *Egypt. J. Otolaryngol.* 29, 130–135. doi: 10.7123/01.EJO.0000426472.83521.26

Schneider, P., Dubé, R. V., and Hayward, D. (2005). The Edmonton narrative norms instrument. Available online at: www.uab.ca/enni (Accessed April 14, 2025).

Schneider, P., Hayward, D., and Dubé, R. V. (2006). Storytelling from pictures using the Edmonton narrative norms instrument. *J. Speech Lang. Pathol. Audiol.* 30, 224–238.

Sheng, L., and McGregor, K. K. (2010). Lexical-semantic organization in children with specific language impairment. *J. Speech Hear. Res.* 53, 146–159. doi: 10.1044/1092-4388(2009/08-0160)

Sheng, L., Shi, H., Wang, D., Hao, Y., and Zheng, L. (2020). Narrative production in mandarin-speaking children: effects of language ability and elicitation method. *J. Speech Lang. Hear. Res.* 63, 774–792. doi: 10.1044/2019_JSLHR-19-00087

Sideridis, G., Antoniou, F., Mouzaki, A., and Simos, P. (2015) Raven's Educational CPM/CVS: motivo.

Soodla, P., and Kikas, E. (2010). Macrostructure in the narratives of Estonian children with typical development and language impairment. *J. Speech Lang. Hear. Res.* 53, 1321–1333. doi: 10.1044/1092-4388(2010/08-0113)

Theodorou, E., and Grohmann, K. K.. (2010). Narratives in Cypriot Greek mono- and bilingual children with SLI. Botinis, A. Proceedings of the third ISCA tutorial and research workshop on experimental linguistics. Athens, Greece, 185–188.

Theodorou, E., Kambanaros, M., and Grohmann, K. K. (2012). Issues in the diagnosis of SLI in Greek Cypriot bilectal children. *Modern Greek Dialects Linguistics Theory* 5, 463–474.

Thomson, J. (2005). Theme analysis of narratives produced by children with and without specific language impairment. *Clin. Linguist. Phon.* 19, 175–190. doi: 10.1080/02699200410001698616

To, C. K. S., Stokes, S. F., Cheung, H. T., and Tsou, B. K. Y. (2010). Narrative assessment of Cantonese-speaking children. *J. Speech Lang. Hear. Res.* 53, 648–669. doi: 10.1044/1092-4388(2009/08-0039)

Tomas, E., and Vissers, C. (2019). Behind the scenes of developmental language disorder: time to call neuropsychology Back on stage. *Front. Hum. Neurosci.* 12:517. doi: 10.3389/fnhum.2018.00517

Torng, P. C., and Sah, W. H. (2020). Narrative abilities of mandarin-speaking children with and without specific language impairment: macrostructure and microstructure. *Clin. Linguist. Phon.* 34, 453–478. doi: 10.1080/02699206.2019.1655097

Tribushinina, E., Dubinkina, E., and Sanders, T. (2015). Can connective use differentiate between children with and without specific language impairment? *First Lang.* 35, 3–26. doi: 10.1177/0142723714566334

Tsimpli, I. M., Peristeri, E., and Andreou, M. (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Appl. Psycholinguist.* 37, 195–216. doi: 10.1017/S0142716415000478

Vandewalle, E., Boets, B., Boons, T., Ghesquiere, P., and Zink, I. (2012). Oral language and narrative skills in children with specific language impairment with and without literacy delay: a three-year longitudinal study. *Res. Dev. Disabil.* 33, 1857–1870. doi: 10.1016/j.ridd.2012.05.004

Vissers, C., and Koolen, S. (2016). Theory of mind deficits and social emotional functioning in preschoolers with specific language impairment. *Front. Psychol.* 7:1734. doi: 10.3389/fpsyg.2016.01734

Vogindroukas, I., Protopapas, A., and Sideridis, G. (2009a). Expressing vocabulary test (in Greek: Dokimasia ekfrastikou lexilogiou). Chania: Glafki Editions. 1–30.

Vogindroukas, I., Protopapas, A., and Stavrakaki, S. (2009b). Action images. Information and grammar proficiency test (in Greek: Ikones drasis. Dokimasia pliroforiakis Kai ekfrastikis eparkias). Chania: Glafki Editions. 1–33.

Wehmeier, C. M. (2019). Development of narrative macrostructure in monolingual preschoolers in Germany and impact of socio-economic status and home literacy environment. *ZAS Papers Linguistics* 62, 52–75. doi: 10.21248/zaspil.62.2019.443

Wellman, R. L., Lewis, B. A., Freebairn, L. A., Avrich, A. A., Hansen, A. J., and Stein, C. M. (2011). Narrative ability of children with speech sound disorders and the prediction of later literacy skills. *Lang. Speech Hear. Serv. Sch.* 42, 561–579. doi: 10.1044/0161-1461(2011/10-0038)

Xue, J., Zhuo, J., Li, P., Liu, J., and Zhao, J. (2022). Characterizing macro- and micro-structures of narrative skills for mandarin-speaking school-age children with specific language impairment. *J. Commun. Disord.* 96:106199. doi: 10.1016/j.jcomdis.2022.106199

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 28 August 2025
DOI 10.3389/fpsyt.2025.1611074

OPEN ACCESS

EDITED BY
Mingshuang Li,
California State University, Northridge,
United States

REVIEWED BY
Bermudez,
Bebv, Federal University of Paraná, Brazil
Can Xu,
Western Washington University, United States

*CORRESPONDENCE
Reem Elbeltagy
ryalbeltagy@pnu.edu.sa

RECEIVED 17 April 2025
ACCEPTED 04 August 2025
PUBLISHED 28 August 2025

CITATION
Alkahtani R, Bakry H, Al Hawi LM, Alshehri JD and Elbeltagy R (2025) Quality of life of mothers of children with Down syndrome: a study comparing those with and without hearing loss.
Front. Psychiatry 16:1611074.
doi: 10.3389/fpsyt.2025.1611074

Quality of life of mothers of children with Down syndrome: a study comparing those with and without hearing loss

Rania Alkahtani[1], Huny Bakry[2], Lama M. Al Hawi[1], Jana D. Alshehri[1] and Reem Elbeltagy[1]*

[1]Department of Health Communication Sciences, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, [2]Department of Health Sciences, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia

Background: Down syndrome is a genetic disorder present from birth, leading to various physical and cognitive challenges. Some children with Down syndrome also experience hearing loss. The combined impact of raising a child with Down syndrome and hearing loss can affect the quality of life (QoL) of mothers. The aim of this study was to examine the differences in QoL between mothers of children with Down syndrome who have hearing loss and those without hearing loss.
Methods: A cross-sectional study was conducted with 103 mothers of children with Down syndrome. Data were collected using the Arabic version of the WHOQOL-BREF, which assesses the QoL across four domains, including physical health, psychological health, social relationships, and environment, with scores ranging from 0-100.
Results: In the studied sample, 16.5% of mothers reported that their child had hearing loss. The mean QoL scores for the total sample were 65.8 $\pm$ 18.6 in physical health, 72.6 $\pm$ 17.5 in psychological health, 65.6 $\pm$ 13.7 in social relationships, and 68.7 $\pm$ 16.8 in environment. There were no significant differences in QoL scores across any domain between mothers of children with and without hearing loss ($p > 0.05$). Mothers' perceptions of their overall QoL and health were high and similar between both groups.
Conclusion: Most mothers in this study reported satisfactory QoL. While HL does not seem to drastically affect overall QoL across various domains, it is evident that social challenges persist.

KEYWORDS

Down syndrome, hearing loss, quality of life, mothers, children

1 Introduction

Down syndrome (DS) is the most prevalent chromosomal disorder. Children diagnosed with DS typically exhibit slower developmental rates and a range of other health problems, including congenital heart disease, endocrine disorders, eye disorders, and obstructive sleep apnea (1). These conditions necessitate heightened parental dedication, impacting the entire family dynamics (2).

Children with DS commonly experience recurrent ear infections and hearing loss (HL), with a reported prevalence of 47% (3). HL in children with DS may have various etiologies. Conductive hearing loss (CHL), the most common type of HL in these children, is often caused by factors such as recurrent wax impaction due to stenosis in the external auditory canal, recurrent otitis media from Eustachian tube dysfunction, and ossicular anomalies in the middle ear (4). Sensorineural hearing loss (SNHL), which is less common in children with DS, can result from abnormalities in the inner ear, such as inner ear hypoplasia, cochlear dysplasia, or cochlear nerve canal anomalies (4). Additionally, SNHL may be influenced by perinatal risk factors or remain idiopathic (3).

Mothers of newly diagnosed children with HL often grapple with persistent feelings of being overwhelmed and inadequately equipped to manage their child's HL effectively, which may manifest as anger (5). Existing literature has shown that mothers of children with DS alone often experience reduced QoL due to increased caregiving demands and concerns about long-term developmental outcomes (6). Similarly, mothers of children with HL alone report emotional stress, social isolation, and challenges accessing early intervention services (7) However, limited research has explored how the co-occurrence of DS and HL may uniquely impact maternal QoL, despite evidence that both conditions independently impose significant caregiving burdens.

When DS is coupled with HL, the challenges can be further compounded. Communication difficulties may be more pronounced, and accessing appropriate interventions, addressing increased social isolation, and ensuring adequate educational advocacy and support services may require additional effort. Mothers may experience heightened levels of stress and anxiety related to societal stigma and concerns about their child's ability to communicate effectively and participate fully in social and educational settings (8, 9).

Given the specialized attention required by children with DS, mothers often find themselves involved in every aspect of their child's activities, which might affect their quality of life (QoL). The World Health Organization (WHO) defined the QoL as the individuals' perceptions of their position in life within the context of their cultures, values, and goals. It is a subjective, multidimensional construct that encompasses both positive and negative elements of evaluation (10).

The present study aimed to assess the QoL among mothers of children with DS, whether their children experience HL or not, to explore if HL adds more burden to their QoL. By examining whether HL further exacerbates maternal burden, the current study aims to fill a critical gap in understanding how dual diagnoses affect maternal well-being. The study distinguishes itself from existing literature through the following. First, while extensive literature exists on the impact of either DS or HL on family well-being, the specific challenges and their combined effect on maternal QoL when these conditions co-occur remain underexplored except by Hussin et al. (11). Second, it specifically examined mothers in Saudi Arabia. This provided an underrepresented perspective on how specific Saudi cultural factors (e.g., family structures, local resources) uniquely influence maternal QoL. Third, this study explored additional variables such as the relationship between QoL and maternal/child age, offering a broader view of influencing factors than prior research. Previous studies have suggested that a caregiver's age may influence coping strategies, stress tolerance, and access to social support, all of which can impact QoL (12). Similarly, the age of the child may reflect different caregiving demands such as early intervention needs in younger children versus long-term planning stress in older children. Including these variables provides a broader understanding of potential factors influencing maternal QoL in the context of raising a child with DS.

By addressing these specific gaps, this study enhanced the understanding of complex caregiving demands and provided insights for targeted support interventions for mothers of children with DS and co-occurring HL in this particular cultural setting.

2 Materials and methods

A cross-sectional study was conducted, including 103 mothers of children with DS. Given the inherent rarity of DS (1.8:1000 in Saudi Arabia and 1:700–1000 globally), achieving a sample size of 103 represents a substantial cohort for this population (13). Participants were recruited through the Down Syndrome Charity in Saudi Arabia. Additionally, recruitment was conducted via social media platforms such as X, WhatsApp, and Telegram to broaden the search area for participants.

Data were collected using the Arabic version of the WHOQOL-BREF (10). The questionnaire comprises 26 questions, including two questions about overall QoL and general health, and 24 questions representing specific facets from the original WHOQOL-100 tool (14). The four domains assessed are physical health, psychological health, social relationships, and environment with scores ranging from 0-100; higher scores indicate better QoL. Categorization of QoL was based on a 60% cutoff, where scores ≥ 60% were considered "good", and scores < 60% were considered "poor" (15).

The WHOQOL-BREF was chosen to assess maternal QoL due to its comprehensive and holistic approach, covering physical, psychological, social, and environmental health, which is crucial for understanding the multifaceted impact of caregiving. Its proven cross-cultural validity, developed through international collaboration and specifically validated in Arabic-speaking populations (16), makes it highly suitable for the study's population. The instrument's brevity (26 items) minimizes respondent burden for busy caregivers, improving

data quality and feasibility. Furthermore, the WHOQOL-BREF possesses established psychometric properties and has been widely validated and utilized in various caregiver populations, including parents of children with chronic conditions and disabilities. This allows for meaningful comparisons not only between the study groups (mothers of children with DS with vs. without HL) but also with broader population norms. While specialized caregiver tools exist, the WHOQOL-BREF's generic nature provides a robust and broadly comparable measure of overall QoL. It was selected over caregiver-specific HRQOL tools because of its strong psychometric foundation, extensive cross-population validation, and its ability to facilitate comparison with both clinical and general population norms, features particularly important for the aims of this study.

2.1 Participants inclusion and exclusion criteria

Inclusion criteria: Mothers of children with DS who reside in Saudi Arabia and are willing to participate in the study by providing informed consent.

Exclusion Criteria:

- Mothers with a diagnosed mental health condition (e.g., severe depression, anxiety disorder, or psychosis) that could significantly impact their self-reported QoL, or other chronic illnesses that might heavily influence their QoL independent of their child's condition. This helps ensure that the measured QoL is primarily related to the caregiving experience for a child with DS.
- Children with DS who have additional significant co-morbidities beyond DS and HL (e.g., severe congenital heart defects requiring ongoing intensive care, cerebral palsy, or autism spectrum disorder). This helps to isolate the impact of DS and HL.
- If the child with DS is not currently living at home with the mother (e.g., institutionalized, living independently), this would significantly alter the daily caregiving experience.
- Inability or unwillingness to provide informed consent.
- Inability to communicate effectively in the study language.

The final sample included 103 mothers of children with DS who were residents in Saudi Arabia and provided informed consent. Among them, 17 mothers reported that their child had co-occurring HL (DS+HL), while 86 mothers reported no known HL (DS-only). Mothers were selected as participants due to their primary caregiving role in the Saudi cultural context, which typically involves direct responsibility for managing the child's medical, developmental, and emotional needs.

HL status was identified based on maternal report, wherein participants were asked whether a healthcare professional had diagnosed their child with HL. No audiological testing was conducted as part of the study. Some demographic variables, such as maternal education level, employment status, and household income were not collected in this study, which limits the ability to explore how these factors may influence QoL outcomes.

2.2 Data analysis

Statistical analysis was performed using IBM SPSS version 29. Descriptive statistics were presented as frequencies, percentages, means, and standard deviations (SD). Differences between means were evaluated using the Welch's t-test, which is appropriate when unequal variances and unequal group sizes are present. This test was selected to account for the imbalance in sample size between the two groups (DS-only vs. DS+HL). Pearson correlation was utilized to determine relationships between quantitative variables. A p-value of less than 0.05 was considered statistically significant.

2.3 Ethical consideration

Ethical approval was obtained from the Institutional Review Board (ID 22-0937). Participants consents were obtained electronically.

3 Results

The demographic characteristics of the studied sample including the age of the mother, the age of the child and the gender of the child are detailed in Table 1. Among the total children, 16.5% (n=17) had HL and 22.3% (n=23) had recurrent ear infection as reported by their mothers. Of those with HL, 11.8% had SNHL, 23.5% had CHL and 64.7% were unsure of the type of HL their child experienced.

TABLE 1 Demographic characteristics of the studied sample (n= 103).

Demographic characteristics	No.	%
Mother's age (years)		
18-25	8	7.8
26-35	10	9.7
36-45	38	36.9
46-55	34	33
≥56	13	12.6
Child's gender		
Male	56	54.5
Female	47	45.6
Child's age (years)		
Less than 1	8	7.8
01-May	27	26.2
06-Oct	22	21.4
Nov-15	25	24.3
≥16	21	20.4
Total	103	100

10.3389/fpsyt.2025.1611074

The mean QoL scores among the studied sample in the physical, psychological, social and environmental domains were 65.8 ± 18.6, 72.6 ± 17.5, 65.6 ± 13.7 and 68.7 ± 16.8 respectively (Table 2). Most participants reported good QoL in the physical, psychological, and environmental domains regardless of their child's hearing status, with ≤30% reporting poor QoL in these domains (Table 3). However, slightly over half of the mothers of children with both DS and HL reported poor QoL scores in the social domain (53%) compared to 29.1% of the mothers of children with DS but without HL (Table 3).

To determine if HL poses more challenge to the QoL of mother of children with DS, the differences in mean scores across different QoL domains between mothers of children with HL and those without HL were assessed using the Welch's t-test. The results revealed no significant differences in any of the domains (Table 4).

The majority of mothers reported a good perception of their overall QoL (82.4% for others of children with HL and 88.4% for mothers of children with normal hearing). Similarly, most mothers reported a good perception of their health status (76.5% for mothers of children with HL and 86.1% for mothers of children with normal hearing). The association between mothers' perceptions of their QoL and health status and whether they had a child with HL or not was examined using the Welch's t-test, which also revealed no significant association (Table 5). Additionally, the correlation between the different QoL domains and both the mother's age and the child's age were investigated using the Pearson correlation test, which indicated no significant correlations (Table 6).

4 Discussion

Mothers play a crucial role in the lives of their children, and this is especially true for children with disabilities, such as DS. Studies have shown that intellectual disabilities and other disorders lead to significant socio-occupational dysfunction and impaired QoL for caregivers (17). Children with DS may also experience HL (3), adding an additional burden to their mothers, who may face unique communication challenges and require additional support services to ensure their child's well-being (8, 9). Caregivers often report physical difficulties, such as insufficient sleep, lack of exercise, irregular and inadequate meals, and neglecting their own medical needs. These challenges can result in clinical depression, social isolation, high stress, and low QoL (18). Therefore, it is important to assess the QoL of mothers of children with disabilities to provide them with better support and interventions, if needed.

The findings of the current study indicate that the majority of mothers reported satisfactory QoL, aligning with results from other studies (11, 19, 20) and even surpassing some (6, 21). However, the social domain had the lowest score, particularly for mothers of children with HL. This may be due to social isolation stemming from feelings of difference from other mothers and limited social networks due to communication barriers, which can lead to loneliness (8). Additionally, isolation could be exacerbated by limited support networks or societal misunderstandings about their children's conditions (22).

TABLE 2 Means and SDs of the scores of the QoL domains among the total studied sample (n= 103).

QoL domain	Mean ± SD
Physical health	65.8 ± 18.6
Psychological health	72.6 ± 17.5
Social relationships	65.6 ± 13.7
Environment	68.7 ± 16.8
Total score	68.2 ± 16.6

TABLE 3 Mothers' perception of their QoL of each domain (n= 103).

Mothers' perception	Mothers of children with DS who experience HL No (%)	Mothers of children with DS who do not experience HL No (%)
Physical health		
Good	12 (70.5)	62 (72)
Poor	5 (29.5)	24 (28)
Psychological health		
Good	15 (88.2)	62 (72)
Poor	2 (11.8)	24 (28)
Social relationships		
Good	8 (47)	61 (70.9)
Poor	9 (53)	25 (29.1)
Environment		
Good	12 (70.5)	60 (69.8)
Poor	5 (29.5)	26 (30.2)

TABLE 4 Means and SDs of the scores of the QoL domains among mothers of children of both groups (with and without HL) (n= 103).

QoL domain	Mothers of children with DS who experience HL (Mean ± SD)	Mothers of children with DS who do not experience HL (Mean ± SD)	P value*
Physical health domain	66.7 ± 17.4	65.6 ± 18.9	0.8
Psychological health domain	76.7 ± 15.6	71.8 ± 17.8	0.3
Social relationships domain	59.8 ± 15.6	66.7 ± 13.1	0.06
Environment domain	67.7 ± 13.7	68.9 ± 17.4	0.8
Total	67.7 ± 15.5	68.2 ± 16.8	0.8

*Welch's t-test.

10.3389/fpsyt.2025.1611074

TABLE 5 Perception of the mothers regarding their overall QoL and health status (n= 103).

Mothers' perception	Mothers of children with DS who experience HL (Mean ± SD)	Mothers of children with DS who do not experience HL (Mean ± SD)	P value*
Mother's Perception regarding their overall QOL	4.3 ± 0.8	4.5 ± 0.7	0.4
Mother's Perception regarding their overall health status	4.4 ± 0.9	4.2 ± 0.8	0.4

*Welch's t-test.

The present study found the psychological domain to have the highest QoL scores, while the social domain was the lowest, a pattern that contrasts with findings from Hussin et al. (11), where the social domain scored better. This notable discrepancy may be explained by several contextual and methodological differences. While previous studies have shown similar findings (19), others reported higher social domain scores (11, 23, 24), attributing this to strong parent-child relationships and familial bonds developed through shared experiences. The differing QoL domain scores between the current study and Hussin et al. (11) can be significantly attributed to cultural and contextual influences. The availability and nature of formal and informal support systems (healthcare, social programs) in each respective context profoundly influence how caregivers experience and perceive their psychological well-being and social connections.

Additionally, the current study did not assess marital status, education level, and income, which presents a limitation when comparing our findings to Hussin et al. (11). Socioeconomic status and education are known to influence access to resources, coping mechanisms, and social networks, thereby impacting perceived QoL. Similarly, marital status and the availability of family support

TABLE 6 Correlation between QoL domains of mothers of children with DS who experience HL and both the age of the mother and the age of the child.

QoL domain	Mother's age		Child's age	
	r*	P value	r*	P value
Mother's Perception regarding their overall QOL	-0.09	0.7	0.1	0.6
Mother's Perception regarding their overall health status	-0.3	0.3	-0.02	0.9
Physical health domain	-0.1	0.6	-0.05	0.8
Psychological health domain	-0.2	0.4	0.06	0.8
Social relationships domain	-0.3	0.3	-0.3	0.2
Environment domain	0.04	0.9	0.2	0.4

*Pearson correlation test.

significantly affect psychological well-being (by reducing isolation and stress) and social domain scores. These unmeasured differences in participant demographics and family structures likely contributed to the observed variations in QoL domain scores between the two studies.

The results suggest that HL in children with DS may not directly impact the overall QoL of their mothers across various domains. This aligns with Hussin et al. (11), who found no significant difference in QoL between mothers of children with DS and HL and those with DS alone, with 60% of mothers reporting satisfaction. This could be due to the typically mild to moderate nature of HL in DS cases and the ability of some children with mild HL to compensate (24).

Although no significant statistical difference was found between mothers of children with HL and those without in the social domain, the difference is nonetheless noteworthy. More than half of the mothers of children with HL (53%) reported poor scores in this domain compared to 29% of mothers of children without HL. This nearly twofold disparity, while not statistically significant, suggests that the co-occurrence of DS and HL may uniquely and substantially burden a mother's social life.

This burden may stem from multiple interrelated factors. Compounded communication challenges can increase time and resource demands on the mother, while limited social opportunities for the child may further restrict the mother's own social engagement. In addition, mothers of children with both conditions may face greater social isolation due to a perceived lack of understanding from single-diagnosis support networks and the presence of societal stigma.

These challenges may be further intensified by the cumulative stress of managing multiple diagnoses, navigating early intervention services, advocating for educational placements, and coordinating ongoing therapy and communication support. Emotional strain, financial pressures, and social stigma could all contribute to the reduced QoL (25) reported by these mothers. Taken together, these findings underscore the importance of the social domain as a critical area for future targeted interventions and research, even if broader QoL differences were not statistically significant.

Contrary to previous studies suggesting that behavioral problems in children with DS increase with age, which could decrease caregivers' QoL (26, 27), the current study found no significant correlation between the child's age and different QoL domains. However, the result is consistent with other studies such as Vadakedom et al. (6). The correlation between mothers' age and QoL is also debated in the literature, with some studies showing negative correlations (20), others showing positive correlations (27), and some showing no correlation (19). Our study supports those reporting no correlation (19), suggesting that mothers of all ages adapt well to their children's conditions, particularly since children with DS often exhibit fewer behavioral problems compared to those with other disabilities (28).

The study has some limitations. The cross-sectional design restricts causal inferences, and the reliance on maternal reports rather than objective audiological assessments could affect the accuracy of the reported prevalence of HL. This approach may result in misclassification or underreporting of HL, especially in cases where hearing issues have not been formally diagnosed. The absence of

audiological confirmation also limits our ability to assess the severity and type of HL, which may have differential impacts on maternal QoL. Additionally, potential confounding factors such as the severity of the child's HL, maternal support systems, and socioeconomic status were not explored. Furthermore, the notable imbalance in sample sizes between the DS-only and DS+HL groups may have limited the statistical power to detect significant differences between groups. While the use of Welch's t-test helped to partially account for this issue, the small number of participants in the DS+HL group (n=17) remains a limitation that may affect the robustness and generalizability of between-group comparisons. Future studies with larger and matched samples are recommended to validate these findings.

5 Conclusions and recommendations

Most mothers in this study reported satisfactory QoL. Notably, the social domain had the lowest scores, particularly for mothers of children with HL. This suggests that social isolation and communication barriers may significantly impact QoL. While HL does not seem to drastically affect overall QoL across various domains, it is evident that social challenges persist.

The study found no significant correlations between the child's or mother's age and QoL, indicating that age might not be a critical factor in determining QoL in this context. This suggests that other variables may be more influential in shaping QoL for mothers of children with DS.

To improve the accuracy of findings, longitudinal studies are needed to explore how HL and DS interact over time and how QoL evolves as children with DS grow older. Such research could offer valuable insights into the long-term effects of HL and other factors on mothers' QoL. Further investigation into the role of maternal support systems and societal understanding is crucial. Understanding these factors could inform the development of targeted support mechanisms to alleviate social isolation and stigma. Importantly, there is a pressing need for structured interventions specifically addressing the social QoL of mothers caring for children with both DS and HL. These may include accessible community-based programs, social engagement initiatives, and tailored peer support networks that can help mitigate isolation and emotional burden.

A focused analysis of the social domain of QoL is warranted. Research should examine specific aspects of social isolation and communication barriers faced by mothers of children with HL. Developing interventions to enhance social integration and support networks could improve QoL for these mothers. Prioritizing this subgroup in intervention design is essential, as they may face compounded challenges due to dual diagnoses. Additionally, future studies should consider socioeconomic status and other potential confounders to provide a comprehensive understanding of the factors influencing QoL. This broader perspective could help tailor support strategies more effectively.

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.

Ethics statement

Ethical approval was obtained from the Institutional Review Board at Princess Nourah bint Abdulrahman University (ID 22-0937). Participants consents were obtained electronically.

Author contributions

RA: Conceptualization, Data curation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. HB: Formal Analysis, Validation, Writing – review & editing. LA: Data curation, Writing – original draft. JA: Data curation, Writing – original draft. RE: Data curation, Methodology, Project administration, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number RI-44-0372.

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

1. Centers of Disease Control and Prevention. Data Statistics on Down Syndrome (2023). Available online at: https://www.cdc.gov/birth-defects/living-with-down-syndrome/?CDC_AAref_Val=https://www.cdc.gov/ncbddd/birthdefects/downsyndrome/data.html (Accessed August 14, 2024).

2. Bourke J, Ricciardo B, Bebbington A, Aiberti K, Jacoby P, Dyke P, et al. Physical and mental health in mothers of children with Down syndrome. *J Pediatr*. (2008) 153:320–6. doi: 10.1016/j.jpeds.2008.02.047

3. De Schrijver L, Topsakal V, Wojciechowski M, Van de Heyning P, Boudewyns A. Prevalence and etiology of sensorineural hearing loss in children with down syndrome: A cross-sectional study. *Int J Pediatr Otorhinolaryngol*. (2019) 116:168–72. doi: 10.1016/j.ijporl.2018.10.048

4. Shott SR. Down syndrome: common otolaryngologic manifestations. *Am J Med Genet C Semin Med Genet*. (2006) 142C:131–40. doi: 10.1002/ajmg.c.30095

5. Kurtzer-White E, Luterman D. Families and children with hearing loss: Grief and coping. *Ment Retard Dev Disabil Res Rev*. (2003) 9:232–5. doi: 10.1002/mrdd.10085

6. Vadakedom SS, Antony JM, Padma BK, Mammen DS, Thankappan BP. Quality of life of mothers of children with Down syndrome. *J Evol Med dent Sci*. (2017) 6:2939–43. doi: 10.14260/Jemds/2017/633

7. Pelarak F, Radmehr M, Sanjari H, Abdalvand A, Forouzesh F. The children with hearing disabilities: Mothers' general health status and children's quality of life. *JundishapurJChronicDisCare*. (2021) 10:e115472. doi: 10.5812/jjcdc.115472

8. Gunjawate DR, Ravi R, Driscoll C. Stress among parents of children with hearing loss and how they deal with it: A systematic review. *Int Arch Otorhinolaryngol*. (2023) 27:166–77. doi: 10.1055/s-0042-1743273

9. Watanabe M, Kibe C, Sugawara M, Miyake H. Courtesy stigma of parents of children with Down syndrome: Adaptation process and transcendent stage. *J Genet Couns*. (2022) 31(3):746–57. doi: 10.1002/jgc4.1541

10. The WHOQOL GROUP. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. *Psychol Med*. (1998) 28:551–8. doi: 10.1017/s0033291798006667

11. Hussin N, Ismail A, Ismail J, Abdullah A. The quality of life of mothers of down syndrome children with and without hearing impairment in Universiti Kebangsaan Malaysia Medical Center. *Indian J Otol*. (2021) 27:189–92. doi: 10.4103/indianjotol.INDIANJOTOL_70_20

12. Alrayes N, Issa NM, Alghubayshi OY, Al-Amaa JY, Alsabban AH, Al Shaer DS, et al. Quality of life in children with Down syndrome and its association with parent and child demographic characteristics: Parent-reported measures. *Mol Genet Genomic Med*. (2024) 12:e2337. doi: 10.1002/mgg3.2337

13. Parveen S, Ahmad A, Reshi AA. Empowering lives: navigating the landscape of down syndrome support in Saudi Arabia. *JDR*. (2024) 3. doi: 10.57197/JDR-2024-0044

14. The WHOQOL GROUP. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med*. (1995) 41:1403–9. doi: 10.1016/0277-9536(95)00112-k

15. Silva PA, Soares SM, Santos JF, Silva LB. Cut-off point for WHOQOL-bref as a measure of quality of life of older adults. *Rev saude publica*. (2014) 48:390–7. doi: 10.1590/s0034-8910.2014048004912

16. Dalky HF, Meininger JC, Al-Ali NM. The reliability and validity of the arabic world health organization quality of life-BREF instrument among family caregivers of relatives with psychiatric illnesses in Jordan. *J Nurs Res*. (2017) 25:224–30. doi: 10.1097/JNR.0000000000000146

17. Malhotra S, Khan W, Bhatia MS. Quality of life of parents having children with developmental disabilities. *Delhi Psychiatry J*. (2012) 15:171–6.

18. Dumas EE, Wolf LC, Fisman SN, Culligan A. Parenting stress, child behavior problems, and dysphoria in parents of children with autism, Down syndrome, behavior disorders, and normal development. *Exceptionality J Spec Educ*. (1991) 2:97–110. doi: 10.1080/09362839109524770

19. AlAhmari FS, Alageel AF, Aldosari MA, Bagha MY. The quality of life of parents of children with Down syndrome in a tertiary care hospital: a qualitative research study at Saudi Arabia. *Ann Intern Med*. (2022) 81:104428. doi: 10.1016/j.amsu.2022.104428

20. Geok CK, Abdullah KL, Kee LH. Quality of life among Malaysian mothers with a child with Down syndrome. *Int J Nurs Pract*. (2013) 19:381–9. doi: 10.1111/ijn.2013.19.issue-4

21. Senses Dinc G, Cop E, Tos T, Sari E, Senel S. Mothers of 0–3-year-old children with Down syndrome: Effects on quality of life. *Pediatr Int*. (2019) 61:865–71. doi: 10.1111/ped.13936

22. Wright-Sexton LA, Compretta CE, Blackshear C, Henderson CM. Isolation in parents and providers of children with chronic critical illness. *Pediatr Crit Care Med*. (2020) 21:e530–7. doi: 10.1097/PCC.0000000000002344

23. Dabrowska AQA, Pisula E. Parenting stress and coping styles in mothers and fathers of pre-school children with autism and Down syndrome. *J Intellect*. (2010) 54:266–80. doi: 10.1111/j.1365-2788.2010.01258.x

24. Moore DR, Zobay OF, Erguson MA. Minimal and mild hearing loss in children: Association with auditory perception, cognition, and communication problems. *Ear Hear*. (2020) 41:720–32. doi: 10.1097/AUD.0000000000000802

25. Fucà E, Costanzo F, Ursumando L, Vicari S. Parenting stress in mothers of children and adolescents with down syndrome. *J Clin Med*. (2022) 11:1188. doi: 10.3390/jcm11051188

26. Norizan A, Shamsuddin K. Predictors of parenting stress among Malaysian mothers of children with Down syndrome. *J Intellect*. (2010) 54:992–1003. doi: 10.1111/j.1365-2788.2010.01324.x

27. Corrice AM, Glidden LM. The Down syndrome advantage: fact or fiction? *A.J.I.D.D*. (2009) 114:254–68. doi: 10.1352/1944-7558-114.4.254-268

28. Eisenhower AS, Baker BL, Blacher J. Preschool children with intellectual disability: syndrome specificity, behavior problems, and maternal well-being. *J Intellect*. (2005) 49:657–71. doi: 10.1111/j.1365-2788.2005.00699.x

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 03 September 2025
DOI 10.3389/feduc.2025.1641303

OPEN ACCESS

EDITED BY
Xinyue Qiu,
Beijing Union University, China

REVIEWED BY
Susana Cardoso,
University of Trás-os-Montes and Alto Douro, Portugal
Giuditta Smith,
University of East Anglia, United Kingdom

*CORRESPONDENCE
Vasiliki Lymperopoulou
✉ vlymperop@uth.gr

RECEIVED 04 June 2025
ACCEPTED 18 August 2025
PUBLISHED 03 September 2025

CITATION
Andreou G, Lymperopoulou V, and Aslanoglou V (2025) Structural language in neurodevelopmental disorders: comparison between autism spectrum disorder (ASD) and developmental language disorder (DLD). *Front. Educ.* 10:1641303. doi: 10.3389/feduc.2025.1641303

Structural language in neurodevelopmental disorders: comparison between autism spectrum disorder (ASD) and developmental language disorder (DLD)

Georgia Andreou, Vasiliki Lymperopoulou*, and Vasiliki Aslanoglou

Department of Special Education, University of Thessaly, Volos, Greece

Introduction: The aim of the present study is to investigate structural language of children with Autism Spectrum Disorder (ASD) and Developmental Language Disorder (DLD) in a Greek-speaking population.

Methods: Three groups participated in the study, matched for age and sex: 25 children with ASD aged 6–8 years (mean age 84.19 months; *SD* = 6.55), 25 children with DLD aged 6–8 years (mean age 84.09 months; *SD* = 6.72), and 25 typically developing children, who served as controls (mean age 84.09 months; *SD* = 6.72). Structural language was examined in all three groups by using standardized tests assessing their skills in phonological awareness, morphosyntax and vocabulary. Phonological skills were evaluated using subscales from the Test for the Detection and Investigation of Reading Difficulties, while scales from the Diagnostic Test of Linguistic Intelligence for school-age children were used for the assessment of morphosyntax. Finally, the lexical scale from WISC-V was used to assess expressive vocabulary.

Results: The findings of the study showed that both children with ASD and children with DLD performed worse on tests examining structural language than their typically developing peers. More specifically, statistically significant differences were observed across all measurements ($p < 0.001$). However, the comparison between the two clinical groups revealed that the performance of children with DLD was worse than that of children with ASD on all tests examining structural language. More specifically, the difference between the two groups in terms of phonology and morphosyntax was $p < 0.001$, while the difference in terms of expressive vocabulary was $p = 0.03$.

Discussion: The findings shed light on important aspects of structural language in both ASD and DLD by providing insights into the common and differential language challenges faced by individuals with these neurodevelopmental disorders. This analysis enhances the understanding of language development in the Greek-speaking population and offers a cross-disorder approach. These findings may contribute to the development of targeted educational strategies to support children with ASD and DLD.

KEYWORDS

structural language, autism spectrum disorder, developmental language disorder, neurodevelopmental disorders, Greek language

1 Introduction

Language is crucial for communication, the components of which are language form, content and language use. Structural language consists of the language form and content, while language use is part of the pragmatic language level (Reindal et al., 2021). More specifically, for a more complete understanding of the language mechanism, language is divided into five interrelated levels: phonology, morphology, syntax, semantics and pragmatics (Reetzke et al., 2015). The term structural language is used to describe phonology, morphosyntax and semantics (Reetzke et al., 2015; Vassiliu et al., 2022). Although, this distinction of language linguistic levels is artificial, it seems useful in order to evaluate these skills, especially when examining atypical populations (Matthews et al., 2018).

Language development is a complex process that progresses through different levels and follows a dynamic trajectory (Hoff, 2009), which is impacted by both biological and environmental factors (Rinaldi et al., 2023). Nevertheless, sometimes this trajectory does not follow the typical development, leading to deviations that may affect different aspects of language competence. Such deviations are often found in the context of neurodevelopmental disorders, mainly in Autism Spectrum Disorder (ASD) and Developmental Language Disorder (DLD) where language development appears with atypical features and may be accompanied by broader difficulties in communication and learning (Conti-Ramsden et al., 2012; Luyster et al., 2011; Whyte and Nelson, 2015).

ASD is a neurodevelopmental disorder which mainly concerns social difficulties, as reported in DSM-5 [American Psychiatric Association (APA), 2013]. The diagnostic criteria for ASD include deficits in social communication and social interaction occurring in multiple social contexts and limited repetitive (stereotypical) patterns of behaviors, interests and activities [American Psychiatric Association (APA), 2013]. Similarly, according to ICD-11, ASD is characterized by persistent deficits in initiating and maintaining reciprocal social interaction and social communication, accompanied by atypical or excessive limited, repetitive and rigid patterns of behavior, interests or activities relative to age and sociocultural context.

Research has shown that ASD presents high heterogeneity, as language and cognitive skills and deficits vary (Girolamo et al., 2024; Henderson et al., 2014; Silleresi, 2023). One aspect of this heterogeneity concerns structural language abilities. Difficulties in structural language as part of communication deficits are not a criterion for the diagnosis of ASD, according to the DSM-5 [American Psychiatric Association (APA), 2013]. However, approximately 65% of children who receive a diagnosis of ASD also have language deficits and receive a concurrent diagnosis of a language disorder (Levy et al., 2010; Schaeffer et al., 2023; Tager-Flusberg et al., 2005). Numerous studies highlight the strong heterogeneity of ASD population regarding language abilities, reporting different types of language difficulties in children with ASD (Rapin and Dunn, 2003). Tager-Flusberg and Joseph (2003) distinguished two language subtypes among verbal children with ASD: those with typical language skills and those with language impairments. Tager-Flusberg (2006) further confirmed this variability by examining phonological processing and grammatical morphology, distinguishing a group with structural language impairments from one with typical structural language abilities. Later studies, also revealed a subgroup with structural language impairments (ASD-LI) and another exhibiting typical structural language development (Georgiou and Spanoudis, 2021; Whitehouse et al., 2008). The group with typical structural language development often demonstrates high level of verbal fluency, performing similarly to typically developing peers on tests of structural language (Tek et al., 2014) and exhibits advanced vocabulary and syntax (Boucher, 2012; Tager-Flusberg and Caronna, 2007).

While early studies describe two main subgroups in ASD, more recent research suggests a greater diversity of language profiles. Vogindroukas et al. (2022) proposed four language profiles: ASD and pragmatic difficulties, without the presence of any other language difficulties; ASD and comorbidity with DLD, or another developmental disorder; ASD and intellectual disability; and ASD and social communication and interaction difficulties. Similarly, Silleresi (2023) proposes three profiles that have been strongly established. The first profile consists of autism with language and intellectual abilities in accordance with the norms (ASD-LN), the second concerns autism with language and intellectual impairments, while the third profile includes autism with language impairments (ASD-LI) without intellectual impairments.

Significant language deficits and atypical language development are also observed in the population with DLD. Although the term DLD is not included in the DSM-5 [American Psychiatric Association (APA), 2013], it is accepted by both the scientific community and clinical practitioners. DSM-5 uses the term "Language Disorder", which includes it among neurodevelopmental disorders, in order to identify persistent difficulties in the acquisition and use of language, both in terms of comprehension and production. In addition, ICD-11 uses the term DLD, which is defined as a neurodevelopmental disorder characterized by persistent deficits in the acquisition, comprehension, production or use of language (World Health Organization, 2019). These deficits arise during the developmental period, usually in early childhood, and cause significant limitations in the individual's ability to communicate (Bishop et al., 2016, 2017).

Children with DLD exhibit delayed language development, as the acquisition of language skills is slower than that of typically developing children. Nonetheless, they exhibit great heterogeneity in their language abilities and weaknesses (Ryder and Leinonen, 2014). The difficulties of children with DLD, which may involve language expression, comprehension, or both (Bishop, 1997), negatively affect all cognitive functions. These difficulties impact all levels of language development, both in oral and written language, in children with DLD (Andreou and Aslanoglou, 2022; Girbau and Schwartz, 2007).

Previous research has shown that individuals with DLD exhibit difficulties in the use of structural language, namely phonology, morphology, syntax and semantics, while those difficulties have also been found present in ASD.

Phonology regards the linguistic sounds transmitted from the speaker to listener during the communication. It examines

the phonemes of the language system and focuses on phonetic categories, phonemes, and intonation (Shakila et al., 2024). Phonemes are the smallest sound units that differentiate words (Barokova and Tager-Flusberg, 2020; Stemberger and Bernhardt, 2023). By the age of 1 year, the child can produce phonetically stable forms for communication varying the tone and volume of the voice, and from this age onwards, the child begins to produce his or her first words (Dore et al., 1976; Shakila et al., 2024). Phonological development leads to phonological awareness, a very important skill for oral and written language. Phonological awareness is divided into phonemic awareness, which concerns the perception of the smallest units of speech (phonemes and speech sounds), and syllabic awareness (larger units of speech, such as syllables). Phonological awareness supports word and sentence production by enabling the perception of phonological units as distinct parts of language and their functional use in speech (Berninger et al., 2010; Sun and Poeppel, 2023). Children with ASD exhibit a delay in the development of phonological skills compared to that of typically developing children (Papoudi and Vakalopoulou, 2022; Schaeffer et al., 2023). However, Wetherby et al. (2004) observed that the order of occurrence of phonemes in children with ASD did not differ from that of typically developing children, and no differences were found in the pattern of expected phonological errors. On the other hand, the phonological processing of children with DLD is similar to that of typically developing younger children (Leonard, 1998). Nevertheless, they often omit or substitute phonemes and have difficulty managing even simple syllabic structures (of the consonant-vowel form) (Aguilar-Mediavilla et al., 2002). Furthermore, they have difficulty in the accurate articulation of laterals, nasals, and stops (Aguilar-Mediavilla et al., 2002), and in producing sibilant (/s/, /z/) and liquid (/l/, /r/) phonemes. In addition, children with DLD make errors in consonant clusters and experience articulation difficulties in the flow of speech, while errors in the articulation of polysyllabic words are also evident. In general, the speech intelligibility of children with DLD is negatively affected by the phonological errors they produce. Finally, there is difficulty in repeating pseudowords (Lalioti et al., 2016), especially when the repetition involves pseudowords with more than two syllables (Mengisidou et al., 2020).

Morphology refers to the individual components of words and the relationships between them. It studies the structure of words and the rules for their analysis and creation (Apel et al., 2013; Barokova and Tager-Flusberg, 2020; James et al., 2021). In particular, it examines the morphemes, which are the minimal units of language that have a meaning, and the way the morphemes are identified, analyzed and described through the structure of words. The ability of an individual to distinguish speech into morphemes is called morphological awareness and is important for language development (Berninger et al., 2010; Carlisle, 1995; James et al., 2021). From the age of 2 or 3 years, children begin to focus on word form through the use of morphemes to assign different functions to the words they produce and are able to identify individual words in the flow of speech (Peters, 2017). Gradually, around the age of three, the child is able to perform a sub-generalization of the rule for forming a word, which is considered an indication of the application of symbolic rules. From the age of 3 years, children gradually master the ability to distinguish the semantic and phonological texture of words and word themes, as well as to follow the rules of language articulation in the words they produce (Hoff, 2009; Peters, 2017). Regarding morphological development in ASD, it has been observed that the use of functional words such as articles and pronouns is limited, while Kelley et al. (2006) found that children with ASD have difficulty in the correct use of tenses in a sentence, as they do not fully understand the temporal order of "now" and "then". In addition, Vogindroukas (2020) noted difficulties in understanding and generalizing grammatical rules. The research of Crandall et al. (2019) confirms the above finding as the researchers also observed difficulties in using grammatical rules. In addition, Terzi et al. (2014) found that inflectional morphology in Greek language constitutes an area of difficulty for children with ASD. The development of morphology is also deficient in children with DLD. Research in the English language reveals difficulties in the use of verbs and specifically in the formation of regular past tense marker *-ed* and regular plural number marker *-s* (Joye et al., 2019), while the use of the passive voice and the formation of questions are also deficient (Andreou et al., 2023; Stavrakaki, 2020). In Greek, difficulties have been reported in the use of the definite article and in the production and understanding of personal pronouns (Tsimpli and Stavrakaki, 1999), as well as difficulties in specific grammatical functions (Stavrakaki, 2006). Furthermore, deficits have been observed in the use of tenses, in the acquisition of the definite article and weak forms of the definite pronoun (Tsimpli and Stavrakaki, 1999), as well as in the perception and use of morphological information in terms of number and case (Stavrakaki et al., 2015). Interestingly, studies investigating the production of clitics in children with DLD in Greek produced controversial results. Tsimpli and Stavrakaki (1999) reported that children with DLD omit direct object clitics while Theodorou and Grohmann (2015) failed to find differences between children with DLD and TD peers on similar tests. Conflicting findings suggest that difficulties at the morphological level may not be homogeneous, but may be related to individual linguistic and cognitive factors. More specifically, it has been suggested that such morphological difficulties may be associated with deficits in grammatical awareness and the ability to process and use grammatical information, as well as with limitations in perceptual ability (Aslanoglou et al., 2023; Lancasterand Camarata, 2019).

Syntax refers to the structure of a sentence, which concerns the order of terms in a sentence. In every language, there is a system of rules that determine the order of terms in a sentence, through which speakers understand the meaning of the sentence (Fromkin et al., 2017). In other words, syntax refers to the hierarchical relationships and rules that regulate how words are connected to form sentences (Shakila et al., 2024). According to Hoff (2009), the way words are combined conveys the speaker's though, and often the meaning of a sentence is understood through its syntax. In particular, in languages with strict word order, two sentences may have the same formulas or words, but when they are in different order, the meaning changes (Lyons, 1995). Greek, by contrast, exhibits a relatively free word order due to its rich inflectional morphology (Alexiadou and Anagnostopoulou, 2000). By the age of six, the child is in the stage of full syntactic and morphological development. The child is able to systematically use functional words, as well as grammatical forms. Gradually,

the acquisition of basic syntactic and morphological structures is completed, and at this age the child's language comes close to the language of adults to a considerable extent (Hoff, 2009). Syntactic development in ASD is also deficient, as their sentences are short and simple, and they exhibit difficulties in both producing and understanding long complex sentences. In addition, the use of stereotypical expressions is still frequently observed, as they reproduce stereotypical expressions mechanically (Papoudi and Vakalopoulou, 2022). Zarokanellou et al. (2025) investigating the narrative skills of Greek-speaking children with ASD indicated that children with ASD exhibit a delay in syntactic development as compared to their TD peers. Nevertheless, Talli and Stavrakaki (2020) reported that syntactic deficits are key clinical features of the DLD population. Specifically, children with DLD produce simpler sentences than their TD peers, and they encounter difficulties in understanding long and complex syntactic sentences as well as in reading comprehension (Talli et al., 2016; Aslanoglou et al., 2023). Difficulties are also observed in sentence repetition and in their ability to understand the grammatical relationship between subject and object (Mengisidou et al., 2020).

Semantics is the level that examines the meaning of forms, words, phrases and sentences. These meanings are conventional or encoded in each language (Hoff, 2009). The morpheme, as mentioned above, is the smallest linguistic unit with a fixed form and meaning, while the word is the basic unit of the semantic level (Lyons, 1995). In a language, words may have a concrete or an abstract meaning, may be used with a literal or figurative meaning or even express more than one meaning (Andreou, 2012). Each person has a "mental lexicon", a repository of information, containing the morphemes and words of their language (Fromkin et al., 2017; Papafragou et al., 2022; Sun and Poeppel, 2023). As early as the 7th month, the child is able to distinguish words produced in his/her native language from words in a foreign language (Höhle and Weissenborn, 2003). By the age of 9 years, children are able to categorize objects and know verbs that mainly denote movement (Andreou, 2012). Children with ASD exhibit difficulties at the semantic level of language. Rapin and Dunn (2003) found deficits on the part of individuals with ASD in understanding deep word meanings, resulting in weaknesses in non-literal language comprehension and vocabulary acquisition. Horvath et al. (2018) confirm the above findings, as they observed difficulties in understanding and generalizing abstract words, as well as in understanding and recalling multiple meanings of a word. The research of Eigsti et al. (2007) showed that the vocabulary of children with ASD includes neologisms, idiomatic words or even meaningless words, which children use for communicative purposes. Naigles and Tek (2017) observed that most children with ASD are able to acquire the rules of speech form more easily than meaning. Furthermore, Kambanaros et al. (2019) investigated children with ASD with low language skills and evaluated the comprehension and production of compound words with two constituents. The results showed that children with ASD were able to identify the two constituents that compose a compound word. However, they presented difficulty in understanding the meaning of this word, and even in their attempts to explain the meaning of compound words, they produced semantically incomprehensible responses. Moreover, Auza-Benavides et al. (2024) found that children with ASD exhibit difficulties in expressive vocabulary, with the greatest difficulties presented in verbs and functional words. Similar findings regarding weaker expressive vocabulary in children with ASD in comparison to their TD peers were also reported by Seol et al. (2014). Furthermore, Liu et al. (2025) found an absence of the use of numerals and a low use of pronouns, while interjections seem to be the most frequent word category in the expressive vocabulary of children with ASD. On the other hand, children with DLD have limited lexical knowledge and poorer vocabulary as compared to their TD peers (Adams, 2002; Mengisidou et al., 2020), while learning nouns is considered easier for them than learning verbs (Alt et al., 2004). Dockrell et al. (2007) attribute deficits in semantics to limited knowledge of word morphology and to difficulty to producing compound words. Greater difficulties are found in words expressing abstract meanings and in multisyllabic words, and they also have difficulty in quickly and accurately recalling words that have an abstract meaning or are of low frequency. Furthermore, poorer expressive vocabulary and low word production have been observed in children with DLD in comparison to their TD peers (Auza-Benavides et al., 2024; Seol et al., 2014) while Jackson et al. (2021) found that children with DLD showed difficulties in expressive vocabulary tasks, such as naming and describing newly learnt words. In addition, difficulties have been reported in understanding proverbs and figurative and ironic expressions (Aslanoglou et al., 2023).

Language difficulties are the main feature of DLD and, as has already mentioned above, language deficits are also frequently found in individuals with ASD (Roberts et al., 2004; Tager-Flusberg and Joseph, 2003). Therefore, many studies have focused on the similarities between children with DLD and children with ASD at all levels of language development (Andreou et al., 2022; Leyfer et al., 2008; Ramírez-Santana et al., 2019). As a result of these similarities, the hypothesis of a common phenotype was set. This hypothesis argues that DLD and ASD are related and are probably different expressions of the same core cause or different parts of a continuum of the same disorder (Bishop, 2010; Kjelgaard and Tager-Flusberg, 2001; Tager-Flusberg and Joseph, 2003).

Research investigating language development through language tests, as well as studies that use neuroimaging and genetic methods, have been conducted in order to clarify the relationship between the two disorders. More specifically, the study of Herbert et al. (2005) regarding ASD and DLD showed similar patterns of asymmetry in the cerebral cortex, while both groups were more similar to each other than to the control group. Furthermore, it was observed that the right-asymmetrical region of the brain was more strongly developed in ASD compared to DLD, but both groups had significantly more right-handed asymmetry compared to the control group. Hodge et al. (2010) studying ASD and DLD found deficits in both populations regarding working memory, attention, language processing and motor control. Regarding language phenotype, Taylor and Whitehouse (2016) observed that children with language disorders met criteria for ASD, supporting the hypothesis that there is phenotypic overlap between ASD and DLD. They therefore concluded that the two disorders may be aspects of a single continuum, manifesting deficits in different domains. Voulgaraki (2023) found, also a high probability of autistic symptomatology in DLD, while, Leyfer et al. (2008) showed

that a significant proportion of children with DLD met the criteria to receive a diagnosis of ASD. In addition, Félix et al. (2024) in a review on the similarities between the two disorders, suggested that the differences in the language development of children with DLD and children with ASD were more evident during preschool age, but these differences decreased during school age.

On the contrary, several studies have identified significant differences between the two disorders in terms of language development (Creemers and Schaeffer, 2022; Schaeffer, 2018; Williams et al., 2008). Tager-Flusberg (2006) claimed that language impairments cannot serve as the sole criterion to differentiate the two disorders, as these commonalities in language impairments of DLD and ASD reflect only apparent distinctions. Therefore, it is not possible to draw definitive conclusions regarding the potential for a shared etiology derived from these similarities (Whitehouse et al., 2007).

However, the relationship between ASD and DLD has not yet been clarified, while various hypotheses have been put forward in recent years as an attempt to develop a well-defined diagnostic criterion that would separate the two disorders (Bishop and Norbury, 2002; Kjelgaard and Tager-Flusberg, 2001; Roberts et al., 2004).

Therefore, based on the above, the aim of the present study is to investigate structural language of children with ASD and DLD in a Greek-speaking population, and compare their abilities and weaknesses in this domain.

The specific objectives set for the present investigation are the following: (a) to assess and compare the performance of Greek-speaking children with ASD, DLD and typically developing children in structural language, namely phonology, morphosyntax and expressive vocabulary; (b) to identify specific areas of structural language strength and weakness in ASD and DLD; (c) to determine whether children with ASD and DLD exhibit overlapping or distinct structural language profiles.

The following research hypotheses were set:

Research hypothesis 1: Children with DLD are expected to perform lower on all tests that examine structural language than typically developing children.

Research Hypothesis 2: Children with ASD are expected to perform lower on all tests examining structural language than typically developing children.

Research Hypothesis 3: Children with ASD are expected to perform higher on tests examining phonology and morphosyntax than children with DLD.

Research Hypothesis 4: Children with ASD are expected to perform similarly to children with DLD on the test examining expressive vocabulary.

2 Materials and methods

2.1 Study design

The approach used for this research was quantitative, non-intrusive and cross-over. More specifically, the research followed a comparative approach to examine structural language skills in different clinical groups (ASD and DLD) and typically developing (TD) children. Language performance was assessed through standardized tests and results were compared using statistical analyses of group means. In addition, a sampling research design was followed (Creswell, 2011). The sample was collected using the non-probability sampling technique, and more specifically using convenience sampling (Etikan et al., 2016).

2.2 Participants

The research was carried out in two phases. In the first phase, the sample simulation was conducted, and in the second phase the main research was carried out, in which participants of typical development (*mean age* = 84.09 months; *SD* = 6.72), participants with ASD (*mean age* = 84.19 months; *SD* = 6.55) and participants with DLD aged 6-8 years (*mean age* = 84.09 months; *SD* = 6.72) were evaluated.

The main research involved 75 students who were divided into two clinical groups and a control group. The first clinical group consisted of 25 children with ASD (17 boys and 8 girls), while the second group consisted of 25 children (17 boys and 8 girls) with official diagnoses including characteristics that consist the linguistic profile of DLD. The control group consisted of 25 typically developing (TD) children (17 boys and 8 girls). Inclusion criteria for all participants concerned age, non-verbal intelligence and language. To confirm comparability and internal validity of the sample, both non-verbal intelligence and Mean Length of Utterance(w) (MLUw) were measured to ensure that the participants belonged to the developmental category declared. In addition, the Greek version of the Children's Communication Checklist (CCC-2, Georgiou and Spanoudis, 2021; Bishop, 2003) was completed by the parents of all participants in order to further confirm DLD and ASD diagnoses.

The selection of participants of clinical groups was based on the current diagnostic framework in Greece, as implemented by the competent public institutions (e.g., Center for interdisciplinary assessment—counseling and support, Medical and Pedagogical Centers), which assign diagnoses according to official taxonomic systems (e.g., ICD-10, DSM-5) using relevant psychometric tools. However, for research validity, we ensured that only children with a clear and distinct diagnostic profile, without comorbidities, according to health professionals' reports and relevant assessments, were included in this study. To ensure relative homogeneity in language profiles, we included only children without intellectual disability, with no history of minimally verbal development and verbal language was the primary mode of communication.

Thus, the inclusion criteria for all participants in all groups were as follows: (a) monolingual Greek speakers; (b) aged 6–8 years; (c) non-verbal intelligence 85 and above. Participants in all groups were matched for age and sex.

Regarding the TD group, participants came from primary schools. Also, (a) their MLU(w) was greater than or equal to 6.0 (Rice et al., 2010); (b) General Communication Composite and Social-Interaction Deviance Composite based on the CCC-2 were above 55 and above 8 respectively (Norbury et al., 2004); (c) they did not have any special educational needs; (d) they performed well in the subjects of the school curriculum overall, according to their teachers.

10.3389/feduc.2025.1641303

TABLE 1 Socio-demographic profile of mothers in the sample (%)

Parental educational level	Percentage %	Parental occupation	Percentage %
Primary Education	10.7	Public sector employees	28
Secondary Education	50.7	Private sector employees	24
University degree (Bachelor)	36	Self-employed/ Farmers/ Business owners	16
Postgraduate degree (Master/PhD)	2.7	Unemployed/ Homemakers/ Other	32

Regarding the ASD group, participants came from special kindergartens, primary schools and special education centers and had an official diagnosis for ASD, High Functioning Autism, Non-typical Autism. They also followed a needs-based intervention program. High-functioning autism and non-typical autism were subsumed under the ASD label provided that participants met the language criteria set in the study. Specifically, (a) Mean Length of Utterance (MLUw) (Rice et al., 2010) was greater than or equal to 4.0; (b) General Communication Composite and Social-Interaction Deviance Composite based on the CCC-2 were below 55 and below 8 respectively (Norbury et al., 2004).

Regarding the DLD group, participants came from primary schools and special education centers and had an official diagnosis for language disorders with characteristics that compose the DLD profile. They also followed a needs-based intervention program. Inclusion followed ICD-11 criteria and required a diagnosis based on persistent language difficulties not explained by other neurodevelopmental conditions. To reduce internal heterogeneity, only children with combined expressive and receptive disorders were included. Additionally, the following were also taken into account: (a) Mean Length of Utterance (MLUw) (Rice et al., 2010) was greater than or equal to 4.0; (b) General Communicative Composite and Social-Interaction Deviance Composite based on the CCC-2, were below55 and above 8 respectively (Norbury et al., 2004).

The exclusion criteria were as follows:

Participants were excluded from the study if (a) they did not meet the age, language or cognitive inclusion criteria; (b) they had comorbidity with other developmental or mental disorders; (c) they had a history of neurological impairment or severe sensory impairment; (d) they had articulation or voice disorders that could affect the results. In addition, for both clinical groups, participants were excluded if they did not have a diagnosis of the disorder (ASD, DLD) from an official public agency.

In terms of socio-demographic characteristics, all participants lived in small towns and villages in Greece. The families had similar socioeconomic and educational levels, as shown in the questionnaires completed by the participants' parents. Therefore, we consider the sample to be relatively homogeneous in terms of socioeconomic background, which limits the influence of possible confounding factors (Table 1).

2.2 Measures

For the purposes of this study, measures were initially administered for the selection and matching of the sample, followed by the measures used for the main research.

2.2.1 Measures for selection and matching of the sample

MLU(w) and non-verbal intelligence were measured, while the CCC-2 was completed by the parents of participants in order to simulate the sample. The number of morphemes or words that children use in each spontaneous expression is one of the most reliable indicators of language acquisition and is called Mean length of Utterance (MLU) (Ezeizabarrena and Garcia Fernandez, 2018). The MLU measurement demonstrates the language level at which the individual is at and is used to diagnose language disorders in children either as a measure to evaluate the effects of an intervention aimed at addressing language difficulties (Eisenberg et al., 2001; Tager-Flusberg et al., 2009) or as a factor for matching clinical groups in research studies (Rice et al., 2010). The MLU value is obtained by calculating the total number of produced morphemes (MLUm) or words (MLUw) divided by the total number of words in the utterance (Rice et al., 2010). For the present study conducted in Greek, the measurement of Mean Length of Utterance- Word (MLUw) was chosen, as in languages with high morphosyntactic complexity, such as the Greek language, it is considered more appropriate than the measurement of Mean Length of Utterance- Morpheme (MLUm) (Arif and Bol, 2008). For the purpose of the study, 50 sentences of the participants were collected through free discussion with them and divided by the words used in each utterance.

Additionally, to measure non-verbal intelligence, the Raven's Colored Progressive Matrices (CPM) (Raven et al., 1998) was used. Raven's Colored Progressive Matrices is addressed to children aged 4 to 12 years, is not influenced by cultural or linguistic factors, as no language responses are required and the verbal instructions given by the examiner are limited, and is standardized for the Greek population (Sideridis et al., 2015). In addition, Raven's Colored Progressive Matrices (Sideridis et al., 2015) showed high internal consistency (Cronbach's a = 0.90). Reliability was measured by the test-retest method and showed high levels of shared variance between the two measurements ($R^2 = 73.4\%$)

Finally, CCC-2 (Bishop, 2003) is a 70-question questionnaire completed by parents or teachers who have known the child for at least 6 months. The questionnaire has been used in order to identify different communication profiles in children with DLD and ASD (Andrés-Roqueta et al., 2021; Creemers and Schaeffer, 2022; Gorman et al., 2016; Schaeffer, 2018) as it has been shown to effectively differentiate between structural language disorders and pragmatic difficulties (Norbury et al., 2004). It concerns children aged 4 to 16 years. The 70 questions of the questionnaire are grouped into 10 subscales concerning: (A) Speech, (B) Syntax, (C) Semantics, (D) Coherence, (E) Inappropriate onset, (F) Stereotyped language, (G) Use of context, (H) Non-verbal communication, (I) Social relationships and (J) Interests and assess General Communication Competence and Social-Interaction Deviance

Composite. Respondents are asked to respond about how often they observe various linguistic, pragmatic and social behaviors in children. On the General Communicative Composite, typically developing children receive a score >55, while a score <55 probably indicates difficulties in structural language. Regarding, the Social-Interaction Deviance Composite, a negative score (<0) indicates pragmatic and social difficulties. In particular, according to Norbury et al. (2004), the performance of children with ASD does not exceed 8 points, while the performance of children with ASD on this scale is usually above 8. The internal consistency of the questionnaire is Cronbach' a = 0.80–0.87, while the reliability is r~0.80 (Norbury et al., 2004).

2.2.2 Measures of the main research

Regarding the main research, structural language was examined in all three groups by using standardized tests assessing their skills in phonological awareness, morphosyntax and expressive vocabulary. Phonological skills were assessed using subscales from the Test for the Detection and Investigation of Reading Difficulties (Porpodas, 2007). For the assessment of morphology and syntax, scales from the Diagnostic Test of Linguistic Intelligence (DTGL) for school-age children (Stavrakaki and Tsimpli, 2000) were administered, while expressive vocabulary was assessed using the lexical scale from WISC-V.

The Test for Detection and Investigation of Reading Difficulties (Porpodas, 2007) is a detective and investigative test that addresses reading difficulties during the most critical period of a child's age for the acquisition of reading. The purpose of this test is both to identify children who are likely to have difficulties in learning to read (when administered to kindergarten children) and to investigate the level of the individual main cognitive-linguistic factors of reading that are likely to be related to reading difficulties (when administered to children in the first two grades of primary school). The Test for Detection & Investigation of Diagnostic Difficulties (Porpodas, 2007) is standardized and it showed internal consistency (Cronbach's a = 0.60–0.79). It consists of 9 scales, some of which are administered only to primary school students, some of which are administered only to kindergarten students and some of which are administered to both age groups. The test may be administered as a whole or partially, depending on the circumstances of the test and the skills that need to be assessed. For the purposes of this study, the two scales assessing phonological awareness (phoneme segmentation and phoneme deletion) were administered. Each of the two scales consists of 24 pseudowords of increasing difficulty. The pseudowords have been selected in such a way that they include the phonemes of the Greek language in various simple combinations at the syllabic level. In the phoneme segmentation scale, pseudowords of two to seven phonemes and pseudowords of one, two or three syllables are included. The examiner reads out each pseudoword and the candidate is asked to break it down into phonemes by tapping his/her pencil on the table. The phoneme deletion scale includes pseudowords of one syllable, each of which contains between 2 and 4 phonemes. The examiner reads a pseudoword and the candidate is asked to delete either the initial or the final phoneme and to pronounce the remaining part of the pseudoword (excluding the deleted phoneme).

The Diagnostic Test of Language Intelligence is a test that assesses the language skills of children in terms of levels of language development, determining their language age and identifying cases of deviations from normal language development. The test assesses three levels of language performance: production, comprehension and repetition. The test is designed for school-age children (6–12 years old), testing the production of clitic and productive morphology and syntax, the understanding of meta-linguistic concepts, syntactic structures and thematic roles, text comprehension and repetition of syntactic structures (Stavrakaki and Tsimpli, 2000). In the present study, scales related to morphology and syntax were administered. Specifically, the following were used: (a) The scale of clitic morphology; (b) the scale of productive morphology; (c) the scale of syntax; (d) the scale of understanding morphology/syntax; and (e) the scale of recalling syntactic structures. This test was chosen because it is standardized, it is intended for the age groups studied in the present research and provides a full assessment of the morphological and syntactic language development of the examinees. The internal consistency of the test, as calculated through Cronbach's a, was 0.78, while the Guttman split-half reliability coefficient was 0.75, indicating good consistency of the individual questions. In terms of validity, Spearman's correlation coefficient was used between the performance of the groups in the different sections of the test. A statistically significant positive correlation was found ($p < 0.001$), supporting the structural validity of the instrument.

Finally, the WISC-V lexical scale was chosen because it examines the depth and breadth of vocabulary, while it is standardized to the Greek population. WISC-V demonstrated high reliability (internal consistency: subtests 0.79–0.91) and validity through confirmatory factor analysis, supporting the factorial structure of the test. It is also important to note that Weschler's scales have a dominant role in assessing special populations, such as individuals with language disorders and pervasive developmental disorders (Vogindroukas and Zikopoulou, 2009).

2.3 Procedure

For the implementation of this research, a certificate of approval was granted by the Internal Ethics Committee of the Department of Special Education of the University of Thessaly (protocol number 937), as well as the required permission from the Institute of Educational Policy and from the Ministry of Education, Religious Affairs and Sports separately from each competent Department (Special and General Education) (protocol number: Φ15/83926/ΕΚ/100881/Δ1). Additionally, permission was secured from the school principals for the researcher's access to public schools and the administration of research measures to participants. Furthermore, written consent was obtained from parents, both for their children's participation in the study and for the publication of the research findings.

The evaluation process included three visits to schools and special education centers. At the first meeting, the school principal or the special education center manager was informed about the aims of the study, the tests to be administered and the duration of each evaluation. After obtaining the consent of the teachers'

association, parents were informed and, if they agreed to their children's participation in the study, they signed a consent form and a form with their demographic data.

Then, the day and time of the next meeting was set, during which the Raven's Colored Progressive Matrices (CPM) test (Sideridis et al., 2015) was administered individually and the MLU(w) was measured through free discussion with the participants. For those students who met the criteria set in terms of non-verbal intelligence and MLU(w), a third session was set, during which the main research tests were administered. An additional session was scheduled for the administration of the main research test for participants who were observed to exhibit signs of fatigue.

Participants were assessed individually in a silent classroom in one or two sessions. The total duration of the assessment was approximately 1 h for each participant. Prior to administration, clear instructions in simple words and examples were given to the participant so that they could understand what each test was asking for. No additional information or assistance was given during the administration. However, comments were made encouraging the participation of children, and at the end of each test, the participants were asked if they would like to take a short break. At the end of the assessment, the score that the participants obtained on each test was calculated and standard scores were calculated for the tests where this was possible.

Statistical analysis was carried out using the statistical analysis software Statistic Package for Social Sciences (SPSS v25). Initially, descriptive analyses of the quantitative data were conducted and the Mean Score (M), the Standard Deviation (SD) and the range of variables were calculated. Subsequently, a normality test (One Sample Kolmogorov Smirnov Test) was carried out separately for each group of children (TD, ASD, DLD) in order to check whether the variables met the conditions of normal distribution or not, in order to select the appropriate statistical test to determine the presence or absence of statistical significance in the research data. Specifically, as the variables did not meet the conditions of normal distribution, the non-parametric Kruskal-Wallis test was chosen to compare the performance of the three groups. In addition, pairwise comparisons of the groups were then made to check for statistical significance in each pair. For this purpose, variables that met normal distribution were subjected to the parametric *t*-test, while those that did not meet normal distribution were subjected to the non-parametric Mann-Whitney *U* test.

3 Results

The present study involved 75 children, divided into three groups; 25 children with ASD; 25 children with DLD; and 25 children of typical development. The groups were matched based on the age and sex of the participants. Each group consisted of 25 children aged 6–8 years (mean = 84.10 months), while 68% of the participants were boys ($N = 17$) and 32% were girls ($N = 8$).

The mean age of participants with TD was 84.09, the mean age of participants with DLD was 84.09 and the mean age of participants with ASD was 84.19. In terms of non-verbal intelligence, children with TD had a mean score of 89.40, children with DLD had a mean score of 88.20 and children with ASD had a mean score of 89.80. Finally, regarding MLU(w), children with TD had higher mean score ($M = 6.56$) than children with DLD ($M = 4.37$) and children with ASD ($M = 5.79$) (Table 2).

TABLE 2 Means and Standard Deviations of age, non-verbal intelligence and MLUw of the participants.

		TD[a] ($n = 25$)		ASD[b] ($n = 25$)		DLD[c] ($n = 25$)	
		M	SD	M	SD	M	SD
Age (months)		84.09	6.716	84.19	6.55	84.09	6.72
Non-verbal intelligence		89.40	3.905	89.80	4.20	88.20	4.54
MLU(w)[d]		6.56	0.25	5.79	0.82	4.37	0.10
CCC-2[e]	GCC	77.08	4.85	37.52	11.38	35.40	3.81
	SIDC	6.36	9.35	−7.08	9.35	7.16	1.16

[a]TD = typical development.
[b]ASD = autism spectrum disorder.
[c.] DLD = developmental language disorder.
[d]MLUw = mean length of utterance (word).
[e]CCC-2 = children's communication checklist; GCC = general communication composite; SIDC = social-interaction deviance composite.

Kruskal-Walli's analysis was performed to compare the three groups with each other. Table 3 presents the descriptive data on the performance of the children in the three groups on structural language, as well as the statistical significance between the groups on this measurement. Specifically, the results of the study as presented in Table 3, showed that there was a statistically significant difference in phonological awareness between the three groups (TD $M = 20.80$, $SD = 3.12$; DLD $M = 12.32$, $SD = 3.86$; ASD $M = 18.20$, $SD = 4.01$). Also, there was a statistically significant difference regarding morphosyntax between the three groups (TD $M = 82.96$, $SD = 4.23$; DLD $M = 31.00$, $SD = 7.84$; ASD M= 45.88, SD= 13.54). Finally, there was a statistically significant difference concerning expressive vocabulary (TD $M = 11.24$, $SD = 1.67$; DLD $M = 6.52$, $SD = 1.78$; ASD $M = 7.84$, $SD = 2.84$). As presented above analytically statistical analysis showed a statistically significant difference between the three groups on all measures related to structural language ($p < 0.001$) with the clinical groups exhibiting lower performance than the control group (TD children). A comparative representation of the scores of participants across the three measurements and between the groups is presented in Figure 1.

In addition, pairwise comparisons between groups were performed using independent samples *t*-tests and Mann-Whitney U tests, depending on the normality of the data.

Regarding the first hypothesis, children with DLD are expected to perform lower on all tests that examine structural language than typically developing children. The results indicated lower performance for DLD group across all measurements. Specifically, statistically significant differences were observed regarding phonology (TD $M = 20.80$, $SD = 3.12$; DLD $M = 12.32$, $SD = 3.86$), morphosyntax (TD $M = 82.96$, SD=.23; DLD $M = 31.00$, $SD = 7.84$), and expressive vocabulary (TD $M = 11.24$, $SD = 1.67$; DLD $M = 6.52$, $SD = 1.78$) (Table 4).

Regarding the second hypothesis, children with ASD are expected to perform lower on all tests examining structural language than typically developing children. The comparison of the

TABLE 3 Performance of the TD, DLD and ASD groups in structural language.

	TD[a] ($n = 25$)		DLD[b] ($n = 25$)		ASD[c] ($n = 25$)		p	H
	M	SD	M	SD	M	SD		
Phonology	20.80	3.12	12.32	3.86	18.20	4.01	**<0.001*** **	37.63
Morphosyntax	82.96	4.23	31.00	7.84	45.88	13.54	**<0.001*** **	57.07
Expressive vocabulary	11.24	1.67	6.52	1.78	7.84	2.84	**<0.001*** **	39.52

[a]TD = typical development.
[b]DLD = Developmental language disorder.
[c]ASD = autism spectrum disorder.
*** $p < 0.001$.

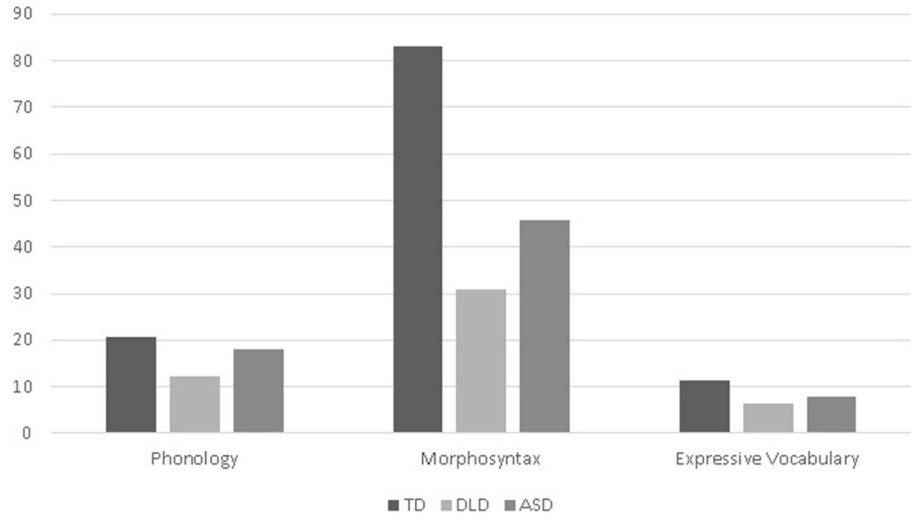


FIGURE 1
Performance of the ASD, DLD and TD groups on tests examining structural language.

TD group and the group with ASD revealed lower performance for children with ASD, while also showed statistically significant differences between the performance of the two groups, as shown in Table 5. Specifically, in the performance of Morphosyntax (TD $M = 82.96$, SD = 4.23; ASD M = 45.88, SD = 13.54) and expressive vocabulary (TD $M = 11.24$, $SD = 1.67$; ASD $M = 7.84$, $SD = 2.84$) the statistical significance was $p < 0.001$, while that of Phonological Awareness was $p = 0.02$ (TD $M = 20.80$, $SD = 3.12$; ASD $M = 18.20$, $SD = 4.01$).

According to the third hypothesis, children with ASD are expected to perform higher on tests examining phonology and morphosyntax than children with DLD. As shown in Table 6, the comparison of the two clinical groups also showed statistically significant differences in phonology and morphosyntax with the group with ASD performing higher than the group with DLD. More specifically, the measure for Phonological Awareness (DLD $M = 12.32$, $SD = 3.86$; ASD $M = 18.20$, $SD = 4.01$) and Morphosyntax (DLD $M = 31.00$, $SD = 7.84$; ASD $M = 45.88$, $SD = 13.54$) gave a statistical significance of $p < 0.001$.

Finally, as regards the fourth hypothesis, children with ASD are expected to perform similarly to children with DLD on the test examining expressive vocabulary. The results indicated that ASD group performed higher than the DLD group (DLD $M = 6.52$, $SD = 1.78$; ASD $M = 7.84$, $SD = 2.84$) with a statistical significance of $p = 0.03$.

4 Discussion

The purpose of this study was to investigate and compare abilities and weaknesses in structural language of Greek-speaking children with ASD and DLD. The specific aims set were to identify the strengths and deficits of children with ASD and children with DLD in structural language (phonology, morphosyntax and expressive vocabulary) as compared to typically developing children in this area, and to compare the performance of the two clinical groups in structural language skills. The results of this study indicate that DLD population performed poorer in almost all measurements compared to ASD population. In addition, both children with ASD and children with DLD exhibited impairments in structural language as compared to TD children.

Our first research hypothesis states that children with DLD will present lower performance on all tests that examine structural language than typically developing children. The results of the study showed deficits for the DLD group as compared to their TD peers on phonology, morphosyntax and expressive vocabulary and

TABLE 4 Comparison of the DLD and TD groups regarding structural language.

	DLD[a] ($n = 25$)		TD[b] ($n = 25$)		p	U-test	t
	M	SD	M	SD			
Phonology	12.32	3.86	20.80	3.12	**<0.001*** **	24.00	
Morphosyntax	31.00	7.84	82.96	4.23	**<0.001*** **	0.000	
Expressive Vocabulary	6.52	1.78	11.24	1.67	**<0.001*** **		9.68

[a]DLD = developmental language disorder.
[b]TD = typical development.
***p < 0.001.

TABLE 5 Comparison of the ASD and TD groups regarding structural language.

	ASD[a] ($n = 25$)		TD[b] ($n = 25$)		p	U-test
	M	SD	M	SD		
Phonology	18.20	4.01	20.80	3.12	**0.02***	188.00
Morphosyntax	45.88	13.54	82.96	4.23	**<0.001*** **	0.00
Expressive Vocabulary	7.84	2.84	11.24	1.67	**<0.001*** **	92.50

[a]ASD = autism spectrum disorder.
[b]TD = typical development.
*p ≤ 0.05
***p < 0.001.

TABLE 6 Comparison of the ASD and DLD groups regarding structural language.

	ASD[a] ($n = 25$)		DLD[b] ($n = 25$)		P	U-test
	M	SD	M	SD		
Phonology	18.20	4.01	12.32	3.86	**<0.001*** **	88.00
Morphosyntax	45.88	13.54	31.00	7.84	**<0.001*** **	99.50
Expressive Vocabulary	7.84	2.84	6.52	1.78	**0.03***	19.50

[a]ASD = autism spectrum disorder.
[b]DLD = developmental language disorder.
*p ≤ 0.05.
***p < 0.001.

are consistent with previous studies examining structural language in the DLD population. More specifically, Botting (2020) found that children with DLD have difficulties in structural language, while the development of structural language is slower in the DLD population as compared to typically developing children. Similarly, Andrés-Roqueta et al. (2021) found that children with DLD had more structural language difficulties than their TD peers. Regarding the phonological development of the DLD population, Georgiou and Theodorou (2023) investigating Greek-speaking children with DLD, concluded that children with DLD had significant difficulties in discriminating voicing contrasts, while Moraleda-Sepúlveda et al. (2022), who investigated the phonological awareness of children with DLD, found that they had difficulties in both phonemic and syllabic awareness. In a recent study in the Greek language, Georgiou and Theodorou (2025) presented evidence that young children with DLD exhibited more phonological difficulties than children with TD, while Mengisidou and Marshall (2019) suggested that children with DLD show clear deficits in phonological processing skills, particularly in phonological processing and phonological representations.

In terms of morphological and syntactic development, Stanford and Delage (2020) concluded that children with DLD had more deficits in morphosyntactic skills than their TD peers, while Deevy and Leonard (2018) claimed that children with DLD exhibited weaker knowledge of tense/agreement forms in their speech. Our results also align with those of Abdalla and Mahfoudhi (2023) who found morphological deficits in children with DLD as compared to a language-matched TD group and a chronologically age-matched TD group, as children with DLD presented difficulties regarding the correct use of third-person verb agreement. In addition, Calder et al. (2022) revealed the presence of difficulties in morphosyntax in children with DLD, while Georgiou and Theodorou (2023) found that children with DLD exhibited deficits also in grammar.

Finally, concerning vocabulary of the DLD population, our findings agree with those of Pijnacker et al. (2017) who showed that the expressive vocabulary of children with DLD is deficient as compared to that of their TD peers. Similarly, Jackson et al. (2021) observed a clear deficit in expressive vocabulary and in word learning in children with DLD, a finding that has also been identified by Ghawi-Dakwar and Saiegh-Haddad (2024) who investigated word learning in Arabic-speaking children with DLD. Difficulties in the vocabulary of children with DLD were also noted by Sandgren et al. (2021) studying their lexical knowledge in comparison to typically developing peers. Therefore, the first research hypothesis was confirmed, reinforcing the findings of previous studies that documented deficits in phonology, morphology, syntax and vocabulary in the DLD population.

The second hypothesis states that children with ASD are expected to perform lower on all tests examining structural language than typically developing children. The results showed that children with ASD exhibited more deficits in structural language as compared to their TD peers and therefore are in line with previous research investigating the structural language of children with ASD. In particular, Boo et al. (2022) found that children with ASD demonstrated lower complexity in structural language as compared to children with typical development. Regarding phonological abilities, Zarokanellou et al. (2023) concluded that Greek-speaking children with ASD produced more errors in their speech than typically developing children, and deficits were observed in terms of phonological representations. In addition, Alnemr (2022) showed that children with ASD presented difficulties in phonological awareness, while Dynia et al. (2019)

also observed impaired phonological awareness skills in children with ASD.

Regarding morphosyntactic abilities of ASD population, our findings agree with those of Meir and Novogrodsky (2020) who found deficits in the syntactic skills of children with ASD and those of Durrleman et al. (2017b) who showed difficulties in morphosyntax for children with ASD. In addition, Ramírez-Santana et al. (2019) observed morphosyntactic deficits in children with ASD, while Al-Hassan and Marinis (2021) found that children with ASD exhibited deficits in grammatical abilities as compared to their TD peers.

In terms of vocabulary, our results align with those of Liu et al. (2025) who found that children with ASD exhibited lower overall language production and weaker expressive vocabulary that their TD peers. Kover et al. (2013) found that children with ASD showed deficits in terms of perceptual vocabulary as compared to their TD peers. Deficits in the semantic development of children with ASD were also highlighted by Di Stefano et al. (2019) in their research. In addition, Hart and Curtin (2023) observed that children with ASD exhibited slower developmental trajectories of vocabulary and showed significant differences as compared to children with typical development. Thus, the second research hypothesis was confirmed strengthening existing research on phonological, morphosyntactic and semantic difficulties of children with ASD.

Our third hypothesis was that children with ASD would perform higher on tests examining phonology and morphosyntax than children with DLD. The results showed that children with DLD exhibited more deficits in phonology and morphosyntax than children with ASD. The results of the present study are consistent with those of Riches et al. (2011) who found differences in the language skills between children with DLD and children with ASD and they claim that the two populations have different language profile, but do not exclude a small degree of overlap. Moreover, de la Torre Carril et al. (2021) found that the structural language of school-aged children with DLD was less developed than that of age-matched children with ASD. More specifically, the results are in agreement with previous research that studied the phonological developmental in children with ASD in comparison to children with DLD. Ramírez-Santana et al. (2019) concluded that children with DLD exhibited more profound phonological deficits. Same results were found by Hill et al. (2015), investigating phonological skills, as they observed differences between ASD and DLD in terms of verbal memory and pseudoword repetition, possibly indicating that different mechanisms are involved in language learning in the two disorders. Different cognitive profiles, which probably also affect language skills, are also suggested by Taylor et al. (2014). In addition, the review by Wolk et al. (2016) showed that findings on the phonological development of children with ASD are conflicting and they concluded that some children with ASD use typical phonological processes, while others do not show phonological deficits.

In terms of morphosyntactic abilities, the results of the present study are in line with Craig and Trauner (2017), who showed significant differences between the two groups, with individuals with DLD making more grammatical errors than children with ASD. Additionally, Sukenik and Friedmann (2018) investigated the syntactic skills of children with ASD and DLD and, although the results revealed similar performance in terms of overall scores for children in the two groups across all tests, however, when examined in terms of the type and form of incorrect responses made by participants, significant differences were found between the two groups. Specifically, the two groups differed in terms of the type of syntactic errors as well as in terms of consistency in performance. The errors of the children with DLD were observed in specific sentence types, whereas the errors of the children with ASD were not consistent in terms of sentence types. Furthermore, Creemers and Schaeffer (2022) found differences between children with ASD and children with DLD on tests that examined grammar, while Schaeffer (2018) did not find any similarities in the structural language of children with ASD and children with DLD and concluded that there is no overlap in the language profile of the two disorders.

On the other hand, the results of the present study are not in line with Huang and Finestack (2020), who investigated the morphosyntactic skills of children with DLD and children with ASD and did not find any differences between them. The differences in the findings between our study and those of the study of Huang and Finestack (2020) could be attributed to the fact than in their study they only included children with ASD and language difficulties. In addition, Durrleman et al. (2017a) concluded that there are similarities in morphosyntax between individuals with DLD and individuals with ASD. However, in the study by Durrleman et al. (2017a), participants also had similar performances on tests that examined Theory of Mind, a skill that, according to Spanoudis (2016), is related to morphosyntactic development. Therefore, the third research hypothesis was confirmed, reinforcing the existing research on the phonological and morphosyntactic skills of the two clinical populations.

According to the fourth hypothesis, children with ASD are expected to perform similarly to children with DLD on the test examining expressive vocabulary. According to the results, the performance of children with ASD was not similar to that of children with DLD. This finding contradicts findings from previous studies examining the semantic abilities of the two clinical groups. In particular, Félix et al. (2024) suggested that there is an overlap of language phenotypes in terms of lexical knowledge. Similar conclusions were reached by Haebig et al. (2015) who found similarities in the lexical-semantic knowledge of children with ASD and children with DLD. In addition, McGregor et al. (2012) suggested similarities in semantic development between the two populations. Furthermore, de la Torre Carril et al. (2021) found similar performance on semantics between children with ASD and children with DLD. Whitehouse et al. (2008) also found similar performance between the two populations in terms of semantic skills, as did Georgiou and Spanoudis (2021) in their study in the Greek language. Nevertheless, despite the observed similarities, it has been found that children with ASD showed stronger performance in lexical depth, word associations, and structures as compared to children with DLD (Lloyd et al., 2006; Loucas et al., 2013; Manolitsi and Botting, 2011), while Bekmurat et al. (2024) suggested that there is variation in the semantic skills of children with ASD. Specifically, some children with ASD indicated high proficiency and other children with ASD presented lower levels of vocabulary. In addition, Auza-Benavides et al. (2024)

confirm the variation observed across these populations, while simultaneously they found distinct vocabulary profiles across DLD and ASD. Therefore, the fourth research hypothesis regarding expressive vocabulary of children with DLD and children with ASD was not confirmed.

The different results between our study and previous research concerning expressive vocabulary of children with DLD and children with ASD are probably due to the heterogeneity in the language profiles of the children of the two populations in our study, especially in the ASD group. The groups were matched for age and non-verbal intelligence. MLU(w) was measured in all participants to prevent large variations within groups, however there was no matching of participants in terms of MLU(w). Consequently, the ASD group had a higher mean score on MLU(w) than the DLD group, which probably affected the ASD group's performance on verbal knowledge.

The findings of the present study revealed impairments in structural language for both clinical populations as compared to their typically developing peers. In addition, considerable differences in their deficits were observed between children with ASD and children with DLD. Both clinical groups indicated deficits in phonology, morphosyntax and expressive vocabulary as compared to the TD group. However, children with DLD appear to have more deficits than children with ASD in terms of structural language. Importantly, the deficits of children with DLD and children with ASD in the Greek language are in line with research that has been conducted in other languages in previous studies.

The differences in language skills between ASD and DLD can be attributed to the fact that children with DLD usually show broader structural language impairments, whereas some children with ASD, particularly those without a co-occurring language disorder, may show relatively preserved structural language skills despite pragmatic difficulties. In the present study, comorbidity with another disorder was an exclusion criterion for both clinical groups.

Furthermore, according to the findings, the present study does not support the hypothesis of a common etiology or overlap between the two clinical populations. On the other hand, previous research has observed a common biological basis and genetic overlap between the neurodevelopmental disorders (Nisiotou and Vlachos, 2014; van Wijngaarden et al., 2024; Vernes et al., 2008). Consequently, investigating the language phenotype of ASD and DLD through the assessment of language skills is probably not able by itself to provide a response to the hypothesis regarding a common etiology between the two clinical populations.

4.1 Implications

Through the present study, an effort was made to strengthen the findings on structural language in ASD and DLD. The investigation of structural language in Greek populations with neurodevelopmental disorders, specifically in children with ASD and DLD, as well as the comparison of the two clinical groups, provides important data that can contribute to a clearer understanding of their language profiles. Specifically, due to the particular characteristics of the Greek language, the investigation of language skills in phonology, morphosyntax and expressive vocabulary offers a cross-linguistic perspective, allowing comparison with other languages and contributing to the understanding of the common and differentiated language difficulties faced by the ASD and DLD populations. This approach can enhance diagnosis and intervention, taking into account the linguistic specificities of each language.

More specifically, given that children with DLD performed significantly lower not only compared to typically developing peers, but also compared to children with ASD, assessment strategies aimed at detecting DLD should focus on key structural language domains, including phonological awareness, morphosyntactic abilities and expressive vocabulary. Hence, diagnostic protocols and measures for DLD should incorporate more accurate and sensitive tasks examining multiple levels of language structure, so that a detailed profiling of the children's language profile is possible for the early diagnosis of DLD and its differentiation from other neurodevelopmental disorders.

Additionally, given some common elements that ASD and DLD present in terms of language profile, their separation is often difficult, resulting in the diagnoses given being inaccurate and, thus, the interventions followed being inappropriate. Therefore, the findings of the present study can be used to create axes of observation or diagnostic tools for the assessment of language skills in the Greek language, while simultaneously suggesting guidelines for the differentiation of the two disorders, reducing the risk of diagnostic confusion.

In addition, the findings provide information which can be used for the design and implementation of interventions tailored to the language needs of each population. These interventions can enhance structural language, improving both comprehension and production of speech, while also contributing to the improvement of the communicative competence and social interaction of individuals with ASD and DLD.

4.2 Limitations and future research

There are some limitations in the present study that need to be considered. A limitation of the present study that the three groups of participants (ASD, DLD, and TD) were matched for chronological age and sex, but not matched for language abilities. In the study, participants' MLU(w) were measured to limit large language variations between groups and for both clinical groups, MLU(w) was defined as 4 or higher. However, no one-to-one matching of participants was applied, nor were children with ASD and language difficulties separated from those with ASD without language difficulties. Future research could include comparisons between four distinct groups: children with ASD and language difficulties, children with ASD without language difficulties, children with DLD, and typically developing children, allowing for a more detailed analysis of linguistic differences and similarities between them.

The small number and the limited age range of the participants in all three groups could also be considered as a limitation of the study. In general, small samples do not allow generalization of the findings which need to be replicated with larger samples in order to be confirmed. Further research on language development

in ASD and DLD populations could be conducted with a larger number of participants, so that the results are more reliable and generalizable. Also, a broadening of the age range of participants would allow the trajectory of language development to be observed at different developmental stages, providing valuable information on the differences between the two populations in language ability as they grow older. Finally, it would be useful to conduct comparative research with other neurodevelopmental disorders, in order to better understand the common and differentiated language difficulties observed in these populations.

Finally, a limitation of this research is the reliance on accuracy scores without qualitative analysis of error patterns. Qualitative differences in language behavior may reveal strategies or vulnerabilities of specific groups in structural language. Future research should incorporate qualitative analyses and, where feasible, dynamic assessment procedures to better capture underlying language profiles.

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 humans were approved by Ethics Committee of University of Thessaly. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

GA: Writing – review & editing, Writing – original draft, Supervision, Methodology. VL: Data curation, Methodology, Formal analysis, Conceptualization, Project administration, Writing – original draft, Software, Writing – review & editing. VA: Formal analysis, Data curation, Writing – original draft, Software, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Abdalla, F., and Mahfoudhi, A. (2023). Verb agreement production in Arabic-speaking children with developmental language disorder. *Lang. Acquis.* 31, 224–245. doi: 10.1080/10489223.2023.2231924

Adams, C. (2002). Practitioner review: the assessment of language pragmatics. *J. Child Psychol. Psychiatry* 43, 973–987. doi: 10.1111/1469-7610.00226

Aguilar-Mediavilla, E. M., Sanz-Torrent, M., and Serra-Raventós, M. (2002). A comparative study of the phonology of pre-school children with specific language impairment (SLI), language delay (LD) and normal acquisition. *Clin. Ling. Phonetics* 16, 573–596. doi: 10.1080/02699200210148394

Alexiadou, A., and Anagnostopoulou, E. (2000). Greek syntax: a principles and parameters perspective. *J. Greek Ling.* 1, 171–222. doi: 10.1075/jgl.1.09ale

Al-Hassan, M. A., and Marinis, T. (2021). "Sentence repetition in children with autism spectrum disorder in Saudi Arabia: an investigation of morphosyntactic abilities", in *Experimental Arabic Linguistics* (Amsterdanm: John Benjamins Publishing Company), 143–176. doi: 10.1075/sal.10.06alh

Alnemr, M. A. (2022). Profiles of phonological awareness skills in young children with autism spectrum disorders, language impairment, and their typically developing peers. *J. Posit. Sch. Psychol.* 6, 842–852.

Alt, M., Plante, E., and Creusere, M. (2004). Semantic features in fast-mapping: performance of preschoolers with specific language impairment versus preschoolers with normal language. *J. Speech Lang. Hear. Res.* 47, 407–420. doi: 10.1044/1092-4388(2004/033)

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edn.* Washington, DC: American Psychiatric Association. doi: 10.1176/appi.books.9780890425596

Andreou, G. (2012). *Language: Theoretical and Methodological Approach.* Athens: Pedio.

Andreou, G., and Aslanoglou, V. (2022). Written language production in children with developmental language disorder. *Front. Psychol.* 13:833429. doi: 10.3389/fpsyg.2022.833429

Andreou, G., Aslanoglou, V., Lymperopoulou, V., and Vlachos, F. (2023). Morphological and phonological errors in the written language production of children with DLD. *Eur. J. Spec. Needs Educ.* 39, 235–248. doi: 10.1080/08856257.2023.2207055

Andreou, G., Lymperopoulou, V., and Aslanoglou, V. (2022). Developmental language disorder (DLD) and autism spectrum disorder (ASD): similarities in

pragmatic language abilities. A systematic review. *Int. J. Dev. Disabilities* 1–15. doi: 10.1080/20473869.2022.2132669

Andrés-Roqueta, C., Garcia-Molina, I., and Flores-Builes, R. (2021). Association between CCC-2 and structural language, pragmatics, social cognition, and executive functions in children with developmental language disorder. *Children* 8:123. doi: 10.3390/children8020123

Apel, K., Diehm, E., and Apel, L. (2013). Using multiple measures of morphological awareness to assess its relation to reading. *Topics Lang. Disord.* 33, 42–56. doi: 10.1097/TLD.0b013e318280f57b

Arif, H., and Bol, G. W. (2008). Counting MLU in morphemes and MLU in words in a normally developing child and child with a language disorder: a comparative study. *Dhaka Univ. J. Ling.* 1, 167–182.

Aslanoglou, V., Andreou, G., Vlachos, F., and Lymperopoulou, V. (2023). "Writing skills of individuals with developmental language disorder (DLD)", in *Developmental Language Disorders in Childhood and Adolescence*, ed. D. Katsarou (Hershet, PA: IGI Global), 323–337. doi: 10.4018/979-8-3693-0644-4.ch024

Auza-Benavides, A., Márquez-Caraveo, M. E., Murata, C., and Perez-Barron, V. (2024). Profiles of early expressive vocabulary in children with typical and atypical language development. *Front. Commun.* 9:1368076. doi: 10.3389/fcomm.2024.1368076

Barokova, M., and Tager-Flusberg, H. (2020). Commentary: measuring language change through natural language samples. *J. Autism Dev. Disord.* 50, 2287–2306. doi: 10.1007/s10803-018-3628-4

Bekmurat, A., Yessengulova, M., Mirzakhmedova, U., Ayaganova, A., and Kulbayeva, B. (2024). Evaluating the lexico-semantic aspects of the spoken language of preschool aged children with autism spectrum disorder (ASD). *Eur. J. Appl. Ling.* 10, 126–137. doi: 10.32601/ejal.10211

Berninger, V. W., Abbott, R. D., Nagy, W., and Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *J. Psycholing. Res.* 39, 141–163. doi: 10.1007/s10936-009-9130-6

Bishop, D.V. (2003). *The Children's Communication Checklist-2*. San Antonio, TX: Psychological Corporation.

Bishop, D. V. (1997). *Uncommon Understanding: Development and Disorders of Language Comprehension in Children*. San Antonio, TX: Psychology Press.

Bishop, D. V. (2010). Overlaps between autism and language impairment: phenomimicry or shared etiology?. *Behav. Genet.* 40, 618–629. doi: 10.1007/s10519-010-9381-x

Bishop, D. V., and Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: a study using standardised diagnostic instruments. *J. Child Psychol. Psychiatry* 43, 917–929. doi: 10.1111/1469-7610.00114

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T. and Catalise-2 Consortium (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: terminology. *J. Child Psychol. Psychiatry Allied Disciplines* 58, 1068–1080. doi: 10.1111/jcpp.12721

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., and The CATALISE Consortium (2016). CATALISE: a multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS One* 11:e1986v1. doi: 10.7287/peerj.preprints.1986v1

Boo, C., Alpers-Leon, N., McIntyre, N., Mundy, P., and Naigles, L. (2022). Conversation during a virtual reality task reveals new structural language profiles of children with ASD, ADHD, and comorbid symptoms of both. *J. Autism Dev. Disord.* 52, 2970–2983. doi: 10.1007/s10803-021-05175-6

Botting, N. (2020). Language, literacy and cognitive skills of young adults with developmental language disorder (DLD). *Int. J. Lang. Commun. Disord.* 55, 255–265. doi: 10.1111/1460-6984.12518

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

Calder, S. D., Claessen, M., Leitão, S., and Ebbels, S. (2022). A profile of expressive inflectional morphology in early school-age children with developmental language disorder. *Clin. Ling. Phonetics* 36, 341–358. doi: 10.1080/02699206.2021.1931454

Carlisle, J. (1995). "Morphological awareness and early reading achievement," in *Morphological Aspects of Language Processing*, ed. L. B. Feldman (Mahwah, NJ: Lawrence Erlbaum Associates), 189–209.

Conti-Ramsden, G., St Clair, M. C., Pickles, A., and Durkin, K. (2012). Developmental trajectories of verbal and nonverbal skills in individuals with a history of specific language impairment: from childhood to adolescence. *J. Speech, Lang. Hear.* 55, 1716–1735. doi: 10.1044/1092-4388(2012/10-0182)

Craig, M., and Trauner, D. (2017). Comparison of spontaneously elicited language patterns in specific language impairment and high-functioning autism. *Pediatr. Neurol.* 79, 53–58. doi: 10.1016/j.pediatrneurol.2017.09.017

Crandall, M. C., Bottema-Beutel, K., McDaniel, J., Watson, L. R., and Yoder, P. J. (2019). Children with autism spectrum disorder may learn from caregiver verb input better in certain engagement states. *J. Autism Dev. Disord.* 49, 3102–3112. doi: 10.1007/s10803-019-04041-w

Creemers, A., and Schaeffer, J. C. (2022). Grammatical and pragmatic properties of the DP in children with specific language impairment (SLI) and in children with high functioning autism (HFA). *Ling. Netherlands* 32, 16–32. doi: 10.1075/avt.32.02cre

Creswell, J. W. (2011). *Educational Research: Planning, Conducting, and Evaluating Quantative and Qualitative Research*. Boston, MA: Pearson Education International.

de la Torre Carril, A., Durán-Bouza, M., and Pérez-Pereira, M. (2021). Capacity of the CCC-2 to discriminate ASD from other neurodevelopmental disorders. *Children* 8:640. doi: 10.3390/children8080640

Deevy, P., and Leonard, L. B. (2018). Sensitivity to morphosyntactic information in preschool children with and without developmental language disorder: a follow-up study. *J. Speech Lang. Hear. Res.* 61, 3064–3074. doi: 10.1044/2018_JSLHR-L-18-0038

Di Stefano, C., Senturk, D., and Jeste, S. S. (2019). ERP evidence of semantic processing in children with ASD. *Dev. Cognit. Neurosci.* 36:100640. doi: 10.1016/j.dcn.2019.100640

Dockrell, J.E., Lindsay, G., Connelly, V., and Mackie, C. (2007). Constraints in the production of written text in children with specific language impairment. *Exceptional Child.* 73, 147–164. doi: 10.1177/001440290707300202

Dore, J., Franklin, M.B., Miller, R.T., and Ramer, A.L. (1976). Transitional phenomena in early language acquisition. *J. Child Lang.* 3, 13–28. doi: 10.1017/S0305000900001288

Durrleman, S., Burnel, M., and Reboul, A. (2017a). Theory of mind in SLI revisited: links with syntax, comparisons with ASD. *Int. J. Lang. Commun. Disord.* 52, 816–830. doi: 10.1111/1460-6984.12317

Durrleman, S., Delage, H., Prévost, P., and Tuller, L. (2017b). The comprehension of passives in autism spectrum disorder. *Glossa* 2, 1–30. doi: 10.5334/gjgl.205

Dynia, J. M., Bean, A., Justice, L. M., and Kaderavek, J. N. (2019). Phonological awareness emergence in preschool children with autism spectrum disorder. *Autism Dev. Lang. Impairments* 4. doi: 10.1177/2396941518822453

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

Eisenberg, S. L., Fersko, T. M., and Lundgren, C. (2001). The use of MLU for identifying language impairment in preschool children: a review. *Am. J. Speech Lang. Pathol.* 10, 323–342. doi: 10.1044/1058-0360(2001/028)

Etikan, I., Musa, S. A., and Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *Am. J. Theor. Appl. Stat.* 5, 1–4. doi: 10.11648/j.ajtas.20160501.11

Ezeizabarrena, M. J., and Garcia Fernandez, I. (2018). Length of utterance, in morphemes or in words?: MLU3-w, a reliable measure of language development in early Basque. *Front. Psychol.* 8:2265. doi: 10.3389/fpsyg.2017.02265

Félix, J., Santos, M. E., and Benitez-Burraco, A. (2024). Specific language impairment, autism spectrum disorders and social (pragmatic) communication disorders: is there overlap in language deficits? A review. *Rev. J. Autism Dev. Disord.* 1, 86–106. doi: 10.1007/s40489-022-00327-5

Fromkin, V., Rodman, R. and Hyams, N. (2017). *An Introduction to Language, 11th Edn*. Boston, MA: Cengage Learning.

Georgiou, G. P., and Theodorou, E. (2023). Abilities of children with developmental language disorders in perceiving phonological, grammatical, and semantic structures. *J. Autism Dev. Disord.* 53, 4483–4487. doi: 10.1007/s10803-022-05548-5

Georgiou, G. P., and Theodorou, E. (2025). Voicing discrimination as a diagnostic marker of developmental language disorder. *J. Neuroling.* 73:101228. doi: 10.1016/j.jneuroling.2024.101228

Georgiou, N., and Spanoudis, G. (2021). Developmental language disorder and autism: commonalities and differences on language. *Brain Sci.* 11:589. doi: 10.3390/brainsci11050589

Ghawi-Dakwar, O., and Saiegh-Haddad, E. (2024). Word learning in Arabic diglossia in children with typical language development and developmental language disorder. *J. Speech Lang. Hear. Res.* 68, 1533–1551. doi: 10.1044/2024_JSLHR-23-00618

Girbau, D., and Schwartz, R.G. (2007). Nonword repetition in Spanish speaking children with specific language impairment (SLI). *Int. J. Lang. Commun. Disord.* 42, 59–75. doi: 10.1080/13682820600783210

Girolamo, T., Shen, L., Monroe Gulick, A., Rice, M. L., and Eigsti, I. M. (2024). Studies assessing domains pertaining to structural language in autism vary in reporting practices and approaches to assessment: a systematic review. *Autism* 28, 1602–1621. doi: 10.1177/13623613231216155

Gorman, K., Olson, L., Hill, A. P., Lunsford, R., Heeman, P. A., and van Santen, J. P. (2016). Uh and um in children with autism spectrum disorders or language impairment. *Autism Res.* 9, 854–865. doi: 10.1002/aur.1578

Haebig, E., Kaushanskaya, M., and Ellis Weismer, S. (2015). Lexical processing in school-age children with autism spectrum disorder and children with specific language impairment: the role of semantics. *J. Autism Dev. Disord.* 45, 4109–4123. doi: 10.1007/s10803-015-2534-2

Hart, C. M., and Curtin, S. (2023). Trajectories of vocabulary development in children with autism spectrum disorder across multiple measures. *J. Autism Dev. Disord.* 53, 1347–1362. doi: 10.1007/s10803-021-05379-w

Henderson, L., Powell, A., Gareth Gaskell, M., and Norbury, C. (2014). Learning and consolidation of new spoken words in autism spectrum disorder. *Dev. Sci.* 17, 858–871. doi: 10.1111/desc.12169

Herbert, M. R., Ziegler, D. A., Deutsch, C. K., O'Brien, L. M., Kennedy, D. N., Filipek, P. A., et al. (2005). Brain asymmetries in autism and developmental language disorder: a nested whole-brain analysis. *Brain* 128, 213–226. doi: 10.1093/brain/awh330

Hill, A. P., Santen, J. V., Gorman, K., Langhorst, B. H., and Fombonne, E. (2015). Memory in language impaired children with and without autism. *J. Neurodevelop. Disord.* 1, 7–19. doi: 10.1186/s11689-015-9111-z

Hodge, S. M., Makris, N., Kennedy, D. N., Caviness, V. S., Howard, J., McGrath, L., et al. (2010). Cerebellum, language, and cognition in autism and specific language impairment. *J. Autism Dev. Disord.* 40, 300–316. doi: 10.1007/s10803-009-0872-7

Hoff, E. (2009). *Language Development*. Wadsworth Cengage Learning.

Höhle, B., and Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Dev. Sci.* 6, 122–127. doi: 10.1111/1467-7687.00261

Horvath, S., McDermott, E., Reilly, K.M., and Arunachalam, S. (2018). Acquisition of verb meaning from syntactic distribution in preschoolers with autism spectrum disorder. *Lang. Speech Hear. Serv. Sch.* 49, 668–680. doi: 10.1044/2018_LSHSS-STLT1-17-0126

Huang, T., and Finestack, L. (2020). Comparing morphosyntactic profiles of children with developmental language disorder or language disorder associated with autism spectrum disorder. *Am. J. Speech Lang. Pathol.* 29, 714–731. doi: 10.1044/2019_AJSLP-19-00207

Jackson, E., Leitão, S., Claessen, M., and Boyes, M. (2021). Word learning and verbal working memory in children with developmental language disorder. *Autism Dev. Lang. Impairments* 6. doi: 10.1177/23969415211004109

James, E., Currie, N. K., Tong, S. X., and Cain, K. (2021). The relations between morphological awareness and reading comprehension in beginner readers to young adolescents. *J. Res. Read.* 44, 110–130. doi: 10.1111/1467-9817.12316

Joye, N., Broc, L., Olive, T., and Dockrell, J. (2019). Spelling performance in children with developmental language disorder: a meta-analysis across European languages. *Sci. Stud. Read.* 23, 129–160. doi: 10.1080/10888438.2018.1491584

Kambanaros, M., Christou, N., and Grohmann, K. K. (2019). Interpretation of compound words by Greek-speaking children with autism spectrum disorder plus language impairment (ASD-LI). *Clin. Ling. Phonetics* 33, 135–174. doi: 10.1080/02699206.2018.1495766

Kelley, E., Paul, J., Fein, D., and Naigles, L. R. (2006). Residual language deficits in optimal outcome children with a history of autism. *J. Autism Dev. Disord.* 36, 807–828. doi: 10.1007/s10803-006-0111-4

Kjelgaard, M. M., and Tager-Flusberg, H. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Lang. Cognit. Process.* 16, 287–308. doi: 10.1080/01690960042000058

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.* 2013, 2696–2709. doi: 10.1007/s10803-013-1823-x

Lalioti, M., Stavrakaki, S., Manouilidou, C., and Talli, I. (2016). Subject-verb agreement and verbal short-term memory: a perspective from Greek children with specific language impairment. First *Lang.* 36, 279–294. doi: 10.1177/0142723716648844

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

Leonard, L. B. (1998). *Children with Specific Language Impairment*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/1810.001.0001

Levy, S. E., Giarelli, E., Lee, L. C., Schieve, L. A., Kirby, R. S., Cunniff, S., et al. (2010). Autism spectrum disorder and co-occurring developmental, psychiatric, and medical conditions among children in multiple populations of the United States. *J. Dev. Behav. Pediatr.* 31, 267–275. doi: 10.1097/DBP.0b013e3181d5d03b

Leyfer, O. T., Tager-Flusberg, H., Dowd, M., Tomblin, J. B., and Folstein, S. E. (2008). Overlap between autism and specific language impairment: comparison of autism diagnostic interview and autism diagnostic observation schedule scores. *Autism Res.* 1, 284–296. doi: 10.1002/aur.43

Liu, M., Luo, Y., Bai, X., Wang, Y., Hu, X., Sun, M., et al. (2025). Expressive vocabulary in mandarin-speaking autistic, developmentally delayed, and typically developing children: a cross-sectional study. *J. Autism Dev. Disord.* 1–17. doi: 10.1007/s10803-025-06875-z

Lloyd, H., Paintin, K., and Botting, N. (2006). Performance of children with different types of communication impairment on the clinical evaluation of language fundamentals (CELF). *Child Lang. Teach. Ther.* 22, 47–67. doi: 10.1191/0265659006ct297oa

Loucas, T., Riches, N., Baird, G., Pickles, A., Simonoff, E., Chandler, S., et al. (2013). Spoken word recognition in adolescents with autism spectrum disorders and specific language impairment. *Appl. Psycholing.* 34, 301–322. doi: 10.1017/S0142716411000701

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. Disabilities Res. Rev.* 17, 151–159. doi: 10.1002/ddrr.1109

Lyons, J. (1995). *Linguistic Semantics: An Introduction*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511810213

Manolitsi, M., and Botting, N. (2011). Language abilities in children with autism and language impairment: using narrative as an additional source of clinical information. *Child Lang. Teach. Ther.* 27, 39–55. doi: 10.1177/0265659010369991

Matthews, D., Biney, H., and Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: a review of associations with formal language, social cognition, and executive functions. *Lang. Learn. Dev.* 14, 186–223. doi: 10.1080/15475441.2018.1455584

McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., et al. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *J. Autism Dev. Disord.* 42, 35–47. doi: 10.1007/s10803-011-1210-4

Meir, N., and Novogrodsky, R. (2020). Syntactic abilities and verbal memory in monolingual and bilingual children with high functioning autism (HFA). First *Lang.* 4, 341–366. doi: 10.1177/0142723719849981

Mengisidou, M., Marshall, C., and Stavrakaki S. (2020). Semantic fluency difficulties in developmental dyslexia and developmental language disorder (DLD): poor semantic structure of the lexicon or slower retrieval processes? *Int. J. Lang. Commun. Disord.* 55, 200–215. doi: 10.1111/1460-6984.12512

Mengisidou, M., and Marshall, C. R. (2019). Deficient explicit access to phonological representations explains phonological fluency difficulties in Greek children with dyslexia and/or developmental language disorder. *Front. Psychol.* 10:638. doi: 10.3389/fpsyg.2019.00638

Moraleda-Sepúlveda, E., López-Resa, P., and Pulido-García, N. (2022). Difficulties in phonological awareness in children and adolescents with developmental language disorder (DLD). *Eur. J. Educ. Pedagogy* 3, 110–113. doi: 10.24018/ejedu.2022.3.5.462

Naigles, L. R., and Tek. S. (2017). 'Form is easy, meaning is hard' revisited: (re) characterizing the strengths and weaknesses of language in children with autism spectrum disorder. *WIREs Cognit. Sci.* 8:e1438. doi: 10.1002/wcs.1438

Nisiotou, I., and Vlachos, F. (2014). Neurodevelopmental disorders: is there a common biological basis? *Child Adolesc. Psychiatry* 2, 31–41. doi: 10.4324/9780203582107-8

Norbury, C. F., Nash, M., Baird, G., and Bishop, D. V. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the children's communication checklist-2. *Int. J. Lang. Commun. Disord.* 39, 345–364. doi: 10.1080/13682820410001654883

Papafragou, A., Trueswell, J. C., and Gleitman, L. R. (2022). *The Oxford Handbook of the Mental Lexicon*. Oxford: Oxford University Press doi: 10.1093/oxfordhb/9780198845003.001.0001

Papoudi, D., and Vakalopoulou, M. (2022). "Autism - The role of joint attention in the language development of children with autism", in *Language and Special Education: A Psycholinguistic Approach, 2nd Edn*., ed. G. Andreou, (Greece: Yahoudis Publications), 111–133.

Peters, A. M. (2017). "Strategies in the acquisition of syntax," in *The Handbook of Child Language*, eds. P. Fletcher, and B. MacWhinney (Hoboken, NJ: Wiley), 462–482. doi: 10.1111/b.9780631203124.1996.00019.x

Pijnacker, J., Davids, N., van Weerdenburg, M., Verhoeven, L., Knoors, H., and van Alphen, P. (2017). Semantic processing of sentences in preschoolers with specific language impairment: evidence from the N400 effect. *J. Speech Lang. Hear. Res.* 60, 627–639. doi: 10.1044/2016_JSLHR-L-15-0299

Porpodas, K. (2007). *Detection and Investigation Test of Diagnostic Difficulties in the First and Second Grade*. Greece: YPEPTH – EPEAEK.

Ramírez-Santana, G. M., Acosta Rodríguez, V. M., and Hernández Expósito, S. (2019). A comparative study of language phenotypes in autism spectrum disorder and specific language impairment. *Psicothema* 31, 437–442. doi: 10.7334/psicothema2019.92

Rapin, I., and Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain Dev.* 25, 166–172. doi: 10.1016/S0387-7604(02)00191-2

Raven, J., Raven, J. C., and Court, J. H. (1998). *Coloured Progressive Matrices*. Oxford: Oxford Psychologists Press.

Reetzke, R., Zou, X., Sheng, L., and Katsos, N. (2015). Communicative development in bilingually exposed Chinese children with autism spectrum disorders. *J. Speech Lang. Hear. Res.* 58, 813–825. doi: 10.1044/2015_JSLHR-L-13-0258

Reindal, L., Nærland, T., Weidle, B., Lydersen, S., Andreassen, O. A., and Sund, A. M. (2021). Structural and pragmatic language impairments in children evaluated for autism spectrum disorder (ASD). *J. Autism Dev. Disord.* 51,701–719. doi: 10.1007/s10803-020-04853-1

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)

Riches, N. G., Loucas, T., Baird, G., Charman, T., and Simonoff, E. (2011). Non-word repetition in adolescents with specific language impairment and autism plus language impairments: a qualitative analysis. *J. Commun. Disord.* 44, 23–36. doi: 10.1016/j.jcomdis.2010.06.003

Rinaldi, P., Pasqualetti, P., Volterra, V., and Caselli, M. C. (2023). Gender differences in early stages of language development. Some evidence and possible explanations. *J. Neurosci. Res.* 101, 643–653. doi: 10.1002/jnr.24914

Roberts, J. A., Rice, M. L., and Tager-Flusberg, H. (2004). Tense marking in children with autism. *Appl. Psycholing.* 25, 429–448. doi: 10.1017/S0142716404001201

Ryder, N., and Leinonen, E. (2014). Pragmatic language development in language impaired and typically developing children: incorrect answers in context. *J. Psycholing. Res.* 43, 45–58. doi: 10.1007/s10936-013-9238-6

Sandgren, O., Salameh, E. K., Nettelbladt, U., Dahlgren-Sandberg, A., and Andersson, K. (2021). Using a word association task to investigate semantic depth in swedish-speaking children with developmental language disorder. *Logopedics Phoniatrics Vocol.* 46, 134–140. doi: 10.1080/14015439.2020.1785001

Schaeffer, J. (2018). Linguistic and cognitive abilities in children with specific language impairment as compared to children with High-Functioning Autism. *Lang. Acquis.* 25, 5–23. doi: 10.1080/10489223.2016.1188928

Schaeffer, J., Abd El-Raziq, M., Castroviejo, E., Durrleman, S., Ferré, S., Grama, I., et al. (2023). Language in autism: domains, profiles and co-occurring conditions. *J. Neural Transm.* 130, 433–457. doi: 10.1007/s00702-023-02592-y

Seol, K.I., Song, S.H., Kim, K.L., Oh, S.T., Kim, Y.T., Im, W.Y., et al. (2014). A comparison of receptive-expressive language profiles between toddlers with autism spectrum disorder and developmental language delay. *Yonsei Med. J.* 55, 1721–1728. doi: 10.3349/ymj.2014.55.6.1721

Shakila, I., Alyani, E., and Wulandari, S. (2024). Phonetic and syntactic language development of 3 years old childhood. *Variable Res. J.* 1, 566–574.

Sideridis, G., Antoniou, F., Mouzaki, A., and Simos, P. (2015). *RAVEN'S: Colour Progressive Matrices and Vocabulary Tests*. Greece: Motivo.

Silleresi, S. (2023). *Developmental Profiles in Autism Spectrum Disorder: Theoretical and Methodological Implications, Vol. 68*. Amsterdam: John Benjamins Publishing Company. doi: 10.1075/lald.68

Spanoudis, G. (2016). Theory of mind and specific language impairment in school-age children. *J. Commun. Disord.* 61, 83–96. doi: 10.1016/j.jcomdis.2016.04.003

Stanford, E., and Delage, H. (2020). Executive functions and morphosyntax: distinguishing DLD from ADHD in French-speaking children. *Front. Psychol.* 11:551824. doi: 10.3389/fpsyg.2020.551824

Stavrakaki, S. (2006). Developmental perspectives on specific language impairment: evidence from the production of wh-questions by Greek SLI children over time. *Adv. Speech Lang. Pathol.* 8, 384–396. doi: 10.1080/14417040600880714

Stavrakaki, S. (2020). Introduction to the special issue on syntax and verbal short term/working memory in developmental disorders. First *Lang.* 40, 335–340. doi: 10.1177/0142723720933770

Stavrakaki, S., Tasioudi, M., and Guasti, T. (2015). Morphological cues in the comprehension of relative clauses by Greek children with specific language impairment and typical development: a comparative study. *Int. J. Speech Lang. Pathol.* 17, 617–626. doi: 10.3109/17549507.2015.1048826

Stavrakaki, S., and Tsimpli, I. M. (2000). "Diagnostic test of language intelligence for school and preschool children: standardization, statistical analysis, psychometric properties", in *Proceedings of the 8th Panhellenic Conference of Speech and Language Therapists 2000* (Greece: Hellenic Letters), 95–106.

Stemberger, J. P., and Bernhardt, B. H. (2023). *Handbook of Phonological Development: From the Perspective of Constraint-Based Nonlinear Phonology*. Leiden: Brill.

Sukenik, N., and Friedmann, N. (2018). ASD is not DLI: individuals with autism and individuals with syntactic DLI show similar performance level in syntactic tasks, but different error patterns. *Front. Psychol.* 9:317206. doi: 10.3389/fpsyg.2018.00279

Sun, Y., and Poeppel, D. (2023). Syllables and their beginnings have a special role in the mental lexicon. *Proc. Natl. Acad. Sci.* 120:e2215710120. doi: 10.1073/pnas.2215710120

Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clin. Neurosci. Res.* 6, 219–224. doi: 10.1016/j.cnr.2006.06.007

Tager-Flusberg, H., and Caronna, E. (2007). Language disorders: autism and other pervasive developmental disorders. *Pediatr. Clin. North Am.* 54, 469–481. doi: 10.1016/j.pcl.2007.02.011

Tager-Flusberg, H., and Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philos. Trans. R. Soc. London Ser. B. Biol. Sci.* 358, 303–314. doi: 10.1098/rstb.2002.1198

Tager-Flusberg, H., Paul, R., and Lord, C. (2005). "Language and communication in autism", in *Handbook of Autism and Pervasive Developmental Disorders: Diagnosis, Development, Neurobiology, and Behavior, 1st Edn.*, eds. F. R. Volkmar, R. Paul, A. Klin, and D. Cohen (Hoboken, NJ: John Wiley and Sons Inc), 335–364. doi: 10.1002/9780470939345.ch12

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)

Talli, I., Sprenger-Charolles, L., and Stavrakaki, S. (2016). Specific language impairment and developmental dyslexia: what are the boundaries? Data from Greek children. *Res. Dev. Disabilities* 49–50, 339–353. doi: 10.1016/j.ridd.2015.12.014

Talli, I., and Stavrakaki, S. (2020). Short-term memory, working memory and linguistic abilities in bilingual children with developmental language disorder. First *Lang.* 40, 437–460. doi: 10.1177/0142723719886954

Taylor, L. J., Maybery, M. T., Grayndler, L., and Whitehouse, A. J. (2014). Evidence for distinct cognitive profiles in autism spectrum disorders and specific language impairment. *J. Autism Dev. Disord.* 44, 19–30. doi: 10.1007/s10803-013-1847-2

Taylor, L. J., and Whitehouse, A. J. (2016). Autism spectrum disorder, language disorder, and social (pragmatic) communication disorder: overlaps, distinguishing features, and clinical implications. *Aust. Psychol.* 51, 287–295. doi: 10.1111/ap.12222

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

Terzi, A., Marinis, T., Kotsopoulou, A., and Francis, K. (2014). Grammatical abilities of Greek-speaking children with autism. *Lang. Acquis.* 21, 4–44. doi: 10.1080/10489223.2013.855216

Theodorou, E., and Grohmann, K. K. (2015). Object clitics in cypriot Greek children with SLI. *Lingua* 161, 144–158. doi: 10.1016/j.lingua.2014.11.011

Tsimpli, I. M., and Stavrakaki, S. (1999). The effects of a morphosyntactic deficit in the determiner system: the case of a Greek SLI child. *Lingua* 108, 31–85. doi: 10.1016/S0024-3841(98)00041-2

van Wijngaarden, V., de Wilde, H., Mink van der Molen, D., Petter, J., Stegeman, I., Gerrits, E., et al. (2024). Genetic outcomes in children with developmental language disorder: a systematic review. *Front. Pediatr.* 12:1315229. doi: 10.3389/fped.2024.1315229

Vassiliu, C., Mouzaki, A., Antoniou, F., Ralli, A. M., Diamanti, V., Papaioannou, S., et al. (2022). Development of structural and pragmatic language skills in children with attention-deficit/hyperactivity disorder. *Commun. Disord. Q.* 44, 207–218. doi: 10.1177/15257401221114062

Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., et al. (2008). A functional genetic link between distinct developmental language disorders. *New Engl. J. Med.* 359, 2337–2345. doi: 10.1056/NEJMoa0802828

Vogindroukas, I. (2020). Pragmatic abilities in children with autism. *Psychol. J. Hell. Psychol. Soc.* 12, 276–292. doi: 10.12681/psy_hps.23959

Vogindroukas, I., Stankova, M., Chelas, E. N., and Proedrou, A. (2022). Language and speech characteristics in autism. *Neuropsychiatr. Dis. Treat.* 18, 2367–2377. doi: 10.2147/NDT.S331987

Vogindroukas, I., and Zikopoulou, O. (2009). Comparative results on the WISC-III in children with specific developmental reading disorder - dyslexia. *Psychol. J. Hell. Psychol. Soc.* 16, 99–114. doi: 10.12681/psy_hps.23806

Voulgaraki, M. (2023). *Detection of autistic symptomatology in other neurodevelopmental disorders: data from Williams syndrome and developmental language disorder [Doctoral dissertation]*. Faculty of Philosophy, Department of Italian Language and Literature, Aristotle University of Thessaloniki, Thessaloniki.

Wetherby, A., Woods, J., Allen, L., Cleary, J., Dickinson, H., and Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *J. Autism Dev. Disord.* 34, 473–493. doi: 10.1007/s10803-004-2544-y

Whitehouse, A. J., Barry, J. G., and Bishop, D. V. (2007). The broader language phenotype of autism: a comparison with specific language impairment. *J. Child Psychol. Psychiatry* 48, 822–830. doi: 10.1111/j.1469-7610.2007.01765.x

Whitehouse, A. J., Barry, J. G., and Bishop, D. V. (2008). Further defining the language impairment of autism: is there a specific language impairment subtype?. *J. Commun. Disord.* 41, 319–336. doi: 10.1016/j.jcomdis.2008.01.002

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

Williams, D., Botting, N., and Boucher, J. (2008). Language in autism and specific language impairment: where are the links?. *Psychol. Bull.* 134:944. doi: 10.1037/a0013743

Wolk, L., Edwards, M. L., and Brennan, C. (2016). Phonological difficulties in children with autism: an overview. *Speech Lang. Hear.* 19, 121–129. doi: 10.1080/2050571X.2015.1133488

World Health Organization (2019). *International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision)*. Geneva: World Health Organization.

Zarokanellou, V., Gryparis, A., and Papanikolaou, K. (2025). Exploring narrative ability in greek-speaking children with high-functioning ASD: associations with memory and attention. *Brain Sci.* 15:73. doi: 10.3390/brainsci15010073

Zarokanellou, V., Kotsopoulos, A., Tafiadis, D., Prentza, A., Kolaitis, G., and Papanikolaou, K. (2023). Specificity of phonological representations in school-age high-functioning ASD children. *Int. J. Speech Lang. Pathol.* 25, 608–618. doi: 10.1080/17549507.2022.2065030

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 10 September 2025
DOI 10.3389/feduc.2025.1651335

OPEN ACCESS

EDITED BY
Xinyue Qiu,
Beijing Union University, China

REVIEWED BY
Tiago Ferrão Venâncio,
Polytechnic Institute of Portalegre, Portugal
Maria Micheliana Da Costa Silva,
Universidade Federal de Viçosa, Brazil

*CORRESPONDENCE
Omonlola K. Elvire Assogba
✉ elvireassogba4@gmail.com

RECEIVED 21 June 2025
ACCEPTED 18 August 2025
PUBLISHED 10 September 2025

CITATION
Assogba OKE, Madodé YE, Francisco AME, Kossou J, Lokonon J, Luginaah I and Amoussa Hounkpatin W (2025) Factors associated with academic underachievement: a cross-sectional study in Atacora Northern Benin republic.
Front. Educ. 10:1651335.
doi: 10.3389/feduc.2025.1651335

Factors associated with academic underachievement: a cross-sectional study in Atacora Northern Benin republic

Omonlola K. Elvire Assogba[1,2]*, Yann Emeric Madodé[3], Anelia Merveille Ebilola Francisco[1], Jahdiel Kossou[1], Jaurès Lokonon[1], Isaac Luginaah[2] and Waliou Amoussa Hounkpatin[1,3]

[1]Laboratoire de Nutrition Humaine et de Valorisation des Bio-ingrédients Alimentaires (NutriFood-Lab), Faculty of Agronomics Sciences, University of Abomey-Calavi, Abomey-Calavi, Benin, [2]Department of Geography and Environment, University of Western Ontario, London, ON, Canada, [3]Laboratoire de Sciences et Technologies des Aliments, University of Abomey-Calavi, Abomey-Calavi, Benin

Aim: This study aimed to identify factors associated with low academic performance among public primary pupils in Atacora Department, Benin Republic.

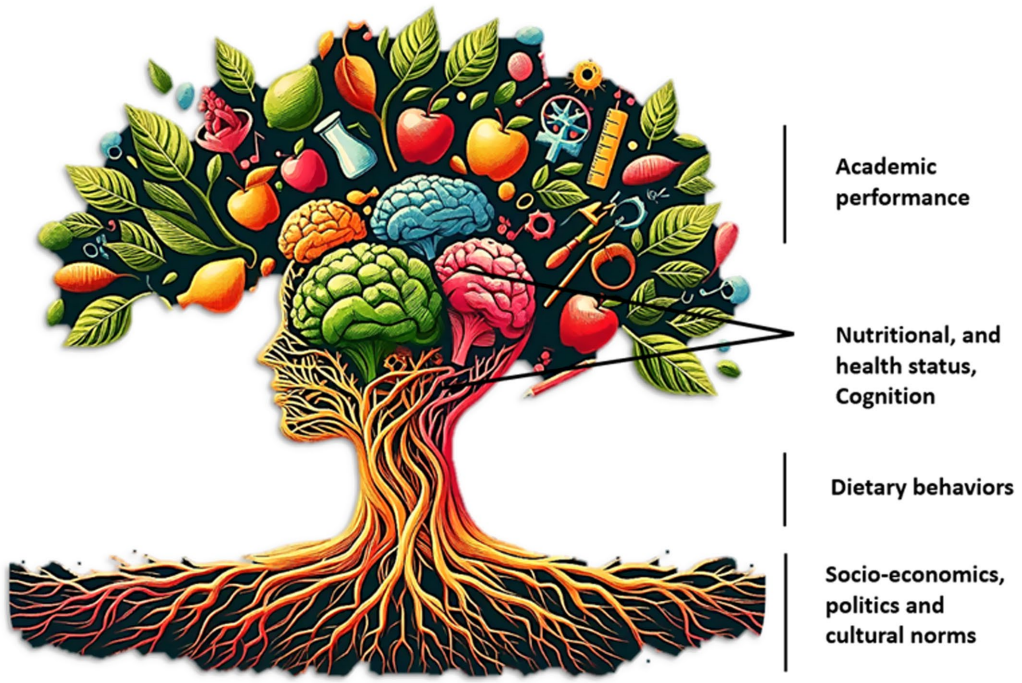
Methods: This cross-sectional study is conducted among pupils aged 8–14 years from public primary schools. Data on dietary diversity were collected using a 24-h dietary recall tool adapted from the FAO guidelines. Nutritional status was assessed through anthropometric measures and hemoglobin level, while cognitive abilities were assessed using Digit Span and Verbal fluency tests. The socio-economic, demographic and health characteristics were collected through digitalized questionnaires administered to pupils. Physical activity levels were measured using the Physical Activity Questionnaire for children (PAQ-C) Form. Academic performance was measured using end-of-month examination results provided by school boards.

Results: findings revealed that almost half (46.34%) of the pupils scored less than 10 out of 20 and were in fail category. Among pupils with normal growth according to WHO standards, cognitive factors as low Verbal fluency (OR = 3.119, $p < 0.05$); low Digit span (OR = 2.623, $p < 0.05$); nutritional factors as low dietary diversity (OR = 2.283, $p < 0.05$); socioeconomic conditions including paternal illiteracy (OR = 1.422, $p < 0.05$), and lack of household electricity (OR = 2.009, $p < 0.05$), and school related factors as long distance to school (OR = 3.187, $p < 0.05$), high level of absenteeism (OR = 1.052, $p < 0.05$), are predictors of academic underachievement.

Conclusion: Overall, Cognition, dietary diversity, access to electricity, pupils' gender, distance to school, father's literacy, are predictors of school performance in the study area. Integrated, context-sensitive policy interventions—spanning early childhood education, rural electrification, gender equity, parental engagement, school attendance, teacher training, nutritional support, and improved food accessibility—are crucial for enhancing academic performance in food-insecure regions of Northern Benin.

KEYWORDS

academic performance, dietary diversity, cognition, nutritional status, parent literacy, physical activity



GRAPHICAL ABSTRACT
Factors influencing academic performance in the department of Atacora in Benin Republic.

1 Introduction

Educational performance is a cornerstone of personal development and a driver of economic progress, influencing individual opportunities and national growth trajectories. Countries with high-performing education systems often experience accelerated economic development due to the cultivation of a skilled and innovative workforce (OECD, 2025). On the other hand, children in low- and middle-income countries (LMICs) face persistent barriers to academic achievement, including systemic inequalities, resource constraints, and socio-economic challenges. Sub-Saharan Africa, including Benin, exemplifies these disparities, where limited access to educational resources and persistent food insecurity significantly impact learning outcomes (WorldBank, 2025). For instance, studies have shown that food insecurity and malnutrition negatively impact cognitive function, reducing students' ability to concentrate and retain information in class (Black et al., 2017). Despite global efforts to improve education, regional inequalities persist, with rural areas often lagging far behind urban centers (UNICEF for every child, 2025). These disparities underline the importance of understanding the unique challenges of food-insecure and underserved rural communities as Northern Benin to design effective interventions.

In food-insecure regions, the intersection of poverty, malnutrition, and inadequate educational infrastructure poses severe challenges to school performance. Poverty affects over 35% of households in Atacora region, reducing access to adequate nutrition and essential educational resources (FAO, 2022). Nutritional deprivation, particularly deficiencies in micronutrients such as iron, zinc, and iodine, has been shown to impair cognitive development and learning capacity (Prado and Dewey, 2014). In addition, families living in poverty often struggle to afford learning materials, school fees, and transportation, further compounding educational disadvantages (UNICEF for every child, 2025). Poor school infrastructure and high student-to-teacher ratios exacerbate these issues, creating unconducive learning environments for academic success. For example, many schools in Northern Benin lack adequate classroom space, functional sanitation, and qualified teachers, all of which contribute to low academic performance (WorldBank, 2025). These interrelated factors highlight the need for a multi-dimensional approach to tackling educational disparities in such settings.

Beyond nutrition and poverty, a range of cognitive, household, and school-related factors influence academic achievement. Cognitive functioning is a critical predictor, with strong evidence linking working memory, attention, and language skills to school performance (Glewwe and Muralidharan, 2016; Alloway and Alloway, 2010). For instance, children with lower verbal fluency and digit span scores often struggle more in classroom settings. Yet, cognitive assessments in LMICs may underestimate children's capabilities if conducted in a language they do not fully master, particularly when local dialects differ from the language of instruction, as is the case in Benin (Piper et al., 2016; Baloubi, 2024).

Household characteristics, such as parental education and household size, also shape academic trajectories. Children of literate parents are more likely to benefit from academic support at home, while those in large families often experience resource dilution, with less time and fewer materials allocated to each child (Black et al., 2017; Filmer and Pritchett, 1999). Other household-level determinants include access to electricity, which affects study time, and parental involvement in school monitoring. These social factors interact with

children's motivation, attendance, and ultimately their academic success.

School-related factors such as absenteeism, long distances to school, and lack of school meals also play a significant role. Long commutes have been linked to fatigue and lateness, reducing classroom engagement, particularly in rural areas where infrastructure is poor (UNESCO, 2023). High absenteeism, whether due to illness, household responsibilities, or lack of interest, directly affects the continuity of learning and performance. Moreover, overburdened classrooms and low teacher-to-student ratios reduce the quality of instruction and individualized support (WorldBank, 2025). Schools that lack essential infrastructure or learning support services often report lower achievement rates and higher dropout rates.

School feeding programs have emerged as a promising intervention to address food insecurity and improve educational outcomes. By providing regular meals to students, these programs aim to alleviate short-term hunger, enhance nutritional status, and improve cognitive function (Kristjansson et al., 2007). In Northern Benin, where food insecurity is pervasive, such programs are particularly relevant and have shown positive effects on attendance, concentration, and enrollment rates (FAO, 2022). However, their direct impact on academic performance remains underexplored in this context. The effectiveness of school feeding programs depends on several factors, including the quality and nutritional value of meals, the frequency of distribution, and the socio-economic conditions of beneficiaries (Black et al., 2017). For instance, schools that consistently provide balanced meals report higher attendance and engagement levels than those with irregular or low-quality meals distributions (Albright and Bundy, 2018). Despite these potential benefits, gaps in implementation and limited evaluations hinder the scalability and optimization of these programs. Understanding their specific impact and factors that are associated to academic performance in Northern Benin could provide valuable insights for improving program effectiveness.

While existing research has examined the role of individual factors such as nutrition (Zaini et al., 2005; Abebe et al., 2017; Akubuilo et al., 2020; Hafte Teklay and Verstegen, 2023) and socio-economic status (Tomul et al., 2021) in academic performance, there remains a gap in understanding their combined effects in food-insecure rural contexts. Most studies tend to focus on urban settings or generalized populations, often overlooking the unique challenges of rural, food-insecure, and underserved communities like Toukountouna and Boukoumbé (UNICEF for every child, 2025). Furthermore, the interaction between nutritional status, feeding practices, and school-specific characteristics, such as distance to school, home tutoring, remain poorly understood. Addressing these gaps is crucial for developing targeted interventions that address the root causes of low academic performance in such contexts. This study aims to identify the factors associated with low school performance in Northern Benin and hypothesizes that child cognition, parental education and distance to school are associated with school performance. This study contributes to the broader discourse on achieving Sustainable Development Goal 4 (SDG 4), which seeks to ensure inclusive and equitable quality education for all by 2030 (United Nations Development Programme, 2024).

2 Theoretical framework

This study applies Bronfenbrenner's Ecological Systems Theory (EST) (Bronfenbrenner, 2013) and Becker's Human Capital Model (HCM) (Lindahl et al., 2014) to analyze factors influencing academic performance in Atacora North, Benin. EST highlights that child development results from interactions between biological traits and environmental contexts. HCM emphasizes investments in health and education as drivers of cognitive capacity and economic productivity (Bailey et al., 2016). Combining these frameworks enables a comprehensive understanding of how nutrition, health, socioeconomic status, and educational inputs jointly affect academic outcomes.

At the microsystem level, biological factors such as anemia and dietary quality have demonstrated strong links to cognitive function and learning. Studies in Sub-Saharan Africa revealed that anemia and micronutrient deficiencies adversely impact attention, memory, and school performance (Appiah et al., 2023; Boivin and Giordani, 1993). In rural settings in Sub-Saharan Africa, poor dietary diversity correlates with undernutrition and impaired child growth, which can negatively influence cognitive development (Hadidjaja et al., 1998; La Rue et al., 1997; Mantzorou et al., 2020). Parental involvement and early education programs further support cognitive and academic development by providing stimulating environments (Chang et al., 2009; Ma et al., 2016).

At the exosystem and macrosystem levels, household poverty and food insecurity limit access to adequate nutrition and schooling. Food insecurity is associated with increased school absenteeism and poorer academic outcomes across Africa, (Buthelezi et al., 2025; Jyoti et al., 2005; Tamiru et al., 2017). From the HCM perspective, improvements in nutrition and education represent investments that generate long-term cognitive and economic returns. Nutritional interventions targeting anemia and dietary quality have been shown to improve cognitive function and school performance in children in SSA (Kedir et al., 2024; Kyere et al., 2020; Neumann et al., 2002). Likewise, quality early childhood education improves foundational skills, increasing future educational attainment and productivity (Heckman, 2012). Integrating these theories offers a multidimensional approach to understanding and identifying the factors associated with low school performance in Atacora North.

3 Methods

3.1 Study context

This study is a cross-sectional study that took place in public primary schools in Atacora Department in Benin republic a Sub-Saharan Africa country. The Republic of Benin is a West African country bordered by Togo to the west, Nigeria to the east, Burkina Faso and Niger to the north, and the Atlantic Ocean to the south. Benin spans 114,763 km^2 and has a population of 13,712,828 (Benin Demographics, 2023). Atacora is currently ranked among the lowest two departments in the country in terms of pupils' academic performance in their final primary school completion certificate examination (La Nouvelle Tribune, 2024). In addition, Atacora has been identified as having a high prevalence of food insecurity, especially in the municipalities of Boukoumbe and Toukountouna (INSAE_AGVSA, 2018). Toukountouna and Boukoumbe are two rural municipalities with limited infrastructure and poor access

to public services. Known for their cultural richness and traditions, the majority of their population are engaged in agricultural and cultural craft activities.

The Beninese National Integrated School Feeding Program (PNASI) significantly improves educational outcomes and alleviates food insecurity for public primary schoolchildren particularly in rural regions like Atacora, where food insecurity and poverty are rampant (INSAE_AGVSA, 2018). Launched in 2017, the program provides daily hot meals to public primary school students, benefiting over 1.1 million children by 2022 across all 77 communes. This initiative not only addresses malnutrition but also boosts enrollment and attendance, reducing absenteeism and dropouts by mitigating hunger among schoolchildren (Amoussa Hounkpatin et al., 2024). Further, research highlights that school canteens encourage parents to keep their children, especially girls, in school, thereby promoting gender equity in education (Djagba et al., 2023). In addition, local management committees and parental contributions strengthen the program, although challenges such as delayed food supply and overcrowded classrooms persist (Djagba et al., 2023). Overall, the PNASI exemplifies a holistic approach to addressing educational and nutritional challenges, offering immediate benefits to children while fostering community development.

3.2 Subjects

Sample size estimation was performed using Charan and Biswas's (2013) sample size formula for quantitative cross-sectional studies. We assumed 95% for confidence interval, considered food insecurity prevalence in Atacora department (INSAE_AGVSA, 2018) and 10% drop out rate (Charan and Biswas, 2013). As a result, a total of 304 pupils were recruited for the study. Participants were selected from four (4) schools in each municipality. Those four schools were randomly selected from the municipality school list. In each school and in each class, the number of participants recruited were proportional to number of students, and gender. Pupils were eligible if they: (1) provided written informed consent from a parent/guardian and gave their assent to participate, (2) were in good health with no contraindications regarding meals served in the school canteen, (3) had no physical limitations, and (4) were 8–14 years of age and in grade 4 or 5. Children were excluded if they: (1) were found sick the day of data collection, (2) were taking medications (3) did not provide signed informed consent from their legal guardians or did not give their assent to participate in the study. The schools administrations and teachers were informed of the study and its objectives.

The data collection was held in October 2024. On the day of data collection, 5 selected pupils were absent, and 7 were reported sick. A total 5 pupil's data was incomplete, they withdrew from the study because of time constraints. Overall, 287 pupils participated in the study.

3.3 Ethical clearance

The study protocol was approved by the institutional review board of Ethics Research Committee of Applied Biomedical Sciences (N° 260 du 26/07/2024. Date of approval July 26 2024) and was conducted in accordance with the Declaration of Helsinki. The study design and objectives were communicated to the school administration and teachers. Before the study began, students and their parents or legal guardians were informed about the study objectives and procedures. Participant gave their assent and their parents or legal guardians gave their written informed consent.

3.4 Data collection

3.4.1 Academic performance

The grade point average scores in all studied subjects in October month for the academic years 2024–2025 represent the academic performance of the participants. The grading categories included: High performance (pass), score ranged from 10 to 20 and Low performance (fail) category, score ranged from 0 to 9.99. The subjects studied at school expand in Mathematics, life sciences, social sciences, languages, physical education and creative activities (as drawing) domains. The average grade in all subjects studied was used for consistency with current assessment practices in the Beninese primary school system to generate final grades and determine whether a pupil succeeded and will advance to the next grade the following year or failed. Given that this study is cross-sectional, we only used the October 2024 exam results.

3.4.2 Cognition

Cognition was assessed across three cognitive domains: attention, short-term memory retention, and thinking. Two cognitive tests were used: the Verbal fluency test (Regard et al., 1982) and the Digit Span test (Jasinski et al., 2011; Wechsler, 2008). The Digit Span test assessed how many items a pupil could recall immediately, in the order showed in the test. The performance on this test provides valuable information about an individual's short-term memory span and their ability to process and temporarily store information (Jasinski et al., 2011). The Wechsler Intelligence Scale for Children Digit Span Revised tasks have been proven reliable and enable valid measures of short-term memory capacity in several studies (De Paula et al., 2016; Conway et al., 2005). The test was administered according to Wechsler guidelines. Based on the methodological review and user's guide for memory span tasks (Conway et al., 2005), and the mean of score in the current study (5.88 rounded to 6), we categorized the pupils' scores into two groups: "low scores" = 0 to 5 lines correct, and "high scores" = 6 to 16 lines correct.

Additionally, the Verbal fluency test assessed the speed and ease with which pupils can use words, and their ability to think and organize information in a limited timeframe. The Verbal fluency test was evaluated in French according to Regard et al. (1982). Participants were asked to generate as many words as possible that began with a specific letter. The detailed description of the test can be found in prior literature (Sauzéon et al., 2004; Regard et al., 1982; Zorza et al., 2016). Two categories were considered in this study: "low scores" = 0 to 2 words, and "high score" = 2 to 16 words.

3.4.3 Anthropometry and hemoglobin measures

Body mass index for age was used to assess the pupils' nutritional status according to the World Health Organization standard operating procedures (WHO, 2006) using pupils' height and weight measured according to the WHO's guidelines. The *z*-scores were calculated via WHO's Anthro Plus software, following standard operating procedures. In the same way, hemoglobin level was measured according to World Health Organization standard operating procedures (Weltgesundheitsorganisation, 2006). Portable

hemoglobinometer (HemoCue AB) was used to determine hemoglobin level from a capillary blood sample collected from the fingertip of each child aseptically, using sterile single-use disposable lancet. It was done by trained and experienced laboratory technicians. The necessary safety measures were taken during blood collection. A child was identified as anemic if the hemoglobin concentration was <11.5 g/dL for children (5–11 years) and < 12 g/dL for children older than 12 years of age. These indicators are globally accepted and used in several studies to assess children's nutritional status (Assefa et al., 2014; Correa-Burrows et al., 2016; Sun et al., 2024).

3.4.4 Physical activity questionnaire

Physical activity was measured using the Physical Activity Questionnaire for children (PAQ-C) appropriate for school-aged children (4–14 years old) (Kowalski et al., 2004). Each of the first 9 (PAQ-C) questions is scored between 1 (low) and 5 (high physical activity), and a mean score of all items constitutes the overall PAQ score. The reliability and validity of the PAQ-C have been reported in different study populations and countries (Benítez-Porres et al., 2016; Janz et al., 2008; Kowalski et al., 2004; Voss et al., 2017).

3.4.5 Dietary diversity

Dietary diversity was assessed using a 2 weekday 24-h dietary recall (24HDR). A detailed description of all the food and drink consumed by pupils was recorded. According to FAO guidelines (Kennedy et al., 2011), the number of different food groups consumed over 24-h was generated to reflect their dietary diversity. A pupil's dietary food diversity was constructed based on 10 food groups. A child was considered as having a good dietary diversity when he had eaten 5 or more food groups, and a child's dietary diversity was considered poor when he reported fewer than 5 food groups (Kennedy et al., 2011; Nago et al., 2009; Coates et al., 2007).

3.4.6 Distance to school

The information was gathered on the distance to school by asking participants to record the time spent walking from their house to school. The variable distance to school were categorized based on empirical evidences (d'Aiglepierre, 2012; Oneya and Onyango, 2021), into three groups: pupils close (less than 15 min to school), pupils living far (between 15 min and 29 min) and pupils living very far from school (30 min and above).

3.4.7 Socioeconomics and other educational characteristics

The socio-economic characteristics of the study participants were collected using a digitalized questionnaire. The data collected included pupils' family wealth assessed through household asset ownership, household ownership of a bike, cell phone, radio, TV, stove, access to electricity, and access to running water in the house. Principal Component Analysis (PCA) was used to reduce the socioeconomic variables into two categories. Pupils from "poor households" and pupils from "non-poor households." Based on the PCA, the poverty line separating the poor and non-poor groups was −0.00477. Using PCA is a more pragmatic alternative to categorize household wealth in contexts where both money flow and household membership are volatile, seasonally dependent, and difficult to track. Empirically, this approach was considered reliable for predicting household wealth (Armah and Luginaah, 2012). The socio-demographic data included pupils' gender, age, parent literacy, and household size. Health issue data were gathered by administering questions to pupils to determine whether during the term they experienced any illness including endemic diseases such as malaria and anemia. Pupils were asked to bring to school their medical records to verify. All diseases, infections, or pain were recorded through the digitalized questionnaire. Other educational characteristics of the participants were collected, including absenteeism, whether they had attended pre-school program, and had a home tutor.

3.5 Statistical analysis

All statistical analyses were performed using Stata version 18.0. The data analysis involved a three-stage approach (univariate, bivariate, and multivariate). The univariate approach examined the descriptive statistics to understand characteristics and distribution of each variable before conducting further analysis. It provided valuable insights into the individual variables, which helped identify outliers, understand the central tendency, and assess the overall shape of the data distribution. We used binary logistic regression model in the second and third stages since the dependent variable (pupil's school performance) is binary. This allowed us to investigate the association between school performance and independent variables (Hedeker et al., 2000). The second stage involved bivariate ordered logistic regression to understand the relationship between school performance and each independent variable. This helped assess each variable's impact on the likelihood of passing or failing at school. The final stage involved investigating the association of the combined independent variables and their impact on the dependent variable (Hedeker et al., 2000). We used a likelihood estimation method to calculate the odds ratios for pupils having low school performance (Murad et al., 2003). The odds ratios exceeding one (OR>1) indicate a higher likelihood of low school performance, while those below one (OR<1) demonstrate a lower likelihood. Statistical significance was set at $p < 0.05$. The reliability of the model was assessed using R-squared, the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC).

4 Results

4.1 Univariate results

The univariate analysis is presented in Table 1. The survey revealed that 46.34% scored low academic achievement and 53.66% scored high. Only 23.34% of the pupil's had attended a preschool program. Most pupils' (51.57%) were reported to have moderate physical activity. Only one fourth (26.13%) of the pupils' household were reported to have access to electricity.

The survey revealed that among the participants, the majority of pupils were aged 11–14 years (66.55%). Only 45.64% of pupils had never been absent in the considered term. Dietary diversity, was limited for most pupils, with 58.54% reporting inadequate dietary diversity. Physical activity levels varied, with most pupils reporting low (51.57%) or high (35.54%) activity; boys were more likely to

TABLE 1 Descriptive analysis of the sample.

Variables categories	Variables	Percentage (%)	Frequency
Dependent variable	School performance	10.37 (mean), SD: 2.53	Min:0; Max:16.56
	Low (fail)	46.34	133
	High (succeed)	53.66	154
Cognition	Digit span	5.88 (mean), SD: 2.89	Min: 0; Max: 10
	Low	53.31	153
	High	46.69	134
	Verbal fluency	2.36 (mean), SD: 2.66	Min: 0; Max: 16
	Low	88.85	255
	High	11.15	32
Other school data	Absenteeism		
	Never been absent	45.64	133
	have been absent	54.36	154
	Class		
	Grade 4	46.04	64
	Grade 5	53.96	75
	Pupil's age	11.34 (mean), SD: 1.52	Min: 8; Max: 14
	8 to 10	33.45	96
	11 to 14	66.55	191
	Preschool		
	Yes	23.34	67
	No	76.66	220
	Having a tutor		
	Yes	34.15	98
	No	65.85	189
	Distance to school		
	Close (*less than 15 min of walk*)	29.27	84
	Far (*between 15-30 min of walk*)	54.70	157
	Very far (*more than 30 min by walk*)	16.03	46
	Teacher level of education		
	High school	50	8
	College	31.25	5
	University degree	18.75	3
Nutritional status	Hemoglobin level (dL)	11.49 (mean), SD: 1.24	Min: 6.4; Max: 14.5
	Anemic pupils	52.61	151
	Non anemic pupils	47.39	136
	BMI-age		
	Severe	0.7	2
	Moderate	6.62	19
	Mild	27.18	78
	Normal	62.72	180
	Over	2.79	8
	DDS		
	≥5	68.99	198
	<5	31.01	89

(Continued)

TABLE 1 (Continued)

Variables categories	Variables	Percentage (%)	Frequency
Physical activity	Pupil's physical activity		
	Very low	11.50	33
	Moderate	51.57	148
	High	35.54	102
	Very high	1.39	4
Socio-economic characteristics	Electricity		
	Yes	26.48	76
	No	73.52	211
	Household wealth		
	Poor	26.13	75
	Non poor	73.87	212
	Household size	7.06 (mean), SD: 2.53	Min: 3; Max: 15
	3–7	62.37	179
	8–15	37.63	108
	Father work status		
	Working father	90.59	260
	Jobless father	9.41	27
	Mother or caregivers working status		
	Working caregivers	85.71	246
	Jobless caregivers	14.29	41
	Father literacy		
	Literate	49.13	141
	Illiterate	50.87	146
	Caregivers literacy		
	Literate	28.92	83
	Illiterate	71.08	204

**p-value<0.05; ***p-value<0.001; ^p-value < 0.10; OR: odd ratio; SE: Standard Error; Cont.: continuous variable; ref: group reference; Low performance (fail): score range 0–9.99; High performance (succeed): score range 10–20); Low digit span score: score range 0–6 lines correct; High digit span score: score range 6–16 lines correct; Low verbal fluency score: score range 0–2 correct words uttered; high verbal fluency score: score range 0–16 correct words uttered.

report high activity (38.85%) than girls (32.43%), whereas very low activity was more frequent among girls (14.86%) compared to boys (7.91%). Access to electricity was low overall, with 73.52% of households lacking it and almost three fourth pupils' were from non-poor household (73.87%).

4.2 Bivariate results

Table 2 shows the results of the bivariate binary logistic regression analysis. The findings showed that verbal fluency (OR = 4.300, $p < 0.05$) and digit span (OR = 3.023, $p < 0.001$) impacted significantly school achievements, with pupils showing lower scores being at least three time more likely to score low school performance. In addition, an increase in age (OR = 0.831, $p < 0.05$) reduced the odds of having low academic performance. Female pupils were significantly more likely to perform lower than their male counterparts (OR = 1.802, $p < 0.05$). Lack of electricity were a negative factor for academic performance, as pupils from households without electricity were nearly two time more likely to perform lower in school (OR = 1.832, $p < 0.05$). Pupils' not having a minimum dietary diversity were two time likely to fail at school (OR = 2.042, $p < 0.05$). In the same line, having suffered from malaria in the considered term (OR = 1.895, $p < 0.05$), were positively associated with low performance. Lacking preschool program (OR = 1.895, $p < 0.05$) negatively impacted outcomes. The relationship between distance to school and performance was inconclusive, although pupils living very far (OR = 2.065, $p < 0.05$) from school approached significance for poorer performance.

4.3 Binary logistic regression to predict low academic performance

Table 3 presents binary logistic regression analysis exploring the associations between nutritional, cognitive, socio-economic, and educational factors and low school performance, for both the full sample and pupils with normal BMI-for-age. Cognitive factors as verbal fluency (OR = 3.401, $p < 0.05$) and digit span

TABLE 2 Bivariate binary logistic regression analysis predicting low school performance among pupils (n =287).

Bivariate binary logistic regression analysis			
Variables	OR	SE	*p*-value
Verbal fluency *(ref. low)*	4.300	2.020	0.002**
Digit span *(ref. low)*	3.023	0.748	0.000***
Pupil's Age *(cont.)*	0.831	0.066	0.02**
Pupil's Gender (*ref male*)	1.802	0.432	0.014**
Mothers or caregivers working status (*ref working*)	1.121	0.378	0.735
Fathers Working status (*ref working*)	0.777	0.319	0.541
Mothers or caregivers literacy (*ref. literate*)	1.359	0.358	0.245
Paternal literacy (*ref. literate*)	0.687	0.163	0.115
Household wealth (*ref. non-poor*)	1.176	0.316	0.546
Electricity (*ref. having electricity*)	1.832	0.506	0.028**
Household size (cont)	1.192	0.291	0.471
Anemic status (*ref. Normal hb stores*)	0.800	0.190	0.346
Dietary Diversity Score (*ref. DDS≥5*)	2.042	0.530	0.006**
Malaria (*ref. had not suffer of malaria*)	1.895	0.615	0.049**
BAZ (*ref. normal BMI-Age*)	0.953	0.136	0.739
Home tutor (*ref. had home tutor*)	0.8	0.200	0.371
Preschool (*ref. attended preschool*)	0.536	0.151	0.027**
Absence (*ref. had never been absent*)	1.026	0.018	0.144
Distance to school (*ref. close*)	1.321	0.239	0.124
Far	0.877	0.239	0.633
Very far	2.065	0.776	0.054*
PA (ref very low + low PA)	1.027	0.030	0.357
High (high + very high)	1.341	0.329	0.232
Teacher's level of education *(ref. secondary school diploma)*			
College	0.88	0.400	0.779
University degree	0.618	0.272	0.275

**p-value < 0.05; ***p-value < 0.001; *p-value < 0.10; OR: odd ratio; SE: Standard Error; Cont.: continuous variable; ref: group reference; low performance (fail): score range 0–9.99; High performance (succeed): score range 10–20; Low digit span score: score range 0–6 lines correct; High digit span score: score range 6–16 lines correct; Low verbal fluency score: score range 0–2 correct words uttered; high verbal fluency score: score range 0–16 correct words uttered; PA: Physical Activity.

(OR = 2.431, $p < 0.05$) were significantly associated with higher odds of low school performance, underscoring the importance of cognitive skills. Among nutritional factors, higher food diversity was inversely associated (OR = 2.015, $p < 0.05$) with the odds of having low school performance. In contrast, hemoglobin levels and malaria showed no significant associations. Gender (OR = 1.938, $p < 0.05$) also played a significant role, showing a marginal association with increased risks for girls. Father's literacy (OR = 1.533, $p < 0.05$) emerged as a protective factor, suggesting that paternal literacy could reduce the risk of low performance. Additionally, considerable long distance to school (OR = 2.813, $p < 0.05$) was found to increase the risk of academic underachievement.

Among pupils with normal BMI for their age, similar patterns were observed, though with some variations. Cognitive performance in verbal fluency (OR = 3.119, $p < 0.05$) and digit span (OR = 2.623, $p < 0.05$) continued to be significant predictors. Food diversity remained a significant protective factor for academic performance (OR = 2.283, $p < 0.05$), reinforcing the importance of dietary diversity in supporting cognitive and academic outcomes. Socio-economic and demographic variables such as access to electricity (OR = 2.009, $p < 0.05$) showed similar trends as in the full model, with electricity access being a significant factor. Gender exhibited no effect. Paternal literacy continued to be protective (OR = 1.422, $p < 0.05$), highlighting the role of parental education in shaping academic success. School-related factors such as school absence (OR = 1.052, $p < 0.05$) were significantly associated with low performance, revealing the importance of regular school attendance. Additionally, considerable long distance to school (OR = 3.187, $p < 0.05$) was found to increase the risk of academic underachievement.

These results emphasize the continued relevance of cognitive performance, food diversity, and socio-economic context, in influencing school performance, particularly in pupils with normal BMI-for-age.

5 Discussion

This study investigated the factors associated with school performance in a food-insecure context in Northern Benin, specifically in Toukountouna and Boukoumbé. The findings highlight the complex interplay of cognitive abilities, nutritional aspects, socio-economic factors, and school patterns in influencing school performance.

Consistent with our hypothesis, we found that cognitive abilities especially attention and working memory are essential for academic success. This aligns with evidences suggesting that attention, working memory and ability to organize information are core skill for school performance (Lemos et al., 2025; Demetriou et al., 2020; Peng and Kievit, 2020; Bettini and Giuliani, 2016). Peng and Kievit went further, and showed that working memory is longitudinally linked to school performance throughout childhood (Peng and Kievit, 2020). In other words, pupils with good executive functions will perform better at school than those with cognitive disorders. Other authors even claimed that cognitive abilities are the best predictor of school performance (Karbach et al., 2013). Cognitive development in early childhood and throughout childhood and early adolescence is crucial and closely tied to educational outcomes.

One crucial aspect of our study is that, low dietary diversity influenced negatively school performance. This means that pupils not meeting the minimum dietary diversity have higher chances of exhibiting low school performance. Indeed, pupils with inadequate dietary diversity may fail to provide their bodies essential macronutrients like protein, micronutrients like calcium magnesium that promote optimum growth and brain development resulting later in lower academic results (Lee, 2022; Neumann et al., 2007; Stephenson et al., 2023). This finding aligns with broader agreement that proper feeding practices including dietary diversity are crucial for adequate cognitive development and better academic performance in literature (Beckmann et al., 2021; Uzosike et al., 2020; Yeh et al., 2025). This evidence adds to those in Eastern Morocco (Bouchefra et al., 2023) and Port Hacourt in Nigeria (Uzosike et al., 2020) highlighted low dietary diversity as one of the main predictor of poor academic performance. In contrast, Beressa and collaborators revealed a mixed effect

TABLE 3 Binary logistic regression analysis predicting Low school performance among all the pupils and among pupils with normal Body mass index for age.

Variables group	Variables	Model 1: binary logistic regression (all)			Model 2: binary logistic regression (BMI-age normal pupils)		
		Mixed multiple model			Logistic regression model		
		OR	SE	*p*-value	OR	SE	*p*-value
Cognition	Verbal fluency *(ref. high)*	3.401	1.808	0.021**	3.119	1.689	0.036**
	Digit span *(ref. high)*	2.431	0.686	0.002**	2.623	0.791	0.001**
Nutritional status and health	Hemoglobin level *(cont)*	1.091	0.125	0.448	1.066	0.129	0.599
	Malaria *(ref. normal pupils)*	1.637	0.632	0.202	2.149	0.903	0.069
	Dietary diversity Score *(ref. DDS≥5)*	2.015	0.625	0.024**	2.283	0.757	0.013**
Socio-economic characteristics	Pupils' age *(cont)*	0.881	0.088	0.204	0.888	0.096	0.272
	Gender *(ref. boys)*	1.938	0.557	0.021**	1.749	0.542	0.071
	Electricity *(ref. having electricity)*	1.710	0.556	0.099	2.009	0.691	0.042**
	Household size *(Cont)*	1.245	0.356	0.442	1.523	0.471	0.174
	Household wealth *(ref. non poor)*	1.155	0.368	0.652	1.281	0.441	0.473
	Home tutor *(ref. having a home tutor)*	1.134	0.334	0.670	1.104	0.351	0.755
	Mothers or caregivers working status *(ref. working mother)*	1.076	0.428	0.853	1.381	0.590	0.450
	Mothers or caregivers literacy *(ref. literate mother)*	1.638	0.550	0.142	1.402	0.495	0.338
	Fathers working status *(ref. working mother)*	0.769	0.380	0.596	1.035	0.539	0.947
	Paternal literacy *(ref. literate mother)*	1.533	0.159	0.035**	1.422	0.134	0.006**
Other school patterns	Preschool *(ref. Yes)*	0.667	0.229	0.240	0.710	0.254	0.338
	Absenteeism *(cont)*	1.039	0.022	0.077	1.052	0.025	0.032**
	Distance to school *(ref. close)*						
	Far	0.952	0.308	0.881	1.066	0.369	0.853
	Very far	2.813	1.241	0.019**	3.187	1.540	0.016**
Physical activity	Physical activity *(ref low +moderate physical activity)*	1.345	0.393	0.311	1.273	0.400	0.442
	constant	0.025	0.054	0.080	0.014	0.031	0.056
		R^2= 0.178			R^2=0.207		
AIC		367.810			333.504		
BIC		444.659			408.757		

**p-value < 0.05; ***p-value < 0.001; *p-value < 0.10; OR: odd ratio; SE: Standard Error; Cont.: continuous variable; ref: reference group; R^2: Rsquare.

between dietary diversity and academic performance, even though they concluded that dietary diversity positively influences growth measured with height-for-age z-scores (HAZ) (Beressa et al., 2024). The differences in these findings may be due to bias related to participants' health status, or population specific characteristics and/or geographic location. In fact, in Beressa and collaborators' study, 91.6% of the participants had experienced an illness (such as fever, pneumonia), and the study was conducted in pastoral communities in Southeast Ethiopia, which may exhibit specific context dependent patterns. Higher dietary diversity, by providing various essential building blocks for optimal growth, gives greater opportunities for better brain maturation and development of executive functions – core skills for high cognition and better educational outcomes later in life.

We also found that regarding socio-demographic patterns, electricity in households was significantly associated with better school performance, likely reflecting that improved study environments positively affect educational outcomes. This finding is consistent with studies in other low-income settings that link access to electricity to higher educational attainment (Khandker et al., 2012). Having electricity in a household can be a main factor in a pupils' ability to study at home. Indeed, some pupils have to walk a long distance to reach school and then return home at nighttime or dawn. In addition, once at home, pupils – and mainly girls – are required to help with household chores, and only once these tasks are completed can they reach back to their books, By this time, darkness may already completely fallen because class ends around five (5) in the afternoon. As a result, without electricity at home, they will either not study or will tend to seek other sources outside their houses, sometimes at gas stations or under street lamps, in order to have enough light to complete their homework or assignments. By relying on these outside sources, their studying time cannot be extended. This situation can, in the long run, negatively impact school performance and leads to school dropouts (Lee and Guadagno, 2015).

Moreover, as expected, parental education levels play a significant role in their children's academic achievement. Our

findings indicate that fathers' illiteracy is positively associated with pupils' low performance. In contrast, mothers' literacy did not show a significant influence on school performance. This may be attributed to the fact that in this specific location, the number of literate female caregivers were very few. As a result, the literate group was not substantial enough to allow for meaningful comparisons with the larger group of illiterate mothers. Literate parents with higher educational attainment often provide enriched home environments, including access to educational resources, and structured routines, which are conducive to better cognitive development and better school performance (Davis-Kean, 2005). They also tend to engage more actively in their children's education, promoting behaviors that support academic success in contexts where parents' daily occupations do not prevent them from being involved in their children's lives. Consequently, interventions to promote female education should be emphasized in rural location in Sub-Saharan African region and specifically in northern Benin.

Another important finding of this study is the impact of distance to school on pupils' academic achievement. The distance between pupils' home and school, is prone to affect attendance, punctuality, and access to educational resources. Long commutes often result in fatigue, reduced study time, and lower engagement with school activities, which cumulatively hinder academic performance (Frempong et al., 2011). Additionally, children in rural or remote areas with extended travel distances to school frequently face infrastructure and transportation challenges, leading absenteeism increase and lower academic outcomes (Dubow et al., 2009).

Therefore, School attendance is a fundamental factor influencing normal (normal growth according to WHO) children's academic success, as consistent classroom engagement provides essential opportunities for learning and development. Gottfried confirms that regular attendance is strongly associated with better academic outcomes, as it ensures exposure to the curriculum and fosters social and cognitive growth. Chronic absenteeism has been shown to disrupt learning continuity (Gottfried, 2010). However, the strength of this relationship has been contested in some settings. For example, evidence from South Africa suggests that effective teaching quality (Wills and Hofmeyr, 2019), implementing school feeding programs (Abotsi, 2013; Amoussa Hounkpatin et al., 2024; Cohen et al., 2021) in disadvantaged areas can mitigate the negative impact of absenteeism. Several studies investigated the relationship between teacher demographic characteristics, teaching quality, and students' academic performance by shaping teaching approaches, classroom dynamics, and role modeling. Evidences showed that teachers with higher educational qualifications tend to adopt more effective instructional strategies, that impact positively academic achievements (Clotfelter et al., 2007). Additionally, teacher gender (Antecol et al., 2012), marital status (Harris, 2011), and level of education influence academic achievements but their effect are often mediated by broader contextual factors, such as school resources and teacher training quality, which must be considered in educational policy and practice.

6 Limitation

In the current study, while hemoglobin levels were assessed using a precise tool and cognition and school performance were measured with standardized methods, several variables were self-reported by the schoolchildren. This reliance on pupil's self-reported data may have introduced errors due to potential recall bias and / or socially desirable responses. In addition, this study only assessed cognitive abilities using Digit span and Verbal fluency in French language, which is the national language of Benin but not the traditional mother tongue of the pupils. This could potentially result in lower scores for pupils who have limited comprehension of French. To mitigate this limitation we selected pupils in grades 4 and 5, assuming a higher proficiency in French at this stage. Additionally, only two week-day 24-h qualitative food recalls were done in this study, which may have limited the assessment of the dietary diversity of the participants and estimating the quantity of nutrients consumed. All schools in the study area were enrolled in the school feeding programs, which would make a comparative study challenging; however, a longitudinal study would be useful in revealing the overall impact of the school feeding program on pupils' cognition. Future research in this area could focus on diverse possibilities to improve the school feeding programs through the formulations of more nutritious menus to raise dietary diversity based on local dishes and evaluating their effects on cognition and school performance. Future research should consider employing bioelectrical impedance analysis (BIA) to precisely assess pupils' nutritional status by distinguishing body fat from lean mass, as z-score measurements may not accurately capture associations with school performance (Wu and Billard, 2021). In addition, future research may explore other precise methods to measure pupil's physical activity as Accelerometers or pedometers (Trost et al., 2005).

7 Conclusion and policy implications

This study highlights the complex interplay of factors influencing school performance among pupils benefiting from the National School Feeding Program in Northern Atacora, Benin. Key determinants to academic underachievement are low cognitive abilities, lack of electricity in households, paternal illiteracy, school absenteeism, low dietary diversity, and long distance to school. These findings underscore the need for targeted, multi-sectoral interventions to enhance educational outcomes in this food-insecure region. Policy efforts should prioritize the promotion of early childhood education programs, particularly preschool programs, to support cognitive development from an early age. Expanding rural electrification is crucial to improve home study conditions and facilitate the use of modern teaching tools in schools. Gender equity should be addressed through scholarships for girls, gender-sensitive teacher training, and community advocacy to dismantle cultural barriers to education. Enhancing parental literacy and engagement in education programs can create supportive home environments that foster learning. To ensure consistent school attendance and better nutritional status, strengthening school feeding programs is imperative, throughout efforts to provide balanced diverse meals and promote home and/or school gardens. Building schools closer to communities and providing safe, reliable transportation options are essential to reducing travel

barriers and improving access to education. Finally, investing in teacher training programs will strengthen teaching quality and contribute to better academic outcomes. By implementing these integrated and context-sensitive measures, policymakers can create a sustainable framework for improving school performance and addressing systemic disparities in education.

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 humans were approved by Ethics Research Committee of Applied Biomedical Sciences of the Faculty of Health Science University of Abomey-Calavi. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

OA: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. YM: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing. AF: Data curation, Methodology, Resources, Software, Supervision, Writing – original draft. JK: Conceptualization, Investigation, Methodology, Validation, Writing – review & editing. LJ: Conceptualization, Formal analysis, Investigation, Methodology, Writing – review & editing. IL: Conceptualization, Formal analysis, Validation, Writing – review & editing, Writing – original draft. WAH: Conceptualization, Methodology, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

We sincerely thank the International Atomic Energy Agency (IAEA) for biomedical equipment supply. We also thank all involved pupils, educators, principals and administrators of the selected school and all the enumerators involved in this study fieldwork.

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Abotsi, A. K.. (2013). Expectations of school feeding programme: impact on school enrolment, attendance and academic performance in elementary Ghanaian schools. Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2322392 (accessed May 25, 2025).

Abebe, F., Geleto, A., Sena, L. (2017). Predictors of academic performance with due focus on undernutrition among students attending primary schools of Hawa Gelan district, Southwest Ethiopia: a school based cross sectional study. *BMC Nutrition* 3:30.

Alloway and Alloway. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology* 106: 20–29.

Akubuilo, U. C., Iloh, K. K., Onu, J. U., Iloh, O. N., Ubesie, A. C., and Ikefuna, A. N. (2020). Nutritional status of primary school children: association with intelligence quotient and academic performance. *Clin. Nutr. ESPEN* 40, 208–213. doi: 10.1016/j.clnesp.2020.09.019

Albright, A., and Bundy, D. A. P. (2018). The global Partnership for Education: forging a stronger partnership between health and education sectors to achieve the sustainable development goals. *Lancet Child Adolesc. Health* 2, 473–474. doi: 10.1016/S2352-4642(18)30146-9

Amoussa Hounkpatin, W, Fanou Fogny, N, and Bodjrenou, S (2024) Etude de cas sur l'alimentation scolaire: Bénin

Antecol, H, Eren, O, and Ozbeklik, S (2012). The effect of teacher gender on student achievement in primary school: Evidence from a randomized experiment. 6453, Working Paper. IZA Discussion Papers. Available online at: https://www.econstor.eu/handle/10419/58932 (Accessed May 15, 2025)

Appiah, A. O., Tandoh, M. A., Puotege, P. S., and Edusei, A. K. (2023). The effect of a Turkey berry (*Solanum torvum*)-fortified biscuit on the hemoglobin level and cognitive performance of adolescent females in the Ahafo region of Ghana: a pilot study. *Int. J. Food Sci.* 2023:1–9. doi: 10.1155/2023/1388782

Armah, FA, and Luginaah, I (2012). Water quality index in the Tarkwa gold mining area in Ghana. Available online at: https://ir.ucc.edu.gh/xmlui/handle/123456789/5486 (accessed May 25 2025).

Assefa, S., Mossie, A., and Hamza, L. (2014). Prevalence and severity of anemia among school children in Jimma town, Southwest Ethiopia. *BMC Hematol.* 14:1–9. doi: 10.1186/2052-1839-14-3

Bailey, R., Cope, E. D., and Dan, P. (2016). "The human capital model: realising the benefits of sport and physical activity" in Sport, education and social policy (Routledge), Eds. Gudrun Doll-Tepper, Katrin Koenen, Richard Bailey. 155–179.

Baloubi, D. (2024). "Language policy in Benin and rights to education in local mother tongues" in The Palgrave handbook of language policies in Africa, Cham: Springer International Publishing. 301–319.

Benin Demographics. (2023). Available at: https://www.worldometers.info/demographics/benin-demographics/ (accessed May 25, 2025).

Beckmann, J., Lang, C., du Randt, R., Gresse, A., Long, K. Z., Ludyga, S., et al. (2021). Prevalence of stunting and relationship between stunting and associated risk factors with academic achievement and cognitive function: a cross-sectional study with south African primary school children. *Int. J. Environ. Res. Public Health* 18:4218. doi: 10.3390/ijerph18084218

Benítez-Porres, J., López-Fernández, I., Raya, J. F., Álvarez Carnero, S., Alvero-Cruz, J. R., and Álvarez Carnero, E. (2016). Reliability and validity of the PAQ-C questionnaire to assess physical activity in children. *J. Sch. Health* 86, 677–685. doi: 10.1111/josh.12418

Beressa, G., Biratu, A., Lencha, B., Sahiledengle, B., Zenbaba, D., Bekele, D., et al. (2024). Association between dietary diversity, nutritional status, and academic performance of school-age children in Southeast Ethiopia using structural equation modelling. *J. Health Popul. Nutr.* 43:188. doi: 10.1186/s41043-024-00687-0

Bettini, V, and Giuliani, M (2016) Performances scolaires: investigation des facteurs cognitifs, psychologiques et sociodémographiques. Available online at: https://serval.unil.ch/resource/serval:BIB_S_000000024327.P001/REF.pdf (accessed May 25, 2025).

Black, M. M., Walker, S. P., Fernald, L. C. H., et al. (2017). Early childhood development coming of age: science through the life course. *Lancet* 389, 77–90.

Boivin, M. J., and Giordani, B. (1993). Improvements in cognitive performance for schoolchildren in Zaire, Africa, following an iron supplement and treatment for intestinal parasites. *J. Pediatr. Psychol.* 18, 249–264. doi: 10.1093/jpepsy/18.2.249

Bouchefra, S., El Ghouddany, S., Ouali, K., et al. (2023). Is good dietary diversity a predictor of academic success? *Acta Biomedica Atenei Parmensis* 94:e2023014. doi: 10.23750/abm.v94i2.13940

Bronfenbrenner, U (2013) Bronfenbrenner's ecological systems theory. The Psychology Notes HQ Online Resources for Psychology Students

Buthelezi, M. M., Ncisana, L., and Ntshangase, M. X. (2025). Food insecurity versus learners' performance in the classroom: a case of South African rural school. *Res. Soc. Sci. Technol.* 10, 130–151. doi: 10.46303/ressat.2025.30

Chang, M., Park, B., Singh, K., and Sung, Y. Y. (2009). Parental involvement, parenting behaviors, and children's cognitive development in low-income and minority families. *J. Res. Childhood Educ.* 23, 309–324. doi: 10.1080/02568540909594663

Charan, J., and Biswas, T. (2013). How to calculate sample size for different study designs in medical research? *Indian J. Psychol. Med.* 35, 121–126. doi: 10.4103/0253-7176.116232

Clotfelter, CT, Ladd, HF, and Vigdor, JL (2007). How and Why do Teacher Credentials Matter for Student Achievement? Working Paper Series 12828, Working Paper. National Bureau of Economic Research. Available online at: https://www.nber.org/papers/w12828 (Accessed 15 May 2025)

Coates, J, Swindale, A, and Bilinsky, P (2007). Household Food Insecurity Access Scale (HFIAS) for Measurement of Food Access: Indicator Guide: Version 3: (576842013–001). Available online at: https://doi.apa.org/doi/10.1037/e576842013-001 (Accessed July 1, 2024).

Cohen, J. F. W., Hecht, A. A., McLoughlin, G. M., Turner, L., and Schwartz, M. B. (2021). Universal school meals and associations with student participation, attendance, academic performance, diet quality, food security, and body mass index: a systematic review. *Nutrients* 13:911. doi: 10.3390/nu13030911

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychon. Bull. Rev.* 12, 769–786. doi: 10.3758/bf03196772

Correa-Burrows, P., Burrows, R., Blanco, E., Reyes, M., and Gahagan, S. (2016). Nutritional quality of diet and academic performance in Chilean students. *Bull. World Health Organ.* 94, 185–192. doi: 10.2471/BLT.15.161315

d'Aiglepierre, R. (2012). Primary school exclusion and ways to improve inclusion in Madagascar. New York: UNICEF.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *J. Fam. Psychol.* 19, 294–304. doi: 10.1037/0893-3200.19.2.294

De Paula, J.J., Malloy-Diniz, L.F., and Romano-Silva, M. A. (2016). Reliability of working memory assessment in neurocognitive disorders: a study of the Digit Span and Corsi Block-Tapping tasks. *Revista Brasileira de Psiquiatria* 38: 262–263.

Demetriou, A., Kazali, E., Kazi, S., and Spanoudis, G. (2020). Cognition and cognizance in preschool predict school achievement in primary school. *Cogn. Dev.* 54:100872. doi: 10.1016/j.cogdev.2020.100872

Djagba, A., Abodohoui, A., Nouatin, G. S. (2023). Fonctionnement des cantines scolaires installées dans le cadre des projets de développement dans le département de l'Atacora au Bénin. Revue Française d'Economie et de Gestion. 4:3.

Dubow, E. F., Boxer, P., and Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: mediation by family interactions, child aggression, and teenage aspirations. *Merrill-Palmer Q.* 55, 224–249. doi: 10.1353/mpq.0.0030

FAO. (2022). The State of Food Security and Nutrition in the World 2022: Repurposing Food and Agricultural Policies to Make Healthy Diets More Affordable. Food and Agriculture Organization of the United Nations. FAO. Available online at: http://www.fao.org/documents/card/en/c/cc0639en (Accessed February 17, 2025).

Filmer, D., and Pritchett, L. (1999). The effect of household wealth on educational attainment: evidence from 35 countries. *Popul. Dev. Rev.* 25, 85–120. doi: 10.1111/j.1728-4457.1999.00085.x

Frempong, G., Reddy, V., and Kanjee, A. (2011). Exploring equity and quality education in South Africa using multilevel models. *Compare J. Comp. Int. Educ.* 41, 819–835. doi: 10.1080/03057925.2011.607488

Glewwe, P., and Muralidharan, K. (2016). "Improving education outcomes in developing countries" in Handbook of the economics of education (Amsterdam, Netherlands: Elsevier), 653–743.

Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: an instrumental variables approach. *Am. Educ. Res. J.* 47, 434–465. doi: 10.3102/0002831209350494

Hadidjaja, P., Bonang, E., Suyardi, M. A., Abidin, S. A., Ismid, I. S., and Margono, S. S. (1998). The effect of intervention methods on nutritional status and cognitive function of primary school children infected with Ascaris lumbricoides. *Am. J. Tropical Med. Hygiene* 59, 791–795. doi: 10.4269/ajtmh.1998.59.791

Hafte Teklay, G., and Verstegen, D. (2023). Assessing predictors of students' academic performance in Ethiopian new medical schools: a concurrent mixed-method study. *BMC Med. Educ.* 23:1–9. doi: 10.1186/s12909-023-04372-4

Harris, J. R. (2011). Peer assessment in large undergraduate classes: an evaluation of a procedure for marking laboratory reports and a review of related practices. *Adv. Physiol. Educ.* 35, 178–187. doi: 10.1152/advan.00115.2010

Heckman, JJ. (2012). Invest in early childhood development: Reduce deficits, strengthen the economy. Available online at: https://heckmanequation.org/wp-content/uploads/2013/07/F_HeckmanDeficitPieceCUSTOM-Generic_052714-3.pdf (accessed May 25, 2025).

Hedeker, D., Siddiqui, O., and Hu, F. B. (2000). Random-effects regression analysis of correlated grouped-time survival data. *Statistical Methods in Medical Research* 9:161–179.

INSAE_AGVSA (2018) Analyse Globale de la Vulnérabilité et de la Sécurité Alimentaire 2017_Rapport final AGVSA, 2018. 2017_Rapport final AGVSA_2018.pdf

Janz, K. F., Lutuchy, E. M., Wenthe, P., et al. (2008). Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. *Med. Sci. Sports Exerc.* 40, 767–772. doi: 10.1249/mss.0b013e3181620ed1

Jasinski, L. J., Berry, D. T. R., Shandera, A. L., and Clark, J. A. (2011). Use of the Wechsler adult intelligence scale digit span subtest for malingering detection: a meta-analytic review. *J. Clin. Exp. Neuropsychol.* 33, 300–314. doi: 10.1080/13803395.2010.516743

Jyoti, D. F., Frongillo, E. A., and Jones, S. J. (2005). Food insecurity affects school children's academic performance, weight gain, and social skills. *J. Nutr.* 135, 2831–2839. doi: 10.1093/jn/135.12.2831

Karbach, J., Gottschling, J., Spengler, M., Hegewald, K., and Spinath, F. M. (2013). Parental involvement and general cognitive ability as predictors of domain-specific academic achievement in early adolescence. *Learn. Instr.* 23, 43–51. doi: 10.1016/j.learninstruc.2012.09.004

Kedir, S., Abate, K. H., Mohammed, B., et al. (2024). Weekly iron-folic acid supplementation and its impact on children and adolescents iron status, mental health and school performance: a systematic review and meta-analysis in sub-Saharan Africa. *BMJ Open* 14:e084033. doi: 10.1136/bmjopen-2024-084033

Kennedy, PG, Ballard, T, and Dop, M (2011). Guidelines for measuring household and individual dietary diversity. Available online at: https://openknowledge.fao.org/items/9bb3c3ce-66cf-4b00-b460-de29822bbd75 (accessed May 25, 2025).

Khandker, S. R., Barnes, D. F., and Samad, H. A. (2012). The welfare impacts of rural electrification in Bangladesh. *Energy J.* 33, 187–206. doi: 10.5547/ISSN0195-6574-EJ-Vol33-No1-7

Kowalski, KC, Crocker, PR, and Donen, RM (2004). The physical activity questionnaire for older children (PAQ-C) and adolescents (PAQ-A) manual. College of Kinesiology, University of Saskatchewan. https://journal.ummat.ac.id/journals/5/articles/2361/supp/2361-7046-1-SP.pdf (accessed May 25, 2025).

Kristjansson, B., Petticrew, M., MacDonald, B., et al. (2007). School feeding for improving the physical and psychosocial health of disadvantaged students. *Cochrane Database Syst. Rev.* 2, 1–189. doi: 10.1002/14651858.CD004676.pub2

Kyere, P., Veerman, J. L., Lee, P., and Stewart, D. E. (2020). Effectiveness of school-based nutrition interventions in sub-Saharan Africa: a systematic review. *Public Health Nutr.* 23, 2626–2636. doi: 10.1017/S1368980020000506

La Nouvelle Tribune. (2024). CEP au Bénin: voici le lien pour consulter les résultats (89.67% de réussite, taux par département). Available online at: https://lanouvelletribune.info/2024/06/cep-2024-au-benin-voici-le-lien-pour-consulter-les-resultats-8964-de-taux-de-reussite/ (Accessed 19 May 2025).

La Rue, A., Koehler, K. M., Wayne, S. J., et al. (1997). Nutritional status and cognitive functioning in a normally aging sample: a 6-y reassessment. *Am. J. Clin. Nutr.* 65, 20–29. doi: 10.1093/ajcn/65.1.20

Lee, M. K. (2022). Factors affecting the change in fruit and vegetable intake and adherence to a diet plan in patients with gastric cancer. *Asia Pac. J. Oncol. Nurs.* 9:100062. doi: 10.1016/j.apjon.2022.04.001

Lee, JJH, and Guadagno, L (2015). World migration report 2015: migrants and cities. Available online at: https://policycommons.net/artifacts/1218434/world-migration-report-2015/1771516/ (Accessed 14 May 2025).

Lemos, G. C., Guisande, M. A., and Almeida, L. S. (2025). Cognitive abilities and school achievement: addressing challenges across adolescence. *J. Intelligence* 13:21. doi: 10.3390/jintelligence13020021

Lindahl, M., Palme, M., Sandgren-Massih, S., and Sjögren, A. (2014). A test of the Becker-tomes model of human capital transmission using microdata on four generations. *J. Hum. Cap.* 8, 80–96. doi: 10.1086/674104

Ma, X., Shen, J., Krenn, H. Y., Hu, S., and Yuan, J. (2016). A meta-analysis of the relationship between learning outcomes and parental involvement during early childhood education and early elementary education. *Educ. Psychol. Rev.* 28, 771–801. doi: 10.1007/s10648-015-9351-1

Mantzorou, M., Vadikolias, K., Pavlidou, E., Serdari, A., Vasios, G., Tryfonos, C., et al. (2020). Nutritional status is associated with the degree of cognitive impairment and depressive symptoms in a Greek elderly population. *Nutr. Neurosci.* 23, 201–209. doi: 10.1080/1028415X.2018.1486940

Murad, H., Fleischman, A., and Sadetzki, S. (2003). Small Samples and Ordered Logistic Regression: Does it Help to Collapse Categories of Outcome? *The American Statistician* 57: 155–160.

Nago, E. S., Lachat, C. K., Huybregts, L., Roberfroid, D., Dossa, R. A., and Kolsteren, P. W. (2009). Food, energy and macronutrient contribution of out-of-home foods in school-going adolescents in Cotonou, Benin. *Br. J. Nutr.* 103, 281–288. doi: 10.1017/S0007114509991668

Neumann, C., Harris, D. M., and Rogers, L. M. (2002). Contribution of animal source foods in improving diet quality and function in children in the developing world. *Nutr. Res.* 22, 193–220. doi: 10.1016/S0271-5317(01)00374-8

Neumann, C. G., Murphy, S. P., Gewa, C., Grillenberger, M., and Bwibo, N. O. (2007). Meat supplementation improves growth, cognitive, and behavioral outcomes in Kenyan Children1. *J. Nutr.* 137, 1119–1123. doi: 10.1093/jn/137.4.1119

OECD (2025). The organisation for economic co-operation and development. https://www.oecd.org/en.html (Accessed May 17, 2025).

Oneya, L., and Onyango, D. (2021). Perception of school stakeholders on the effect of school- home distance on students' academic performance among community secondary schools in Rorya District, Tanzania. *East Afr. J. Educ. Soc. Sci.* 2, 76–81. doi: 10.46606/eajess2021v02i01.0068

Peng, P., and Kievit, R. A. (2020). The development of academic achievement and cognitive abilities: a bidirectional perspective. *Child Dev. Perspect.* 14, 15–20. doi: 10.1111/cdep.12352

Piper, B., Zuilkowski, S. S., and Ong'ele, S. (2016). Implementing mother tongue instruction in the real world: results from a medium-scale randomized controlled trial in Kenya. *Comp. Educ. Rev.* 60, 776–807. doi: 10.1086/688493

Prado, EL, and Dewey, KG (2014). Nutrition and brain development in early life. Nutrition reviews 72(4). Oxford university press Oxford, UK: 267–284. Available online at: https://academic.oup.com/nutritionreviews/article-abstract/72/4/267/1859597 (accessed May 25, 2025).

Regard, M., Strauss, E., and Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Percept. Mot. Skills* 55, 839–844. doi: 10.2466/pms.1982.55.3.839

Sauzéon, H., Lestage, P., Raboutet, C., et al. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language* 89: 192–202.

Stephenson, K. B., Wegner, D. R., Hershey, T. G., Doty, T., Davis, E., Steiner-Asiedu, M., et al. (2023). Effect of Peanut paste-based ready-to-use school meals with and without Milk on fluid cognition in northern Ghana: a randomized controlled trial. *Am. J. Clin. Nutr.* 118, 782–791. doi: 10.1016/j.ajcnut.2023.08.001

Sun, Z., Yuan, Y., Xiong, X., Meng, S., Shi, Y., and Chen, A. (2024). Predicting academic achievement from the collaborative influences of executive function, physical fitness, and demographic factors among primary school students in China: ensemble learning methods. *BMC Public Health* 24:274. doi: 10.1186/s12889-024-17769-7

Tamiru, D., Melaku, Y., and Belachew, T. (2017). Food insecurity and its association with school absenteeism among rural school adolescents in Jimma zone, Ethiopia. *Asia Pacific J. Public Health* 29, 114–121. doi: 10.1177/1010539517691606

Tomul, E., et al. (2021). The relative effect of student, family and school-related factors on math achievement by location of the school. *Large-Scale Assess Educ.* 9:1–29. doi: 10.1186/s40536-021-00117-1

Trost, S. G., Mciver, K. L., and Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Med. Sci. Sports Exerc.* 37, S531–S543. doi: 10.1249/01.mss.0000185657.86065.98

UNESCO, 2023: GEM Report UNESCO (2023). Global Education Monitoring Report 2023: Technology in Education: A Tool on Whose Terms? 1st ed. GEM Report UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000385723 (accessed 5 May 2025).

UNICEF for every child. (2025). Available at: https://www.unicef.org/ (Accessed 17 May 2025).

United Nations Development Programme. (2024). Human Development Report 2024: Uncertain times, unsettled lives – shaping our future in a transforming world. New York: United Nations Development Programme. Available online at: https://hdr.undp.org/reports (Accessed February 23, 2025).

Uzosike, T. C. J., Okeafor, I., and Mezie-Okoye, M. (2020). Dietary diversity, nutritional status and academic performance of pupils in public primary schools in Port Harcourt metropolis. *J. Commun. Med. Prim. Health Care* 32, 42–56. doi: 10.4314/jcmphc.v32i2.4

Voss, C., Dean, P. H., Gardner, R. F., Duncombe, S. L., and Harris, K. C. (2017). Validity and reliability of the physical activity questionnaire for children (PAQ-C) and adolescents (PAQ-A) in individuals with congenital heart disease. *PLoS One* 12:e0175806. doi: 10.1371/journal.pone.0175806

Weltgesundheitsorganisation (2006). Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age; methods and development (deMOnis). WHO child growth standards. Geneva: WHO Press.

Wills, G., and Hofmeyr, H. (2019). Academic resilience in challenging contexts: evidence from township and rural primary schools in South Africa. *Int. J. Educ. Res.* 98, 192–205. doi: 10.1016/j.ijer.2019.08.001

WorldBank (2025). World Bank Group - International Development, Poverty and Sustainability. Available online at: https://www.worldbank.org/ext/en/home (Accessed May 17, 2025).

Wu, R., and Billard, A. (2021). Learning from demonstration and interactive control of variable-impedance to cut soft tissues. *IEEE/ASME Trans. Mechatron.* 27, 2740–2751. doi: 10.1109/TMECH.2021.3123356

Wechsler, D. (2008) *Wechsler Adult Intelligence Scale--Fourth Edition (WAIS-IV)*. APA PsycTests.

Yeh, C.-W., Lee, M.-S., Lo, Y.-T. C., Chen, Y. C., Chen, W. C., and Huang, Y. C. (2025). Relationship between dietary diversity, eating behavior, and cognitive performance in economically disadvantaged school children. *Nutrition* 131:112658. doi: 10.1016/j.nut.2024.112658

Zaini, A., et al. (2005). Effects of nutritional status on academic performance of Malaysian primary school children. *Asia Pacific J. Public Health* 17, 81–87. doi: 10.1177/101053950501700204

Zorza, J. P., Marino, J., and Acosta Mesas, A. (2016). Executive functions as predictors of school performance and social relationships: primary and secondary school students. *Span. J. Psychol.* 19:E23. doi: 10.1017/sjp.2016.37

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 02 October 2025
DOI 10.3389/feduc.2025.1574844

OPEN ACCESS

EDITED BY
Xinyue Qiu,
Beijing Union University, China

REVIEWED BY
Carla Trevisan Martins Ribeiro,
Fernandes Figueira Institute, Brazil
Christine Da Silva-Genest,
Université de Lorraine, France

*CORRESPONDENCE
Esther Moraleda Sepulveda
✉ esmora01@ucm.es

RECEIVED 24 April 2025
ACCEPTED 17 September 2025
PUBLISHED 02 October 2025

CITATION
Moraleda Sepulveda E, Loro Vicente N, Durán-Heras P and Chaves-Herráiz P (2025) Analysis of the pragmatic competence profile in the population with 22q11.2 syndrome: a comparison between syndromic presentations. *Front. Educ.* 10:1574844. doi: 10.3389/feduc.2025.1574844

Analysis of the pragmatic competence profile in the population with 22q11.2 syndrome: a comparison between syndromic presentations

Esther Moraleda Sepulveda*, Nadia Loro Vicente, Patricia Durán-Heras and Patricia Chaves-Herráiz

Faculty of Psychology, Complutense University of Madrid, Madrid, Spain

Introduction: Various studies confirm that individuals with 22q11.2 syndrome exhibit communicative alterations that affect their social interactions. However, few compare the pragmatic characteristics of these individuals based on the type of syndrome (either deletion or duplication).
Method: This research aims to address the pragmatic skills of 10 participants with 22q11.2 syndrome, comparing the types 22q11.2DupS and 22q11.2DS, within an age range of 7 to 17 years, thereby confirming or denying communicative differences between syndromic presentations. The evaluation was conducted using the Revised Quick Pragmatic Assessment Protocol, allowing observation of the communicative characteristics of participants at the textual, utterance, and interactive levels. Video recordings of conversational samples with close interlocutors were used for the evaluation.
Results: The results show that, in general, there are differences between the two types of the syndrome, with more pragmatic difficulties in the case of people 22q11.2DS.
Conclusion: These differences do not appear to be related to age, as the percentage of different items is similar across both age ranges. Some limitations of the study are discussed.

KEYWORDS

22q11.2, pragmatics, deletion, duplication, PREP-R

Introduction

22q11.2 syndrome, also known as velocardiofacial syndrome or DiGeorge syndrome (Conley et al., 1979; DiGeorge, 1965), is a genetic condition characterized by the loss (22q11.2DS) or duplication (22q11.2DupS) of genetic material in the q11.2 region of chromosome 22 (Bartik et al., 2022; Driscoll, 2001). While 22q11.2DS is the more prevalent form, with an estimated incidence of between 1 in 2,000 and 1 in 4,000 live births (Blagojevic et al., 2021; Devriendt et al., 1998; Driscoll, 2001), there is some uncertainty regarding the actual prevalence of 22q11.2DupS, which is considerably less studied due to lower medical detection rates and fewer apparent congenital anomalies in these individuals (Olsen et al., 2018). Both syndromes are characterized by the presence of congenital heart defects, dysmorphic facial features, cleft palate, and linguistic, educational, and psychological-psychiatric difficulties that significantly reduce quality of life (Álvarez et al., 2009; Cortés-Martín et al., 2022; Driscoll, 2001; Yu et al., 2019). Additionally, cognitively,

individuals with 22q11.2 exhibit intellectual capabilities in the borderline range, with an IQ between 70 and 85, and there may be cases where IQ falls between 55 and 70 (Drmic et al., 2022; Gerdes et al., 1999).

Studies on 22q11.2 syndrome suggest difficulties in the lexical-semantic and morphosyntactic components in both populations (Solot et al., 2000; Verbesselt et al., 2023). However, in the population with 22q11.2DS, there are more problems in designating and defining concepts, affecting lexical-semantic skills, while individuals with 22q11.2DupS exhibit morphosyntactic alterations, characterized by short and less complex sentences (Verbesselt et al., 2023).

Lastly, regarding pragmatics, the linguistic alterations present in this population tend to transform into pragmatic difficulties that negatively impact social interactions (Álvarez et al., 2009). Consequently, it has been observed that school-aged children with 22q11.2 syndrome show an inability to use contextual information for understanding, organizing, and expressing language appropriately, often focusing on secondary issues or details instead (Van Den Heuvel et al., 2018).

Non-verbal communication is also affected, as children with 22q11.2 tend to engage in editing tasks in close environments (Sebastián-Lázaro et al., 2020) and do not accurately interpret the facial expressions of their interlocutors to understand meaning in context (Murphy, 2004; Sebastián-Lázaro et al., 2020). Similarly, alterations in prosodic nuances are noted, resulting in reduced and disharmonious speech rhythm (Sebastián-Lázaro et al., 2022; Solot et al., 2000; Van Den Heuvel et al., 2017a).

These contextual difficulties, along with severely impaired intelligibility, result in lower communicative intent, expressed through less frequent communication and fewer statements (Van Den Heuvel et al., 2017a). In conclusion, it can be inferred that the enunciative pragmatics are altered due to a violation of the maxims proposed by Grice (1975).

Individuals with 22q11.2 deletion syndrome exhibit significant pragmatic impairments that affect their communicative competence in social contexts. These difficulties include challenges in interpreting tone of voice, facial expressions, abstract language, and emotional cues from interlocutors, as well as maintaining discourse coherence and conversational turn-taking. The literature suggests that these pragmatic disorders are secondary to an altered linguistic profile rather than primary social deficits, as difficulties are observed in semantic fluency, discourse organization, and lexical retrieval (Sebastián-Lázaro et al., 2020; De Smedt et al., 2007). The range of observed disorders includes problems with non-verbal communication, irrelevant or out-of-context verbal interventions, limited use of discourse connectors, and unnecessary visual details. Research primarily focuses on expressive and receptive language, semantic fluency, and emotional comprehension, assessed through standardized psychometric tests and parent questionnaires, although complementary methods such as clinical observation and spontaneous discourse analysis are increasingly recommended (Van Den Heuvel et al., 2017b). Participants are typically children and adolescents aged 5 to 21 years, often with borderline intellectual functioning or mild intellectual disability, which influences their pragmatic performance.

Regarding interactive pragmatics, individuals with 22q11.2 often have difficulties respecting turn-taking due to anxiety and time pressure, leading to multiple overlaps and interruptions during interactions (Sebastián-Lázaro et al., 2020). Theoretically, pragmatics refers to the use of language according to context and speaker intention, while interactive pragmatics emphasizes the dynamic aspects of interaction, such as meaning negotiation, conversational cooperation, and adaptation to the interlocutor. This distinction is supported by Grice's cooperative principles, and Sperber and Wilson's relevance theory (Escandell, 2006). As a result, individuals with 22q11.2 seem to struggle with positioning themselves within the adjacent pair during the communicative process (Solot et al., 2019). Additionally, concerning textual pragmatics, they experience challenges in recalling plots, resulting in narratives with incoherent content structure, thematic leakage, and few cohesive elements (Persson et al., 2006; Van Den Heuvel et al., 2017b).

Although scientific literature has highlighted the linguistic difficulties experienced by people with 22q11.2 syndrome in recent years, little research has been done on language differences based on the type of syndromic presentation in this population. In these investigations, Verbesselt et al. (2023) showed that children with 22q11.2DS displayed linguistic difficulties that began at the word level; the most common linguistic problems in children with 22q11.2DS began at the sentence level. Notably, both expressive and receptive language, as well as lexical-semantic and morpho-syntactical domains, were affected in both types of syndromic presentations. However, no studies have been found that specifically focus on the development of the pragmatic component based on the existence of duplication or deletion.

Therefore, the aim of this research is to analyze the communicative profile of individuals with 22q11.2 in a natural context, in order to determine whether there are differences between the presentations of 22q11.2DupS and 22q11.2DS.

Method

Design

A cross-sectional study with a quasi-experimental design was conducted to compare pragmatic language performance between participants with different types of 22q11.2 deletion.

This non-randomized comparison of pre-existing groups was complemented by descriptive and observational strategies to address the research questions. In this way, we observe a dependent variable, a continuous quantitative type corresponding to the level of pragmatic skill, which has three levels: general pragmatic skill, specific pragmatic skill, and grammatical base pragmatic skill. Additionally, two independent variables are observed. First, there is a dichotomous nominal variable corresponding to the type of 22q11.2 syndrome, divided into two levels: deletion and duplication. Second, a second variable is noted, which is of an ordinal quantitative type, corresponding to age, with two levels: between 7 and 12 years, and between 12 and 17 years. Finally, it is worth noting that the various items of the evaluation tool (see Instruments section) were also analyzed individually, which

TABLE 1 Participants divided by groups and mean age.

Groups	22q11.2DupS (7–12 years)	22q11.2DS (7–12 years)	22q11.2DupS (12–17 years)	22q11.2DS (12–17 years)
Number of participants	2	2	3	3
Mean age	9.09	10.1	15.9	15.5
Standard deviation (SD)	2.64	2.68	1.73	2.3

constitute nominal qualitative variables with three response levels: yes, no, and not assessable.

Participants

The initial sample consisted of 12 participants. However, 2 of them were excluded because they did not meet the established inclusion criteria for the study, which will be explained later. Thus, the final sample for this study comprised 10 participants (5 males and 5 females), aged between 7 and 17 years, with a mean age of 13.3 (SD = 3.69). This group was divided based on the syndromic presentation.

Recruitment was conducted through convenience sampling in collaboration with 22q11.1 Spanish Association, which limited the pool of eligible participants. The sex ratio in our sample (1:1) does not necessarily reflect the exact distribution in the general population with 22q11.2DS, but rather the composition of the accessible sample during the recruitment period.

The first group consisted of 5 individuals with 22q11.2DupS (3 females and 2 males), with a mean age of 13.2 (SD = 4.17), while the second group comprised 5 individuals with 22q11.2DS (2 females and 3 males), with a mean age of 13.7 (SD = 4.12). Additionally, each group was further subdivided into 2 age ranges (7 to 12 years and 12 to 17 years), resulting in 4 participants (2 from each group) in the 7–12 age range, with a mean age of 9.59 (SD = 2.25), and 6 participants (3 from each group) in the 12–17 age range, with a mean age of 15.7 (SD = 1.84). This information is presented visually in the following table (Table 1).

To obtain a comparable analysis of results, a comparison was made between each type (participants with 22q11.2DS and users with 22q11.2DupS), further divided into the age range of 7 to 12 years and the age range of 12 to 17 years. Given the broad age range, there may be differences at the cognitive, linguistic, and emotional levels, with participants in the first group having a simpler language and social environment than those in the second group, who communicate in a wider and more complex context. For these reasons, it was decided to divide the participants into two age ranges to try to control this extraneous variable.

To collect the sample, the inclusion criteria established were that participants had a diagnosis of 22q11.2 made by a medical specialist, had a minimum verbal competence to be evaluated, were Spanish speakers, that the videos displayed the characteristics explained when contacting the participants (detailed further in the Procedure section), absence of comorbidities, and that both their legal guardians and the participants themselves provided consent to conduct the research. Exclusion criteria included participants who did not have a clear syndromic specificity in the diagnosis of 22q11.2 Syndrome or the presence of other disorders and comorbidities. It is important to note that all participants had an Intelligence Quotient (IQ) in the range of 70–79, ensuring that they were equated based on IQ (p s> 0.05).

Instruments

The Revised Quick Pragmatic Assessment Protocol (PREP-R) (Gallardo Paúls et al., 2015) was used, an instrument composed of items divided and explained according to the classification of pragmatic types proposed by Gallardo-Paúls (2009) and discussed in the theoretical framework (enunciative pragmatics, textual pragmatics, and interactive pragmatics). It also allows for differentiation between communicative problems arising from deficits in language components (grammatical base pragmatic deficits) and communicative problems resulting from specific pragmatic deficits. Overall, the global assessment refers to the general pragmatic skill of each evaluated subject but allows for the calculation of percentages of preserved specific pragmatic skill and grammatical base pragmatic skill.

This qualitative test enables the analysis and evaluation of the skills and/or difficulties an individual may present in interactions within their everyday ecological environment, recording difficulties and the strategies or behaviors the speaker uses to compensate for or mask these communicative limitations. It is divided into 18 items organized into three levels of pragmatic analysis: 6 items for enunciative pragmatic evaluation, 5 items for textual pragmatic evaluation and 7 items for interactive pragmatic evaluation.

The enunciative and textual levels are grouped into sublevels within each level. In this regard, the enunciative level has three sublevels (Speech Acts, Editing Tasks, and Inferences), and the textual level has two (Coherence and Cohesion), while the interactive level is not divided into sublevels. Furthermore, when a single item requires examination of several aspects, these are organized into sub-items, and each item and sub-item includes a brief explanation to guide the evaluator and remind them of the behaviors to observe in each case.

The evaluation is conducted through systematic analysis of language samples, typically collected in naturalistic or semi-structured contexts. Each item is scored based on observed communicative behaviors, and the results are expressed as percentages of preserved ability in three domains: general pragmatic ability, specific pragmatic ability, and grammar-based pragmatic ability. These percentages, as shown in Tables 2, 3, are calculated by dividing the number of items scored positively in each domain by the total number of items assessed, providing a profile of strengths and weaknesses.

TABLE 2 Percentages of pragmatic skills in the age range of 7 to 12 years.

Group	General pragmatic skill (GPS)	Average percentage of GPS	Specific pragmatic Skill (SPS)	Average percentage of SPS	Grammatical base pragmatic skill (GBPS)	Average percentage of GBPS
22q11.2DupS	67%	68%	59%	66%	86%	80%
	69%		74%		75%	
22q11.2DS	88%	85%	88%	84%	87%	87%
	81%		80%		86%	

TABLE 3 Percentages of pragmatic skills in the age range of 12 to 17 years.

Group	General pragmatic skill (GPS)	Average percentage of GPS	Specific pragmatic skill (SPS)	Average percentage of SPS	Grammatical base pragmatic skill (GBPS)	Average percentage of GBPS
22q11.2DupS	65%	75%	50%	70%	100%	92%
	87%		82%		100%	
	72%		79%		75%	
22q11.2DS	71%	60%	76%	60%	57%	61%
	85%		79%		100%	
	24%		26%		25%	

Given its central role in ensuring the accuracy and reliability of the data, transcription was undertaken in accordance with the main conventions of the PerLa Corpus (Fernández-Urquiza and Gallardo-Paúls, 2015). These conventions comprise the verbatim reproduction of utterances, the use of square brackets to indicate overlapping speech, ellipses for pauses or unfinished utterances, capitalization to mark prosodic emphasis, and standardized symbols for non-verbal elements. Adhering to these criteria reinforced the methodological rigor of the study and facilitated the interpretation of the illustrative examples.

The PREP-R has been applied primarily in clinical populations with language impairments, including children with neurodevelopmental disorders. In our study, it was used to explore pragmatic profiles in individuals with 22q11.2 deletion and duplication syndromes. We will revise the methodology section to include a clearer description of the scoring system, the rationale for using this tool, and its application context. The use of this test has been studied to assess pragmatics in other populations with intellectual disabilities, such as Down syndrome (Moreno and Díaz, 2014) and Williams syndrome (Shiro et al., 2016).

Procedure

First, a review of the existing scientific literature was conducted to establish the theoretical framework for the research, corresponding with the introduction part. This analysis confirmed the need to study pragmatics in individuals with 22q11.2 syndrome, specifically in comparison between the two existing types. Following the information search, the documentation was sent to the ethics committee, and upon approval, data collection for the participants began along with the distribution of informed consent.

Subsequently, a text message was drafted for the families of the subjects, detailing the procedures to be followed and explaining the nature of their participation. The message was disseminated via email to potential participants who met the selection criteria, and whose information was provided by principal investigator. It is important to note that prior to the data transfer, consent was signed regarding data confidentiality and the use of the research data.

Video recordings lasting between 10 and 15 min were requested, in which the individual with 22q11.2 should be seen interacting with a family member or someone with whom they spent the most time. The decision to use video recording was made as it is the most recommended method in the protocol guidelines (Bertrán et al., 2018). The use of video recordings to capture spontaneous language samples was essential for ensuring the accuracy, richness, and reliability of the pragmatic analysis. Unlike live observation or audio-only formats, video allowed for the detailed review of both verbal and non-verbal communicative behaviors, such as facial expressions, gestures, eye contact, and turn-taking dynamics. This multimodal perspective is particularly important when evaluating pragmatic competence, as it provides contextual cues that are crucial for interpreting the speaker's intentions and interactional strategies. Additionally, video recordings enabled repeated viewing and collaborative coding among researchers, which strengthened the consistency of the evaluations and facilitated a more nuanced understanding of the participants' communicative profiles.

Furthermore, the recordings took place at home, as this was where the participants felt comfortable conducting conversational activities. Specifically, individuals were to engage in conversations that were as natural as possible with an interlocutor. Once the families completed the videos, they sent the corresponding material to the study's principal investigator via an encrypted link. The

recordings were taken into account in their entirety in an attempt to establish a more comprehensive and complete pragmatic profile.

Previously, socio-family and educational data were also collected from the families. All families belonged to a middle-income bracket, and the parents' educational attainment was middle-to-high.

Ethical aspects

This study has been approved by the Social Research Ethics Committee of UCLM under reference CAU-683200-X6H7. Thus, the study adheres to the ethical principles outlined in the Declaration of Helsinki regarding research with human beings, as well as to current Spanish legislation which stipulates that participants must have the necessary information about the project to decide whether or not they wish to participate. In this regard, all subjects were thoroughly and properly informed about the process, ensuring that their participation in the study was voluntary. However, since all participants were minors, authorization from their parents or legal guardians was required, who were also properly informed.

Regarding confidentiality and data protection, informed consent was developed and signed, taking into account the confidentiality guarantees established by the Spanish's laws. Only the research team had access to the collected and pseudonymized data, which were maintained anonymously, meaning that no names or identifying details were included other than the age of each participant. In this way, the confidentiality of all study participants is fully guaranteed.

Data analysis

Ten video recordings were analyzed, one for each case. The total analysis period was 8 weeks, with the first 2 weeks dedicated to data transcription following the conventions established by the PerLa corpus (Gallardo Paúls and Veyrat Rigat, 2004), which provided uniformity in the transcribed representation of the data for subsequent evaluation. In the following 5 weeks, each video underwent individual analysis using the PREP-R protocol by each evaluator, assessing both each item proposed in the protocol and the percentages of pragmatic skills. Finally, in the last week of analysis, an inter-rater agreement process was conducted.

Due to the qualitative and subjective nature of the PREP-R, an inter-rater evaluation was conducted with three members. This methodology involves comparing and averaging the results of one evaluator with those of another. This approach helps to avoid biases when comparing the assessment of each criterion by three different evaluators. The Kappa coefficient of Cohen was used to analyze the degree of agreement among the evaluations, given that the variables were nominal. Following a thorough individual analysis of each conversational sample, a 69% agreement was obtained. After this, a meeting was held with the three evaluators to discuss the disagreements, resulting in a final agreement of 97% in the outcomes. Lastly, to establish a comparison between the groups, and as shown subsequently in Tables 1, 2, a weighted average of the percentages of pragmatic skills described in the PREP-R was calculated, visually displaying the differences and indicating which aspects each group scored higher in comparison to their counterparts with the other type of syndrome.

Results

The results have revealed that in both general and specific pragmatic skills, the group of individuals with 22q11.2DS scored higher than the group with 22q11.2DupS in the age range of 7 to 12 years, while the percentages of grammatical base pragmatic skills were very similar (Table 2).

In the age range of 12 to 17 years, a superiority is observed in the percentages of all evaluated pragmatic skills among individuals with 22q11.2DupS compared to participants with 22q11.2DS (Table 3).

The results obtained are a product of the analysis using the PREP-R. However, it is important to note that there were items that could not be evaluated in all situations due to the characteristics of the interaction not allowing for the specific aspect to be assessed. The following discusses the different items based on the type of pragmatics they evaluate.

Firstly, the results indicate an uneven profile between individuals with 22q11.2DupS and 22q11.2DS in the components of enunciative pragmatics. In this regard, it can be observed that the group with 22q11.2DupS in the age range of 7 to 12 years shows a higher percentage in the production of enunciative acts. In contrast, there are no differences between both groups in the comprehension and/or production of propositional acts, pauses and intraturn silences, direct speech acts, indirect speech acts, locutive acts, or draft acts, as shown in Extract 1.

The following are several examples between the reference adult (A) and the person with 22q11.2 (Q).

Extract 1

Example of the use of a verbal strategy that allows for gaining time for the construction of utterances (draft act).

User 12, 8 years old, with 22q11DS

- A: What about the noodles?
- Q: These ramen ones that come in a—one—uh, this thing I don't know what it's called, a container.
- A: What are those noodles like?
- Q: They're noodles that are like this long.

In the age range of 12 to 17 years, the results show that the group consisting of individuals with 22q11.2DupS also achieves a higher percentage in the production of enunciative acts and in the use of pauses and intraturn silences. Additionally, unlike the previous age range, in this range we find that the group with 22q11.2DS has a higher percentage in the comprehension and/or production of direct speech acts compared to the group with 22q11.2DupS. In the age range of 12 to 17 years, as in the previous age interval, no differences are found in the comprehension and/or production of propositional acts, indirect speech acts, as shown in Extract 2, locutive acts, and draft acts.

Extract 2

Example of correct production of an indirect speech act.
User 2, 17 years old, with 22q11DS

- *A: The Power Rangers, and the shows you used to watch on those videos you had, remember?*
- *Q: Yes, on Boing—they always played them in the summer because every time I watched Boing in the summer, I would always tell my brother, "COME ON ALONSO, LET'S PUT ON BOING because they're showing the Power Rangers now, I don't know what..." Well, every time we traveled—every time we were coming back from a trip—we would always either stay up late traveling or we would always take out the tablet and watch YouTube or whatever.*

Continuing with the paralinguistic elements and editing tasks, the results do not indicate any differences between individuals with 22q11.2DupS and 22q11.2DS in the age range of 7 to 12 years, except in the use of compensatory gesturing, where the group with 22q11.2DupS achieved a higher percentage. In contrast, it was the group with 22q11.2DS that obtained the highest percentage in rectification capacity and metapragmatic awareness (Extract 3).

Extract 3

Example of rectification in the production of a statement.
User 4, 11 years old, with 22q11DS
Q: Eeh/ I'm 9 years old, no - I'm 10 years old - 11 years and 9 months and ///

On the other hand, in the age range of 12 to 17 years, the use of compensatory gesturing is the only item in the editing tasks where differences are found, with the group with 22q11.2DS standing out compared to the group with 22q11.2DupS (Extract 4).

Extract 4

Example of a gesture that regulates verbal production (compass regulator).
User 2, 17 years old, with 22q11DS

- *A: Yesterday was the presentation and today you start it.*
- *Q: And every week we cover a topic (makes a gesture with one finger) from that course.*
- *A: And you have to do...*
- *Q: An exam and some activities on that topic.*

To conclude, regarding the principle of conversational cooperation, the results indicate that in the age range of 7 to 12 years, the group with 22q11.2DS commits fewer violations of the maxims of quality, manner, and relevance, as well as of particularized implicatures (Extract 5). No differences were observed in the maxim of quantity or in conventional implicit acts, also known as lexicalized expressions or idioms.

Extract 5

Example of a violation of the maxim of manner through ambiguous verbal production.
User 12, 8 years old, with 22q11DS

- *A: The Great Wall of China! Uh, how do you play that? I really don't know anything at all.*
- *Q: Look, one person has to stand in the middle, and everyone else—the one in the center has to say "wall," and the others have to say "China," and we have to start running so they don't catch us // (gestures exhaustion).*
- *A: And what?*
- *Q: And you have to get to the other side. You have to say "Great Wall of China" every time and then run away, and that's it.*

However, in the age range of 12 to 17 years, we find a less homogeneous profile. While the group with 22q11.2DupS commits fewer violations of the maxims of quality, manner, and relation, they produce more violations of the maxim of quantity and particularized implicatures compared to the group with 22q11.2DS (Extract 6). On the other hand, no differences were found in lexicalized expressions or idioms.

Extract 6

Example of a violation of the maxim of quantity due to insufficiency.
User 7, 14 years old, with 22q11DupS

- *A: Okay, how's school going // in high school?*
- *Q: Good.*
- *A: Do you have many friends?*
- *Q: Yes.*
- *A: And girl friends?*
- *Q: No.*

In Table 4, the enunciative differences are presented visually, showing the percentages of each participant group for each item.

Continuing with textual pragmatics, in the age range of 7 to 12 years, specifically within the coherence sublevel, the group of individuals with 22q11.2DS demonstrates a higher percentage in the narrative superstructure item (Extract 7). On the other hand, there are no differences between the two groups regarding argumentative superstructure, recognition of a new topic, or the introduction of a thematic shift.

Extract 7

Example of Incorrect Use of Narrative Superstructure Due to Inadequate Character Presentation.
User 1, 12 years old with 22q11DS

- *Q: No, it's Raquel; she loves you more, Barbie.*
- *A: Who is Raquel? I don't know who Raquel is.*

TABLE 4 Characteristics of enunciative pragmatics according to PREP-R items.

Ítems	Group 22q11.2DupS (7–12 years)	Group 22q11.2DS (7–12 years)	Group 22q11.2DupS (12–17 years)	Group 22q11.2DS (12–17 years)
Enunciative acts	100%	0%	100%	66.7%
Propositional acts	100%	100%	100%	100%
Intra-turn pauses and silences	100%	100%	100%	66.7%
Direct speech acts	50%	50%	33%	67%
Indirect speech acts	50%	50%	67%	67%
Locutionary acts	100%	100%	100%	100%
Erasing acts	100%	100%	100%	100%
Compensatory gestures	100%	50%	67%	100%
Rectification and metapragmatic awareness	50%	100%	67%	67%
Quality implicature	50%	100%	67%	33%
Quantitative implicature	0%	0%	0%	33%
Manner implicature	0%	100%	67%	33%
Relation implicature	50%	100%	67%	33%
Specific implicatures	50%	100%	33%	67%
Lexicalized expressions or idioms	50%	50%	67%	67%

- *Q: A friend of Barbie's who has a brother named Ryan, who is Ken's friend.*
- *A: Oh, okay, okay, okay, do we have that Raquel with us?*
- *Q: No.*

Unlike the previous age interval, in the group aged 12 to 17 years, individuals with 22q11.2DupS show a higher percentage compared to their peers with 22q11.2DS in items related to narrative superstructure, recognition of a new theme or thematization, and appropriate thematic change. In contrast, the results do not show differences between the two groups in the item related to argumentative superstructure.

Regarding the sublevel of cohesion, in the age group of 7 to 12 years, there is a higher percentage of lexical effectiveness in the 22q11.2DS group compared to individuals with 22q11.2DupS (Extract 8). However, with respect to morphology, word formation, syntax, and grammatical construction, no differences are found between the two groups.

Extract 8

Example of correct lexical effectiveness, without repetitions or empty words.

User 4, 11 years old with 22q11DS

- *A: What's it about? Because Minecraft is a big topic to discuss.*
- *Q: (LAUGHS) Uh, about Minecraft—well, Minecraft is mainly for building, for imagination, and of course, the game is… is it normal? No, it's all made of squares. I mean, yes, although some things can be round, but no—everything has square pixels. So there's nothing round, or rectangular, or anything.*

On the other hand, the group of individuals with 22q11.2DupS in the age range of 12 to 17 years shows a higher percentage in all items related to textual cohesion, that is, in lexical effectiveness, morphology and word formation, as well as syntax and grammatical construction. Table 5 shows the percentages of groups of individuals with 22q11.2DupS and 22q11.2DS in both age ranges for each of the items related to textual pragmatics.

Lastly, regarding interactive pragmatics, in the age range of 7 to 12 years, we find differences in turn-taking fluency, conversational participation, predictability, and the communicative use of eye contact, where the group of individuals with 22q11.2DS shows a higher percentage. In contrast, in the item related to natural gesturing to complement language (as shown in Extract 9), individuals with 22q11.2DupS scored better compared to participants with 22q11.2DS.

Extract 9

Example of correct use of natural gesturing.

User 6, 7 years old with 22q11DupS

- *Q: I can escape from Manuel.*
- *A: Really?*
- *Q: Yes, I'm running away (gesture of running)*

On the other hand, the results do not show differences between the two types of syndrome concerning the agility and speed of turn-taking and the design of turns based on conversational priority.

Unlike the previous age range, in the interval of 12 to 17 years, there is a higher percentage in the group of individuals with 22q11.2DupS regarding items related to turn agility, turn-taking,

TABLE 5 Characteristics of textual pragmatics according to the items of the PREP-R.

Items	Group 22q11.2DupS (7–12 years)	Group 22q11.2DS (7–12 years)	Group 22q11.2DupS (12–17 years)	Group 22q11.2DS (12–17 years)
Narrative superstructure	0%	50%	33%	0%
Argumentative superstructure	50%	50%	67%	67%
Thematization	100%	100%	100%	100%
Thematic change	50%	50%	67%	33%
Lexical effectiveness	50%	100%	67%	33%
Morphology and word formation	100%	100%	100%	67%
Syntax and grammatical construction	100%	100%	100%	67%

TABLE 6 Characteristics of interactive pragmatics according to the items of the PREP-R.

Items	Group 22q11.2DupS (7–12 years)	Group 22q11.2DS (7–12 years)	Group 22q11.2DupS (12–17 years)	Group 22q11.2DS (12–17 years)
Agility of turn	100%	100%	100%	67%
Turn-taking	50%	100%	67%	33%
Conversational participation	0%	50%	33%	33%
Predictability	50%	100%	33%	0%
Priority	50%	50%	33%	67%
Natural gesturing	100%	50%	67%	67%
Communicative use of eye contact	50%	100%	67%	67%

and predictability. In contrast, individuals with 22q11.2DS design more turns according to the principles of conversational priority than their peers with 22q11.2DupS. However, the results do not show differences between the two types in this age range regarding the conversational participation index, the use of natural gesturing, and the communicative use of eye contact.

Finally, Table 6 shows the percentages for each item related to interactive pragmatics.

Discussion

The study aims to provide novel insights into the pragmatic profile of individuals with 22q11.2 syndrome, establishing differences between the two typologies: 22q11.2DupS and 22q11.2DS. The study highlights that children with 22q11.2DS not only present articulation difficulties, consistent with previous findings on speech sound disorders in this population (Everaert et al., 2023; Persson et al., 2003; Solot et al., 2019), but also exhibit pragmatic impairments that affect their ability to use language effectively in social contexts. These challenges are likely influenced by underlying facial and palatal anomalies (Goldmuntz, 2020; Solot et al., 2001; Yu et al., 2019), which contribute to both structural speech deficits and limitations in expressive communication, further impacting pragmatic functioning. Furthermore, while research by Van Den Heuvel et al. (2017b) indicates that individuals with 22q11.2SD require training to avoid inappropriate pauses in conversations, the results of this study show that most participants manage pauses and intra-turn silences appropriately.

On the other hand, the results align with research indicating that individuals with 22q11.2 syndrome often struggle with tasks involving "theory of mind," which could explain the disparate scores in tasks such as indirect speech acts or the use of lexicalized expressions and idioms (Laorden et al., 2019; Niklasson et al., 2002, 2009). In line with this, Van Den Heuvel et al. (2017a) suggest that individuals with 22q11.2 have difficulty incorporating such speech acts appropriately within context.

Persson et al. (2006) proposed that individuals with 22q11.2 employ editing strategies, such as gestures and paralinguistic cues, due to delays in oral language acquisition. This theory could explain the findings of this study, as most participants consistently and correctly used compensatory behaviors, metapragmatic awareness, and natural gesturing. These results also relate to studies highlighting expressive language poverty compared to receptive language across all ages in individuals with 22q11.2 (Roche et al., 2020; Solot et al., 2019), thus transgressing the maxim of quantity and compensating for difficulties with gestures.

Regarding textual pragmatics, the study shows that a low percentage of individuals can produce coherent narratives, which is consistent with studies indicating that individuals with 22q11.2 exhibit similarities to those with language disorders concerning narrative difficulties (Selten et al., 2021). For instance, Boerma et al. (2023) reported weak narrative skills in children with 22q11.2, particularly regarding macrostructure, which could

explain their challenges in presenting events, characters, and, generally, information in a logical and coherent order.

Regarding thematic management, the results indicate differences between the two types of 22q11.2 syndrome in the age range of 12 to 17 years, with individuals with 22q11.2DS facing more difficulties. These findings do not support the hypothesis of Wenger et al. (2016), who assert that individuals with 22q11.2DupS exhibit restricted interests, which could lead users to repeatedly return to the same topic of conversation (Spiker et al., 2012). In terms of cohesion, the evidence from this research shows that most participants use words and word constructions appropriately, although some specific deficits were noted. This may be attributed to a good command of vocabulary and grammar among participants (Louwerse, 2004), contradicting previous studies that reported difficulties in syntax and discourse organization (Solot et al., 2000; Verbesselt et al., 2023).

The study's results reveal that the majority of participants do not demonstrate adequate conversational engagement, which may be due to a passive and withdrawn conversational style, particularly highlighting the lack of conversational participation by individuals with 22q11.2DS in the 7 to 12 age range (Van Den Heuvel et al., 2017b). Several studies also report challenges in initiating conversations and taking turns (Angkustsiri et al., 2014; Van Den Heuvel et al., 2017b). While the results vary concerning age and type, individuals with 22q11.2DS appear to experience increasing difficulties over time with rising social demands (Sebastián-Lázaro et al., 2020), whereas individuals with 22q11.2DupS may improve their skills as they grow, presenting milder symptoms (Verbesselt et al., 2022). Communicative participation may also be influenced by personal and environmental factors, as well as the interest in the interlocutor (Blum-Kulka et al., 2010; Shea, 2022). Thus, the study's results, in alignment with various studies, suggest that children tend to interact more extensively and confidently with their peers compared to adults, facilitating a reciprocal exchange of idea. Finally, difficulties in the communicative use of gaze were observed in participants with 22q11.2DupS, which could be related to an alteration in the non-linguistic elements of communication, similar to individuals with ASD (Fernández and García, 2020; Wenger et al., 2016).

In summary, the study has allowed us to conclude that there are pragmatic differences between individuals with 22q11.2DupS and 22q11.2DS. However, certain limitations of the study necessitate further investigation in this area, such as the limited sample size due to the low prevalence of the syndrome and the geographic distribution of participants. In relation to this limitation, it would be valuable to evaluate a broader sample concerning age and social context, taking into account differences in family or educational environments.

In conclusion, the results reveal differences between individuals with 22q11.2DupS and 22q11.2DS across most items in the PREP-R protocol. However, these differences do not align with the theory that individuals with 22q11.2DupS present fewer difficulties due to less severe symptoms than those with 22q11.2DS, as the findings do not allow us to reach a consensus regarding the pragmatic severity of one profile compared to the other. Therefore, we can conclude that the main hypothesis posited is fulfilled, establishing differences between both subtypes in the 7 to 12-year age range as well as in the 12 to 17-year interval, with no significant variations between these two age ranges. These results underscore the need for speech therapy intervention within this population to enhance communicative proficiency and, consequently, to improve social skills and the quality of life for both individuals with 22q11.2 syndrome and their regular interlocutors. Pragmatic research on the syndrome is limited, highlighting an area of knowledge that requires further study to facilitate early diagnosis and the initiation of effective treatments, thereby improving the quality of life of individuals and their communities, opening up potential new avenues for research.

Limitations of the study

As possible limitations of the present study, it should be noted that the sample size could have been larger and that the results are primarily based on Spanish speakers, which does not allow us to generalize these findings to other languages. In addition, the duration of the video recordings may not have been sufficient to capture the full range of pragmatic abilities and difficulties, which could have limited the depth of the analyses. Nevertheless, this study provides new evidence regarding pragmatic processing in individuals with 22q11.2DS, although the underlying nature and extent of these alterations remain to be fully elucidated.

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 humans were approved by Social Research Ethics Committee of UCLM under reference CAU-683200-X6H7. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

EM: Data curation, Project administration, Supervision, Writing – original draft, Writing – review & editing. NL: Writing – original draft, Writing – review & editing. PD-H: Writing – original draft, Writing – review & editing. PC-H: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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

Álvarez, D., Palomares, M., and Villena, C. (2009). Descripción de características clínicas del síndrome velocardiofacial. *Rev. Chil. Fonoaudiol.* 9, 41–61. doi: 10.5354/0719-4692.2009.21093

Angkustsiri, K., Goodlin-Jones, B., Deprey, L., Brahmbhatt, K., Harris, S., and Simon, T. J. (2014). Social impairments in chromosome 22q11.2.2 deletion syndrome (22q11.2.2DS): autism spectrum disorder or a different endophenotype? *J. Autism Dev. Disord.* 44, 739–746. doi: 10.1007/s10803-013-1920-x

Bartik, L. E., Hughes, S. S., Tracy, M., Feldt, M. M., Zhang, L., Arganbright, J., et al. (2022). 22q11.2.2 duplications: expanding the clinical presentation. *Am. J. Med. Genet. A* 188, 779–787. doi: 10.1002/ajmg.a.62577

Bertrán, M., Tagle, F. P., and Irarrázaval, M. (2018). Psychiatric manifestations of 22q11.2.2 deletion syndrome: a literature review. *Soc. Esp. Neurol.* 33, 121– 128. doi: 10.1016/j.nrl.2015.07.007

Blagojevic, C., Heung, T., Theriault, M., Tomita-Mitchell, A., Chakraborty, P., Kernohan, K., et al. (2021). Estimate of the contemporary live-birth prevalence of recurrent 22q11.2.2 deletions: a cross-sectional analysis from population-based newborn screening. *CMAJ Open* 9, E802–E809. doi: 10.9778/cmajo.20200294

Blum-Kulka, S., Hamo, M., and Habib, T. (2010). Explanations in naturally occurring peer talk: conversational emergence and function, thematic scope, and contribution to the development of discursive skills. *First Lang.* 30, 440–460. doi: 10.1177/0142723710370528

Boerma, T., Everaert, E., Vlieger, D., Steggink, M., Selten, I., Houben, M., et al. (2023). Grammatical skills of Dutch children with 22q11.2.2 deletion syndrome in comparison with children with developmental language disorder: evidence from spontaneous language and standardized assessment. *Front. Commun.* 8, 1–16. doi: 10.3389/fcomm.2023.1111584

Conley, M. E., Beckwith, J. B., Mancer, J. F., and Tenckhoff, L. (1979). The spectrum of the DiGeorge syndrome. *J. Pediatr.* 94, 883–890. doi: 10.1016/S0022-3476(79)80207-3

Cortés-Martín, J., Peñuela, N. L., Sánchez-García, J. C., Montiel-Troya, M., Díaz-Rodríguez, L., and Rodríguez-Blanque, R. (2022). Deletion syndrome 22q11.2.2: a systematic review. *Children* 9:1168. doi: 10.3390/children9081168

De Smedt, B., Devriendt, K., Fryns, J. P., Vogels, A., Gewillig, M., and Swillen, A. (2007). Intellectual abilities in a large sample of children with velo–cardio–facial syndrome: an update. *J. Intellect. Disabil. Res.* 51, 666–670. doi: 10.1111/j.1365-2788.2007.00955.x

Devriendt, K., Fryns, J. P., Mortier, G., Van Thienen, M. N., and Keymolen, K. (1998). The annual incidence of DiGeorge/velocardiofacial syndrome. *J. Med. Genet.* 35, 789–790. doi: 10.1136/jmg.35.9.789-a

DiGeorge, A. M. (1965). Discussions on a new concept of the cellular base of immunology. *J. Pediatr.* 67:907. doi: 10.1016/S0022-3476(65)81796-6

Driscoll, D. A. (2001). Prenatal diagnosis of the 22q11.2.2 deletion syndrome. *Genet. Med.* 3, 14–18. doi: 10.1097/00125817-200101000-00004

Drmic, I. E., MacKinnon Modi, B., McConnell, B., Jilderda, S., Hoang, N., Noor, A., et al. (2022). Neurodevelopmental functioning in probands and non-proband carriers of 22q11.2.2 microduplication. *Am. J. Med. Genet. A* 188, 2999–3008. doi: 10.1002/ajmg.a.62916

Escandell, M. V. (2006). *Introducción a la Pragmática*. Barcelona: Ariel.

Everaert, E., Selten, I., Boerma, T., Houben, M., Vorstman, J., de Wilde, H., et al. (2023). The language profile of preschool children with 22q11.2.2 deletion syndrome and the relationship with speech intelligibility. *Am. J. Speech Lang. Pathol.* 32,128–144. doi: 10.1044/2022_AJSLP-21-00328

Fernández, L. C., and García, R. C. (2020). Development and improvement of the pragmatic competence in children and adolescents with autistic spectrum disorder and functional language. *Normas* 10, 1–20. doi: 10.7203/Normas.v10i1.19279

Fernández-Urquiza, M., and Gallardo-Paúls, B. (2015). Adaptación del formato CHAT para la transcripción y codificación de los datos de lenguaje patológico del corpus PerLA. *Oralia* 18, 81–102. doi: 10.25115/oralia.v18i.7315

Gallardo Paúls, B., Moreno Campos, V., Lázaro López-Villaseñor, M., Simón López, T., and Díaz Martínez, F. (2015). *Protocolo Rápido de Evaluación Pragmática-Revisado*. Valencia: Universit de València.

Gallardo Paúls, B., and Veyrat Rigat, M. (2004). *Estudios de lingüistica clínica: lingüística y patología*. Valencia: Universit de València.

Gallardo-Paúls, B. (2009). Valoración del componente pragmático a partir de datos orales. *Rev. Neurol.* 48, 57–61. doi: 10.33588/rn.48S02.2008736

Gerdes, M., Solot, C., Wang, P. P., Moss, E., LaRossa, D., Randall, P., et al. (1999). Cognitive and behavior profile of preschool children with chromosome 22q11.2.2 deletion. *Am. J. Med. Genet.* 85, 127–133.

Goldmuntz, E. (2020). 22q11.2.2 deletion syndrome and congenital heart disease. *Am. J. Med. Genet. C* 184, 64–72. doi: 10.1002/ajmg.c.31774

Grice, H. P. (1975). Logic and conversation. *Syntax Semant.* 3, 41–58. doi: 10.1163/9789004368811_003

Laorden, A. T., Carballo, J. J., María, B., Aragón, M., Fraguas, D., and Madrid, H. (2019). *Caracterización del rendimiento neuropsicológico en personas con síndrome de deleción 22q11.2.2.* Madrid: Universidad Complutense de Madrid.

Louwerse, M. M. (2004). Un modelo conciso de cohesión en el texto y coherencia en la comprensión. *Rev. Signos* 37, 41–48. doi: 10.4067/S0718-09342004005600004

Moreno, E., and Díaz, F. (2014). Evaluación del componente pragmático en el Síndrome de Down a través del Protocolo Rápido de Evaluación Pragmática. *Rev. Investig. Logop.* 4, 1–27. doi: 10.5209/rlog.58659

Murphy, K. C. (2004). The behavioural phenotype in velo-cardio-facial syndrome. *J. Intellect. Disabil. Res.* 48, 524–530. doi: 10.1111/j.1365-2788.2004.00620.x

Niklasson, L., Rasmussen, P., Óskarsdóttir, S., and Gillberg, C. (2002). Chromosome 22q11.2 deletion syndrome (CATCH 22): neuropsychiatric and neuropsychological aspects. *Dev. Med. Child Neurol.* 44, 44–50. doi: 10.1111/j.1469-8749.2002.tb00258.x

Niklasson, L., Rasmussen, P., Óskarsdóttir, S., and Gillberg, C. (2009). Autism, ADHD, mental retardation and behavior problems in 100 individuals with 22q11.2 deletion syndrome. *Res. Dev. Disabil.* 30, 763–773. doi: 10.1016/j.ridd.2008.10.007

Olsen, L., Sparsø, T., Weinsheimer, S. M., Dos Santos, M. B. Q., Mazin, W., Rosengren, A., et al. (2018). Prevalence of rearrangements in the 22q11.2.2 region and population-based risk of neuropsychiatric and developmental disorders in a Danish population: a case-cohort study. *Lancet Psychiatry* 5, 573–580. doi: 10.1016/S2215-0366(18)30168-8

Persson, C., Lohmander, A., Ènsson, J., Ålveig, S., Skarsdo Âttir, O. Â., and Èderpalm, E. S. (2003). A prospective cross-sectional study of speech in patients with the 22q11.2 deletion syndrome. *J. Commun. Disord.* 36, 13–47. doi: 10.1016/S0021-9924(02)00133-8

Persson, C., Niklasson, L., Óskarsdóttir, S., Johansson, S., Jönsson, R., and Söderpalm, E. (2006). Language skills in 5-8-year-old children with

22q11.2 deletion syndrome. *Int. J. Lang. Commun. Disord.* 41, 313–333. doi: 10.1080/13682820500361497

Roche, L., Campbell, L., and Heussler, H. (2020). "Communication in 22q11.2.2 deletion syndrome: a brief overview of the profile, intervention approaches, and future considerations. *Curr. Dev. Disord. Rep.* 7, 124–129. doi: 10.1007/s40474-020-00208-8

Sebastián-Lázaro, D., Brun-Gasca, C., and Fornieles, A. (2022). Communication in people with 22q11.2 deletion syndrome: voice and speech. *Rev. Logop. Fon. Audiol.* 42, 238–249. doi: 10.1016/j.rlfa.2021.12.001

Sebastián-Lázaro, D., Brun-Gasca, C., and Fornieles-Deu, A. (2020). The language of children and young people with 22q11.2 deletion syndrome. *Rev. Neurol.* 71, 99–109. doi: 10.33588/rn.7103.2020145

Selten, I., Boerma, T., Everaert, E., Vansteensel, M. J., Vorstman, J., and Wijnen, F. (2021). Narrative comprehension and production abilities of children with 22q11.2.2 deletion syndrome. *Res. Dev. Disabil.* 119, 104–109. doi: 10.1016/j.ridd.2021.104109

Shea, D. P. (2022). Perspective and production. Pragmatics. *Q. Pub. Int. Pragmat. Assoc.* 357–389. doi: 10.1075/prag.4.3.06she

Shiro, M., Diez-Itza, E., Viejo, A., and Fernández-Urquiza, M. (2016). "Pragmática evaluativa de las narraciones en el síndrome de Williams," in *VIIIth International Conference of Language Acquisition* (Palma de Mallorca: Universidad de las Islas Baleares), 31–32.

Solot, C. B., Gerdes, M., Kirschner, R. E., McDonald-McGinn, D. M., Moss, E., Woodin, M., et al. (2001). Communication issues in 22q11.2.2 deletion syndrome: children at risk. *Genet. Med.* 3, 67–71. doi: 10.1097/00125817-200101000-00015

Solot, C. B., Knightly, C., Handler, S. D., Gerdes, M., Mcdonald-Mcginn, D. M., Moss, E., et al. (2000). Communication disorders in the 22q11.2.2 microdeletion syndrome. *J. Commun. Disord.* 33, 187–203. doi: 10.1016/S0021-9924(00)00018-6

Solot, C. B., Sell, D., Mayne, A., Baylis, A. L., Persson, C., Jackson, O., et al. (2019). Speech-language disorders in 22q11.2.2 deletion syndrome: best practices for diagnosis and management. *Am. J. Speech Lang. Pathol.* 28, 984–999. doi: 10.1044/2019_AJSLP-16-0147

Spiker, M. A., Lin, C. E., Van Dyke, M., and Wood, J. J. (2012). Restricted interests and anxiety in children with autism. *Autism 16,* 306–320. doi: 10.1177/1362361311401763

Van Den Heuvel, E., Botting, N., Boudewijns, I., Manders, E., Swillen, A., and Zink, I. (2017b). Developmental course of conversational behaviour of children with 22q11.2.2 deletion syndrome and Williams syndrome. *First Lang.* 37, 583–611. doi: 10.1177/0142723717713890

Van Den Heuvel, E., Manders, E., Swillen, A., and Zink, I. (2018). Atypical language characteristics and trajectories in children with 22q11.2.2 deletion syndrome. *J. Commun. Disord.* 75, 37–56. doi: 10.1016/j.jcomdis.2018.06.001

Van Den Heuvel, E., ReuterskiöLd, C., Solot, C., Manders, E., Swillen, A., and Zink, I. (2017a). Referential communication abilities in children with 22q11.2.2 deletion syndrome. *Int. J. Speech Lang. Pathol.* 19, 490–502. doi: 10.1080/17549507.2016.1221456

Verbesselt, J., Solot, C. B., Van Den Heuvel, E., Crowley, T. B., Giunta, V., Breckpot, J., et al. (2023). Language profiles of school- aged children with 22q11.2.2 copy number variants. *Genes* 14:679. doi: 10.3390/genes14030679

Verbesselt, J., Van Den Heuvel, E., Breckpot, J., Zink, I., and Swillen, A. (2022). Parent- reported social-communicative skills of children with 22q11.2.2 copy number variants and siblings. *Genes* 13:1801. doi: 10.3390/genes13101801

Wenger, T. L., Miller, J. S., DePolo, L. M., De Marchena, A. B., Clements, C. C., Emanuel, B. S., et al. (2016). 22q11.2.2 duplication syndrome: elevated rate of autism spectrum disorder and need for medical screening. *Mol. Autism* 7:27. doi: 10.1186/s13229-016-0090-z

Yu, A., Turbiville, D., Xu, F., Ray, J. W., Britt, A. D., Lupo, P. J., et al. (2019). Genotypic and phenotypic variability of 22q11.2.2 microduplications: an institutional experience. *Am. J. Med. Genet. A* 179, 2178–2189. doi: 10.1002/ajmg.a.61345

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 27 November 2025
DOI 10.3389/fpsyg.2025.1612803

OPEN ACCESS

EDITED BY
Mingshuang Li,
California State University, Northridge,
United States

REVIEWED BY
Mohamed Taiebine,
UEMF, Morocco
Mengxuan Wu,
University of Texas–Pan American,
United States

*CORRESPONDENCE
Ying Hao
✉ ying.hao@njnu.edu.cn

RECEIVED 16 April 2025
REVISED 29 October 2025
ACCEPTED 11 November 2025
PUBLISHED 27 November 2025

CITATION
Hao Y, Park H and Webb J (2025) Digital bonds: patient and therapist factors influence telehealth rapport building in speech-language services.
Front. Psychol. 16:1612803.
doi: 10.3389/fpsyg.2025.1612803

Digital bonds: patient and therapist factors influence telehealth rapport building in speech-language services

Ying Hao[1]*, Hyejin Park[2] and Jaret Webb[3]

[1]School of Chinese Language and Literature, Nanjing Normal University, Nanjing, China, [2]Department of Communication Sciences and Disorders, University of Mississippi, Oxford, MS, United States, [3]Baptist Health Medical Center, Little Rock, AR, United States

Introduction: Rapport building is challenging in pediatric telehealth speech-language services, yet factors influencing it remain unclear. The study specified how patient and clinician factors contributed to the issue.
Methods: Generally, the two disorders reflect different levels of behavioral, sensory, and cognitive challenges, and the two ages index different developmental stages. This study surveyed 207 speech-language pathologists (SLPs) about three aspects (i.e., importance, strategy use, and achievement) of rapport when working remotely with children diagnosed with speech sound disorders (SSD) or autism spectrum disorder (ASD) at two ages (0–3 years, 4–8 years). Clinician factors included clinician age, telehealth experience, and digital literacy.
Results: Regarding patient factors, SLPs rated the ASD group higher importance, more strategy use, but lower achievement than the SSD group. Clinicians reported more strategy use and lower achievement of rapport when serving younger children, but a main effect of patient age was not found in the perceived importance of rapport. Regarding clinician factors, older SLPs tended to place higher importance, use more strategies, and feel more achieved on telehealth rapport than younger SLPs. While SLPs without telehealth experience reported similar levels of strategy use for SSD and ASD groups, those with experience, regardless of the diversity of disorder-age combinations, used strategies more frequently for the ASD group. Finally, digital literacy was significantly related to SLPs' perceived levels of achievement.
Conclusion: Overall, these findings underscore the importance of both patient and clinician factors when establishing rapport in telehealth, which may be implicational for other professionals who provide services to individuals with disabilities.

KEYWORDS

telehealth, rapport building, speech-language pathologist, telehealth experience, digital literacy

Introduction

Telehealth employs a range of telecommunication technologies, such as real-time audiovisual conferencing and asynchronous electronic transmissions of therapy materials, which allows therapists to reach their patients from distance (World Health Organization, 2017). Different terms are used when referring to remote communication and its uses to serve patients, for example, telehealth, telemedicine, and telepractice. Commonly, telehealth is a

broader term which can be used to describe the synchronous or asynchronous variations of videoconferencing and electronic transmission of therapy services. In the opposition, telemedicine is more focused on the delivery of medical services, which is in the process of being phased out in favor of telehealth, as the term likely restricts the scope to services involving only medical management (Center for Connected Health Policy, n.d.). The American Speech-Language Hearing Association (ASHA) adopted the term telepractice to emphasize that speech-language/audiology services delivered remotely span a range of educational settings, in addition to healthcare settings (American Speech-Language-Hearing Association, n.d.-a). Currently, as telehealth is being used most widely by various healthcare disciplines (e.g., dentistry, counseling, physical and occupational therapy, home health, chronic disease monitoring and management, and disaster management), it is used throughout the current study.

Telehealth brings about benefits by offering increased access to services for rural and under-served patients, and increased flexibility of time for both patients and therapists. It is easier to involve caregivers, interpreters, or remote specialists in real time. It also facilitates clinician training and supervision, supports asynchronous tools (e.g., recorded practice, messaging) to extend therapy between sessions, and enhances service resilience during public health emergencies (Zilliacus et al., 2010). Despite its many advantages, the service delivery method contains unique challenges which may not be as evident as in conventional in-person therapy (Tucker, 2012a). A commonly cited challenge of telehealth is building rapport between therapists and patients, possibly due to reduced nonverbal channels, limited visual field, and technical problems. Rapport has been defined as the establishment and maintenance of an interactive, harmonious, and communicative relationship between the therapist and the patient (Pattison and Powell, 1989). Recent studies share a similar definition, that is, both sides share mutual feelings of trust, respect, connection, and agreement on targeted goals (Freckmann et al., 2017; Duffy et al., 2023).

Perceptions of rapport building in telehealth speech-language services

Speech-language pathologists (hereafter SLPs) have been increasingly utilizing telehealth during the COVID-19 pandemic (Hao et al., 2021; Sylvan et al., 2020) and continued to do so in the post-pandemic era (Van Echo et al., 2023). While delivering telehealth services, rapport has been widely perceived by SLPs to be challenging (e.g., Akamoglu et al., 2018; Retamal-Walter et al., 2022a). This difficulty can be especially pronounced in pediatric settings, where children's limited cooperation could impede engagement and participation.

Three clinician-reported aspects were identified from the extant literature, aligning with the notion that rapport is multi-dimensional rather than unitary (Tickle-Degnen and Rosenthal, 1990). The first aspect relates to the perceived importance of rapport building. Tucker (2012b) and Akamoglu et al. (2018) conducted interviews with SLPs who had varying levels of telehealth experience, primarily in school settings. A majority of the participants emphasized the importance of rapport building in the success of telehealth services, citing challenges such as the lack of physical contact or proximity with students and the increased effort and collaboration required. Clinicians also reported that they prioritize a harmonious relationship with the family and the child, so clients are more likely to buy-in and carry over outside the therapy sessions. In addition, it was reported that training on rapport building skills is necessary prior to the start of telehealth, especially for children with more sensory and behavior issues.

The second aspect emerging from the literature, achievement of rapport building, was rated lower in telehealth than in-person therapy. Tucker (2012a) conducted a survey among 170 school-based SLPs in one northeastern state of the United States. Most participants disagreed with the statement that *rapport between the SLP and student can be established during speech-language telehealth just as strongly as during in-person speech-language therapy*. It is important to note, however, that only seven of these SLPs (4.1%) had experience with telehealth at the time of the survey, while the others had only conducted in-person therapy. Hines et al. (2015) noted that SLPs had reservations before the start of telehealth services but developed more positive views after gaining experience with it. Additionally, the speed at which SLPs build rapport in telehealth, which belongs to the achievement of rapport, was reported to take longer in telehealth than expected when compared to face-to-face sessions (Pitt et al., 2018; Anderson et al., 2015).

The third aspect, strategies used by SLPs to facilitate remote rapport building, has been understudied. By referring to findings from other disciplines (e.g., psychology, nursing), some common strategies were summarized. Verbal cues, such as reflections of emotions, restatements, reinforcements, descriptions, and explanations, were believed to form a stronger emotional bond, hold interests and focus, and allow the therapist more control during telecommunication (Sucala et al., 2013). Nonverbal cues are important during in-person interactions, for instance, smiling, directed gazing, head nodding, leaning forward (Tickle-Degnen and Rosenthal, 1990). In telehealth, adaptations of nonverbal cues were reported to be used by SLPs (Grillo, 2017), but specifics in regards to these adaptations were unclear to the reader. One possibility is that SLPs could use reduced nonverbal cues because of a restricted view for patients to observe these cues. The other possibility could be that SLPs exaggerate their use of nonverbal cues to compensate for the limited view, making nonverbal cues more obvious for patients to detect (Retamal-Walter et al., 2022b). Beyond verbal and nonverbal cues, e-helpers (e.g., teachers or parents sitting with the student to facilitate the online session) have been brought up widely in the literature. An e-helper could aid the student by logging on at scheduled time, obtaining and maintaining attention, and working with the student to complete the homework assigned (Tucker, 2012b).

In light of the view that rapport is dimensional, the three aspects map onto Tickle-Degnen and Rosenthal (1990)'s framework: the importance and achievement aspects reflect the affective facet of the bond (i.e., the feelings of the participants during the experience of rapport), while strategy use captures its behavioral facet (i.e., actions that express the affective connection between the participants). The two facets are interdependent rather than strictly separable. Higher perceived importance likely motivates clinicians to deploy more rapport-building strategies, but greater strategy use does not guarantee a high level of rapport achievement (Wolk et al., 2016; Zilcha-Mano et al., 2016). In some cases (e.g., children who have severe behavioral or sensory difficulties), clinicians may use many strategies yet still report low levels of achievement. This potential mismatch underscores the need to consider client and clinician factors when evaluate rapport, a point that will be expanded in the next section.

Potential influences of patient and clinician factors

Bordin's (1979) working alliance framework is a classic and highly influential model that defines psychotherapy alliance in terms of three components, including bond, tasks, and goals. While both patients and clinicians need to agree on the targeted goals and tasks, the bond of the two parties fosters open communication and emotional engagement. Rapport building, the focus of the current study, occupies the bond component, the interpersonal connection, warmth, trust, and engagement that makes collaborative work possible. Bond (or rapport) emerges from a good "fit" between patient and therapist, necessitating the readiness from both parties. In pediatric speech-language services, patient-side readiness to form a bond is largely shaped by a child's developmental level and severity of challenging behaviors, while therapist-side attributes are related to clinical experience and their age, which may determine how well the clinician can echo and scaffold the child's needs in treatment. Crucially, the digital environment adds layers by incorporating interactivity and adaptability of online communication. In particular, clinicians' telehealth experience and digital literacy influence the formation of a strong bond remotely (Tremain et al., 2020). Overall, the bidirectional fit, patient readiness and therapist responsiveness, supports agreement on tasks and goals and thus strengthens the overall working alliance, which contribute to treatment outcomes. However, existing research provides only a broad understanding of this issue, which did not address contextual factors relevant to patients and therapists that may significantly affect telehealth rapport in pediatric speech-language services.

Patient factors: child diagnosis and age

As reported by therapists in interviews, children who are not likely to engage during telehealth have severe behavioral, sensory, and cognitive issues (e.g., Akamoglu et al., 2018; Retamal-Walter et al., 2022b). These symptoms are in general related to a child's diagnosis and age. For example, children with autism spectrum disorder (ASD) are likely harder to be engaged compared to children with speech sound disorder (SSD). Both disorders make up the bulk of an SLP's caseload (Broomfield and Dodd, 2004; Gillon et al., 2017): the overall prevalence of ASD is estimated to be about 3.2% (1 out of 31 children) (Shaw et al., 2025), and the prevalence of SSD is found to be even higher, around 3.5% (Eadie et al., 2015; Hambly et al., 2013).

To be more specific, the two disorders make a difference in how easily a digital bond forms. ASD affects the way a child communicates due to its common symptoms. Rigid and repetitive language, narrow interests and exceptional abilities, uneven language development, and poor nonverbal conversation skills, are all contributing factors that may affect how easy it is for one to build rapport with clients who have ASD (American Psychiatric Association, 2013; National Institute of Deafness and Other Communication Disorders, 2020). Effective rapport building requires clinicians to address these ASD-specific behaviors. In contrast, the behaviors of those with SSD are not as limiting. Most of the behaviors displayed by children with SSD are typically-developing behaviors and are on track with those of children who do not have SSD.

While the ASD group generally represents children with more severe behavioral, sensory, and cognitive issues, the SSD group represents those with mild or no such behaviors. As it was not feasible to exhaustively assess all disorders that SLPs may encounter in clinical settings, the two disorders were chosen based on their distinct behaviors, which may influence how rapport is established in telehealth. The contrast in their behavioral presentations served as a basis for examining how SLPs might adapt their perceptions of rapport across different types of clients. It is important to note that there was no intention to ignore the heterogeneity within each disorder. Rather, the focus was on identifying general patterns, not individual differences.

Differing levels of assistance required by ASD and SSD may alter dynamics among key stakeholders, including children, parents, clinicians, etc., thereby affecting rapport building in remote sessions. For example, children with SSD can be more independent and have more opportunities to directly interact with the clinician to build a strong rapport. However, children with ASD may heavily rely on their parent(s)/e-helper to communicate with the clinician and to assist in completing the tasks in a telehealth session. This reduces opportunities for direct clinician-child interaction and increases the burden on parents, which may undermine the formation of a strong therapeutic bond between the clinician and the child.

In addition, the diagnosis may affect patients' digital literacy differently, which in turn influences the formation of a digital rapport. Digital literacy is defined as the ability to adapt, access, and learn technology to contribute to own community (International Society for Technology in Education, 2007; Wing, 2006). Due to the impaired communication and cognitive skills, children with ASD in general are expected to have lower digital literacy than children with SSD. Though evidence indicates telehealth can be effective for children with ASD (Christopoulou et al., 2022; Sundarrajan and Franco, 2024), it was recommended that support personnel be included for troubleshooting and that backup modalities be secured to ensure reliable therapy delivery for this population (Boisvert et al., 2010). Typically, parents take the role of the support personnel or e-helper, whose digital literacy skills have an influence on rapport in telehealth.

Child age broadly indexes the developmental level, which may influence a child's capacity to form a digital bond. The first 3 years is the most intensive period for acquiring speech and language skills (Hoff, 2013; National Institute of Deafness and Other Communication Disorders, 2017), when early intervention services occur in the home setting (American Speech-Language-Hearing Association, n.d.-b). In contrast, at later developmental stages (e.g., ages 4–8 years), children experience their first introduction to educational settings. While at the younger age communication mainly takes place between the child and the parent, at the older age it shifts to communication among multiple communication partners such as teachers and classmates. Children who are older develop longer attention spans, resulting in increased engagement and inquisitive learning. In contrast, children who are younger possibly pose a greater threat to remote rapport due to short attention spans, requiring the SLP to implement more strategies in order to engage the client and to keep them focused. In addition, younger children tend to shy away from strangers and unfamiliar settings, whereas older children's language skills and attention spans are more mature (Woods et al., 2013), resulting in increased engagement in different settings.

Parent–child bonding patterns vary by diagnosis and developmental level, which may shape how successfully a clinician establishes rapport remotely with the family. The first 8 years of life are the most crucial time for the parent–child relationship, influencing the quality of the family environment and the child's development

(Breiner et al., 2016). However, family members with severe disabilities, such as ASD, may negatively impact the family relationship and introduce an additional source of stress (Schieve et al., 2007). Because rapport is typically built on a positive environment among team members (e.g., SLP, child, and parent or support personnel), the additional time and assistance that parents need to provide in telehealth likely increases parental stress and burden. This may impede clinicians from building a successful rapport with the family, particularly for parents who have lower digital literacy.

Clinician factors: clinician age, telehealth experience, and digital literacy

Based on the existing literature, clinician factors mainly included their age, telehealth experience, and digital literacy. Tucker (2012a) looked at how the overall work duration affected perceptions of telehealth use. A group of SLPs ($n = 25$) with 1–5 years of work experience expressed greater interest in the use of telehealth, while a group of SLPs ($n = 74$) with 25 + years of experience showed reduced interest. The finding was interpreted in relation to clinician age. SLPs who were older displayed lower acceptance of telehealth than SLPs who were younger. Although the divided perceptions were not about rapport building per se, they could influence how SLPs perceive rapport building. In addition, telehealth experience appears to significantly shape SLPs' perceptions of telehealth. Those who have adopted telehealth often develop more positive views, whereas SLPs with no telehealth experience may maintain more negative perceptions (Hines et al., 2015).

Digital literacy refers to the competence in using digital technologies, such as tablets, smartphones, apps, the internet, and digital cameras, independent of any specific health-related purpose (Longhini et al., 2022).[1] Computer glitches and/or internet disconnection could cause sound distortion and overlapped segments of speech, making patients frustrated within their sessions (Retamal-Walter et al., 2022b). This highlights the importance of clinicians' digital literacy in building rapport with their clients in telehealth settings. Limited digital literacy among therapists could hinder child and family participation and could make rapport building difficult. If this occurs, individual therapist's levels of rapport achievement could be lower.

The current study

The study explored the influences of patient and therapist factors on SLPs' perceptions of rapport with pediatric patients in telehealth. Regarding patient factors, it surveyed digital bond when serving children with different diagnosis (i.e., ASD or SSD) and different ages (i.e., 0–3 years old or 4–8 years old). Regarding therapist factors, it assessed how SLPs' digital literacy, age, and telehealth experience correlated with perceptions of rapport in telehealth. Rapport was treated as a multi-dimensional construct and was evaluated from three aspects (i.e., importance, levels of achievement, and strategy use).

Drawing on Bordin (1979)'s framework and the extant literature in telehealth speech-language services, we derived hypotheses addressing patient and clinician factors, respectively. Generally, patients who are younger or who present greater behavioral and sensory challenges are expected to cause more difficulties for clinicians to build rapport remotely. Thus, clinicians were predicted to perceive digital rapport as more important, apply more strategies, but feel less achieved when working remotely with younger children and the ASD group, compared to older ones and the SSD group. Regarding clinician factors, those who are younger in age, more experienced with telehealth, and more skilled in digital literacy would perceive rapport more important, use more strategies, and feel more achieved when building rapport remotely. Regarding correlations among the three aspects of rapport building, we predicted that clinicians' ratings of perceived importance would be positively and significantly associated with reported strategy use, whereas strategy use would not be significantly correlated with their achievement of rapport.

Methods

Participants

Ethical approval was obtained from one of the affiliated universities. Eligible SLPs were those with experience treating children with both disorders, ASD and SSD, and at least one age range, 0–3 or 4–8 years old. As the two age ranges typically relate to different work settings (e.g., Early Intervention versus School), it is hard to have participants to cover both ages. Participants were not required to have experience in telehealth in order to participate in the study, allowing us to investigate how varying levels of telehealth experience affected participants' perceptions of rapport building. Finally, participants were informed that ASD or SSD had to be the primary diagnosis, rather than secondary to any other condition.

A power analysis was conducted to determine the minimum number of participants that should be recruited. Generalized Linear Mixed Modeling (GLMM) was planned to be used for data analysis (specified below). We implemented power analysis for repeated measures ANOVA which is regarded as most similar to the GLMM.[2] With a

1 There is no single consensus definition of "digital literacy" in the literature. For example, Mainz et al. (2024) treated it as an umbrella construct that includes technical competence (e.g., basic computer and software use, internet skills), methodological competence (e.g., data-handling and lifelong learning), social competence (e.g., digital engagement within teams and communities), and personal competence (e.g., reflective awareness of one's digital skills). By contrast, Longhini et al. (2022) used the term more narrowly to denote competence in using technology (tablets, mobile phones, apps, internet), which corresponds closely to the technical competence component in Mainz et al. (2024). In this study, we followed the narrower definition, using "digital literacy" to refer primarily to practical skills related to internet connection and device operation.

2 Power analysis for GLMM was explored extensively, but there was a lack of guidance to address it (Johnson et al., 2015). It is worth noting that GLMM leads to greater power than ANOVA. Each participant responded to 32 questions regarding their perceived importance, strategy use, and achievement of rapport building, yielding 5,920 observations. All these observations were entered into the GLMM, whereas only averages of responses to each aspect of rapport building would be used in ANOVA.

significance criterion of $\alpha = 0.05$, power = 0.95, and the effect size of 0.25, a minimum of 126 SLPs were required. Response rate was unknown to us, so we were being conservative (i.e., 1%). Therefore, invitations were sent to 350 SLPs in each of the 50 states (17,500 invitations in total), which could bring about 175 participants (above the minimum sample size requirement). In practice, there were nine states that did not have 350 SLPs (i.e., Delaware, Hawaii, Alaska, Montana, Wyoming, North Dakota, South Dakota, Rhode Island, Vermont). In this case, survey invitations were sent to all eligible SLPs matching the search criteria for that state. In total, 16,165 invitations were sent out, with a total of 221 SLPs participating in the survey, yielding a response rate of 1.4%.

Dissemination of survey started from the selection of participants using the ASHA Community Directory, which included a private message feature for communication among ASHA members. Filters were applied to help target eligible SLPs, including *autism spectrum disorder, articulation disorders*, and *phonological disorders*. An initial screening link was sent to each potential eligible SLP, which asked if the participant was a licensed SLP who had experience working with children with ASD and SSD. In addition, SLPs were informed via the screening link that incentives ($10 Amazon e-gift card) would be provided to randomly selected participants who completed the survey. If an SLP believed that they were eligible, they were asked to leave their names and email addresses, which were used to generate a unique personal link to take the survey. This approach minimized invalid responses and prevented any ineligible participants from taking the survey in order to gain incentives. Both the initial screening survey and the subsequent research survey were developed and delivered through Qualtrics.

Consent was obtained from the participants before they proceeded to the research survey. Survey responses were collected from October to December 2021. Once the unique survey link was sent, the participant was given 2 weeks to complete it. A reminder email was sent approximately 1 week after the link was initially sent to remind interested SLPs to complete the survey. At the end of December 2021, one final reminder was sent to all interested SLPs who had not initiated or completed the research survey. In January 2022, the survey was closed, and incentives were sent.

Research survey

The study employed an e-survey, following Consensus-Based Checklist for Reporting of Survey Studies (CROSS) (Sharma et al., 2021). The research survey was organized into four sections. The first section included demographic questions (i.e., gender, age, ethnicity, etc.), years holding CCC-SLP certification, years of using telehealth in a pediatric setting, and prior telehealth training. In addition, SLPs were asked to report their telehealth caseloads and their entire working history caseloads, with regard to each of the disorders (ASD and SSD) and each of the age ranges (0–3 and 4–8). Altogether, the information in this section helped define the sample of SLPs. Digital literacy was put into two questions (i.e., internet competence and computer competence), as they are different concepts. Internet competence refers to their ability to connect to the internet and troubleshoot common connectivity issues. Computer competence refers to their ability to work, maneuver, or troubleshoot issues related to the use of a computer, tablet, smart phone, and other devices.

In the remaining three sections, the three aspects of rapport building in telehealth, including perceived importance, strategies employed, and perceived achievement, were evaluated. The questions within each section were based on the literature that have been reviewed. Each question asked the participant to notate their response for one disorder (ASD or SSD) and for one age range (0–3 or 4–8). Rankings were placed on a five-point scale, with lower values representing lower perceptions and higher values representing higher perceptions. See Appendix 1 for the complete survey questionnaire.

After the questions were drafted, the authors collected perspectives and feedback from two researchers who had published in the area of telehealth and two seasoned SLPs who had experience with the two disorders and the two ages via telehealth. They returned the survey with comments and edits, focusing on improving the clarity of the questions and the rewording of vague language. For example, per a comment from one expert, *caseload* was specified as *caseload for the entire working history* versus *caseload for telehealth only*. The survey questions were finalized when the comments and edits were fully addressed. Upon the completion of this step, two graduate students majoring in speech-language pathology tested the initial screening survey and the follow-up research survey, ensuring that both surveys and their links could be used reliably. The process provides content and face validity and pilots the survey before it was disseminated to a large cohort.

Results

The current sample of SLPs

Eligibility criteria were applied to finalize the sample: eight participants were excluded as their responses to the survey were incomplete; six participants were excluded as they had experience with neither disorder throughout their entire career. Altogether, 14 participants were excluded, yielding 207 eligible participants with complete data being included in the analysis.

Table 1 presents the results of the background section of the survey. In the current sample, most SLPs were female and Non-Hispanic White, with around 80% between 30 and 59 years old. Most worked in schools, followed by private practices, early intervention, and non-residential health care, while each of the remaining settings accounted for fewer than 5%.

Participants reported their state and county of practice, allowing researchers to determine geographic region and rurality. Using National Geographic's boundaries (O'Connor, 2012), participants were identified in all five U.S. regions (Northeast, Southeast, West, Southwest, Midwest). Rurality status was determined through Health Resources and Services Administration database (Department of Health and Human Services, n.d.), and 12.1% of counties were unverifiable. Among the verified counties, over three-fifths of SLPs practiced in urban areas, and the remainder practiced in rural areas.

Most participants had held certification for 0–10 years, 11–20 years, or 21–30 years. By contrast, only 15.4% had provided telehealth in pediatric settings for more than two years. Possibly relating to the overall short experience with telehealth, more than half (58.5%) had no formal training on telehealth. Meanwhile, SLPs rated

10.3389/fpsyg.2025.1612803

TABLE 1 Participant background information (n = 207).

Characteristics	Number	Percentage (%)
Gender		
Female	*201*	97.1
Male	*5*	2.4
Other (no-binary)	*1*	0.5
Age		
20–29	*16*	7.7
30–39	*57*	27.5
40–49	*58*	28.0
50–59	*45*	21.7
60–69	*26*	12.6
70–79	*5*	2.4
Ethnicity		
Asian	*5*	2.4
Black or African American	*8*	3.9
Hispanic	*2*	1.0
Non-Hispanic White	*184*	88.9
Mixed or Other (e.g., Non-Hispanic White and Native American, Jewish, European)	*8*	3.9
Highest education		
Master's degree	*200*	96.6
Doctorate degree (PhD, EdD, or SLPD)	*7*	3.4
Work setting (participants could choose more than one)		
School	*109*	52.7
Early intervention	*49*	23.7
College/University	*9*	4.3
Hospital	*11*	5.3
Residential health care	*4*	1.9
Non-residential health care	*20*	9.7
Private practice	*68*	32.9
Other (e.g., behavior day treatment, Head Start)	*7*	3.4
Geographic region of practice[a]		
Northeast	*35*	16.9
Southeast	*47*	22.7
Southwest	*22*	10.6
Midwest	*53*	25.6
West	*50*	24.2
Rurality of practice[b]		
Rural	*42*	20.3
Urban	*140*	67.6
No data found	*25*	12.1
Years of holding certificate		
0–10 years	*62*	30.0
11–20 years	*56*	27.1
21–30 years	*52*	25.1
31–40 years	*28*	13.5

(Continued)

TABLE 1 (Continued)

Characteristics	Number	Percentage (%)
More than 40 years	*9*	4.3
Years spent using telepractice in a pediatric setting		
Less than 1 year	*52*	25.1
1–2 years	*123*	59.4
2–3 years	*15*	7.2
More than 3 years	*17*	8.2
Telepractice training		
Formal training (e.g., courses in college or CEU)	*78*	37.7
No formal training (e.g., self-taught)	*121*	58.5
No training	*8*	3.8
Internet competence (e.g., dealing with internet disconnection)		
1. Not at all competent	*1*	0.5
2. Low competence	*5*	2.4
3. Neutral	*18*	8.7
4. Competent	*131*	63.3
5. Very competent	*52*	25.1
Average (*SD*)	*4.0 (0.7)*	
Computer competence (e.g., computer, tablet, smart phone)		
1. Not at all competent	*0*	0
2. Low competence	*2*	1.0
3. Neutral	*13*	6.3
4. Competent	*122*	58.9
5. Very competent	*70*	33.8
Average (*SD*)	*4.0 (0.6)*	

[a]Based on the state reported by an SLP, researchers referred to National Geographic (O'Connor, 2012) to categorize the SLP into one of the five regions.
[b]Based on the county reported by an SLP, researchers referred to Health Resources and Services Administration (Department of Health and Human Services, n.d.) to define whether the SLP was practicing in a rural or urban area.

both their internet competence and computer competence[3] at 4 out of 5, suggesting competence in digital literacy.

Figure 1 presents the reported caseload demographics for the two disorders and ages. They were separated into four categories: (1) entire work history for 0-3-year-olds, (2) entire work history for 4–8-year-olds, (3) telehealth only for 0-3-year-olds, and (4) telehealth only for 4–8-year-olds. Overall, participating SLPs reported a higher caseload of SSD than ASD and a higher caseload of 4-8-year-olds than 0-3-year-olds for both entire work history and telehealth only. Regarding telehealth, among the 207 participants, 54 (26.1%) had experience across all the four categories, 32 (15.5%) in three of the four categories, 82 (39.6) in two, 17 (8.2%) in only one, and 22 (10.6%) had no telehealth experience in any of the four categories.

Instrument validation

The telehealth rapport building instrument initially included eight questions (three assessing perceived importance, three assessing strategy use, and two assessing perceived achievement). Each question was completed separately in four contexts (SSD at 0–3, SSD at 4–8, ASD at 0–3, ASD at 4–8), resulting in $8 \times 4 = 32$ items. See Appendix 1 for details. The eight questions were based upon a review of the extant literature. Expert review was conducted after the questions were drafted. Two experienced telehealth researchers and two seasoned telehealth clinicians evaluated each of these questions for relevance and clarity, and they further recommended changes in wording (also see "Research Survey" section). This process provided content and face validity.

Based on the data from the 207 participants, Cronbach's α for the 32 items was 0.85 (McDonald's $\omega = 0.86$), indicating good internal consistency. Confirmatory factor analysis (CFA) was conducted on three latent factors (the three aspects of telehealth rapport building). The mean score of the four context-specific responses was used for each question to reduce strong context clustering. One question (Q17 in Appendix 1) exhibited a weak loading and thus was removed. The final model therefore included

3 The two questions captured different aspects of digital literacy (connecting to online services and operating devices such as computers, tablets, and smartphones), which has been brought up most frequently in the literature. In data analysis, as clinicians' ratings of the two questions were strongly correlated, we averaged the scores and formed a single composite measure of overall digital literacy.

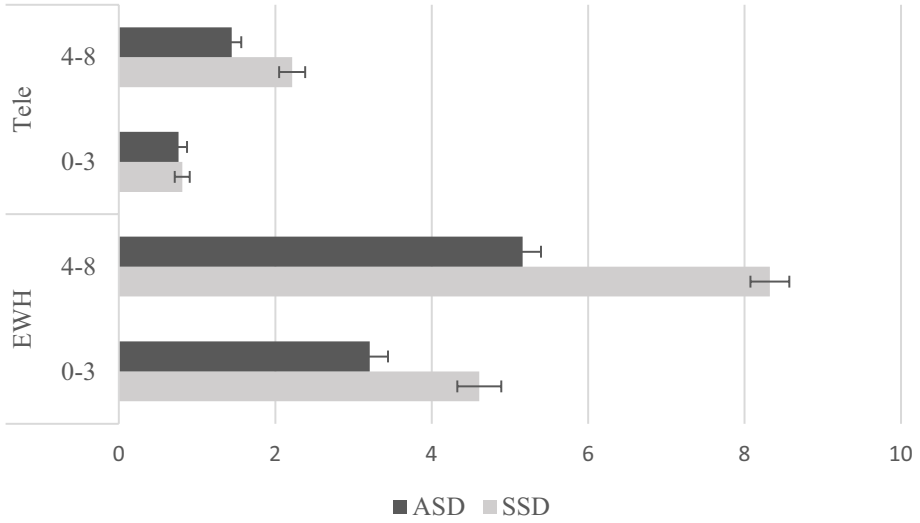


FIGURE 1
Participating SLPs' caseloads with the two diagnoses and ages, separating entire work history and telehealth only. Tele refers to telehealth. EWH refers to entire work history. Caseload: 0 = no experience, 1 = 1–10 children, 2 = 11–20 children, 3 = 21–30 children, 4 = 31–40 children, 5 = 41–50 children, 6 = 51–60 children, 7 = 61–70 children, 8 = 71–80 children, 9 = 81–90 children, 10 = 91–100 children, 11 = 100 + children.

seven aggregated questions. All the indicators loaded significantly on their intended factors (moderate to strong standardized loadings ranging from 0.435 to 0.949). Fit indices were mixed but overall consistent with the three-factor interpretation: CFI = 0.927 (acceptable), TLI = 0.860 (borderline), RMSEA = 0.086 (90% CI [0.047, 0.127], moderate), $\chi^2(11) = 27.953$, $p = 0.003$ (highly sensitive to sample size), AIC = 2811.775. For comparison, a single-factor CFA on the seven aggregated questions fit substantially worse: CFI = 0.532 (poor); TLI = 0.298 (poor); RMSEA = 0.193 (90% CI [0.162, 0.225]) (poor); $\chi^2(14) = 122.026$, $p < 0.001$; AIC = 2899.849. Loadings were weak to moderate ranging from 0.078 to 0.688. Taken together, these results offer support for the three-factor structure while indicating room for improvement.

Influences of patient and clinician factors

The analysis was conducted using R Version 4.4.3 (R Core Team, 2025) and the "ordinal" package (Christensen, 2023). We fitted cumulative link mixed models (CLMMs) with a logit link function. The model had the maximal random effects structure justified by the data that would converge, including a random intercept for participants. Items were not treated as random effects, as they represent different aspects of rapport building, rather than interchangeable stimuli. We began with the most complex by-participant random effects structure (1 + child_disorder + child_age | participant) and then sequentially reduced it to the simplest one (1 | participant). The fitted models were compared in terms of AIC, with a smaller value indicating a better model fit. This was supplemented by likelihood ratio tests conducted to determine whether the inclusion of a predictor significantly improved the model fit.

Independent variables included patient factors and clinician factors. Patient factors, child disorder and age, were treated as within-clinician repeated measures.[4] Disorder and age were coded as categorical (SSD = 1 and ASD = 2; 0–3 years = 1 and 4–8 years = 2). Three clinician factors were included. Clinician age (20–39 = 1, 40–59 = 2, 60–79 = 3) and telehealth experience were categorized (no telehealth experience = 1; lower diversity (worked with 1 or 2 diagnosis-age groups) = 2; higher diversity (worked with 3 or 4 diagnosis-age groups[5]) = 3). All the categorical factors were dummy-coded with 1 set as the reference category. As scores of internet competence (Q11) and computer competence (Q12) were highly correlated ($r = 0.767$, $p < 0.01$), the average scores were entered into the model, indexing overall digital literacy. Interactions were included in the model, including two-way interactions of child diagnosis and child age, clinician experience diversity and child diagnosis, clinician experience diversity and child age, and a three-way interaction among clinician experience diversity, child diagnosis, and child age. Dependent variables were responses to questions regarding perceived importance, strategy use, and achievement of rapport building in telehealth, respectively. Appendix 2 displays means and standard deviations for each question related to the three aspects, distinguishing disorder-age groups.

Regarding the perceived importance of rapport building in telehealth, the analysis were based on a model including a by-participant intercept and child_disorder slope (AIC = 3729.85,

4 Note that each clinician provided rating of telehealth rapport building on 0–3 years old SSD, 4–8 years old SSD, 0–3 years old ASD, and 4–8 years old ASD, respectively.

5 The higher-diversity clinicians would have responded to the questions mostly based on their telehealth experience, whereas the lower-diversity clinicians may have partially predicted how they would establish rapport in telehealth based on limited experience with only one or two of these groups of children. The no telehealth experience group would have completely used predictions while responding to the questions.

$\chi^2(2) = 9.16$, $p = 0.01$). It was shown that ASD was associated with higher ratings than SSD, with an effect that approached significance ($\beta = 0.43$, SE = 0.23, z = 1.93, $p = 0.054$). Clinician age was positively associated with higher perceived importance of rapport: relative to clinicians aged 20–39 years, clinicians aged 40–59 years ($\beta = 0.393$, SE = 0.190, z = 2.064, $p = 0.039$) and 60–79 years ($\beta = 1.389$, SE = 0.287, z = 4.833, $p < 0.001$) showed higher odds of endorsing greater importance. Child age, clinician telehealth experience, clinician digital literacy and all tested interactions were not significant.

Regarding the achievement of rapport building in telehealth, a model with by-participant intercepts and random slopes for child_disorder and child_age provided the best fit (AIC = 3709.6, $\chi^2(3) = 8.82$, $p = 0.03$). Achievement ratings were lower for children with ASD than children with SSD ($\beta = -0.74$, SE = 0.22, $z = -3.29$, $p < 0.01$), higher for older children than younger children ($\beta = 1.15$, SE = 0.22, $z = 5.30$, $p < 0.01$), and higher among clinicians with better digital literacy ($\beta = 0.60$, SE = 0.21, z = 2.85, $p < 0.01$). In addition, compared to clinicians aged 20–39 years, both clinicians aged 40–59 years ($\beta = 0.57$, SE = 0.28, $z = 2.04$, $p = 0.04$) and clinicians aged 60–79 years rated significantly higher achievement ($\beta = 0.75$, SE = 0.39, $z = 1.89$, $p = 0.059$). The remaining main effects and interactions were not significant.

Regarding strategies of rapport building in telehealth, a model with by-participant intercepts and random slopes for child_disorder and child_age provided the best fit (AIC = 4652.0, $\chi^2(3) = 45.54$, $p < 0.01$). ASD received higher ratings than SSD ($\beta = 0.73$, SE = 0.23, $z = 3.16$, $p < 0.01$), and older children received lower ratings than younger ones ($\beta = -1.32$, SE = 0.22, $z = -6.10$, $p < 0.01$). In addition, clinicians aged 60–79 years rated higher in strategy use than 20–39 years ($\beta = 0.69$, SE = 0.29, $z = 2.29$, $p = 0.02$). Interestingly, there was a significant interaction between child disorder and clinician telehealth experience: the disorder effect (ASD > SSD) was present in the higher-telehealth-experience group ($\beta = -0.75$, SE = 0.16, z = −4.78, $p < 0.01$) and the lower-telehealth-experience group ($\beta = -0.59$, SE = 0.15, z = −4.02, $p < 0.01$), but was absent in the no-telehealth-experience group ($\beta = 0.02$, SE = 0.28, z = 0.06, $p = 0.95$). Figure 2 illustrates the interaction. All the other main effects and interactions were not significant.

Correlations among aspects of rapport building

Each aspect score (importance, strategy use, achievement) was computed by averaging its constituent items, and each item score, in turn, was the mean of the four disorder-age conditions. It was found that clinicians' ratings of rapport-building importance were positively and significantly correlated with reported strategy use ($r = 0.384$, $p < 0.001$), whereas strategy use did not significantly correlate with rapport-building achievement ($r = -0.034$, $p = 0.626$). Perceived importance was weakly but significantly associated with perceived achievement ($r = 0.140$, $p = 0.044$).

Discussion

The study advanced the understanding of rapport building in telehealth speech-language services by specifying the influences of patient and therapist factors. Regarding patient factors, clinicians rated the ASD group higher importance, more strategy use, but lower achievement, compared to the SSD group. Clinicians reported more strategy use and lower achievement of rapport when serving younger

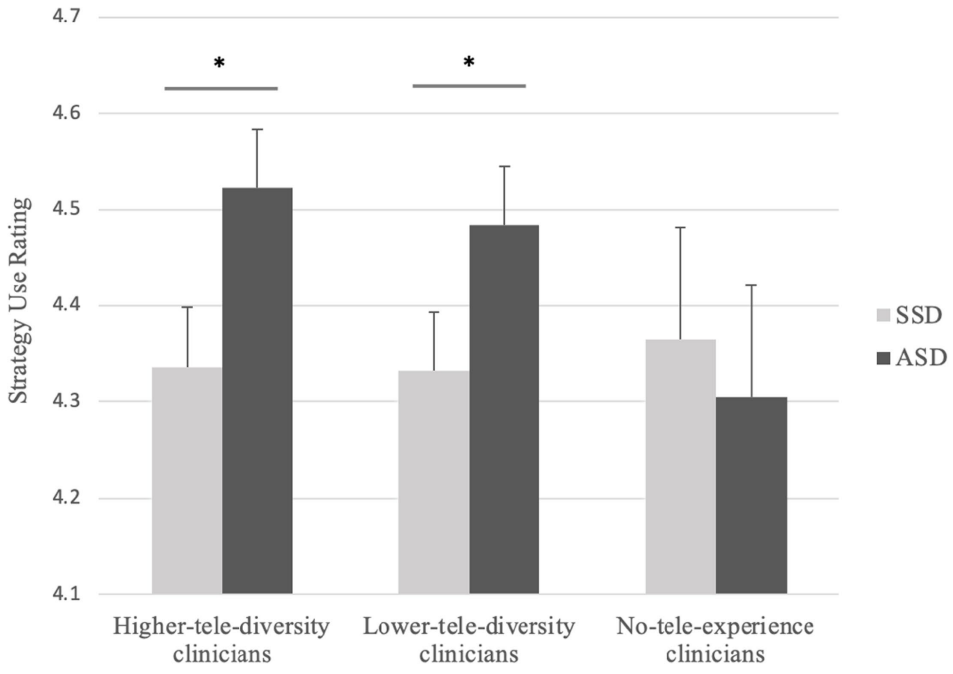


FIGURE 2
Interaction between clinician telehealth experience diversity and child diagnosis for strategy use in telehealth. *indicates $p < 0.01$.

children, but a main effect of child age was not evident in the importance of rapport. Regarding therapist factors, clinician age was related to the three aspects of rapport: older clinicians tended to rate rapport more important, reported higher achievement, and used more strategies than younger clinicians. Clinicians having telehealth experiences with the targeted children, regardless of higher or lower diversity, displayed ASD-SSD distinction in strategies employed, whereas such distinction was not found among those with no telehealth experience. Digital literacy was only related to perceived achievement of rapport building.

Prevalent influences of child diagnosis and age on telehealth rapport building

As expected, clinicians rated children with ASD as requiring greater emphasis on rapport, using more strategies yet achieving lower levels of rapport than children with SSD. Younger children were reported to elicit more frequent strategy use, but a lower sense of achievement than older children, yet perceived importance was not significantly different between younger and older children. In general, these patterns are consistent with the diagnosis-specific and developmental profiles of pediatric clients. In particular, the diagnosis-related behaviors appeared to have more prevalent influences on clinicians' rating of rapport. The increased cognitive and linguistic demands among children with ASD may complicate direct clinician-child interaction and increase reliance on caregivers as e-helpers. Further, digital literacy is likely required to overcome limitations in telecommunication, such as constrained visual field and signal latency in telehealth, reduced availability of subtle nonverbal cues, making it harder for clinicians to form a bond with the ASD group. Overall, these findings highlight not only the differing natures of ASD and SSD, but also the importance of developmental considerations when building rapport remotely.

Regarding the correlations among the three aspects of rapport building, clinicians who rated rapport as more important reported using more strategies, but greater strategy use may not translate into higher achievement in telehealth rapport. This is consistent with our prediction and underscores that rapport is a muti-faceted construct. The finding also suggests that the affective and behavioral facets of rapport may be selectively related rather than uniformly connected. This nuances Tickle-Degnen and Rosenthal (1990)'s account, indicating that links between behavioral and affective facets may depend on which specific aspect of rapport is under consideration. Though not a main focus of the study, the weak but significant correlation between rapport-building importance and achievement likely reflects that both aspects tap the affective facet of the bond.

Future studies may go beyond the current focuses on client age and diagnosis to culturally and linguistically diverse populations with disabilities and explore how this factor influences telehealth rapport building. Currently, limited literature documented differences of perceptions of telehealth between monolingual and bi/multilingual children with communication disorders and their families. Davis et al. (2024) reported different perspectives between clients and clinicians: families of a bilingual background considered SLPs as a driving force during the therapy processes and heavily relied on SLPs for decision making, whereas SLPs considered their role as assistive. Though it is not specific to telehealth, this gap may cause a negative impact on rapport building. Nevertheless, telehealth offers unique opportunities to improve services for culturally and linguistically diverse populations who warrant greater attention (Taiebine and Keegan, 2024). They are frequently under- or over-referred in assessment, as monolingual clinicians often lack familiarity with home languages and cultural practices as well as with appropriate assessment tools and intervention materials (McLeod and Verdon, 2017; Scharff-Rethfeldt et al., 2020). Telehealth likely increases the access to bilingual clinicians and trained interpreters who are competent but not locally available. It also facilitates inclusion of family members as active e-helpers who can support interpretations of a child's speech-language profile during assessment and intervention. In fact, positive evidence, using telehealth for bilingual children with ASD to complete a narrative language sample task (del Hoyo Soriano et al., 2021) and to provide AAC treatment for bilingual children (King et al., 2022), may support the acceptability of telehealth in this population.

The COVID-19 pandemic rapidly accelerated the use of telehealth in speech-language services (Learnihan et al., 2025). Though telehealth continued to be used, many services have since reverted to in-person care post-pandemic, so long-term effects of telehealth remain underexplored (Christopoulou et al., 2022). Insufficient clinician training in telehealth and limited experience in partnering effectively with parents continue to impede optimal delivery. Nonetheless, telehealth is well suited for disseminating brief, targeted training to clinicians, caregivers, educators, and policymakers, helping close research-practice gaps (Kim et al., 2024) and promote evidence-based, culturally responsive care (Petinou et al., 2024). For example, growing evidence indicates that bilingual exposure does not harm language development in children with ASD (Howard et al., 2024; Garrido et al., 2024), counters deficit-based recommendations to use only English at home (Martin Loya and Meadan, 2024; Pang, 2024), and may even confer some advantages associated with multilingualism (Gilhuber et al., 2023). Telehealth platforms are therefore well suited not only for service delivery but also for brief and targeted training to different stakeholders, thereby promoting more consistent care. Though our results provide limited insight into cultural and contextual influences on telehealth rapport building, future research should examine potential benefits and barriers across diverse cultural and linguistic contexts to inform more equitable and effective telehealth implementation.

Different influences of therapist age, telehealth experience, and digital literacy

Compared to younger SLPs, older ones tended to rate higher in the importance, frequency of strategy use, and achievement of rapport in telehealth. The result appeared to contradict our prediction based on Tucker (2012a), which showed that older clinicians held more negative views of telehealth. However, in Tucker's study, most SLPs had not yet started telehealth, whereas a majority of the clinicians in the present study had already adopted telehealth for the targeted children. We speculate that prior to the start of telehealth, older clinicians hold more negative views, which may prompt them to place greater emphasis and put more efforts into rapport building when they begin offering telehealth. In addition, the hands-on experience with telehealth could attenuate negative attitudes and increase their sense of achievement in rapport building. Future research may study the dynamic process by comparing clinicians' perceptions of rapport before and after the initiation of telehealth.

Whether or not clinicians had telehealth experience significantly correlated with the frequency of their rapport building strategy use. Clinicians without telehealth experience with the targeted children reported similar levels of strategy use for both SSD and ASD, indicating a more uniform approach without emphasizing the influence of child diagnosis. In contrast, clinicians who had engaged in telehealth consistently reported higher strategy use for the ASD group compared to the SSD group. Clinicians with experience in telehealth should have experienced more difficulties with children with ASD than children with SSD, forcing them to develop more strategies that could be efficient in coping with the characteristics of ASD. By contrast, clinicians without hands-on telehealth experience lacked opportunities to test and consolidate rapport building strategies, so they may not differentiate their expected strategy use for ASD versus SSD. Also, the findings suggest that telehealth experience could be a key factor in developing strategies to build a strong rapport with children with ASD in telehealth, but not much with children with SSD. Interestingly, the ASD-SSD distinction pattern held irrespective of how many telehealth "cells" (i.e., 0–3-year-old SSD, 4–8-year-old SSD, 0–3-year-old ASD, or 4–8-year-old ASD) a clinician had worked with. Even those with fewer telehealth experiences matched the strategy use patterns reported by clinicians with more diverse telehealth backgrounds, indicating that even limited experience may yield insights into the need for enhanced or reduced frequency of strategies.

Clinician digital literacy was significantly correlated with their feelings of achievement, indicating that the competence with the internet and digital devices may translate into greater achievement of rapport in virtual sessions. Specifically, those who rated themselves higher in digital literacy possibly are better able to navigate telehealth platforms and troubleshoot technical issues, which enable them to devote more attention to client engagement. The client-centered interaction may help them feel more successful in fostering rapport and patient progress. By contrast, digital literacy did not have significant main effects on the perceived importance or strategy use of rapport in telehealth. One plausible explanation is that valuing rapport and selecting strategies may stem more from clinicians' therapeutic beliefs, training, and clinical experience than from their comfort with digital interfaces.

Recommendations to support digital literacy and telehealth rapport building

Given that only about 38% of respondents reported formal telehealth training, clinicians, especially those without hands-on telehealth experiences, are recommended to improve digital literacy and telehealth rapport-building skills through training that pairs technical skills (e.g., computer setup, platform features) with rapport-building techniques (e.g., brief caregiver coaching scripts, use of e-helpers). Unless having prior telehealth experience, clinicians should not provide telehealth guidance to caregivers. Clinicians may participate in supervised mock-sessions and peer observations by watching experienced telehealth clinicians and receiving constructive feedback, before independently implementing telehealth. In our study, verbal and non-verbal cues were used between frequently and always in both disorders and ages, and rating for non-verbal cues was slightly higher than verbal cues across all the disorder-age groups (Appendix 2). Grillo (2017) mentioned adaptations for non-verbal cues in telehealth, but it was unclear that the non-verbal cues were increased or decreased during telehealth. Presently, the results support the former, suggesting more frequent and exaggerated use of non-verbal cues to be detectable in telehealth that allows limited view.

Children with greater sensory and behavioral needs and of younger ages in particular require more intensive rapport-building strategies in telehealth and may nonetheless show lower immediate engagement. Caregivers therefore play a key facilitative role to help build rapport with therapists who work remotely. They are encouraged to prepare the child's environment before sessions (e.g., quiet space, consistent seating, simple visual schedule), act as an active e-helper during sessions (e.g., follow the clinician's cues, prompt turn-taking). Clinicians are encouraged to partner with caregivers by providing short, user-friendly guides (e.g., pictorial checklists, brief demo videos) that show exactly how to set up the camera and offer simple engagement prompts caregivers can use during the session.

For policymakers, community organizations, and patient-advocacy groups, system-level supports are needed to ensure access to telehealth services, especially in low-resourced areas. Policymakers and payers may consider reimbursement models and incentives that fund telehealth training and allow reimbursable preparatory time (e.g., caregiver coaching prior to a therapy session). Supported training should go beyond general platform use and basic telehealth knowledge to include focused, practice-based modules on clinically relevant, "trivial" aspects of remote care (e.g., rapport building techniques), so that clinicians can gain the specific skills needed to deliver high-quality telehealth services. Community organizations and advocacy groups can facilitate access by providing training to translate practical guides into local languages and low-literacy formats. It is important to prioritize early-intervention populations and children who have more severe behavioral and sensory challenges for enhanced supports.

Limitations and future directions

At the time the survey was administered, we could not identify a validated instrument that measured the same construct in pediatric telehealth for speech-language service. Therefore, direct assessment of concurrent validity was not possible. The survey was derived from the extant literature, the best available evidence, and our clinical experience, and then refined through expert review. The instrument demonstrated good internal consistency. Confirmatory factor analysis of the prespecified three-aspect structure (importance, strategy use, achievement) offered modest support and suggests that further item refinement is needed to better distinguish among the three aspects. We acknowledge that the instrument requires additional psychometric validation, and that this is a limitation of the study that warrants continued efforts.

Digital literacy was operationalized narrowly, using two items that assessed clinicians' competence of internet connection and device operation. These items provided limited information without capturing the broader and multi-dimensional competencies related to digital literacy. Therefore, future studies are warranted to employ validated and multi-item instruments to more comprehensively measure clinicians' digital literacy and its relationship to telehealth. In addition, the low response rate (1.4%) may relate to selection bias which could limit the generalizability of the current findings.

Clinicians who chose to participate in the survey may differ from non-responders. For example, the responders may have greater interests in telehealth, stronger buy-in to telehealth-based services, or higher digital literacy, which may not represent how the broader population perceives telehealth rapport building.

This study sampled SLPs practicing in the U.S., whereas many of the core lessons about remote rapport building (e.g., the importance of clinician digital literacy, the facilitative role of caregivers, and the need for rapport-focused clinical skills) are likely to resonate in other high-resource settings that have established telehealth infrastructure. Compared to major cities where resources are more abundant, telehealth has been less likely to be adopted in rural areas (Learnihan et al., 2025). In lower-resource settings, the same principles may still apply but require adapted delivery models, for example, low-bandwidth or phone-first workflows, community telehealth hubs or device-loan schemes, stronger emphasis on caregiver-mediated approaches, and medical insurance reimbursement covering telehealth. Overall, there is a lack of studies exploring telehealth in low-resource areas (Nizeyimana et al., 2022), warranting continued efforts in the generalizability of these findings in future research, particularly across different healthcare systems, resource settings, and cultural and linguistic contexts.

Conclusion

The current study focused on how patient and therapist factors shaped clinicians' perceptions of telehealth rapport building in pediatric speech-language services. The prevalent influences of child diagnosis and age highlight behavioral and developmental considerations pertaining to the client when clinicians need to build rapport remotely. Clinician factors influenced the three aspects of telehealth rapport differently. While clinician age was related to perceived importance, strategy use, and achievement of rapport, telehealth experience was only associated with the frequency of strategy use, and digital literacy was linked specifically to perceived levels of achievement. The three aspects capture different facets of the bond, and the affective and behavioral facets may not be uniformly connected. Together, the findings underscore the need to contextualize telehealth rapport by considering child and clinician factors, ultimately implying future training and practice in remote speech-language services and relevant disciplines where telehealth is frequently used to serving patients with disabilities.

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 humans were approved by IRB Office of University of Mississippi. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YH: Supervision, Methodology, Validation, Investigation, Writing – review & editing, Data curation, Conceptualization, Formal analysis, Resources, Funding acquisition, Writing – original draft, Project administration, Visualization. HP: Methodology, Conceptualization, Writing – review & editing. JW: Data curation, Project administration, Writing – review & editing, Writing – original draft, Investigation, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors wish to express sincere appreciation to the University of Mississippi Graduate School Council Research Grant for funding participant recruitment.

Acknowledgments

The authors are grateful for the speech-language pathologists who kindly participated in the survey.

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1612803/full#supplementary-material

References

Akamoglu, Y., Meadan, H., Pearson, J. N., and Cummings, K. (2018). Getting connected: speech and language pathologists' perceptions of building rapport via telehealth. *J. Dev. Phys. Disabil.* 30, 569–585. doi: 10.1007/s10882-018-9603-3

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders. 5th Edn. Washington: American Psychiatric Association.

American Speech-Language-Hearing Association. (n.d.-a). Telepractice. Available online at: https://www.asha.org/practice-portal/professional-issues/telepractice/ (Accessed October 1, 2025).

American Speech-Language-Hearing Association. (n.d.-b). Early intervention. Available online at: https://www.asha.org/practice-portal/professional-issues/early-intervention/ (Accessed December 27, 2022).

Anderson, K. L., Balandin, S., and Stancliffe, R. J. (2015). Alternative service delivery models for families with a new speech generating device: perspectives of parents and therapists. *Int. J. Speech-Language Pathol.* 17, 185–195. doi: 10.3109/17549507.2014.979876

Boisvert, M., Lang, R., Andrianopoulos, M., and Boscardin, M. L. (2010). Telepractice in the assessment and treatment of individuals with autism spectrum disorders: a systematic review. *Dev. Neurorehabil.* 13, 423–432. doi: 10.3109/17518423.2010.499889

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychother. Theory Res. Pract.* 16, 252–260. doi: 10.1037/h0085885

Breiner, H., Ford, M., and Gadsden, V. L. (2016). Parenting matters: Supporting parents of children ages 0–8. Washington, DC: National Academies Press.

Broomfield, J., and Dodd, B. (2004). The nature of referred subtypes of primary speech disability. *Child Lang. Teach. Ther.* 20, 135–151. doi: 10.1191/0265659004ct267oa

Center for Connected Health Policy (n.d.). What is telehealth? Available online at: https://www.cchpca.org/what-is-telehealth/ (Accessed December 1, 2020).

Christensen, R. H. B. (2023). Ordinal: Regression models for ordinal data (R package) [computer software]: Comprehensive R Archive Network (CRAN). Available at: https://CRAN.R-project.org/package=ordinal (Accessed September 10, 2025).

Christopoulou, M., Drosos, K., and Petinou, K. (2022). Recent advances of telepractice for autism spectrum disorders in speech and language pathology. *Neuropsychiatr. Dis. Treat.* 18, 2379–2389. doi: 10.2147/NDT.S384300

Davis, R., Zaki, F. B. M., and Sargent, L. (2024). Autism and bilingualism: a thematic analysis of practitioner perspectives in the United Kingdom. *Int. J. Lang. Commun. Disord.* 59, 205–222. doi: 10.1111/1460-6984.12939

Department of Health and Human Services (n.d.). Health Professions Shortage Area Find. Available online at: https://data.hrsa.gov/tools/shortage-area/hpsa-find (Accessed February 21, 2022).

del Hoyo Soriano, L., Bullard, L., Hoyos Alvarez, C., Thurman, A. J., and Abbeduto, L. (2021). Using telehealth-delivered procedures to collect a parent-implemented expressive language sampling narrative task in monolingual and bilingual families with autism Spectrum disorder: a pilot study. *Front. Rehab. Sci.* 2:716550. doi: 10.3389/fresc.2021.716550

Duffy, L. V., Evans, R., Bennett, V., Hady, J. M., and Palaniappan, P. (2023). Therapeutic relational connection in telehealth: concept analysis. *J. Med. Internet Res.* 25:e43303. doi: 10.2196/43303

Eadie, P., Morgan, A., Ukoumunne, O. C., Ttofari Eecen, K., Wake, M., and Reilly, S. (2015). Speech sound disorder at 4 years: prevalence, comorbidities, and predictors in a community cohort of children. *Dev. Med. Child Neurol.* 57, 578–584. doi: 10.1111/dmcn.12635

Freckmann, A., Hines, M., and Lincoln, M. (2017). Therapists' perspectives of therapeutic alliance in face-to-face and telehealth speech–language pathology sessions. *Int. J. Speech Lang. Pathol.* 19, 287–296. doi: 10.1080/17549507.2017.1292547

Garrido, D., López, B., and Carballo, G. (2024). Bilingualism and language in children with autistic spectrum disorder: a systematic review. *Neurología (English Edition)* 39, 84–96. doi: 10.1016/j.nrleng.2023.12.007

Gilhuber, C. S., Raulston, T. J., and Galley, K. (2023). Language and communication skills in multilingual children on the autism spectrum: a systematic review. *Autism* 27, 1516–1531. doi: 10.1177/13623613221147780

Gillon, G., Hyter, Y., Fernandes, F. D., Ferman, S., Hus, Y., Petinou, K., et al. (2017). International survey of speech-language pathologists' practices in working with children with autism spectrum disorder. *Folia Phoniatr. Logop.* 69, 8–19. doi: 10.1159/000479063

Grillo, E. U. (2017). Results of a survey offering clinical insights into speech-language pathology telehealth methods. *Int. J. Telerehabilitation.* 9, 25–30. doi: 10.5195/ijt.2017.6230

Hambly, H., Wren, Y., McLeod, S., and Roulstone, S. (2013). The influence of bilingualism on speech production: a systematic review. *Int. J. Lang. Commun. Disord.* 48, 1–24. doi: 10.1111/j.1460-6984.2012.00178.x

Hao, Y., Zhang, S., Conner, A., and Lee, N. Y. (2021). The evolution of telehealth use during the COVID-19 pandemic: perspectives of pediatric speech-language pathologists. *Int. J. Environ. Res. Public Health* 18:12197. doi: 10.3390/ijerph182212197

Hines, M., Lincoln, M., Ramsden, R., Martinovich, J., and Fairweather, C. (2015). Speech pathologists' perspectives on transitioning to telehealth: what factors promote acceptance? *J. Telemed. Telecare* 21, 469–473. doi: 10.1177/1357633X15604555

Hoff, E. (2013). Language development. 5th Edn. Boston: Cengage Learning.

Howard, K. B., Gibson, J. L., and Katsos, N. (2024). Exploring different stakeholder perspectives on bilingualism in autism. *Language* 9:66. doi: 10.3390/languages9020066

International Society for Technology in Education (2007). National Educational Technology Standards for students (NETS). McDonough: International Society for Technology in Education.

Johnson, P. C., Barry, S. J., Ferguson, H. M., and Müller, P. (2015). Power analysis for generalized linear mixed models in ecology and evolution. *Methods Ecol. Evol.* 6, 133–142. doi: 10.1111/2041-210X.12306

Kim, H., Baker, D., Kim, S., Liu, C., and Cook, K. (2024). The impact of educational and medical systems on autistic children from multilingual American homes: a systematic review. *Autism* 28, 2707–2721. doi: 10.1177/13623613241242839

King, M., Ward, H., Soto, G., and Barrett, T. S. (2022). Supporting emergent bilinguals who use augmentative and alternative communication and their families: lessons in telepractice from the COVID-19 pandemic. *Am. J. Speech Lang. Pathol.* 31, 2004–2021. doi: 10.1044/2022_AJSLP-22-00003

Learnihan, V., Hogg, S.-E., and Davis, K. (2025). The telehealth surge: speech-language pathology mode of service delivery and consumer characteristics during COVID-19. *Int. J. Speech-Language Pathol.* 27, 727–734. doi: 10.1080/17549507.2024.2388061

Longhini, J., Rossettini, G., and Palese, A. (2022). Digital health competencies among health care professionals: systematic review. *J. Med. Internet Res.* 24:e36414. doi: 10.2196/36414

Mainz, A., Nitsche, J., Weirauch, V., and Meister, S. (2024). Measuring the digital competence of health professionals: scoping review. *JMIR Med. Educ.* 10:e55737. doi: 10.2196/55737

Martin Loya, M. R., and Meadan, H. (2024). It's our job to bridge the gap: perspectives of bilingual autism providers on heritage language care. *Autism* 28, 2549–2561. doi: 10.1177/13623613241234413

McLeod, S., and Verdon, S. (2017). Tutorial: speech assessment for multilingual children who do not speak the same language (s) as the speech-language pathologist. *Am. J. Speech-Lang. Pathol.* 26, 691–708. doi: 10.1044/2017_AJSLP-15-0161

National Institute of Deafness and Other Communication Disorders (2017). Speech and language developmental milestones. Available online at: https://www.nidcd.nih.gov/health/speech-and-language (Accessed May 20, 2021).

National Institute of Deafness and Other Communication Disorders (2020). Autism Spectrum disorder: communication problems in children. Available online at: https://www.nidcd.nih.gov/health/autism-spectrum-disorder-communication-problems-children#3 (Accessed April 12, 2021).

Nizeyimana, E., Joseph, C., and Louw, Q. A. (2022). A scoping review of feasibility, cost-effectiveness, access to quality rehabilitation services and impact of telerehabilitation: a review protocol. *Digital Health* 8:205520762110667. doi: 10.1177/20552076211066708

O'Connor, S. P. (2012). United States regions national geographic. Available online at: https://www.nationalgeographic.org/maps/united-states-regions/#:~:text=A common way of referring,WestSoutheastandMidwest (Accessed January 28, 2022).

Pang, Y. (2024). Service for children with ASD from culturally linguistically diverse settings at a glance. *Int. Educ. Stud.* 17, 34–38. doi: 10.5539/ies.v17n3p34

Pattison, G., and Powell, T. (1989). Establishing rapport with young children during speech and language diagnostic evaluations. *Natl. Stud. Speech Lang. Hear. Assoc.* 17, 77–80.

Petinou, K., Christopoulou, M., and Antoniou, K. (2024). "Addressing multicultural and multilingual aspects in the assessment of individuals with autism spectrum disorder" in A guide to global language assessment: A lifespan approach. ed. M. Bortz (London: Routledge), 289–303.

Pitt, R., Hill, A. J., Theodoros, D., and Russell, T. (2018). "I definitely think it's a feasible and worthwhile option": perspectives of speech-language pathologists providing online aphasia group therapy. *Aphasiology* 32, 1031–1053. doi: 10.1080/02687038.2018.1482403

R Core Team (2025). R: A language and environment for statistical computing (version 4.4.3) [computer software]. Vienna: R Foundation for Statistical Computing.

Retamal-Walter, F., Waite, M., and Scarinci, N. (2022a). Exploring engagement in telehealth early intervention for young children with developmental disability and their families: a qualitative systematic review. *Disabil. Rehabil. Assist. Technol.* 18, 1508–1521. doi: 10.1080/17483107.2022.2048098

Retamal-Walter, F., Waite, M., and Scarinci, N. (2022b). Families' and professionals' perspectives of building and maintaining engagement in telehealth early intervention for young children with communication disability. *Disabil. Rehabil.* 45, 1165–1177. doi: 10.1080/09638288.2022.2055161

Scharff-Rethfeldt, W., McNeilly, L., Abutbul-Oz, H., Blumenthal, M., Garcia de Goulart, B., Hunt, E., et al. (2020). Common questions by speech and language therapists/speech-language pathologists about bilingual/multilingual children and informed, evidence-based answers. Available online at: https://image.makewebeasy.net/makeweb/0/z4hIV36bT/Document/Common_Questions_by_SLT_SLP_about_Bilingual_Multilingual_Children_and_Informed_Evidence_based_Answers_2020.pdf (Accessed September 27, 2025).

Schieve, L. A., Blumberg, S. J., Rice, C., Visser, S. N., and Boyle, C. (2007). The relationship between autism and parenting stress. *Pediatrics* 119, S114–S121. doi: 10.1542/peds.2006-2089Q

Sharma, A., Minh Duc, N. T., Luu Lam Thang, T., Nguyen Hai Nam, N., Ng, S. J., Abbas, K. S., et al. (2021). A consensus-based checklist for reporting of survey studies (CROSS). *J. Gen. Intern. Med.* 36, 3179–3187. doi: 10.1007/s11606-021-06737-1

Shaw, K. A., Williams, S., Patrick, M. E., Valencia-Prado, M., Durkin, M. S., Howerton, E. M., et al. (2025). Prevalence and early identification of autism spectrum disorder among children aged 4 and 8 years — autism and developmental disabilities monitoring network, 16 sites, United States, 2022. *MMWR Surveill. Summ.* 74, 1–22. doi: 10.15585/mmwr.ss7402a1

Sucala, M., Schnur, J. B., Brackman, E. H., Constantino, M. J., and Montgomery, G. H. (2013). Therapists attitudes toward therapeutic alliance in e-therapy. *J. Gen. Psychol.* 140, 282–293. doi: 10.1080/00221309.2013.830590

Sundarrajan, M., and Franco, J. (2024). Parent-mediated intervention for autistic children offered through in-person and telepractice modalities. *Perspect. ASHA Spec. Inter. Groups* 9, 1217–1234. doi: 10.1044/2024_PERSP-23-00066

Sylvan, L., Goldstein, E., and Crandall, M. (2020). Capturing a moment in time: a survey of school-based speech-language pathologists' experiences in the immediate aftermath of the covid-19 public health emergency. *Perspect. ASHA Spec. Inter. Groups* 5, 1735–1749. doi: 10.1044/2020_PERSP-20-00182

Taiebine, M., and Keegan, L. C. (2024). Telepractice for bilingual and multilingual people with aphasia: a scoping review. *Aphasiology*. doi: 10.1080/02687038.2024.237343

Tickle-Degnen, L., and Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychol. Inq.* 1, 285–293. doi: 10.1207/s15327965pli0104_1

Tremain, H., McEnery, C., Fletcher, K., and Murray, G. (2020). The therapeutic alliance in digital mental health interventions for serious mental illnesses: narrative review. *JMIR Mental Health* 7:e17204. doi: 10.2196/17204

Tucker, J. K. (2012a). Perspectives of speech-language pathologists on the use of telehealth in schools: quantitative survey results. *Int. J. Telerehabil.* 4, 61–72. doi: 10.5195/ijt.2012.6100

Tucker, J. K. (2012b). Perspectives of speech-language pathologists on the use of telehealth in schools: the qualitative view. *Int. J. Telerehabil.* 4, 47–60. doi: 10.5195/ijt.2012.6102

Van Echo, J., Warren, S., and Manthey, J. (2023). SLPs: prepare for telepractice changes as the pandemic eases. ASHA Leader Live. Available online at: https://leader.pubs.asha.org/do/10.1044/leader.IPP.28052023.post-pandemic-telehealth-slps.34/full/?utm_source=chatgpt.com (Accessed September 25, 2025).

Wing, J. M. (2006). Computational thinking. *Commun. ACM* 49, 33–35. doi: 10.1145/1118178.1118215

Wolk, C. B., Marcus, S. C., Weersing, V. R., Hawley, K. M., Evans, A., Hurford, M., et al. (2016). Therapist and client predictors of use of therapy techniques within the context of implementation efforts in a large public mental health system. *Psychiatr. Serv.* 67, 551–557. doi: 10.1176/appi.ps.201500022

Woods, A. J., Göksun, T., Chatterjee, A., Zelonis, S., Mehta, A., and Smith, S. E. (2013). The development of organized visual search. *Acta Psychol.* 143, 191–199. doi: 10.1016/j.actpsy.2013.03.008

World Health Organization (2017). Global diffusion of eHealth: Making universal health coverage achievable: Report of the third global survey on eHealth. Vienna: World Health Organization.

Zilcha-Mano, S., Muran, J. C., Hungr, C., Eubanks, C. F., Safran, J. D., and Winston, A. (2016). The relationship between alliance and outcome: analysis of a two-person perspective on alliance and session outcome. *J. Consult. Clin. Psychol.* 84, 484–496. doi: 10.1037/ccp0000058

Zilliacus, E., Meiser, B., Lobb, E., Dudding, T. E., Barlow-Stewart, K., and Tucker, K. (2010). The virtual consultation: practitioners' experiences of genetic counseling by videoconferencing in Australia. *Telemed. E-Health* 16, 350–357. doi: 10.1089/tmj.2009.0108

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 20 January 2026
DOI 10.3389/feduc.2025.1696031

OPEN ACCESS

EDITED BY
Mingshuang Li,
California State University, Northridge,
United States

REVIEWED BY
Abel Toledano-González,
University of Castilla-La Mancha, Spain
Ruth Crutchfield,
University of Texas–Pan American,
United States

*CORRESPONDENCE
Yao Du
✉ yaodu@usc.edu

RECEIVED 10 December 2025
REVISED 14 November 2025
ACCEPTED 22 December 2025
PUBLISHED 20 January 2026

CITATION
Du Y, Tang Y, Fong KK, Liu Y, Wang D, Tu-Shea X, Liu Y, Zheng Q, Quan S, Xiong J and Sheng L (2026) A citizen science approach toward parents-administered remote language assessment for bilingual Mandarin-English children: an evaluation of in-person and telehealth settings. *Front. Educ.* 10:1696031. doi: 10.3389/feduc.2025.1696031

A citizen science approach toward parents-administered remote language assessment for bilingual Mandarin-English children: an evaluation of in-person and telehealth settings

Yao Du[1]*, Yueyan Tang[2], Ka Kei Fong[3], Yusa Liu[4], Danyang Wang[5], Xintian Tu-Shea[6], Yilan Liu[7], Qingxiao Zheng[8], Shuang Quan[9], Jinjun Xiong[8] and Li Sheng[10]

[1]USC Tina and Rick Caruso Department of Otolaryngology-Head & Neck Surgery, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States, [2]Department of Cognitive Science, University of California - San Diego, San Diego, CA, United States, [3]SPG Therapy and Education, Walnut Creek, CA, United States, [4]Independent Researcher, Los Angeles, CA, United States, [5]Department of Speech-Language Pathology &Audiology, Towson University, Towson, MD, United States, [6]Institute for Learning Sciences, University of Buffalo, Buffalo, NY, United States, [7]Department of Communication Sciences and Disorders, University of Redlands, Redlands, CA, United States, [8]Department of Computer Science and Engineering, University of Buffalo, Buffalo, NY, United States, [9]Department of Education, Juniata College, Huntingdon, PA, United States, [10]Department of Language Science and Technology, The Hong Kong Polytechnic University, Kowloon/Hong Kong, Hong Kong SAR, China

Introduction: The growing population of bilingual children and lack of bilingual clinicians have created an increased need for reliable and accessible bilingual language assessment to accurately detect language delays and disorders globally. To address this growing need, this study evaluated the Mandarin-English Receptive Language Screener (MERLS), a web-based receptive language assessment designed for bilingual Mandarin-English (ME) speaking children.

Methods: Using a citizen science approach, bilingual ME speaking parents based in the United States served as the test administrators. This two-phase study compared bilingual ME speaking children's performance and parent-child interactions across in-person (n = 16) and telehealth (n = 43) settings. Participants in both phases were typically developing children aged 3–10 years who used Mandarin and English for at least 20% of their daily communication.

Results: In Phase I (in-person), despite variability in parent behaviors during administration, parent-administered assessments demonstrated comparable test-retest reliability (Pearson correlation: $r = 0.95$, $p < 0.01$) and item-by-item agreement (82%) to researcher-administered assessments. These reliability metrics are comparable to those of established standardized child language assessments (e.g., PPVT-5 and the QUILS). In Phase II (telehealth), platform improvements (e.g., educational quizzes and videos on proper test administration) significantly reduced interfering parent behaviors (Mandarin items: $W = 485$, $p = 0.004$; English items: $W = 482$, $p = 0.003$) without affecting children's test performance.

Discussion: These results support the feasibility of using a citizen science approach and a digital assessment platform MERLS for parent-administered language assessments. Such innovative assessment approach has great potentials to increase access to accurate and reliable language assessment services for bilingual ME speaking children in the United States. The findings offer clinical and technical insights for developing bilingual child language assessments across both in-person and telehealth settings.

KEYWORDS

citizen science, bilingual children, Mandarin-English, language assessment, telehealth

1 Introduction

Bilingual children are often misdiagnosed with a language disorder when they are not appropriately and accurately assessed in both languages (Freeman and Schroeder, 2022; Boerma and Blom, 2017; Grimm and Schulz, 2014; Oetting, 2018; Samson and Lesaux, 2009). Approximately 7%–11% of bilingual children learners are at risk of developing a language impairment (Park et al., 2017; Norbury et al., 2016; Tomblin et al., 1997). The scarcity of appropriate bilingual language assessments makes it difficult to accurately assess bilingual children's language skills (Westerveld, 2014; Du et al., 2020). In addition, bilingual children have different language development patterns compared to their monolingual peers (Pearson, 2013; Song et al., 2021, 2022). While bilingual children often know fewer words in each of their languages than monolingual learners, the differences disappear when combining bilingual children's "conceptual vocabulary" across both languages (Marchman et al., 2010; Hyter, 2021). Therefore, clinicians need to accurately assess a bilingual child's language abilities in each of their languages (Gillam et al., 2013; Castilla-Earls et al., 2020; Kritikos, 2003). Given the complex bilingual language profiles, standardized tests that are based on bilingual-specific norms are necessary to collect accurate language assessment profiles (Jasso et al., 2020). Despite advancement in bilingual assessment development (Patterson and Pearson, 2004; Peña et al., 2014; Golinkoff et al., 2017; Jasso et al., 2020; Peña and Sutherland, 2022; Caesar and Kohler, 2007), there is a lack of reliable and valid multilingual assessment tools and a limited linguistic diversity for bilingual children (Kimble, 2013; Peña and Sutherland, 2022; Dollaghan and Horner, 2011; Kan et al., 2020).

English and Chinese are among the world's most widely used languages, with an estimated population of 1.5 billion and 1.1 billion speakers, respectively (Dyvik, 2024). The rapid rise in immigration and globalization has led to a growing population of bilingual English and Chinese speakers in major English-speaking nations, including the United States, Canada, the United Kingdom, and Australia (Grenoble and Osipov, 2023; Gov.UK, 2020). In the U.S., Chinese is the most spoken Asian-Pacific Island language among individuals five years and older (Ryan, 2013; McLeod and Crowe, 2018), with Mandarin being the most prevalent dialect of Chinese, spoken by approximately 3.4 million people (Lesso, 2023; U.S. Census Bureau, 2024). Despite such a significant population need, over the past decade, a standardized and comprehensive child language assessment has not been developed for detecting bilingual Mandarin-English (ME) speaking children who are at risk for language delays and disorders. A lack of bilingual ME-speaking SLPs further exacerbated this gap: as of 2024, there are only 491 bilingual ME speaking SLPs in the U.S., and most of them are located in coastal states such as California and New York which further limited access for care (American Speech-Language-Hearing Association [ASHA], 2025b). ASHA continues to address this service gap in 2025 by defining competencies and providing resources for multilingual service delivery (American Speech-Language-Hearing Association [ASHA], 2025a). However, ASHA does not accredit or approve specialized training programs for multilingual service providers, which means the depth and breadth of training can vary significantly across institutions. This uneven distribution severely limits access to appropriate bilingual assessment services across much of the country. As a result, the majority of ME speaking children are assessed by monolingual English-speaking SLPs, who often rely on interpreters to manually translate or interpret assessment items from standardized English tests into Mandarin (Langdon and Quintanar-Sarellana, 2003). Such an assessment practice may fail to capture linguistic constructs that are unique to Mandarin, and diagnosis can be less reliable as translation errors are introduced in the evaluation process (Sheng et al., 2021; Du et al., 2020). Consequently, bilingual ME-speaking children are prone to receive over- and under-diagnosis of language disorders, ultimately impacting their development and well-being and increasing healthcare cost for the society (Flores and Tomany-Korman, 2008; Dollaghan and Horner, 2011; Yu et al., 2021). This assessment gap has severe implications for bilingual ME children with language disorders, who require accurate identification to receive appropriate special education services and interventions.

Despite the urgent need for creating bilingual language assessments to improve the current standard of care, researchers face another practice challenge during assessment development: collecting a large-scale nationally representative sample of bilingual ME-speaking children to establish a bilingual language norm for standardized bilingual language assessment development (Sheng et al., 2021). Asian populations in the U.S. tend to concentrate in certain metropolitan areas (National Academies of Sciences, Engineering, and Medicine, 2018; Cooc, 2018), and researchers outside these geographical areas do not have ready access to bilingual participants. Typical laboratory or school-based testing requires extensive travel, time commitment, and trained multilingual personnel. These barriers hinder researchers from collecting large-scale data across different developmental age groups across different regions in the U.S. To address the data

collection challenges, it is critical to explore alternative methods, such as citizen science, which refers to a research approach that involves members of the public contributing to data collection and scientific discovery (Bonney et al., 2016). These approaches can enable better access to large bilingual children and easier data collection methodologies for researchers.

In addition to assessment tool limitations, tester effects can also influence diagnostic outcomes. Prior research has shown that young children responded differently when interacting with parents versus unfamiliar testers, especially in tasks involving social-communicative cues. For example, Tang et al. (2023) found significant differences in infants' attention-following responses to joint attention cues depending on whether the cue was provided by a caregiver in a home setting or a tester in a lab setting (Brown and Woods, 2015). When adopting a more accessible telehealth approach for assessment development, it is important to consider a variety of contextualized factors when transitioning from in-person to virtual and computer-administered assessments for young children (Paradis, 2011; Brandone et al., 2008; Werfe et al., 2021; Khoshsima and Toroujeni, 2017; Magasi et al., 2018; Solano-Flores et al., 2019).

Our study aims to resolve these issues by establishing the initial feasibility of a digital bilingual assessment tool for ME speaking children in multiple service delivery modes (in-person and telehealth) to increase access for care (Ciccia et al., 2011). This paper addresses this need by testing the feasibility of parent-administered, telehealth-based assessments within a citizen science framework, with the goal of generating reliable bilingual language data and addressing barriers related to geographical limitations and the shortage of bilingual research personnel.

2 Related work

2.1 A citizen science approach for parent-administered assessment via telehealth

Citizen science, broadly defined as the involvement of the public in scientific research, has gained traction in various domains, particularly in environmental and ecological science (Fraisl et al., 2022; Schmitz et al., 2018; Bhattacharjee, 2005). Prior research showed that citizen science samples are far more diverse than samples from lab-based studies (Gosling et al., 2004; Reinecke and Gajos, 2015). The involvement of citizen scientists could vary from merely helping with labor-intensive data processing to direct involvement in the language assessment process as test administrators. The involvement of parents from diverse racial and ethnic backgrounds is vital to collect large scale diverse language data that are essential to support the norm development of a bilingual child language assessment.

Citizen science is also a promising approach to obtain large scales of diverse samples across time and location. With the proliferation of web-based assessment and increased adoption of telehealth as a service delivery method (Waite et al., 2010; Grillo, 2021; Perrin et al., 2020; Farmer et al., 2020; Lehner et al., 2021; McCrae et al., 2021; Schmitt et al., 2022; Shankar et al., 2022; Farmani et al., 2024), citizen science approach further enable parents to serve as telehealth assistants to support researchers by serving as test administrators virtually through videoconferencing platforms (Klatte et al., 2020; Sutherland et al., 2021; Dekhtyar et al., 2020; Marhefka et al., 2020) for collecting child language data in the home setting. Such an approach is in alignment with the core component of family-centered care services which actively involves caregivers as a part of the assessment process (Crais et al., 2006; Corona et al., 2021; Frigerio et al., 2021; Dodge-Chin et al., 2022).

Additionally, another benefit for utilizing the citizen science approach is to reduce practical limitations (e.g., limited research budget), because citizen science projects do not offer cash or course credit for compensation. In citizen science projects, although some people are motivated to participate in a study that has monetary reward, nearly everyone is motivated to participate in a project that is intrinsically rewarding. For example, birders help with bird surveys, and astronomy enthusiasts categorize images of galaxies (Raddick et al., 2009). This incentive structure is particularly relevant to language assessment data collection, as parents are intrinsically motivated to learn more about their children's bilingual language abilities, which also makes parents more likely to participate in citizen science research (Bonney et al., 2016). However, it is unknown whether relying on parents as citizen scientists to collect data for their own children at a large scale would lead to meaningful and high quality data (Li et al., 2024). The present study directly addresses this gap by examining whether parent-administered, telehealth-based assessments can generate reliable bilingual language data suitable for research and future test development.

2.2 Challenges of parent-administered online assessment

Several barriers on data collection must be taken into consideration when adopting parents to assess their children's bilingual language abilities as citizen scientists. Though previous research has involved parents as assistants to help facilitate language service sessions, enlisting parents as independent test administrators for language assessment is uncommon due to their lack of professional training (Tomlinson et al., 2018; Talbott et al., 2020; Corona et al., 2021). After all, parents would not be compensated as professional testers would be, nor expected for their livelihood to adhere to professional standards. A key concern is that parents have varying language competencies and limited knowledge on language assessment principles. Since parents typically do have expectations or concerns regarding their children's language skills, they may not be as impartial or unbiased, which can compromise the validity and reliability of parents-administered assessments (Sullivan, 2011). When administering bilingual assessments with their children, some parents demonstrated limited ability to comprehend and follow proper test instructions; additionally parental interference behaviors in parents' proficient language but children's less proficient languages have been reported (Du et al., 2020, 2021). In addition, the absence of a trained professional (e.g., researcher, clinician) during home-based assessments raises concerns about the overall quality of the data collected. Together, these challenges

underscore the need to better understand parental behaviors and to design parent-administered assessment protocols that provide clear guidance, support, and safeguards for data quality when implementing a citizen science approach.

Researchers have provided models of training programs to improve test administration skills in citizen scientists and to identify potential solutions to address the challenges of engaging parents. Tomlinson et al. (2018) identified 20 studies on applied behavior analysis that trained individuals (e.g., parents) for assessment, teaching, and intervention purposes, and suggested that citizen scientists with no prior experience in behavior analytic techniques can be trained to adhere to protocols and implement a range of behavioral analysis techniques. All training in the reviewed studies were delivered via videoconferencing with a trainer, who was usually an experimenter/professional with prior experience in behavior analytic approaches. Training sessions usually lasted between 15 min to 3 h, which involved strategies such as direct instruction, modeling, or role playing (Alnemary et al., 2015; Barkaia et al., 2017; Hay-Hansson and Eldevik, 2013). Online modules, written explanations of the techniques, and a supplemental trainee manual were also used in studies to enhance trainees' adherence (Scott et al., 2017; Radville et al., 2022). Therefore, in order for successful online data collection with parents via telehealth, it is critical to evaluate not only children's performance but also parents' behaviors during testing (Molini-Avejonas et al., 2015). The present study contributes to this effort by systematically examining parents' behaviors under both in-person evaluation and remote testing conditions.

2.3 Study aims

This study aims to evaluate the feasibility of using parents as citizen scientists to test their own ME speaking children via a web-based telehealth-friendly Mandarin-English Receptive Language Screener (MERLS). We propose that such an approach can bring two contributions to addressing the current standard of care challenges in bilingual child language assessment by (1) training parents as test administrators for assessment data collection and development in partnership with researchers using an automated web-based language assessment, (2) evaluating parents' assessment process and parent-child outcomes during the telehealth setting. Specifically, we investigate whether parents can be trained to act as competent test administrators by adhering to the test protocols in-person (Study 1) and virtually via telehealth context (Study 2). Specifically, the present paper examines the following research questions:

Study 1:

(1) During in-person assessment, is children's performance comparable between parent- and researcher-administered sessions with adequate test-retest reliability?
(2) What are the characteristics of different parent behaviors (e.g., behavioral types and frequency) during parent-administered in-person sessions?

Based on the parent behaviors observed in study 1, we made adjustments to the assessment and training protocol. After improvement, we ask:

Study 2:

(1) With technical improvement of the assessment, what changes were found in the types and frequency of parent behaviors during the telehealth assessment?
(2) What are the verbal and nonverbal interaction patterns of parent-child dyads during the telehealth assessment?
(3) How did contextual factors (e.g., children's age, frequency of digital device use, and test performance) influence the frequency of parent interference behaviors?

We hypothesize that children's language assessment performance will be consistent across parent- and researcher-administered conditions in Study 1, indicating no significant test differences between the two conditions and feasibility for utilizing parents as test administrators. With improved system design, parents' interference should significantly decrease in the telehealth context in Study 2, offering future directions for utilizing a citizen science approach towards developing MERLS assessment via both in-person and telehealth delivery modalities.

3 Study 1: in-person evaluation of MERLS

3.1 Study 1 materials and methods

3.1.1 Participants

A total of 29 ME speaking parent-child dyads (see Supplementary Table 1 for demographic information) were recruited through advertisements distributed via parent email lists affiliated with local Chinese language schools and bilingual SLP Facebook groups in North America. Participating children ranged in age from 3 to 10 years old. This age range is consistent with standardized language assessments such as the Test of Early Language Development-4 (TOELDS-4, 3; 0–7; 11 years) and is narrower than widely-used tools including the Preschool Language Scales-5 (PLS-5, 0; 0–7; 11 years), the Clinical Evaluation of Language Fundamentals-5 (CELF-5, 5; 0–21;11 years), and the Peabody Picture Vocabulary Test-5 (PPVT-5, 2; 6–90+ years), which assess language constructs across broad developmental periods. For detailed analysis of parental behaviors during testing, a subset of the sample (n = 16) was selected based on parents who provided consent for video recording for further analysis. For the purposes of this study, bilingualism was defined broadly to include both simultaneous bilinguals (exposed to both languages from infancy) and sequential bilinguals (learned one language after the other). The video (Supplementary Video) included demonstrations of both prohibited parent interference behaviors and acceptable supportive behaviors (Supplementary Figure 3). The primary inclusion criterion was that children used both Mandarin and English in at least 20% of their daily lives, as reported by their parents. This inclusive definition was chosen to capture a wide range of bilingual experiences representative of the community.

3.1.2 Materials

Mandarin-English Receptive Language Screener 1.0 is an online receptive language comprehension assessment screening designed for bilingual ME-speaking children. The web interface provides pre-recorded audios for children to select the corresponding picture stimuli; Supplementary Figure 1 provides a visual representation of the test interface and an example test item in English. This test evaluates critical language components in Mandarin with 44 items and in English with 36 items, assessing linguistic constructs including prepositions, classifiers (Mandarin) or plurals (English), quantifiers, passive sentences, and relative clauses. These language components selected for evaluation have been demonstrated to be the particular linguistic weaknesses in children with language disorders and have been utilized in previous related studies (Golinkoff et al., 2017; Peña et al., 2014; Hu et al., 2016; Jia and Aaronson, 2003; Wang et al., 2022; Wong et al., 2004; Zhou and Crain, 2011; Sheng et al., 2011, 2016; Sheng, 2018). Prior work has provided preliminary evidence that the Mandarin-English Receptive Language Screener (MERLS) is an effective bilingual screener. Du et al. (2021) reported high test–retest reliability and strong concurrent validity with established English and Mandarin comprehension measures, supporting its use for receptive language assessment in bilingual children.

The receptive language task employs a sentence-picture matching format. Participants are required to select the appropriate picture from a set of four pictures after listening to a pre-recorded sentence audio in either Mandarin or English. The assessment instructions are provided in both Mandarin and English through audio recordings, ensuring accessibility for parents with varying language proficiency levels, effectively eliminating potential language barriers. The assessment begins with a welcoming message in both English and Mandarin. Two practice items are then presented to acquaint children with the testing format. All children followed the instruction and made selections on the computer, whether with or without parental assistance. The audio clips were played twice during the assessment, with a 15-s interval between items. In cases where the child did not respond within 15 s after the second play, the web page automatically advanced to the next item (Supplementary Figure 1). The audio was played at an approximate volume of 65 dB through the computer's built-in audio system. Once a selection was made, the child could not revisit previous items. This type of closed-set tasks was proven to be reliably tested by monolingual clinicians who cannot speak the language (Cowan et al., 2022).

3.1.3 Procedure

During the administration of MERLS 1.0 by parents, all interactions between parents and children were recorded on video. To ensure comprehensive recording, a video camera was positioned behind the dyads, capturing both the activities on the computer screen and dyadic interactions. This setup served to protect the privacy of the participants while minimizing potential distractions. MERLS 1.0 was first administered either at the children's home or in a laboratory setting, with the assessment conducted by either a caregiver (n = 17) or a trained examiner (n = 12). Parents also completed a pre-assessment questionnaire which included questions on participants' demographic information (age, gender, and level of education), and questions about the parent's English proficiency and child digit media and device use. Subsequently, children underwent a second MERLS 1.0 assessment within 2–4 weeks following the first testing.

To examine the reliability of parent-administered test sessions, in addition to these initial 16 parent-child dyads who had parent-first researcher-next sessions; we also tested another 12 dyads with researcher-first and parent-next sessions and analyzed children's test results for test-retest comparison. The second assessment was administered within 2–4 weeks from the initial testing, with a different administrator. For example, children who completed the first MERLS 1.0 assessment with their parents underwent the second testing with a lab examiner and vice versa. During both testing sessions, the items were displayed on a 15-inch laptop monitor positioned approximately two feet away from the child. The laptop utilized in the assessment was equipped with a touch screen, allowing children to select answers by simply pointing and touching the screen. A brief instruction page was provided before the test started. Children were allowed to take unstructured breaks as needed throughout the Mandarin and English modules, and they were expected to complete all the items. Additionally, parent questionnaires were administered by two trained bilingual (Mandarin-English) research assistants (1 undergraduate student and 1 graduate student in Communication Science and Disorders). All administrators completed a standardized training protocol covering questionnaire content and structure, questioning techniques to avoid leading responses, and data recording procedures. Administrators followed a structured script to ensure consistency across all participants.

3.1.4 Data analysis

All children's performance on the sentence comprehension task were automatically scored and recorded within the MERLS 1.0 system. Two trained bilingual ME-speaking research assistants watched video recordings of 16 parent-child dyads, and then independently transcribed children's utterances and coded parental behaviors during the tests using a clinically informed codebook (Du et al., 2020, 2021). This codebook delineated four categories of interference behaviors including "repeating questions, answering questions, analyzing items, and judging of correctness" and four categories of parent support behaviors including "encouragement, verbal or physical technical support, broadcasting, and miscellaneous" (Supplementary Table 2). Video coding included parent and child verbal and non-verbal behaviors, as well as environmental distractors (Du et al., 2020). An interobserver agreement (IOA) of 97% was reached between two trained video analysts.

3.2 Study 1 results

3.2.1 Children's performance across different administrators in Study 1

The reliability between children's performance on the MERLS 1.0 administered by the parent and the researcher was examined using item-by-item analysis and correlational analysis. The item-by-item analysis was conducted by comparing children's accuracy (0 or 1) on the same item between the first and second testing session. Reliability was calculated by using the number of

consistent items divided by the number of total items. Pearson correlations (Bishara and Hittner, 2012) were also conducted between children's overall performance on parent- and researcher-administered sessions to examine whether the two sessions yielded similar performance on the same tasks.

All 29 children completed the English MERLS 1.0 in both sessions. Five children did not complete the Mandarin MERLS 1.0 in either session. Item-by-item analysis (Cohen et al., 2003) showed that the overall item consistency was 82% (children scored the same on 82% of the total items in the first and second testing session), with similar consistency in the parent-first (n = 17, consistency = 82%) and the researcher-first groups (n = 12, consistency = 82%). Item-by-item consistency was slightly higher for the English (84%) subtest compared to the Mandarin subtest (80%). Pearson correlation results showed that children's performance in the two administration sessions was significantly correlated for the overall group (r = 0.95, p < 0.01), and for the parent-first (English: r = 0.94, p < 0.01; Mandarin: r = 0.85, p < 0.01) and researcher-first groups (English: r = 0.97, p < 0.01; Mandarin: r = 0.91, p < 0.01) separately. A correlation of 0.90 and above is considered excellent; in the 0.80 s is good, and in the 0.70 s is adequate (Cohen et al., 2003); therefore, parents were able to supervise their children in completing our task by eliciting similar performance as compared to performance supervised by trained researchers.

3.2.2 Types and frequency of parent behaviors

The 16 randomly selected parent-child dyads in Study 1 demonstrated a total of 677 behaviors, including 296 interference behaviors and 381 support behaviors while administering MERLS 1.0 to their children (Supplementary Table 5). Eleven out of 16 dyads demonstrated adherence, defined as less than 10 parent interference behaviors (Du et al., 2020; Kelders et al., 2011) to the assessment protocol after viewing the introduction video. Five out of 16 parents failed to adhere to the testing protocol and demonstrated more than 10 interfering behaviors per person. Specifically, these five parents demonstrated a total of 280 out of 296 (95%) interference behaviors across the 16 parent-child dyads. Furthermore, a cross-language variation was found in parent behaviors, characterized by more support and inference behaviors in the Mandarin than English modules. On average, 16 parents interfered in approximately 10 items in Mandarin and five items in English and offered support to 11 items in Mandarin and six items in English. The top two frequent interference behaviors are "Repeating Questions" and "Analyzing Items," whereas the top two frequent support behaviors are "Technical Supports" and "Encouragement."

3.3 Study 1 discussion

In-person evaluation of MERLS 1.0 showed that parents were able to administer language assessments to their children independently using MERLS 1.0, offering additional insights for parent-administered automated web assessment to collect bilingual child language data. The test-retest reliability of children's performance between parent- vs. researcher-administered sessions are consistent, suggesting that child language data collected by parents using MERLS 1.0 were consistent with the data collected by researchers. Furthermore, the test-retest reliability based on a Pearson correlation coefficient of.95 (range = 0.85–0.91) for MERLS 1.0 is comparable to other standardized child language assessments, indicating high quality assessment outcomes for MERLS 1.0. For example, Peabody Picture Vocabulary Test (PPVT-5) has a Pearson correlation coefficient of.93 (range = 0.92–0.96) from 340 subject samples during a 4-week test-retest interval; the Quick Interactive Language Screener (QUILS) which sampled 75 subjects during a 3–5 weeks test-retest duration showed an overall test-retest correlation of 0.83. Furthermore, the item-by-item agreement for MERLS 1.0 was 82% (range = 80%–84%), consistent between parent-administered and researcher-administered sessions for the 11 parents who did not interfere much of their children's sessions and the five parents who showed most interference behaviors. This item-by-item agreement indicated that parent behaviors did not impact children's overall performance on individual assessment test items, and that these sessions can be as reliable as researcher-administered assessment sessions. These findings offered initial feasibility for the citizen science approach using parents to collect bilingual language data to gather receptive language assessment from their own children.

Although high test-retest reliability and item-by-item agreement were observed between sessions administered by parents and researchers, parent interference behaviors were still observed as a potential concern when engaging parents as citizen scientists. Closer examination revealed that interference often arose from a combination of language- and culture-related factors. For example, parents were more likely to intervene in Mandarin modules than English ones, reflecting greater comfort with the home language and a desire to clarify tasks for their children (Du et al., 2020). Natural code-switching practices in bilingual households also contributed to parents repeating or translating questions across languages, inadvertently increasing children's cognitive load. In addition, cultural expectations surrounding parental roles in education may have shaped parents' tendency to confirm or encourage children's answers, as many interpreted their role as co-administrators rather than passive observers. Finally, parental anxiety about their child's performance and desire for success motivated them to repeat or analyze test items, even when explicitly instructed not to. Together, these findings suggest that parent interference behaviors were not random but stemmed from linguistic, psychological, and sociocultural motivations. This prompted us to investigate the parent instruction page for MERLS 1.0, which provided essential education on parental interference behaviors that are prohibited during the assessment. Prior work by Du et al. (2020) suggested that parent behaviors during the administration of MERLS 1.0 might impact children's performance based on a subset of the participating dyads' performance and behaviors in Study 1, indicating ongoing needs to evaluate parents' adherence to the MERLS platform. To resolve this issue of parent interference behaviors, we adjusted MERLS 1.0 by adding new interface features (e.g., break pages and animated pictures to better engage children) and parent education and assessment materials (e.g., instructional video and quiz questions) and developed a new version of MERLS 2.0 (Supplementary Table 3).

4 Study 2: telehealth evaluation of MERLS

Building on the findings from Study 1 which demonstrated the feasibility of parent-administered assessments and highlighted the impact of interference behaviors, we made corresponding adjustments to the testing platform to improve its functionality. In Study 2 we explore how design improvements affect parent behaviors and parent-child interactions in a telehealth setting. By shifting from in-person to virtual testing, Study 2 evaluates whether these interventions can reduce interference, increase the support behaviors, and maintain data quality when parents independently administer the MERLS assessment at home.

4.1 Study 2 materials and methods

4.1.1 Participants

A total of 43 ME-speaking parent-child dyads (see Supplementary Table 4 for demographic information) in North America were recruited in Study 2 through advertisements on social media platforms such as WeChat. Participating children were aged from 3 to 10 years old, were typically developing, had normal or corrected-to-normal vision with no known genetic, neurological, or psychiatric disorders. All children used Mandarin and English in at least 20% of their daily life. This 20% threshold aligns with established bilingual assessment protocols (Hoff et al., 2012; Peña et al., 2014) and ASHA clinical practice guidelines for identifying bilingual status in pediatric populations (De Lamo White and Jin, 2011). Dyads completed the task remotely from their homes via Zoom, without an in-person experimenter presence (Pearson Education Inc., 2020). For a more in-depth analysis of parent-child interaction patterns, a subset of 36 bilingual ME-speaking parent-child dyads from this larger group was selected for detailed analysis on modes of parental-child interactions, including verbal utterances and non-verbal behaviors. This selection was made because only these 36 videos observed parental behavior or verbal utterances during the assessment; the remaining 7 out of 43 dyads showed no observable parental behavior or verbal utterances.

4.1.2 Materials

The MERLS 2.0 was developed as an updated version of MERLS 1.0, incorporating the redesign recommendations outlined in Du et al. (2020) (Supplementary Table 3). A major enhancement in MERLS 2.0 was the addition of a three-minute parent training video that provided a comprehensive orientation to the testing procedure (Supplementary Figure 2). The video included demonstrations of both prohibited parent interference behaviors and acceptable supportive behaviors (Supplementary Figure 3). Additional updates included a brief parent assessment quiz to reinforce understanding of the protocol, revised testing item order, and updated graphic designs to maintain child engagement.

4.1.3 Procedure

Mandarin-English Receptive Language Screener 2.0 was administered once at children's homes by their caregiver. In cases when the internet connectivity was insufficient to support the video conferencing platform (e.g., for P2, P8, P9), an experimenter provided support by screen sharing and granting the child remote control access to complete the task. Families received an online testing preparation sheet one day before the scheduled appointment, outlining the required equipment and environment setup. During the session, the task was presented on the screen of the computer or iPad positioned approximately two feet away from the child. The child was instructed to respond to the questions by selecting the answers via the iPad touchscreen, a mouse, or a touchpad. Parents were allowed to help with technical difficulties, such as helping the child click responses. Audio instructions were played through headphones and/or speakers and also shared with researchers via Zoom. Parents were instructed to adjust the audio volume to a comfortable level during the newly added instructional video, which they viewed before the task began. To ensure comprehension of the test protocols, parents completed a quiz at the end of the instructional video before beginning the assessment. After watching the instructional video, the parent would access the MERLS 2.0 website via a link shared in the Zoom chat. The testing process was recorded on Zoom, capturing both the shared screen (to document children's testing progress), and the webcam video (to observe parent-child interactions during the assessment). Animated break pages were built-in to give children a break during the assessment. The experimenter remained muted throughout the testing process unless there's technical issues that required interventions. Additionally, pre-assessment questionnaires were completed by parents independently via online survey platform (RedCap) with built-in validation checks. Two trained research assistants (1 graduate student, 1 undergraduate research assistant) reviewed all completed questionnaires for completeness and clarity. Follow-up clarification was conducted via email or brief Zoom calls when responses were unclear or incomplete.

4.1.4 Data analysis

All parent-child interactions were video-recorded using the recording function via Zoom. Two ME speaking research assistants transcribed the videos verbatim based on children and parents' verbal communication, and also nonverbal actions visible via the video recording camera through Zoom following the coding categories presented in Supplementary Tables 2, 4. Two research assistants independently coded all 43 videos with an IOA of 86.1%. Transcription was further verified for accuracy using nine randomly selected videos out of the 43 videos. All children's performance during the test was automatically recorded and collected online.

Additionally, within the 43 video data, 36 videos were observed with parent verbal or non-verbal behavior and were transcribed for further analysis. Transcription focused on participating children's verbal utterance (CU) and children's non-verbal behaviors (CB), as well as their parents' verbal utterances (PU) and parents' non-verbal behaviors (PB) in both Mandarin and English sessions. First, PB, PU, CB, and CU were coded and documented to gather the occurrences of these interactions in a spreadsheet with the de-identified participant ID and timestamps. Parent-child interactions were further classified into four types of codes: PB2CB, PB2CU, PU2CB, and PU2CU (Supplementary Table 4). After all behaviors were coded, an inter-rater reliability (IRR) check was conducted by two research assistants, who re-watched and independently coded 20% of randomly selected videos selected. Then, another senior

researcher compared the consistency of the codes between two research assistants. Across the four videos reviewed, 172 instances of children's behavior were identified, with 141 coded consistently. Thus, the IRR for the coding process is 82.0%.

Then we first run descriptive statistics to generate an overall pattern of the parents and children behavior. Each occurrence of PB, PU, CB, and CU were counted as 1. Each PB or PU followed by one occurrence of CB or CU within two timestamps was counted as one parent-child interaction (i.e., PB2CB, PB2CU, PU2CB, PU2CU). Descriptive analysis was conducted for PB, PU, CB, and CU, as well as four types of parent-child interaction in both Mandarin and English sessions. Paired-test was used to examine the differences of occurrences between two language sessions. Additionally, a qualitative interaction analysis (Jordan and Henderson, 1995) was conducted by two authors who analyzed the transcript with most parents' behavior and parents' behaviors that lead to children's utterances and interactions. This qualitative interaction analysis primarily focused on: (1) how parents supported young children, and (2) how parents encouraged young children to engage in the online assessment task.

4.2 Study 2 results

4.2.1 Types and frequency of parent behaviors

The 43 parents in Study 2 demonstrated a total of 795 behaviors, including 50 interference behaviors and 745 support behaviors. A total of 42 out of the 43 parent-child dyads adhered to the assessment protocol and demonstrated less than 10 parental interference behaviors (Kelders et al., 2011). Only one parent demonstrated more than 10 interreference behaviors during the test (Mandarin: $n = 13$; English: $n = 0$). Four dyads experienced technical issues during the assessment, which led to an increase of verbal technical support behaviors. Different types and frequencies of parent behaviors in Study 2 are presented in Supplementary Table 5.

4.2.2 Overall parent behaviors across Study 1 and Study 2

To compare parent interference and support behaviors between Study 1 and Study 2, Shapiro Wilk's tests (Ghasemi and Zahediasl, 2012) were first conducted to check the normality of parent behaviors during Mandarin and English modules. All variables were not normally distributed. Wilcoxon rank sum test was performed to check if there were significant differences in parent behaviors between the two groups. Specifically, parents in Study 2 demonstrated an increase in adherence to the assessment protocol and displayed fewer parent interference behaviors ($W = 485$, $p = 0.007$). On average, each parent demonstrated 18.5 interference behaviors in Study 1 and 1.2 interference behaviors in Study 2.

4.2.3 Parent behaviors across language modules and studies

To examine parent behaviors between English and Mandarin modules, the Wilcoxon rank sum test was conducted on parent interference and support behaviors across Study 1 and Study 2. Initial analyses indicated that parent interference behaviors have decreased significantly in Study 2 compared to Study 1 in Mandarin modules ($W = 485$, $p = 0.004$) and English modules ($W = 482$, $p = 0.003$), especially in "Repeating Question" behavior (Mandarin: $W = 509$, $p < 0.001$; English: $W = 445$, $p = 0.01$) and "Analyzing Items" (Mandarin: $W = 464$, $p = 0.002$; English: $W = 450$, $p < 0.001$). Parents also displayed significantly less "Judging of Correctness" behaviors in Study 2 in English modules ($W = 507$, $p < 0.001$). The decreases were not significant, however, in "Answering Questions" for both language modules (Mandarin: $W = 385$, $p = 0.17$; English: $W = 344$, $p = 0.81$) and "Judging of Correctness" ($W = 432$, $p = 0.07$) in Mandarin modules.

To examine the consistency of parent behavior across languages within a Study, a Wilcoxon signed-rank test was conducted on parent behaviors between Mandarin and English modules. Analyses indicated that there were no significant differences in parent behaviors in Study 1 across language modules for MERLS 1.0 (interference behaviors: $W = 154$, $p = 0.30$; support behaviors: $W = 159$, $p = 0.25$). Similarly, there were no significant differences in parents' behavior patterns across two language modules in Study 2 (interference behaviors: $W = 968$, $p = 1$; support behaviors: $W = 1{,}041$, $p = 0.68$).

The descriptive statistics for different modes of parent-child dyads in the English and Mandarin sessions are presented in Supplementary Table 6. Paired t-tests were conducted to examine differences between language sessions. The results showed that children exhibited significantly more utterances (CU, $t = -3.299$, $p < 0.01$) in Mandarin sessions compared to English sessions. Additionally, more child behaviors following parent verbal utterances (PU2CB, $t = -2.190$, $p < 0.05$) were found in Mandarin than in English sessions. For other types of parent-child interactions, no significant differences were found between the two language sessions.

Building on the quantitative findings, particularly the increased child utterances and parent utterances leading to child behaviors observed in Mandarin sessions, an in-depth qualitative interaction analysis was conducted to examine the specific ways parents supported and encouraged their children. These sessions revealed three primary types of parental support: (1) technical guidance on using the MERLS platform, (2) encouragement and re-engagement strategies, and (3) clarification of meaning when children struggled with Mandarin vocabulary. For example, in one Mandarin session (P5, Supplementary Figure 4), the transcript illustrated parent-child interactions across four test items. When the child encountered the sentence "The calf is carrying a crocodile who is painting," she turned to her mother for assistance (Line 1). The mother leaned in to read the sentence aloud and explained the clicking process (Line 3), providing the child with technical support that enabled her to make the correct selection. However, when the child heard a later Mandarin sentence she did not understand, she became anxious and repeatedly verbalized the item (Lines 6–9). The mother first offered emotional reassurance ("It's okay," Line 8) and encouraged the child to complete the task independently. When confusion persisted, the mother began translating specific Mandarin terms to English to help the child understand the item.

These patterns illustrate that parental support tended to increase from minimal assistance (e.g., guiding technical interaction) to more involved clarification when children showed signs of distress or disengagement. Notably, parents rarely gave direct answers unless the child became visibly frustrated. Instead, they employed prompts like "listen again," "try it yourself," or "calm

down" to help children re-engage. In addition, parents offered affirmative feedback to maintain motivation. Common phrases included "good job," "yes, that's it," and "you're doing great," which often prompted enthusiastic responses from children (e.g., "Oh! Yes!" or "I'm correct!"), suggesting that emotional support played a role in sustaining engagement. These findings highlight how parents in the Mandarin module not only adhered to the MERLS protocol but also actively scaffolded their children's participation using verbal strategies that supported comprehension, emotional regulation, and task persistence, especially when the child encountered linguistic or attentional challenges.

4.2.4 Contextual variables of parent behaviors during MERLS 2.0

The above results indicate that caregivers can be effectively trained by a short instructional video to demonstrate adherence to the test protocols for MERLS 2.0. The following analyses describe three contextual variables: children's age, children's device use, and children's assessment performance, and consider how these contextual factors influence parent interference behaviors while administering MERLS 2.0. Children's age was reported by parents and recorded in months. Children's performance was measured as the percentage accuracy that children obtained in the language tasks, which indicates language proficiency. Children's device use was collected from the digital media questionnaire reported using a Likert scale from 1 to 3, where 1 indicates the child almost never uses computers or electronic devices, 2 indicates at least once a week, and three indicates almost every day (see Supplementary Table 7).

A linear regression model showed that parent interference behaviors decreased with children's age ($ß_{Std.} = -0.388$, 95% CI [−0.838, 0.062], $p = 0.09$), though the correlation was only marginally significant (Supplementary Figure 5). In addition, parent interference behaviors significantly decreased with children's device use ($ß_{Std.} = -1.54$, 95% CI [−2.72, −0.36], $p = 0.012$) (Supplementary Figure 6). A linear regression model also showed that parent interference behaviors decreased with children's performance on MERLS, but the correlation did not reach significance ($ß_{Std.} = -2.90$, 95% CI [−6.96, 1.17], $p = 0.158$) (Supplementary Figure 7). These findings suggest that as children grow older or demonstrate greater capability in completing the language tasks, parents tend to interfere less during the assessment. This pattern may reflect parents adapting their behaviors to match their children's increasing language abilities.

4.3 Study 2 discussion

In this study, by exploring parent-led administered bilingual assessment using the MERLS, we aimed to investigate the feasibility of using parents as citizen scientists to support test administration with their bilingual children in lieu of a researcher/clinician. Correlation analysis of test-retest reliability and item-by-item agreement of parent-administered vs. researcher-administered sessions revealed good to excellent reliability comparable to gold standard clinical assessment and also aligns with prior research suggesting that parents can be trained to assist in developmental and language screenings under supervision (e.g., Crais et al., 2006; Roberts and Kaiser, 2011). The primary difference between Study 1 (in-person assessment) and Study 2 (telehealth assessment) was the addition of an instructional video in Study 2 for parental adherence improvement. While Study 1 involved an in-person setting with researchers present, Study 2 adopted a telehealth format where parents administered the MERLS assessment virtually. The language tasks and test objectives of both studies were identical, which allows a direct evaluation of the video intervention's impact on parents' behaviors and parent-child interaction patterns especially during a telehealth context. However, we acknowledge that the observed reduction in interference behaviors cannot be attributed solely to the instructional video, as the change in modality (in-person vs. telehealth) and the passage of time between studies may have also influenced results. The study findings not only demonstrate the potential of citizen science approach to gather large-scale speech assessment data for establishing robust bilingual language norms but also provide insights on the utility of telehealth service delivery to overcome geographical and access barriers.

While most previous studies explored the possibility of involving parents in in-person test administration, our work highlights the feasibility of telehealth formats, where assessment can occur remotely to circumvent the geographical barriers. This format could help parents monitor their children's progress independently and seek professional support when necessary. However, it is important to clarify that in the current telehealth format, test administration remains supported by researchers or healthcare clinicians, ensuring adherence to protocols and addressing technical challenges. Future adaptations to MERLS could explore fully autonomous parent-administered assessments.

4.3.1 Improvement of parent adherence from Study 1 to Study 2

This study also demonstrated that brief instructional interventions can significantly improve parent adherence to test protocols, which lays the foundation for recruiting caregivers as test administrators in telehealth language assessments. In study 1, parents' ability to administer the assessment was evaluated against professional examiners in a supervised, in-person context. In Study 2, parents administered the test virtually after watching an instructional video, an intervention that improved adherence while reducing interference behaviors. Moreover, our detailed qualitative analysis revealed specific parent behavioral patterns of verbal and nonverbal support during bilingual telehealth assessments.

Our results showed that all four types of parent interference behaviors (Repeating Questions, Answering Questions, Analyzing Items, Judging of Correctness) decreased significantly in Study 2 compared to Study 1, demonstrating the effectiveness of the instructional video intervention method. Notably, parent supporting behaviors, such as encouragement, remained stable, indicating that parents continued to provide motivational scaffolding without intruding on the child's task performance. This balance, which preserves children's task independence while allowing parents to offer appropriate affective support, is important for the validity of language assessment conducted in non-clinical settings. These findings provided robust statistical evidence for the feasibility of the citizen science approach, demonstrating that caregivers can be systematically trained to administer assessments and thereby contribute to large-scale collection of bilingual language data.

4.3.2 Scaffolding and language preferences in bilingual parent-child interaction

In-depth interaction analysis revealed how parent support behaviors interacted with the child during bilingual assessment. Children's increased verbal interactions in Mandarin often stemmed from requests for technical assistance or clarification of meaning, leading parents to provide tailored verbal cues, including technical guidance, encouragement to re-engage, and clarifications. These dynamics highlight the adaptive role parents play as facilitators, adjusting their strategies based on the child's needs. Beyond the overall adherence findings, the analysis of parent-child interaction modes revealed a strong communication preference for Mandarin among ME-speaking children during the telehealth assessment, with children exhibiting significantly more verbal utterances (characterized by statistical significance in CU) especially in Mandarin sessions compared to English sessions. This preference likely stems from the cognitive, cultural, and emotional connections children have with their heritage language (Cummins, 2001; Levey and Polirstok, 2010). Parents play a key role in reinforcing this dynamic by offering greater scaffolding and support in their shared primary language (Yeh, 2019), as evidenced by the qualitative findings of adaptive parental involvement during Mandarin sessions. This preference underscores a crucial consideration for designing technology-mediated parent-child interaction systems. Children's linguistic comfort zones can significantly impact their engagement and performance during structured interactions (Puckett et al., 2009). For bilingual families, platforms should not only support multiple languages but also adapt dynamically to children's language preferences and communication styles (Verhagen et al., 2022; Hoff et al., 2012). Incorporating features like "awareness display" (Gao et al., 2015), which include culturally contextualized prompts, adaptive language scaffolding, and sensitivity to family communication styles, could help ensure the accuracy of telehealth assessments. These features would also enhance the effectiveness of child-focused technologies, supporting children's development, especially for those from diverse linguistic backgrounds.

4.3.3 Contextual variables influencing parent interference behaviors

Children's behaviors and performance often act as mediators of parent interference behaviors. Previous research has found that parent and children's behaviors coregulate and reflect moment-to-moment coordination of goal-oriented behaviors (Calkins, 2010). To investigate the effects of children's behaviors on parent interference behaviors during the test, we examined three variables: age, device use, and test performance. This study includes a wide age range (3–10 years), which encompasses several distinct developmental stages. Cognitive and behavioral characteristics vary significantly across this span (Diamond, 2002). For instance, younger children (e.g., 3–5 years old) typically have shorter attention spans and less experience with digital interfaces (McClelland et al., 2006; Mahone and Schneider, 2012), while older, school-aged children (e.g., 6–10 years old) generally possess greater task autonomy and digital literacy (Liu et al., 2024). As expected, parent interference behaviors decreased as children grew older, likely due to increased linguistic competence and independence. Interestingly, parent interference behaviors also significantly decreased when children spent more time on digital devices, further supported by our qualitative observations that familiarity with the assessment interface reduced the need for direct parental technical assistance, thereby playing a key role in reducing parent involvement. This finding aligns with studies on digital literacy, which emphasize the role of child familiarity in reducing reliance on parental assistance (Neumann and Neumann, 2017). Lastly, while children's test performance was negatively associated with parent interference behaviors, the relationship does not reach statistical significance. Nonetheless, the trend suggests that parents might adapt their behaviors to children's language capabilities, which is an important consideration for designing scalable, parent-led assessments.

5 Study limitations

This research has several limitations which could be addressed in future research. First, the data for the two groups were collected in different modalities: Study 1 was conducted in-person, and Study 2 was conducted virtually via Zoom. This change in testing modality introduces potential confounding factors. While our detailed video analysis in Study 2 aimed to capture both verbal and non-verbal interactions, the virtual setting and limited camera angle inherently constrained the complete observation of all non-verbal parent behaviors occurring outside the video frame. Therefore, it is important to acknowledge that some non-verbal interference behaviors in Study 2 might have gone uncaptured, potentially influencing the observed reduction in overall interference. Future studies should mitigate this limitation by using self-recording devices or multiple camera angles to capture a holistic view of the testing environment. Second, Study 1 had design and data collection shortcomings. Specifically, parents chose the order of completing the English and Mandarin modules based on their preferences, rather than through random assignment. A counterbalanced design, where module order is systematically alternated, would improve the study's internal validity. Third, our sample primarily consisted of middle-class families with highly educated parents, which limits the generalizability of our findings to families from diverse socioeconomic and cultural backgrounds. Socioeconomic factors may influence parent digital literacy, access to reliable internet and devices, availability of quiet testing environments, and cultural beliefs about parent roles in formal assessment. Future studies should prioritize recruitment from diverse socioeconomic groups to evaluate whether parent adherence patterns vary across demographics. Community partnerships with Title I schools, community health centers, and immigrant service organizations may help achieve more representative sampling.

6 Conclusion

Traditional approaches to language assessment face several challenges, including a shortage of bilingual SLPs, limited

availability of bilingual language assessment tools, and insufficient development of bilingual language norms. This study demonstrates how citizen science, an underutilized data collection method, can expand the current assessment paradigm by positioning parents as active contributors to research and service delivery. Beyond its immediate implications for ME assessments, this approach highlights how collaborative, family-centered methods can reshape the way child language data are collected, diversify research samples, and accelerate the development of more equitable assessment tools across languages. The online MERLS test is equipped with an automated scoring system based on prior research (Gale et al., 2021), simplifying the process for parent-child dyads to access via standard telehealth equipment (e.g., such as laptops, videoconferencing software). Our findings from both studies suggested technical design insights to improve parental adherence, and identified qualitative insights regarding contextual factors (e.g., more supportive behaviors in Mandarin session, and more child's utterances following parents' behaviors) particularly in the dominant home language observed from parent-child interaction, adding more recommendations for future language assessment development.

Additionally, this study evaluates the feasibility of telehealth assessment to increase access to bilingual SLPs for ME-speaking children, while also can benefit a larger group of monolingual clinicians to administer tests to bilingual children, addressing the unique service needs of bilingual assessment with children from diverse social cultural backgrounds in the speech-language field (Hyter and Salas-Provance, 2019; De Lamo White and Jin, 2011). By establishing feasibility of remote, parent-administered bilingual assessment, this work provides a foundation for developing scalable approaches to identify language disorders in bilingual children, a population with significant special educational needs arising from systemic assessment barriers and limited access to bilingual clinical services. Future work should continue to examine the utility of the citizen science approach to accommodate a wider range of parent-child profiles across socioeconomic status, geographic location, and cultural backgrounds for more complex language assessment tasks, to ensure comprehensive child language assessment practices and inclusivity and equity during the assessment practice. Large-scale validation studies are needed to compare MERLS outcomes from parent-administered sessions with gold-standard diagnoses made by qualified bilingual speech-language pathologists in real-world clinical settings. Sensitivity, specificity, positive and negative predictive values must be established across different age groups and language proficiency levels. Lastly, more implementation research should examine the scalability of this approach, including cost-effectiveness analyses, integration into clinical workflows, parent satisfaction and retention over time, and quality assurance mechanisms for maintaining data integrity at scale. Only through such comprehensive validation can we move from a promising feasibility study to a clinically viable assessment tool that improves access to equitable language services for bilingual 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 humans were approved by University of Delaware Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

YD: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. YT: Formal analysis, Visualization, Writing – original draft, Writing – review & editing. KF: Formal analysis, Writing – original draft, Writing – review & editing. YL: Formal analysis, Writing – original draft, Writing – review & editing. DW: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing. XT: Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. YL: Formal analysis, Writing – original draft, Writing – review & editing. QZ: Formal analysis, Writing – original draft, Writing – review & editing. SQ: Formal analysis, Writing – original draft, Writing – review & editing. JX: Supervision, Writing – review & editing, Validation. LS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by the Spencer Foundation Research Grant (Small) and ASHA Multicultural Activities Grant (#015703-00001) for funding the development of the platform, research assistantship, and human subject compensation for the data collected.

Acknowledgments

We acknowledge the contributions of the following individuals for their efforts on data collection and analysis of this project: Elena Hu, Sharon Hollenbach, Caila Walsh, Ganya Luo. We especially appreciate the following individuals for their generous feedback for this manuscript: Joseph Hin Yan Lam, Yannan Li, and Gedeon Deák.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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/feduc.2025.1696031/full#supplementary-material

References

Alnemary, F. M., Wallace, M., Symon, J. B., and Barry, L. M. (2015). Using international videoconferencing to provide staff training on functional behavioral assessment. *Behav. Intervent.* 30, 73–86.

American Speech-Language-Hearing Association [ASHA] (2025a). *Multilingual service providers*. Rockville, MD: ASHA.

American Speech-Language-Hearing Association [ASHA] (2025b). *Profile of ASHA multilingual service providers, Year-End 2024*. Rockville, MD: ASHA.

Barkaia, A., Stokes, T. F., and Mikiashvili, T. (2017). Intercontinental telehealth coaching of therapists to improve verbalizations by children with autism. *J. Appl. Behav. Anal.* 50, 582–589.

Bhattacharjee, Y. (2005). Citizen scientists supplement work of Cornell researchers. *Science* 308, 1402–1403. doi: 10.1126/science.308.5727.1402

Bishara, A. J., and Hittner, J. B. (2012). Testing the significance of a correlation with nonnormal data: Comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychol. Methods* 17, 399–417. doi: 10.1037/a0028087

Boerma, T., and Blom, E. (2017). Assessment of bilingual children: What if testing both languages is not possible? *J. Commun. Disord.* 66, 65–76. doi: 10.1016/j.jcomdis.2017.04.001

Bonney, R., Phillips, T. B., Ballard, H. L., and Enck, J. W. (2016). Can citizen science enhance public understanding of science? *Public Understand. Sci.* 25, 2–16. doi: 10.1177/0963662515607406

Brandone, A. C., Golinkoff, R. M., and Hirsh-Pasek, K. (2008). Feasibility of computer-administered language assessment. *Perspect. School-Based Issues* 9, 57–65. doi: 10.1044/sbi9.2.57

Brown, J. A., and Woods, J. J. (2015). Effects of a triadic parent-implemented home-based communication intervention for toddlers. *J. Early Intervent.* 37, 44–68. doi: 10.1177/1053815115589350

Caesar, L. G., and Kohler, P. D. (2007). The state of school-based bilingual assessment: Actual practice versus recommended guidelines. *Lang. Speech Hear. Services Schools* 38, 190–200. doi: 10.1044/0161-1461(2007/020)

Calkins, S. D. (2010). "Caregiving as coregulation: Psychobiological processes and child functioning," in *Biosocial foundations of family processes*, eds A. Booth, S. M. McHale, and N. S. Landale (New York, NY: Springer New York), 49–59.

Castilla-Earls, A., Bedore, L., Rojas, R., Fabiano-Smith, L., Pruitt-Lord, S., Restrepo, M. A., et al. (2020). Beyond scores: Using converging evidence to determine speech and language services eligibility for dual language learners. *Am. J. Speech-Lang. Pathol.* 29, 1116–1132. doi: 10.1044/2020_AJSLP-19-00179

Ciccia, A. H., Whitford, B., Krumm, M., and McNeal, K. (2011). Improving the access of young urban children to speech, language and hearing screening via telehealth. *J. Telemed. Telecare* 17, 240–244. doi: 10.1258/jtt.2011.100810

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*, 3rd Edn. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Cooc, N. (2018). Examining the underrepresentation of Asian Americans in special education: New trends from California school districts. *Exceptionality* 26, 1–19. doi: 10.1080/09362835.2016.1216847

Corona, L. L., Weitlauf, A. S., Hine, J., Berman, A., Miceli, A., Nicholson, A., et al. (2021). Parent perceptions of caregiver-mediated telemedicine tools for assessing autism risk in toddlers. *J. Autism Dev. Disord.* 51, 476–486. doi: 10.1007/s10803-020-04554-9

Cowan, T., Paroby, C., Leibold, L. J., Buss, E., Rodriguez, B., and Calandruccio, L. (2022). Masked-speech recognition for linguistically diverse populations: A focused review and suggestions for the future. *J. Speech Lang. Hear. Res.* 65, 3195–3216. doi: 10.1044/2022_JSLHR-22-00011

Crais, E. R., Roy, V. P., and Free, K. (2006). Parents' and professionals' perceptions of the implementation of family-centered practices in child assessments. *Am. J. Speech Lang. Pathol.* 15, 365–377. doi: 10.1044/1058-0360(2006/034)

Cummins, J. (2001). Bilingual children's mother tongue: Why is it important for education? *Sprogforum* 7, 15–20.

De Lamo White, C., and Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *Intern. J. Lang. Commun. Disord.* 46, 613–627. doi: 10.1111/j.1460-6984.2011.00049.x

Dekhtyar, M., Braun, E. J., Billot, A., Foo, L., and Kiran, S. (2020). Videoconference administration of the western aphasia battery-revised: Feasibility and validity. *Am. J. Speech-Lang. Pathol.* 29, 673–687. doi: 10.1044/2019_AJSLP-19-00023

Diamond, A. (2002). "Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry," in *Principles of frontal lobe function*, eds D. T. Stuss and R. T. Knight (Oxford: Oxford University Press), 466–503.

Dodge-Chin, C., Shigetomi-Toyama, S., and Quinn, E. D. (2022). Teaching parents read, ask, answer, prompt strategies via telepractice: Effects on parent strategy use and child communication. *Lang. Speech Hear. Services Schools* 53, 237–255. doi: 10.1044/2021_LSHSS-21-00075

Dollaghan, C. A., and Horner, E. A. (2011). Bilingual language assessment: A meta-analysis of diagnostic accuracy. *J. Speech Lang. Hear. Res.* 54, 1077–1088. doi: 10.1044/1092-4388(2010/10-0093)

Du, Y., Liu, Y., Tang, Y., Fong, K. K., and Sheng, L. (2021). "Parental adherence to test protocols: Exploratory study on a remote, web-based Mandarin-English receptive language assessment," in *Proceedings of the 30-minute technical research oral presentation accepted at the 2021 American Speech-Language and Hearing Association Annual Convention*, (Washington, DC).

Du, Y., Sheng, L., and Tekinbas, K. S. (2020). "Try your best": Parent behaviors during administration of an online language assessment tool for bilingual mandarin-english children," in *Proceedings of the Interaction Design and Children Conference. Association for Computing Machinery*, (New York, NY), 409–420. doi: 10.1145/3392063.3394441

Dyvik, E. H. (2024). *The most spoken languages worldwide 2023*. Germany: Statista.com.

Farmani, E., Fekar Gharamaleki, F., and Nazari, M. A. (2024). Challenges and opportunities of tele-speech therapy: Before and during the COVID-19 pandemic. *J. Public Health Res.* 13:22799036231222115. doi: 10.1177/22799036231222115

Farmer, R. L., McGill, R. J., Dombrowski, S. C., McClain, M. B., Harris, B., Lockwood, A. B., et al. (2020). Teleassessment with children and adolescents during

the coronavirus (COVID-19) pandemic and beyond: Practice and policy implications. *Prof. Psychol. Res. Pract.* 51, 477–487. doi: 10.1037/pro0000349

Flores, G., and Tomany-Korman, S. C. (2008). The language spoken at home and disparities in medical and dental health, access to care, and use of services in US children. *Pediatrics* 121:e1703-14.

Fraisl, D., Hager, G., Bedessem, B., Gold, M., Hsing, P. Y., Danielsen, F., et al. (2022). Citizen science in environmental and ecological sciences. *Nat. Rev. Methods Prim.* 2:64. doi: 10.1038/s43586-022-00144-4

Freeman, M. R., and Schroeder, S. R. (2022). Assessing language skills in elementary-aged bilingual children: Current trends in research and practice. *J. Child Sci.* 12, e33–e46. doi: 10.1055/s-0042-1743575

Frigerio, P., Monte, L. D., Sotgiu, A., De Giacomo, C., and Vignoli, A. (2021). Parents' satisfaction of tele-rehabilitation for children with neurodevelopmental disabilities during the covid-19 pandemic. *BMC Fam. Pract.* 23, 146–146. doi: 10.1186/s12875-022-01747-2

Gale, R., Bird, J., Wang, Y., van Santen, J., Prud'hommeaux, E., Dolata, J., et al. (2021). Automated scoring of tablet-administered expressive language tests. *Front. Psychol.* 12:668401. doi: 10.3389/fpsyg.2021.668401

Gao, G., Yamashita, N., Hautasaari, A. M., and Fussell, S. R. (2015). "Improving multilingual collaboration by displaying how non-native speakers use automated transcripts and bilingual dictionaries," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, (New York, NY: ACM), 3463–3472.

Ghasemi, A., and Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *Intern. J. Endocrinol. Metab.* 10, 486–489. doi: 10.5812/ijem.3505

Gillam, R. B., Peña, E. D., Bedore, L. M., Bohman, T. M., and Mendez-Perez, A. (2013). Identification of specific language impairment in bilingual children: Assessment in English. *J. Speech Lang. Hear. Res.* 56, 1813–1823. doi: 10.1044/1092-4388(2013/12-0056)

Golinkoff, R. M., De Villiers, J., Hirsh-Pasek, K., Iglesias, A., Wilson, M. S., Morini, G., et al. (2017). *User's manual for the quick interactive Language Screener (QUILS): A measure of vocabulary, syntax, and language acquisition skills in young children*. Baltimore, MD: Paul H. Brookes Publishing Co.

Gosling, S. D., Vazire, S., Srivastava, S., and John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *Am. Psychol.* 59, 93–104. doi: 10.1037/0003-066X.59.2.93

Gov.UK (2020). *English proficiency: Pupils with English as an additional language*. Washington, D.C: Department of Education.

Grenoble, L. A., and Osipov, B. (2023). The dynamics of bilingualism in language shift ecologies. *Ling. Approaches Biling.* 13, 1–39. doi: 10.1075/lab.22035.gre

Grillo, E. U. (2021). Functional voice assessment and therapy methods supported by telepractice, Voiceevalu8, and Etill voice training. *Sem. Speech Lang.* 42, 41–53. doi: 10.1055/s-0040-1722753

Grimm, A., and Schulz, P. (2014). Specific language impairment and early second language acquisition: The risk of over- and underdiagnosis. *Child Indicators Res.* 7, 821–841. doi: 10.1007/s12187-013-9230-6

Hay-Hansson, A. W., and Eldevik, S. (2013). Training discrete trials teaching skills using videoconference. *Res. Autism Spectrum Disord.* 7, 1300–1309.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., and Parra, M. (2012). Dual language exposure and early bilingual development. *J. Child Lang.* 39, 1–27. doi: 10.1017/S0305000910000759

Hu, S., Gavarró, A., Vernice, M., and Guasti, M. T. (2016). The acquisition of Chinese relative clauses: Contrasting two theoretical approaches. *J. Child Lang.* 43, 1–21. doi: 10.1017/S0305000914000865

Hyter, Y. D. (2021). "The power of words: A preliminary critical analysis of concepts used in speech, language, and hearing sciences," in *Critical perspectives on social justice in speech-language pathology*, ed. R. Horton (Hershey, PA: IGI).

Hyter, Y. D., and Salas-Provance, M. B. (2019). *Culturally responsive practices in speech, language, and hearing sciences*. San Diego, CA: Plural Publishing Inc.

Jasso, J., McMillen, S., Anaya, J. B., Bedore, L. M., and Peña, E. D. (2020). The utility of an English semantics measure for identifying developmental language disorder in Spanish–English bilinguals. *Am. J. Speech-Lang. Pathol.* 29, 776–788. doi: 10.1044/2020_AJSLP-19-00202

Jia, G., and Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Appl. Psychol.* 24, 131–161. doi: 10.1017/S0142716403000079

Jordan, B., and Henderson, A. (1995). Interaction analysis: Foundations and practice. *J. Learn. Sci.* 4, 39–103. doi: 10.1207/s15327809jls0401_2

Kan, P. F., Huang, S., Winicour, E., and Yang, J. (2020). Vocabulary growth: Dual language learners at risk for language impairment. *Am. J. Speech-Lang. Pathol.* 29, 1178–1195. doi: 10.1044/2020_AJSLP-19-00160

Kelders, S., Kok, R., and Van Gemert-Pijnen, J. (2011). "Technology and adherence in web-based interventions for weight control: A systematic review," in *Proceedings of the ACM international conference proceeding series*, (New York, NY: ACM), doi: 10.1145/2467803.2467806

Khoshsima, H., and Toroujeni, S. M. H. (2017). Transitioning to an alternative assessment: Computer-based testing and key factors related to testing mode. *Eur. J. English Lang. Teach.* 2, 54–73. doi: 10.5281/zenodo.268576

Kimble, C. (2013). Speech-language pathologists' comfort levels in English language learner service delivery. *Commun. Disord. Quar.* 35, 21–27. doi: 10.1177/1525740113487404

Klatte, I. S., Lyons, R., Davies, K., Harding, S., Marshall, J., McKean, C., et al. (2020). Collaboration between parents and SLTs produces optimal outcomes for children attending speech and language therapy: Gathering the evidence. *Intern. J. Lang. Commun. Disord.* 55, 618–628. doi: 10.1111/1460-6984.12538

Kritikos, E. P. (2003). Speech-language pathologists' beliefs about language assessment of bilingual/bicultural individuals. *Am. J. Speech-Lang. Pathol.* 12, 73–91. doi: 10.1044/1058-0360(2003/054)

Langdon, H. W., and Quintanar-Sarellana, R. (2003). Roles and responsibilities of the interpreter in interactions with speech-language pathologists, parents, and students. *Sem. Speech Lang.* 24, 235–244. doi: 10.1055/s-2003-42826

Lehner, K., Pfab, J., and Ziegler, W. (2021). Web-based assessment of communication-related parameters in dysarthria: Development and implementation of the Kommpas web app. *Clin. Ling. Phonet.* 36, 1093–1111. doi: 10.1080/02699206.2021.1989490

Lesso, R. (2023). *Which countries and regions speak Mandarin Chinese?*. Montreal, QC: TheCollector.

Levey, S., and Polirstok, S. (2010). *Language development: Understanding language diversity in the classroom*. Thousand Oaks, CA: Sage Publications.

Li, W., Germine, L. T., Mehr, S. A., Srinivasan, M., and Hartshorne, J. (2024). Developmental psychologists should adopt citizen science to improve generalization and reproducibility. *Infant Child Dev.* 33:e2348. doi: 10.1002/icd.2348

Liu, S., Reynolds, B. L., Thomas, N., and Soyoof, A. (2024). The use of digital technologies to develop young children's language and literacy skills: A systematic review. *SAGE Open* 14:21582440241230850. doi: 10.1177/21582440241230850

Magasi, S., Harniss, M., and Heinemann, A. W. (2018). Interdisciplinary approach to the development of accessible computer-administered measurement instruments. *Arch. Phys. Med. Rehabil.* 99, 204–210. doi: 10.1016/j.apmr.2017.06.036

Mahone, E. M., and Schneider, H. E. (2012). Assessment of attention in preschoolers. *Neuropsychol. Rev.* 22, 361–383. doi: 10.1007/s11065-012-9217-y

Marchman, V. A., Fernald, A., and Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *J. Child Lang.* 37, 817–840. doi: 10.1017/S0305000909990055

Marhefka, S., Lockhart, E., and Turner, D. (2020). Achieve research continuity during social distancing by rapidly implementing individual and group videoconferencing with participants: Key considerations, best practices, and protocols. *AIDS Behav.* 24, 1983–1989. doi: 10.1007/s10461-020-02837-x

McClelland, M. M., Acock, A. C., and Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Res. Quar.* 21, 471–490. doi: 10.1016/j.ecresq.2006.09.003

McCrae, C. S., Chan, W. S., Curtis, A. F., Nair, N., Deroche, C. B., Munoz, M., et al. (2021). Telehealth cognitive behavioral therapy for insomnia in children with autism spectrum disorder: A pilot examining feasibility, satisfaction, and preliminary findings. *Autism: Intern. J. Res. Pract.* 25, 667–680. doi: 10.1177/1362361320949078

McLeod, S., and Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *Am. J. Speech-Lang. Pathol.* 27, 1546–1571. doi: 10.1044/2018_AJSLP-17-0100

Molini-Avejonas, D. R., Rondon-Melo, S., Amato, C. A., and Samelli, A. G. (2015). A systematic review of the use of telehealth in speech, language, and hearing sciences. *J. Telemed. Telecare* 21, 367–376. doi: 10.1177/1357633x15583215

National Academies of Sciences, Engineering, and Medicine (2018). *Improving Health Research on Small Populations*. Washington, D.C: National Academies Press, doi: 10.17226/25112

Neumann, M. M., and Neumann, D. L. (2017). The use of touch-screen tablets at home and pre-school to foster emergent literacy. *J. Early Childhood Liter.* 17, 203–220. doi: 10.1177/1468798415619773

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* 57, 1247–1257. doi: 10.1111/jcpp.12573

Oetting, J. B. (2018). Prologue: Toward accurate identification of developmental language disorder within linguistically diverse schools. *Lang. Speech Hear. Services Schools* 49, 213–217. doi: 10.1044/2018_LSHSS-CLSLD-17-0156

Paradis, J. (2011). Individual differences in child English second language acquisition comparing child-internal and child-external factors. *Ling. Approaches Biling.* 1, 213–237. doi: 10.1075/lab.1.3.01par

Park, M., O'Toole, A., and Katsiaficas, C. (2017). *Dual language learners: A national demographic and policy profile*. Washington, DC: Migration Policy Institute.

Patterson, J. L., and Pearson, B. Z. (2004). "Bilingual lexical development: Influences, contexts, and processes," in *Bilingual language development and disorders in Spanish-English speakers*, ed. B. A. Goldstein (Baltimore MD: Paul Brookes), 77–104.

Pearson Education Inc. (2020). *In remote assessment set up*. London: Pearson Education Inc.

Pearson, B. Z. (2013). "Distinguishing the bilingual as a late talker from the later talker who is bilingual," in *Late talkers: Language development, interventions, and outcomes*, eds L. Rescorla and P. Dale (Baltimore MD: Paul Brookes), 67–87.

Peña, E. D., and Sutherland, R. (2022). Can you see my screen? Virtual assessment in speech and language. *Lang. Speech Hear. Services Schools* 53, 329–334. doi: 10.1044/2022_LSHSS-22-00007

Peña, E. D., Gillam, R. B., and Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *J. Speech Lang. Hear. Res.* 57, 2208–2220. doi: 10.1044/2014_JSLHR-L-13-0151

Perrin, P. B., Rybarczyk, B. D., Pierce, B. S., Jones, H. A., Shaffer, C., and Islam, L. (2020). Rapid telepsychology deployment during the COVID-19 pandemic: A special issue commentary and lessons from primary care psychology training. *J. Clin. Psychol.* 76, 1173–1185. doi: 10.1002/jclp.22969

Puckett, M. B., Black, J. K., Wittmer, D. S., and Peterson, S. H. (2009). *The young child*. Upper Saddle River, NJ: Pearson.

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Murray, P., Schawinski, K., et al. (2009). Galaxy zoo: Exploring the motivations of citizen science volunteers. *arXiv preprint* arXiv:0909.2925

Radville, K. M., Larrivee, E. C., Baron, L. S., Kelley-Nazzaro, P., and Christodoulou, J. A. (2022). Online training modules for teaching assessment skills to graduate student clinicians. *Lang. Speech Hear. Services Schools* 53, 417–430. doi: 10.1044/2021_LSHSS-21-00068

Reinecke, K., and Gajos, K. Z. (2015). "Labinthewild: Conducting large-scale online experiments with uncompensated samples," in *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, (New York, NY: ACM), 1364–1378.

Roberts, M. Y., and Kaiser, A. P. (2011). The effectiveness of parent-implemented language interventions: A meta-analysis. *Am. J. Speech Lang. Pathol.* 20, 180–199. doi: 10.1044/1058-0360(2011/10-0055)

Ryan, C. L. (2013). *Language use in the United States: 2011*. Washington, DC: U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau.

Samson, J. F., and Lesaux, N. K. (2009). Language-minority learners in special education: Rates and predictors of identification for services. *J. Learn. Disabil.* 42, 148–162. doi: 10.1177/0022219408326221

Schmitt, M. B., Tambyraja, S., Thibodeaux, M., and Filipkowski, J. (2022). Feasibility of assessing expressive and receptive vocabulary via telepractice for early elementary-age children with language impairment. *Lang. Speech Hear. Services Schools* 53, 445–453. doi: 10.1044/2021_LSHSS-21-00057

Schmitz, H., Howe, C. L., Armstrong, D. G., and Subbian, V. (2018). Leveraging mobile health applications for biomedical research and citizen science: A scoping review. *J. Am. Med. Inform. Assoc.* 25, 1685–1695. doi: 10.1093/jamia/ocy130.

Scott, K., Chu, J., and Schulz, L. (2017). Lookit (Part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind* 1, 15–29. doi: 10.1162/OPMI_a_00001

Shankar, V., Ramkumar, V., and Kumar, S. (2022). Understanding the implementation of telepractice in speech and language services using a mixed-methods approach. *Wellcome Open Res.* 7:46. doi: 10.12688/wellcomeopenres.17622.2

Sheng, L. (2018). "Typical and atypical lexical development," in *Handbook of communication disorders: Theoretical, empirical, and applied linguistic perspectives*, 101–116.

Sheng, L., Lam, B. P. W., Cruz, D., and Fulton, A. (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *J. Exp. Child Psychol.* 141, 229–238.

Sheng, L. I., Lu, Y., and Kan, P. F. (2011). Lexical development in Mandarin-English bilingual children. *Bilingualism Lang. Cogn.* 14, 579–587.

Sheng, L., Wang, D., Walsh, C., Heisler, L., Li, X., and Su, P. L. (2021). The bilingual home language boost through the lens of the COVID-19 pandemic. *Front. Psychol.* 12:667836. doi: 10.3389/fpsyg.2021.667836

Solano-Flores, G., Chia, M., and Kachchaf, R. (2019). Design and use of pop-up illustration glossaries as accessibility resources for second language learners in computer-administered tests in a large-scale assessment system. *Intern. Multiling. Res. J.* 13, 277–293. doi: 10.1080/19313152.2019.1611338

Song, L., Luo, R., and Liang, E. (2021). Dual language development of Chinese 3- and 4-year-olds: Associations with the family context and teachers' language use. *Early Educ. Dev.* 33, 219–242. doi: 10.1080/10409289.2020.1865746

Song, L., Sheng, L., and Luo, R. (2022). Comprehension skills of Chinese-English dual language learners: Relations across languages and associations with language richness at home. *Intern. J. Biling. Educ. Biling.* 27, 19–37. doi: 10.1080/13670050.2022.2137386

Sullivan, G. M. (2011). A primer on the validity of assessment instruments. *J. Graduate Med. Educ.* 3, 119–120. doi: 10.4300/JGME-D-11-00075.1

Sutherland, R., Hodge, A., Chan, E., and Silove, N. (2021). Barriers and facilitators: Clinicians' opinions and experiences of telehealth before and after their use of a telehealth platform for child language assessment. *Intern. J. Lang. Commun. Disord.* 56, 1263–1277. doi: 10.1111/1460-6984.12666

Talbott, M. R., Dufek, S., Zwaigenbaum, L., Bryson, S., Brian, J., Smith, I. M., et al. (2020). Brief report: Preliminary feasibility of the TEDI: A novel parent-administered telehealth assessment for autism spectrum disorder symptoms in the first year of life. *J. Autism Dev. Disord.* 50, 3432–3439. doi: 10.1007/s10803-019-04314-4

Tang, Y., Triesch, J., and Deák, G. O. (2023). Variability in infant social responsiveness: Age and situational differences in attention-following. *Dev. Cogn. Neurosci*. 63:101283.

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

Tomlinson, S. R. L., Gore, N., and McGill, P. (2018). Training individuals to implement applied behavior analytic procedures via telehealth: A systematic review of the literature. *J. Behav. Educ.* 27, 172–222. doi: 10.1007/s10864-018-9292-0

U.S. Census Bureau (2024). *Language spoken at home in S1601*. Washington, DC: U.S. Department of Commerce.

Verhagen, J., Kuiken, F., and Andringa, S. (2022). Family language patterns in bilingual families and relationships with children's language outcomes. *Appl. Psycholing.* 43, 1109–1139. doi: 10.1017/S0142716422000297

Waite, M. C., Theodoros, D. G., Russell, T. G., and Cahill, L. M. (2010). Internet-based telehealth assessment of language using the CELF-4. *Lang. Speech Hear. Services Schools* 41, 445–458. doi: 10.1044/0161-1461(2009/08-0131)

Wang, D., Zheng, L., Lin, Y., Zhang, Y., and Sheng, L. (2022). Sentence repetition as a clinical marker for mandarin-speaking preschoolers with developmental language disorder. *J. Speech Lang. Hear. Res. JSLHR* 65, 1543–1560. doi: 10.1044/2021_JSLHR-21-00401

Werfe, K. L., Grey, B., Johnson, M., Brooks, M., Cooper, E., Reynolds, G., et al. (2021). Transitioning speech-language assessment to a virtual environment: Lessons learned from the ELLA Study. *Lang. Speech Hear. Services Schools* 52, 769–775. doi: 10.1044/2021_LSHSS-20-00149

Westerveld, M. F. (2014). Emergent literacy performance across two languages: Assessing four-year-old bilingual children. *Intern. J. Biling. Educ. Biling.* 17, 526–543. doi: 10.1080/13670050.2013.835302

Wong, A. M. Y., Leonard, L. B., Fletcher, P., and Stokes, S. F. (2004). Questions without movement. *J. Speech Lang. Hear. Res.* 47, 1440–1453. doi: 10.1044/1092-4388(2004/107)

Yeh, E. (2019). Parent Matters: The impact of parental involvement on non-native English speakers' postsecondary education enrollment. *School Commun. J.* 29, 39–62.

Yu, B., Epstein, L., and Tisi, V. (2021). "A DisCrit-informed critique of the difference vs. disorder approach in speech-language pathology," in *Critical Perspectives on Social Justice in Speech-Language Pathology*, ed. R. Horton (Hershey, PA: IGI).

Zhou, P., and Crain, S. (2011). Children's knowledge of the quantifier dou in Mandarin Chinese. *J. Psycholing. Res.* 40, 155–176. doi: 10.1007/s10936-010-9161-z

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 30 January 2026
DOI 10.3389/feduc.2025.1641126

OPEN ACCESS

EDITED BY
Yuan Chen,
The Education University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Nirmala Vasudevan,
Amrita Vishwa Vidyapeetham, India
Patricia Makaure,
Stellenbosch University, South Africa

*CORRESPONDENCE
Sharon McMurray
✉ sharon@literacydiffculties.com
Mary Morrow
✉ m.morrow@stran.ac.uk

RECEIVED 25 June 2025
REVISED 01 December 2025
ACCEPTED 16 December 2025
PUBLISHED 30 January 2026

CITATION
McMurray S, Morrow M and Bower P (2026) An examination of the orthographic and phonological spelling knowledge observed in a sample of independent writing completed by 267 children with specific literacy difficulties.
Front. Educ. 10:1641126.
doi: 10.3389/feduc.2025.1641126

An examination of the orthographic and phonological spelling knowledge observed in a sample of independent writing completed by 267 children with specific literacy difficulties

Sharon McMurray *, Mary Morrow* and Paula Bower

SEN Literacy Unit, Stranmillis University College, Belfast, United Kingdom

This paper considers spelling in samples of writing collected in October 2019 (pre COVID-19) from 267 children in the 8–9 age range in 143 mainstream primary schools who were identified by their schools as presenting with the most severe specific literacy difficulties in their age group. They were referred to the Northern Ireland Education Authority Psychology Service for assessment and were formally assessed to provide standardized scores for literacy attainment and cognitive profile. They presented with a cognitive profile which included a standardized score of 90 or above in one or more of the subtests of the Wechsler (2016) Intelligence Scale for Children (WISC). Spelling in independent writing samples was analysed to establish what sources of linguistic knowledge (phonemic, orthographic and morphemic) the children were drawing on to spell words. It was evident from the analysis of these writing samples that children were dependent on phoneme-to-grapheme correspondence when spelling, often selecting letters that did represent the phonemes but it was the wrong selection of letters for phonemes with multiple mappings. The observed pattern of errors indicates that these 267 children had difficulty developing orthographic knowledge resulting in phonologically plausible spelling choices impacting spelling accuracy. The development of orthographic knowledge was limited or had failed to develop for this group of children.

KEYWORDS

orthographic knowledge, orthographic mapping, orthography, phonics, spelling, orthographic processing

1 Introduction

Northern Ireland is one of the four distinct jurisdictions within the United Kingdom (UK), the others being England, Scotland and Wales. In Northern Ireland the primary approach for teaching phonics is a phoneme-to-grapheme correspondence only approach, an alphabetic phonics strategy, with systematic synthetic phonics (SSP) being widely used following the Rose Report (Rose, 2006).

Rose (2006) recommended SSP as the best way for children to learn to read, even though there was no robust research evidence to support this (Bowers, 2020). In SSP teaching must be explicit, structured, and sequential. Children are taught phonemes (smallest units of sound) and graphemes (the letter or letter combinations representing phonemes) and to use these to decode words by identifying and blending each phoneme all through a word to the exclusion of any other strategies. SSP teaches children to segment words into their individual phonemes to spell. Rose (2006) acknowledged the lack of research evidence and chose to consider evidence from practice observed in school inspections. "*...and notwithstanding the uncertainties of research, there is much convincing evidence to show from the practice observed that, as generally understood, 'synthetic' phonics is the form of systematic phonic work that offers the vast majority of beginners the best route to becoming skilled readers*" (p.19, paragraph 47).

By stating that this approach offers "*the vast majority*" of beginners the best route to becoming skilled readers Rose acknowledged that SSP was not going to meet the needs of all children. Yet it is currently mandated in England to the extent that children who are at risk of falling behind may be offered one-to-one support, but it must be using the same SSP programme (DfE, 2023, Note, 9).

Because SSP has been used for 20 years in England, we have evidence of the longer term impact of exclusively using SSP to teach beginning reading and spelling, as defined by the Department for Education (DfE) in England's core phonics criteria published in 2010 and updated in 2021. Publishers of phonics programmes must adhere to this criteria to obtain validation by the DfE for use in schools in England which Wyse and Hacking (2024) point out, forces publishers to meet this criteria. They consider the criteria to be problematic, "*not any particular synthetic phonics scheme*" (p54). Training in SSP was offered to all primary schools in Northern Ireland (NI) by the NI Education Authority from 2008–2012 following the Rose Review (Rose, 2006). The vast majority, but not all schools, availed of this training. In recent years many schools in NI have also availed of the training offered by the DfE validated SSP programme they are using. Phonics programmes available for purchase in NI are those published in England and validated by the DfE, thus impacting classroom practice. The points made in this document may not apply to SSP programmes developed in other countries that differ in content and the strategies used, to the extent that they would not meet the core phonics criteria prescribed by the DfE (2023). This should be taken into account when considering findings from research conducted in other education systems.

When SSP was first introduced by Rose (2006) the early improvements in reading were to be expected because the whole language approach to teaching reading, which preceded SSP, did not include a systematic approach to phonics teaching. From 2016 onward the extent of the difficulties experienced by up to 25% of children became increasingly evident. SSP focuses on the phoneme level of phonological awareness and the link to letters that represent phonemes. SSP does not address the onset and rime level, syllable level and whole word level of the sounds we can hear in spoken language, and the links to the orthographic patterns that represent these larger units of sound.

Teaching to develop children's phonological processing ability: the ability to identify, store, retrieve, and manipulate the sounds in spoken language, has been successful in remediating the phonological deficit and establishing phonological skills: the skills of segmenting, manipulating and deleting phonemes. Many children do, however, have problems with blending due to the working memory demands of the task. These phonological processing skills underpin phonics learning and support the formation of phoneme-to-grapheme links. However, phonological skills at phoneme level do not aid the decision making needed to select the correct letters when spelling words that include phonemes with multiple mappings (phonemes that can be spelled different ways). This requires orthographic knowledge of what the word looks like (McMurray, 2020). Orthographic knowledge is knowledge of spelling patterns that represent larger units of sound: those consisting of more than one phoneme, such as onsets, rimes, syllables, and whole words, and are recognized or recalled as orthographic units without the need to encode phoneme-by-phoneme. Orthographic knowledge is also identifiable in parts of words that cannot be identified by sound: for example, double letters or silent letters. Teaching phoneme-to-grapheme correspondence only, can result in difficulties with sight word recognition, reading fluency and spelling accuracy for a significant minority of children as identified by DfE statistics. SSP has impacted the development of reading skills in England and areas in the wider UK where the method has been used. However, DfE statistics on phonics and reading show that up to 25% of children at the end of Key Stage 1 (Year 2, 6–7-year-olds) and at the end of Key Stage 2 (Year 6, 10–11 year olds) continue to fail to meet expected standards.

The DfE statistics from 2016–2023 make clear the size of the minority Rose recognized would not make adequate progress with SSP. Koutsouris et al. (2021) drawing on the DfE (2016) statistics found that 20% of children entering Key Stage 2 in England (7–8-year-olds) were delayed or non-start readers.

In 2019, prior to the pandemic, 9% of Key Stage 1 children in England (Year 2, 6–7-year-olds) did not meet the standard set for phonics, and 25% failed to meet the standards set for reading at the end of Year 2. This means that slightly over 16% of the total number of children at the end of Key Stage 1 in England still failed to meet the standards set for reading, even though this 16% had demonstrated sufficient knowledge of phonics to pass the phonics test. This 16% despite having met the expected standard set for phonics were not able to meet the standards set for reading. Working at the expected standard includes fluency, comprehension, reading most common exception words and most common suffixes (Standards and Testing Agency, 2018).

These statistics provide evidence that systematic synthetic phonics (SSP) alone, is insufficient for reading success for this group. Furthermore, in 2019, 27% of 10–11-year-old children did not meet the standards set for reading by the end of Key Stage 2 (DfE, 2019b statistics). Post-covid, the DfE statistics published in October 2023 found that 11% of children failed to meet the expected standard in phonics at the end of Key Stage 1 and 32% of children failed to meet the standard set for reading. The DfE statistics published in October 2024 found that 11% of children failed to meet the expected standard in phonics at the end of Key Stage 1. The percentage of children who failed to meet the standard set for reading at the end of Key Stage 1 is not available as these assessments became non-statutory from 2023/2024 onward (DfE, 2024a). However, in the Key Stage 2 statistics published in October

2024, 26% of children failed to meet the standard set for reading (DfE, 2024b).

According to the Education Policy Institute [EPI] (2024) report, since the introduction of the DfE (2019a) Phonics Screening Check (PSC) there is no evidence of improved Key Stage 1 (end of Year 2 England) or Key Stage 2 (end of year 6 England) reading results, and no evidence that the PSC narrowed the attainment gap at the end of Year 2 (children 6–7 years of age) or the end of Year 6 (children 10–11 years of age) In addition to this the EPI report concludes that neither national data from the International Reading Panel Study (PIRLS, 2021), previous research using the NPD (National Pupil Database), or the new analysis in the EPI report, find a discernible positive impact of the Phonics Screening Check on the reading levels of primary aged children in England.

1.1 Current issues

Zarić et al. (2021) research provides evidence that even when phoneme-to-grapheme correspondence is one-to-one, it is insufficient for the development of orthographic knowledge for reading fluency and spelling. This research was conducted in German, a transparent orthography. Orthographic knowledge refers to spelling patterns stored in memory that are recognized on sight without having to decode or encode phoneme-by-phoneme. Orthographic mapping refers to the linkage between these larger units of sound and their orthographic patterns (onsets, rimes, syllables and words). Kilpatrick (2020), a proponent of phoneme-to-grapheme correspondence as the route to orthographic mapping acknowledges that "*letter-sound knowledge and phonemic awareness are not enough- perhaps they are enough for phonic decoding- but not for efficiently remembering words during real world reading*," (p.13). He theorizes that "proficiency" of letter-sound and phonemic knowledge explain orthographic mapping. This theory, however, fails to provide insight into cases where reading and spelling can develop at a normal rate, if the child can develop mappings between groups of letters that represent groups of sounds without being able to identify the individual phonemes within these groups. Stothard et al. (1996) reported the case of LF who was unable to establish phoneme-to-grapheme correspondences and could not read non-words, yet she learned to read and spell at a normal rate. LF learned to read by developing mappings between orthography and larger units of sound. She recognized groups of letters that represented groups of sounds without being able to identify the individual phonemes within these groups. This case demonstrates the importance of teaching orthographic mappings.

Phonemic and letter-sound proficiency as described by Kilpatrick (2020) provides an account of alphabetic mapping (phoneme-to-grapheme links) but not an adequate account of orthographic mapping. This phonemic proficiency theory fails to take account of multiple mappings in English and the orthographic choices that must be made. "For example, the "o" sound can be spelled seven different ways as in, go, know, though, note, toe, boat, sew. Remembering the auditory sequence of letter names as an aid to remembering the visual sequence of letters is a strategy that is needed to learn these spellings. It is not possible to spell these words based on identifying the phonemes only.

This raises the issue of the importance of letter names. In the USA letter names are taught in advance of letter sounds and before formal schooling. Whereas in "*England, children are taught to label letters by their sounds. . .before they learn the letter names*" (Treiman and Wolter, 2020. p.48). It is not until Year 1 in England, when children are 5–6 years of age, that the statutory curriculum requires that the names of the letters of the alphabet are taught. The DfE acknowledge that "*Knowing letter names is necessary to distinguish between alternative spellings of the same sound*" (DfE, 2014). However, for the first year at school children (4–5 years of age) are learning to associate letters with their sound, not letter names. This is confusing for young children when encountering words with the same letter and different sound. It is important that teachers can refer to letter names, from the beginning, so that they can discuss the sound made by the letters with multiple mappings, e.g., "c" at the beginning of cat and at the beginning of circle. Treiman and Wolter (2020) highlight the benefits for 4–5-year-old children, of knowing letter names when attempting to spell words, and as a prerequisite to learning letter sounds. It is important that children are taught the names of the letters of the alphabet in advance of teaching phoneme-to-grapheme correspondences because knowing the letter names helps with the spelling of words containing phonemes with multiple mappings (Treiman and Wolter, 2020).

1.2 The development of orthographic knowledge

Alphabetic mapping refers to the linkage between phonemes and graphemes. Alphabetic knowledge refers to the letters of the alphabet and the phonemes they represent that are stored in memory. English has approximately 44 phonemes and only 26 letters. This means that phoneme-to-grapheme correspondence is not straightforward and is further challenged by multiple mappings. For this reason, orthographic knowledge is required to know if a spelling "looks right." A representation of the word, or key word parts, must already be in memory to know if the correct letters have been chosen from the range of letter(s) that can represent phonemes that have multiple mappings, for example, because not "becos," said not "sed."

Orthographic processing is how you identify and form orthographic units, commit them to memory and retrieve them. It is the cognitive process that enables the detection of letter patterns representing a whole word, patterns within words (onset and rime), syllables, and rules and regularities in print, by attending to each letter whilst also noting letter sequences in parallel as units, thus enabling these orthographic units to be stored in memory. This learning can be acquired implicitly from reading experience. However, children with poor orthographic processing require orthographic mappings to be explicitly taught. McMurray (2004, 2020) provides evidence that spelling is the vehicle that can provide the level of systematic teaching needed.

Mather (2024) notes that orthographic processing facilitates the recall of letters and their sounds, letter combinations and whole words contributing to the development "*of a robust sight word vocabulary and accurate spelling*" (p.22). Good readers who are good spellers are children who have good orthographic processing

ability and can, therefore, detect patterns in print and learn about spelling rules and regularities implicitly from their reading experience, without being taught these rules (McMurray, 2020). Lennox and Siegel (1994) identified three groups of young readers: "good readers/good spellers," "good readers/poor spellers," "poor readers/poor spellers." They found that children with normal spelling abilities (i.e., "good readers and good spellers") develop visual skills and the use of analogy implicitly from exposure to print "in tandem" with phonological skills. However, Lennox and Siegel found that those children who were "poor spellers but good readers," were able to use phonological skills in their spelling but had difficulty choosing the correct orthographic representation of a word from the phonologically accurate alternatives. They considered that this difficulty was due to deficits in visual memory and lack of awareness of orthographic patterns. The result of neglecting to teach phonics explicitly at orthographic levels allows the "good readers/good spellers" group, who can develop orthographic knowledge implicitly from their reading experience, to advance well beyond their peers (McMurray, 2020). For orthographic knowledge to develop the child must be able to map blended units of sound at onset and rime, syllable and whole word level, to the orthographic patterns that represent them, in addition to, and not instead of, establishing phoneme-to-grapheme correspondence (PGC). However, it is stated in explanatory Note 1 of the DfE (2023) 16 essential core criteria for validation of phonics programmes, that "*The focus should be on phonemes [footnote 2], and not on "consonant clusters" (/s/ + /p/ + /l/not/spl/) or "onset and rime" (/c/ + /a/ + /t/not c-at, m-at, b-at)*".

The ability to recognize patterns in print that consist of more than one phoneme is an essential strategy for children with working memory difficulties and/or orthographic processing difficulties (Mather and Jaffe, 2021; McMurray, 2020). The needs of children who fail to make progress when phonics is taught at phoneme-to-grapheme level only, without the inclusion of phonics at the orthographic levels of onset and rime, syllable and whole word, cannot be addressed by the SSP method being repeated for small groups or on a one-to-one basis as recommended by the DfE (2023). According to Adams (1990), children 4–5 years of age should be taught phoneme-to-grapheme correspondences first, but not all phonemes, as the long vowel sounds are difficult for many young children to hear and discriminate. Adams (1990) advises that onset and rimes are important for stabilization of the vowel sounds which she describes as notoriously difficult to learn in isolation. Their pronunciation is more stable within rime patterns. McMurray (2022) advises that when consonant phonemes have been established, teaching should move quicky to establish recognition of initial consonant orthographic patterns also known as onsets, i.e., consonant blends such as cl, bl, fl, gl, sl, and consonant clusters such as spl. These orthographic patterns are recognized on sight as orthographic units that occur at the beginning of many words. This is not an additional memory load but a step in developing effective decoding that does not place unnecessary demands on working memory. Children with orthographic processing difficulties when taught using SSP develop phonemic awareness because this is what they have been taught to focus on. Because they are insensitive to orthographic patterns they cannot acquire orthographic knowledge implicitly and become dependent on decoding phoneme-by-phoneme. Alphabetic mapping is an important first step in the learning process (Ehri, 2020; Mather and Jaffe, 2021; McMurray, 2022), however, due to the complexity of English orthography, and the range of multiple mappings from grapheme-to-phoneme and phoneme-to-grapheme, decoding words for reading and encoding words for spelling involves choices that are dependent on orthographic knowledge being in place. Daffern and Critten (2019) found that the only evidence of orthographic knowledge for low achieving spellers was a small bank of high frequency words. It is, therefore, important to consider if this is the case for the 267 children in this study. In the National Curriculum programmes of study for England the DfE (2013) advises that children "*should also be able to make phonically plausible attempts to spell words they have not yet learnt*"(p.16). However, Treiman et al. (2019) found that phonological plausibility is not a good indicator of later success in spelling. They found that the best predictor of later spelling was orthographic correctness which significantly outperformed phonological plausibility. This was based on a UK sample of children in the reception year (mean age 5 years 1 month at baseline) and their later spelling performance in Year 2 (mean age 7 years 3 months). Treiman et al. (2019) concluded that, "*These findings fit with other evidence that even young children can attend to and remember visual orthographic features of words (Cassar and Treiman, 1997; Martinet et al., 2004; Wright and Ehri, 2007)*," (p.92).

Treiman et al.'s (2019) findings provide evidence of the orthographic processing ability of young children and the positive impact on spelling accuracy. McMurray (2004, 2020) advises that if children know the correct phonemes in a word and they select letters that do represent the phoneme, but it is the wrong choice of letters for words they have encountered when reading (phonological plausibility), then this is a warning signal that the child is unable to acquire orthographic knowledge implicitly from their reading experience (McMurray, 2020). She contends that failing to recognize these orthographic processing difficulties at 5–6 years of age results in spelling difficulties becoming entrenched if intervention, through learning to spell, is not implemented at this early age. Her research found that an integrated approach to teaching spelling to develop phonic, orthographic and morphemic knowledge in parallel from 5–8 years of age, resulted in success for all children with the average post intervention standardized spelling score being 113, almost one standard deviation above the mean ($p < 0.0001$, Effect size 1.19). This programme included a structured approach to teaching patterns and sequences consistent in sound and spelling (onset and rime) and the associated morphemic knowledge for words in the patterns, alongside the teaching of high frequency words and words with a curriculum focus (McMurray, 2004, 2020).

As the difficulties with over-reliance on phoneme-to-grapheme correspondence have become clear, key researchers have emphasized the importance of phonics teaching extending beyond phoneme-to-grapheme correspondence to develop knowledge of orthographic units for automatic sight word reading, reading fluency and spelling (Ehri, 2020, 2024; Mather and Jaffe, 2021). Evidence of good orthographic processing ability is demonstrated by very young children who can learn to read words by sight without having to decode every word

phoneme-by-phoneme (sound by sound), and by a young child's ability to recognize and recall irregular words that have to be learned as single units, for example, the, are, our, were. Miles and Ehri (2019) and Deheane (2009) contend that familiar words are read as single units with each letter processed in parallel rather than sequentially. This is something that children with poor orthographic processing have particular difficulty with.

The significance of orthographic processing difficulties has been highlighted in the meta-analysis conducted by Georgiou et al. (2021). They examined 68 studies published between January 1990 and December 2019, to examine if individuals with dyslexia (DYS) have an orthographic knowledge deficit when compared to their chronological-age (CA)– and reading-level (RL)–matched controls. A random-effects model analysis revealed a large effect size (Cohen's d = 1.17) for the CA-DYS comparison and a small effect size (Cohen's d = 0.18) for the RL-DYS comparison. They concluded that children with dyslexic-type difficulties have a deficit in orthographic knowledge that is as large as the deficits in phonological skills and rapid naming and that this should be assessed, and consideration given to including activities to develop orthographic knowledge in intervention programmes. It is important to establish whether the 267 children in this study have spelling errors that are consistent with Georgiou's claim about the significance of the orthographic deficit. Difficulties in acquiring orthographic skills are now included in the new Delphi definition of dyslexia and the criteria for assessment (Holden et al., 2025). A standardized spelling test is now a compulsory element of the assessment of literacy difficulties and if spelling is poor, this triggers an assessment of orthographic processing using a standardized test. Mather et al. (2024) are the authors of the new TOD (Tests of Dyslexia) which include tests to assess orthographic processing.

Ehri (2024) stresses the importance of explicit spelling instruction explaining that reading is not enough for spelling accuracy. Ehri (2024) provides evidence that recognition of legitimate spelling patterns results in only partial retention of orthographic patterns which is sufficient for word recognition when reading, but is insufficient for the development of the orthographic processing skills needed for spelling accuracy. Ehri (2024) contends that onset and rime patterns such as "bent, tent, spent" and multi-letter units such as the suffix "ing" (p.8) are learned through a process of unitization whereby grapheme-phoneme units enable multi-letter units to form, which in turn connect these larger spelling patterns to the sound units in words they represent, enabling storage in memory. The process Ehri (2024) describes as unitization, does not happen automatically for children with orthographic processing difficulties hence the importance she places on explicit spelling instruction (2020, 2024). This is supported by Fernandez et al. (2011) who found that children with dyslexia have an implicit learning deficit, which may make it difficult to acquire orthographic representations implicitly from reading experience. Because children with orthographic processing difficulties show no benefit from frequency sensitivity when reading (Mather and Jaffe, 2021), they must be explicitly taught patterns such as onset and rime, in a structured and developmental sequence for orthographic knowledge to develop (McMurray, 2020).

1.3 Strategies used in SSP that are problematic for children with orthographic processing difficulties and/or working memory difficulties

1. DfE (2023) essential core criteria for the validation of SSP programmes advises publishers not to include consonant clusters or onset and rime and to teach phoneme-grapheme correspondences only. As a result, the term "phonics" has become synonymous with a phoneme level only approach to phonics teaching. Children are, therefore, being trained to look for phonemes only, and not orthographic patterns. This entrenches the difficulties experienced by children with poor orthographic processing (McMurray, 2020).

2. Children are taught to "*read printed words by identifying and blending (synthesizing) individual phonemes, from left to right all through the word*" (DfE, 2023). This strategy places excessive demands on working memory making blending beyond the capacity of many of 5–8-year-old children. According to Alloway (2011) the average 5-year-old can hold one item in short term temporary storage within working memory, the average 7-year-old two, the average 10-year-old three, and the average 14-year-old four items. Only 5% of one syllable words in English are two phonemes long, 43% are three phonemes long and 52% of one syllable words have more than 3 phonemes (Wyse and Goswami, 2008). This explains why, when learning to read in English, blending phonemes as an activity, is extremely difficult if not impossible for children due to limitations in working memory capacity. This problem is not experienced by children learning to read in Spanish or Italian where syllables are only two phonemes long and there is phoneme-to-grapheme consistency, i.e., one letter-one sound. Difficulties decoding unknown words is not a diagnostic criterion for dyslexia in these languages. This is also because there are no multiple mappings therefore no orthographic choices to be made.

3. Ehri's (2020) error analysis found that breaking the speech stream between phonemes causes students to forget initial phonemes during blending. In this research study 5–6-year-old children, who were taught to "*stretch and pronounce phonemes without breaking the speech stream*" (p.S52), were more successful when decoding than the group who did break the speech stream between phonemes (Ehri, 2020). However, the strategy of taking each sound in turn as recommended by DfE (2023) does not stipulate the importance of ensuring that there is no break between each phoneme. Furthermore, teachers are not advised to ensure that children develop strategies to enable them to cope with the working memory demands of blending phonemes in one syllable words with more than two phonemes.

4. The DfE (2023) note 6 promotes the use of decodable books made up of words that can be decoded using phoneme-to-grapheme correspondences only, to make sure that children "*learn to rely on phonic strategies.*" This reliance on phonemic decoding and encoding can become a major contributor to spelling difficulties and difficulties with reading fluency if children become over-reliant on alphabetic mapping and orthographic knowledge fails to develop (Ehri, 2020, 2024; Mather and Jaffe, 2021, 2024: McMurray, 2020).

SSP works well for 75% of children because good orthographic processing supports the phonemic decoding strategy enabling the retention of words in orthographic memory for fast and efficient

retrieval. Children with poor orthographic processing can learn phoneme-to-grapheme correspondences, but when asked to spell words containing phonemes that can be spelled a number of different ways, they have no orthographic units stored in memory to enable the correct selection of letters and their spellings are phonologically plausible. It is important to consider the extent to which phonologically plausible spellings are evident in the writing samples of the 267 children in this study.

Furthermore, as early as 2012 Bell and McLean pointed out that, "*. . .effective literacy skills specialists experienced in teaching synthetic phonics may not be successful in teaching those skills to those dyslexics whose phonological skills make them unreachable by this approach*" (p.136).

Given that the DfE (2023) recommend more SSP intervention for children experiencing difficulty, and this is common practice in Northern Ireland, have any of the 267 children in this study been unable to develop phoneme-to-grapheme correspondences despite 4 years of SSP intervention.

1.4 The importance of language development

The role of spelling in vocabulary development is supported by evidence presented by Miles and Ehri (2019). They cite research spanning from 2005–2016 in support of the importance of learning spellings. A further advantage of teaching spelling was noted by Ehri (2020) who advises that teaching spelling units enables children to generalize to new words.

McMurray's (2004) Ph.D research provides evidence that learning to spell plays a significant role in language development. McMurray (2004, 2020, 2022) contends that effective, explicit instruction to develop orthographic knowledge, spelling accuracy and language development can only be achieved through an integrated approach to learning to spell because the connection between orthographic units and morphemic units must be taught together. Morphemes are the smallest units of meaning in spoken language.

The selection of schools in McMurray's Ph.D study was randomized using criteria to ensure matched experimental and control schools (RCT) and a 3 year quasi-experimental research design was conducted with children 5–8 years of age. This resulted in post intervention average standardized spelling score for the experimental schools of 113, almost one standard deviation above the mean ($p < 0.0001$). The focus on developing sensitivity to patterns in print linked to meaning and language development, resulted in the very strong effect size of 1.19. All children in the experimental schools increased their standardized spelling score from Year 2, 5–6 years to Year 4, 7–8 years (equivalent Year 1–3 in England) with the average standardized spelling score increase being 19 standard points from baseline in the experimental schools. The standardized spelling score for 24% of children in the control schools decreased over the same period with the average standardized spelling score increase for the control schools being 5 standard points (McMurray, 2006, 2020; Lavan and Talcott, 2020). Phonologically plausible spellings were still evident in independent writing in the control schools at the end of the study (Year 4: children 7–8 years of age).

According to McMurray the extent of the explicit and focused language development required to establish connections between orthographic units and morphemic units is not appropriate for children aged 4–5 years. At 4–5 years of age, the focus should be on teaching letter names, establishing phoneme-to-grapheme correspondence, and strategies for blending phonemes together including recognition of initial consonant blends. McMurray (2004, 2020) advises that teaching spelling with a focus on onset and rime patterns in Year 2 Northern Ireland (5–6 years of age) enables the identification of children with orthographic processing difficulties through focused activities designed specifically for this purpose. This enables informed extensive intervention at a whole class level to be taught commencing in Year 3 (6–7 years of age, equivalent Year 2 in England) (McMurray, 2004, 2020).

At 6–7 years of age onset and rime spelling patterns should be taught with the associated word meanings (morphemic knowledge) for each of the spellings in the pattern. Taking the onset and rime pattern, cap, lap, gap, map, nap, rap, clap, trap, strap, wrap as an example, the spelling "lap" represents 4 free morphemes: 1. a noun – part of your body "*the cat sat on my lap*"; 2. a noun – a circuit in a race "*I ran a lap of the track*" 3. a verb – water breaking gently against the shore "*the water lapped against the shore,*" 4. a verb – animals drinking by collecting liquid on their tongue "*the cat lapped up the milk.*"

The range of meanings for one spelling is discussed with the whole class for language development purposes and to motivate children to think about ways they can use spellings with more than one meaning. Receptive and expressive language can be well in advance of the ability to communicate in written language. For this reason, the differentiation in the written follow up activities takes account of working memory and orthographic processing difficulties and leaves more challenging written morphemic activities until the next stage of the programme, for children with specific processing difficulties.

McMurray (2004, 2020) found that phonemic, orthographic and morphemic knowledge develop in parallel from the beginning for children who do not have orthographic processing difficulties. These children can learn implicitly from their reading experience and become increasingly proficient in their use as demonstrated in the sample of writing in Figure 1.

This sample of writing (Figure 1) was written at the end of the second year of formal schooling (5–6 years of age, Year 2 NI, equivalent Year 1 in England). The only words that were explicitly taught for the purpose of spelling were CVC words and a small number of high frequency words. The following correctly spelled words had not been explicitly taught *(like, new, called, every, time, take, for, near, away, sports, day, nearly, came, first, bean, race, little*). This child has good orthographic processing ability and can abstract the orthographic rules and regularities in print from reading experience as demonstrated by the correct spellings and the spelling errors in this sample. Phonemic, orthographic and morphemic knowledge are developing in parallel because this child can learn implicitly from reading experience. For example, the incorrect spelling "brghte" for "brought," demonstrates developing orthographic knowledge because the child knows that "ght" should be in the spelling. A phonologically plausible spelling would be "brot." Similarly, the incorrect spelling "agian" demonstrates orthographic awareness of the letters that should be in the word. A phonologically plausible spelling would be "agen."

FIGURE 1
May 1999 Child A (Year 2 NI children 5–6 years of age). McMurray, S. (2020). Learning to spell for children 5–8 years of age: The importance of an integrated approach to ensure the development of phonic, orthographic and morphemic knowledge at compatible levels. *Dyslexia* 26, 442–458. doi: 10.1002/dys.1663.

The sample of writing in Figure 2 was written by a child in the same class. There was no statistical difference in the verbal ability of Child B and Child A. There is, however, a significant difference in their orthographic processing ability. All the words that are spelled correctly in this writing sample had been explicitly taught. Child B spelled the words that had not been taught the way he said them, and the incorrect spellings are phonologically plausible (vre -very, tol- tall, dat- dad, ped-played, futbl-football). It is important to establish the extent to which pronunciation impacts spelling in the 267 samples in the current study.

In Figure 3 Child B continued to spell words the way he said them without the orthographic knowledge needed to spell words he had not been taught, for example, just-jesd, lifid- lifted, wes-was, desrod-destroyed, weos-was, sil-stil, sot-shot, cilt-killed.

By the end of the first term in Year 4 (NI, 7–8 years of age, equivalent Year 3 in England) Child B had developed sensitivity to orthographic patterns. This was achieved through the spelling intervention which was designed to ensure the development of phonic, orthographic and morphemic knowledge at compatible levels. The design of the spelling programme was based on McMurray's (2004) hypothesis that it is not until the brain has experience of a significant number of visual patterns and sequences, consistent in sound and spelling, that it can begin to make sense of the common elements in the specific formula (pattern) that make up, for example, rhyme patterns and sequences in general. Each rhyme pattern is different; for example, man, can, ran, fan//got, lot, hot, but rhyme patterns, consistent in sound and spelling, have common elements –that is, same end pattern and sound with changes only in the initial sound. What is sufficient experience for one child, to abstract these statistical relationships between phonology and orthography, may be insufficient experience for another and cannot be achieved via implicit learning when reading for the group of children with orthographic processing difficulties. Repeated attempts to spell the same rhyme pattern can be unproductive for this group. This may be because repeating the same rhyme pattern does not supply the brain with sufficient information regarding common elements relating to rhyme patterns in general. Experience of many patterns is needed to develop sensitivity to, and storage of, orthographic patterns for effective recall.

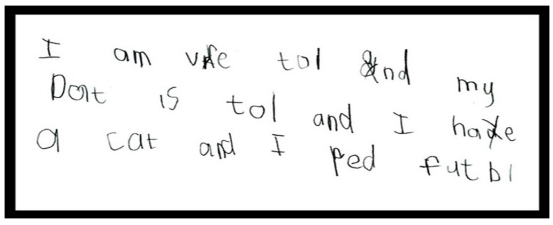


FIGURE 2
May 1999 Child B. McMurray, S. (2020). Learning to spell for children 5–8 years of age: The importance of an integrated approach to ensure the development of phonic, orthographic and morphemic knowledge at compatible levels. *Dyslexia* 26, 442–458. doi: 10.1002/dys.1663.

Child B was the child who took the longest to develop sensitivity to orthographic patterns and when this was achieved there was a significant improvement in his spelling in independent writing as seen in Figure 4. His standardized spelling score improved by 29 standard points from baseline in January 1999 to May 2001.

Child B has only eight spelling errors in a total of 99 words written: (press-prese), (chewing-chuthing & chuwing), (shoe-*shuy*), (thrown- thorn), (shovel- shevel), (lifted- levted), (sting-stingk), McMurray (2004, 2020). His orthographic processing ability has improved significantly as demonstrated by the correct spelling of words that include sounds that can be spelled more than one way, e.g., if "machine" was spelled by sound alone, it could be spelled "masheen," called could be spelled "cawld," adventure could be spelled "advencher."

In the SSP only approach to beginning reading, it has been a flaw to assume that the orthographic knowledge essential for its success is in place or can develop without explicit teaching for all children (McMurray, 2020). Just as phoneme-to-grapheme correspondence must be taught in a systematic and structured way so too must onset and rime patterns. McMurray (2020, 2022) contends that the predominant strategy for

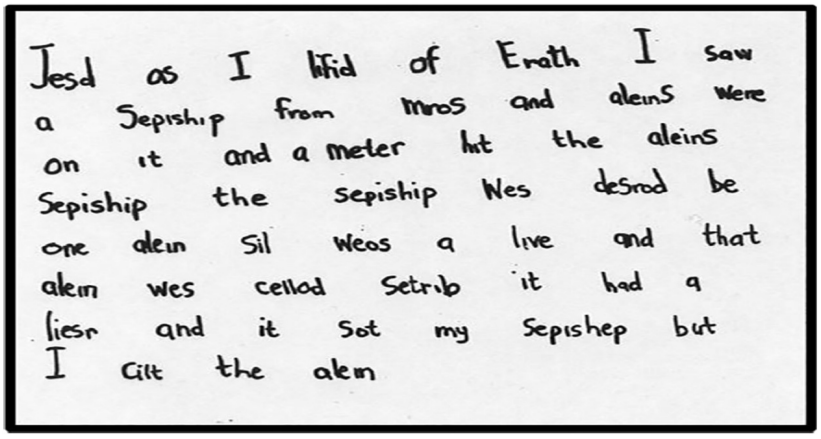


FIGURE 3
January 2000 Child B. McMurray, S. (2020). Learning to spell for children 5–8 years of age: The importance of an integrated approach to ensure the development of phonic, orthographic and morphemic knowledge at compatible levels. *Dyslexia* 26, 442–458. doi: 10.1002/dys.1663.

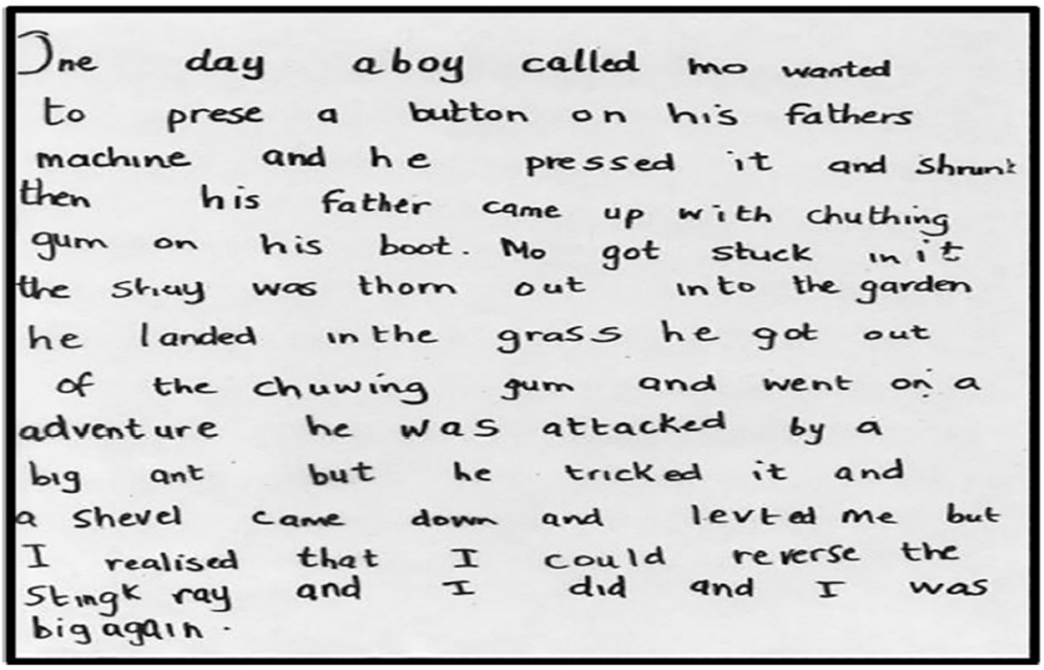


FIGURE 4
December 2000 (7–8 years of age). McMurray, S. (2020). Learning to spell for children 5–8 years of age: The importance of an integrated approach to ensure the development of phonic, orthographic and morphemic knowledge at compatible levels. *Dyslexia* 26, 442–458. doi: 10.1002/dys.1663.

reading unknown words should be decoding using phoneme-to-grapheme correspondence but strategies should also be taught to ensure the development of commonly occurring initial consonant orthographic units to support working memory until the systematic and structured teaching of an integrated approach to learning to spell commences in January of the second year of formal schooling (5–6 years, McMurray, 2020, 2022).

Research evidence to support intervention programmes often focuses on group averages to make claims about effectiveness. Where effect size and statistical significance are moderate or low then it is clear that a group of children may not be progressing at an expected rate for their age, or may even be regressing. The extent of difficulties experienced by children who underachieve is hidden in the average score. Failure to highlight this group results in a lack of transparency and expectations that the intervention will benefit everyone when this is not the case.

There has been no research, up to the time of writing this paper, that has investigated the sources of linguistic knowledge children in mainstream schools with the severest difficulties in literacy development, draw on in their attempts to spell correctly. This research is important given the continued focus in schools to teach phonics at phoneme-to-grapheme level only and the recommendation to use the same method when providing further support for children who struggle. The analysis of spelling

in independent writing provides a window into the child's mind, enabling the assessment of the development of phonemic, orthographic and morphemic knowledge. These three sources of linguistic knowledge are essential for normal literacy development (Daffern, 2017; McMurray, 2004, 2020). McMurray (2020) has highlighted the importance of ensuring that they develop in an integrated way within an optimal timeframe from 5–8 years of age. Given that Treiman et al. (2019) have established that from a very early age (5–6 years) orthographic correctness is the most significant indicator of later spelling ability (at age 7), the sources of linguistic knowledge that children with the most significant literacy difficulties draw on at 8–9 years of age is the focus of this research. This age group was chosen to ensure that the focus was on children with the severest and most persistent specific literacy difficulties. Children aged 8–9 years (Year 5 in Northern Ireland) are beyond the optimum period for the development of orthographic knowledge identified by McMurray (2004, 2006, 2020) which occurs between 5 and 8 years of age (Years 2 to 4 in Northern Ireland).

Treiman et al., 2019 (p.92) stated that although phoneme based measures do not advance orthographic knowledge for the majority of children "*It is possible that phoneme-based measures would be significantly better predictors of future performance than letter-based measure for children who are even less advanced than those studied here. The analysis we conducted with the poorest spellers in the study provide a hint of such an effect, and further research is needed to examine the possibility*." The research in this paper considers whether the poorest spellers, do or do not, benefit from phoneme based measures as hinted by Treiman et al. (2019). Furthermore, Carroll et al. (2025) suggested that further research is needed to investigate the role of orthographic processing, which is a key consideration based on the findings of this research.

1.5 Aim of the study and research questions

The overarching aim of this study was to examine spelling attempts in writing samples from a large group of children with significant specific literacy difficulties of a dyslexic-type, who were beyond the optimum period for the development of orthographic knowledge identified by McMurray (2004, 2006, 2020) and Treiman et al. (2019). These children were 8–9 years of age and at the start of their fifth year of formal schooling. The research aimed to establish the sources of linguistic knowledge these children were using to spell when writing independently and the extent to which they could identify the phonemes in the words they were attempting to spell.

Research Questions (RQ)
The research questions to address the stated aim are:

RQ1. Is there evidence to support Bell and McLean's (2012) claim that some dyslexic children have severe phonological difficulties that teaching phoneme-to-grapheme correspondence cannot address?

RQ2. Is a small bank of high frequency words the only evidence of orthographic knowledge for poor spellers? (Daffern and Critten, 2019)

RQ3. Do the spelling errors in the writing samples support Georgiou et al.'s (2021) claim that children with dyslexic-type difficulties have a deficit in orthographic knowledge?

RQ4. Does the teaching of phoneme-to-grapheme correspondences only, enable the development of orthographic knowledge or are children's spelling errors phonologically plausible?

RQ5. Is spelling impacted by how a child pronounces words?

2 Methodology

2.1 Participants

All of the children who participated in this study were formally assessed by the Education Authority psychology service. These children demonstrated a cognitive profile which included a standardized score of 90 or above in one or more of the subtests of the Wechsler (2016) Intelligence Scale for Children (WISC). Standardized scores for single word reading, reading comprehension and spelling were also obtained using the Wechsler (2017) Individual Achievement Test (WIAT-111UK). To meet regional criteria for additional support from the literacy service, standardized test scores had to be in the below average range and a discrepancy between actual and expected attainments in standardized tests of word reading, reading comprehension or spelling large enough to be expected to occur in only 2% or fewer of pupils of that age. Although this profile may indicate specific literacy difficulties of a dyslexic-type, no formal clinical diagnoses of dyslexia were obtained.

The literacy support service teachers also gathered baseline data which included assessment of reading rate using the York Assessment of Reading for Comprehension YARC (2011). The presentation of reading difficulties for most of the children referred to the EA literacy service includes very slow reading rate and overdependence on decoding phoneme-by-phoneme.

Reading rate could not be calculated for 32.58% (n = 87) of the children due to exceeding the number of errors permitted for a reading rate to be calculated on this test. 24.7% (n = 66) children scored between 70 and 79 and 24% (n = 64) scored between 80 and 84.

18.72% (n = 50) had scores in the average range (85–115). Twenty-five of these children had scores between 85 and 89 and twenty four children had scores between 90 and 99. One child obtained a score of just over 100. These scores demonstrate that a significant number of children fall within the lower end of the average range for reading rate relative to the national average for their age. It implies that these children demonstrate a degree of slow and dysfluent reading. They are spending time decoding words which can negatively impact on understanding (YARC, 2011, p2) at the cost of time and effort (p73).

Information about how spelling was taught in the child's school was provided on the data collection cover sheet which was completed by the Education Authority Literacy support teacher and

submitted with the writing sample. All of the children were being taught to learn to spell using a phoneme-grapheme approach.

2.2 Selection of the sample (September 2019)

Inclusion in this research was on a voluntary basis and permission was sought at all levels by the Education Authority. In Northern Ireland there are five regional literacy support services. Four of the five regions agreed to participate, and one region declined. As this research was above and beyond the normal workload for the Education Authority literacy service teachers, agreement was sought from each individual teacher in the four Education Authority regions that agreed to participate. Permission was then sought from the primary schools the children attended. With the agreement of the school, permission was sought from the children in Year 5 (NI 8–9 years-of-age), who were about to commence support from the Education Authority literacy service teacher. The child's parents or guardians were also asked to give permission. If any child or parent declined at the outset or during the study, then their writing sample was not sent to the researchers. No child or parent declined participation during the study. The writing samples were anonymized by the Education Authority prior to being sent to the researchers. Agreement to participate was achieved for 267 children in Year 5 in 143 schools.

2.3 Ethical approval

Ethical Approval was obtained from the Ethics Committee within the authors' university. Ethical approval was also obtained from the Education Authority who gathered all baseline assessment data including the samples of independent writing and provided access to the researchers, with strict adherence to data protection legislation and the secure storage of data.

2.4 Data collection

In October 2019 baseline data on spelling was obtained from 267 children in Year 5 (8–9 years of age) who each completed a sample of independent writing under the supervision of the Education Authority literacy service specialist teacher assigned to them. The independent writing samples were collected as part of their baseline assessment process prior to commencing support. These writing samples were completed in a one-to-one setting in the child's school. All literacy service teachers were briefed by the lead author of this paper in August 2019 about the controlled conditions for collecting the writing sample (for example, an attractive environment with no words on the walls that might be copied, encouragement given to persist but no help with spelling), and the brief that should be given to the children to stimulate interest in writing. The children were asked to write about their best day. The words in the title (My best day) were not counted as correct spellings because they could be copied from the title at the top of the children's response page. Up to 30 minutes was allowed for this independent writing activity, however, children completed in 10–15 minutes and spent the remaining time drawing a picture. When the child finished writing he/she immediately read back what they had written and the teacher transcribed what the child read back (the transcript). This transcript was used to list the correct and incorrect spellings.

2.5 Data analysis

The authors assigned samples of writing to 4 categories. The categories were developed from the authors' experience working with children with extensive literacy difficulties. Categorization was discussed and agreed by the four moderators who had previously been Special Educational Needs Coordinators in mainstream schools. They all had a master's degree in teaching and assessing children with dyslexia and extensive teaching experience. 100% agreement was required for each sample. Each sample was independently assessed by each moderator and in cases where there was not 100% agreement the sample was discussed and 100% agreement reached by strict adherence to the categories. The best sample from each category was used in the paper to allow the boundaries to be clear. To be included in a category the samples had to meet all the criteria for that category.

Category 1: Unreadable

Samples are unreadable and look like a jumble of letters.

Attempts to denote individual words are limited to the occasional simple (2 or 3 letter) high frequency words.

Evidence of initial consonant sounds being correct for some, but not all of the attempts to spell words within the sample of writing.

Insufficient evidence of phonemic knowledge beyond the initial sound in the word with the occasional final sound identifiable in some of the samples.

Insufficient evidence of the development of phoneme-to-grapheme correspondence because so little is written in the sample.

Category 2: Readable with the transcript

Small bank of simple correctly spelled high frequency words.

Some correctly spelled phonemically regular words.

Evidence of words being spelled the way the child says them.

Attempts to spell all or some of the phonemes in a word and the word is readable but spelling is incorrect.

Reading the transcript enables identification of words with the incorrect letter or letters for the vowel sounds.

Unusual spelling that is not close to how some of the words sound or should look.

There is limited (i.e., one or two words) or no evidence of correct orthographic choices for words with multiple mappings.

Category 3: Transcript needed to read some words

The text is easily read because the reader can draw on contextual clues made possible because of the greater number of correctly spelled words.

The phonemic knowledge demonstrated in misspellings allows incorrectly spelled words to be read as they sound, even though they don't look like the word.

Evidence of words being spelled the way the child says them.

The transcript is needed to decipher a small number of words that are not identifiable because the letters used do not represent the sound sufficiently well or because the reader feels a check is needed to ensure the correct interpretation of what is written.

Correct spellings are predominantly high frequency words or words that are phonemically regular.

There is limited evidence of orthographic knowledge being used to make correct spelling choices for phonemes that can be spelled a number of different ways.

Category 4: Transcript not needed to read words

The writing samples can be read without reference to the transcript.

All misspellings are readable because they represent the phonemes in the words sufficiently well to make recognition possible.

Some evidence of orthographic knowledge but children in this category did not write at length

2.6 Limitations of the study

The sample consists of children receiving support through four of the five regional literacy support services in Northern Ireland. Findings may not generalize to all children with literacy difficulties, particularly those not accessing such services.

The study used a single writing prompt ("My best day"); patterns may differ with other writing tasks.

Readability categories were determined by consensus among moderators; while agreement was high, some degree of subjective judgment in classification is inevitable.

3 Results

The writing samples were classified based on their readability. Of the 267 samples of independent writing examined, only 23 could be read without reference to the transcript (8.61% of 267). Three children were reluctant to write anything, and their submissions were blank (1.12% of 267). Seventeen children's writing was unreadable (6.36% of 267). The remaining samples had varying degrees of readability, and the transcript was required to allow the reader to read and understand the writing in full (83.9% of 267). Three categories of readability were identified (categories 2–4: 247 samples in total).

The title "My best day" was written on the sheet directly above where the children wrote their stories. "My," "best" and "day" are not counted if spelled correctly because they can be copied.

3.1 Category 1

Samples in this category are unreadable (see Figure 5). Six of the 17 scripts had no identifiable words and in a further 6 the only identifiable word was "I" (4 scripts) or "to" (2 scripts), 5 scripts had two identifiable words (1, in), (to, the),(is, dog), (a, big), (I, to, see Figure 5). Based on the transcripts the average number of words attempted was eleven. Figure 5 is a typical example. There is no viable evidence of the development of an effective phoneme-to-grapheme strategy or the acquisition of orthographic knowledge.

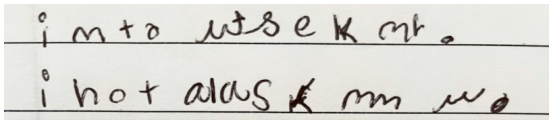


FIGURE 5
Category 1 Unreadable (17 writing samples: 6.36% of 267). Transcript: I went to Newcastle. I got ice cream.

The percentage of correct spellings demonstrating orthographic knowledge (that is, words that require spelling choices for phonemes that can be spelled a number of different ways) is given in categories 2–4 below. The high frequency word "was" if spelled correctly is included in the count of orthographically correct words (misspellings include "woz" or "wus"). The correct spelling of the high frequency word "the" is also included in the count of orthographically correct words as it has to be learned as an orthographic unit. In all cases where the word "the" is used, it is spelled correctly.

3.2 Category 2

In Figure 6 some simple correctly spelled high frequency words are evident: "is, I, in, he, to, get, and, a, go, for, us" providing support for RQ2. The words dog, pond, can be spelled phonemically. This provides evidence of phonemic knowledge being correctly applied when mappings are one-to-one.

There is evidence of attempts to spell words the way they are pronounced by the child (RQ5), for example, "olir" for other. Olir is a commonly occurring mispronunciation of the word "other" by young children in NI. Substitution of the "l" sound for the "th" in the word other, can occur due to poor articulation of the "th" sound or because the child has poor auditory discrimination and cannot distinguish the sound within the word. With regard to the misspelling of the word "doesn't", it is likely that the child does not pronounce the "n" sound, hence the spelling "dosit." Had the child been able to draw on orthographic knowledge of what the word looks like, the "e" and the "n" would not have been omitted. Although best is in the title it has not been spelled correctly. This may be because the child says "besdie" in his speech and therefore attempted to spell it as he says it. This a common colloquial pronunciation of the word best in areas of Northern Ireland.

"Wok" (walk) is an example of a phonologically plausible spelling; an attempt to spell the phonemes in the word and the word is readable, but the spelling is incorrect. The transcript is needed to enable identification of words with the incorrect letter or letters for the vowel sounds: "wein" (when), "hes" (his), "sad" (side), "lied" (lead). In the misspelling "wein" (when) the child has not heard the "h" sound following the "w." Many children with dyslexic-type difficulties cannot hear the difference between "w" and "wh." It is not uncommon for children in this group to either leave the "h" out of words starting with "w" or add it in when it is not needed.

10.3389/feduc.2025.1641126

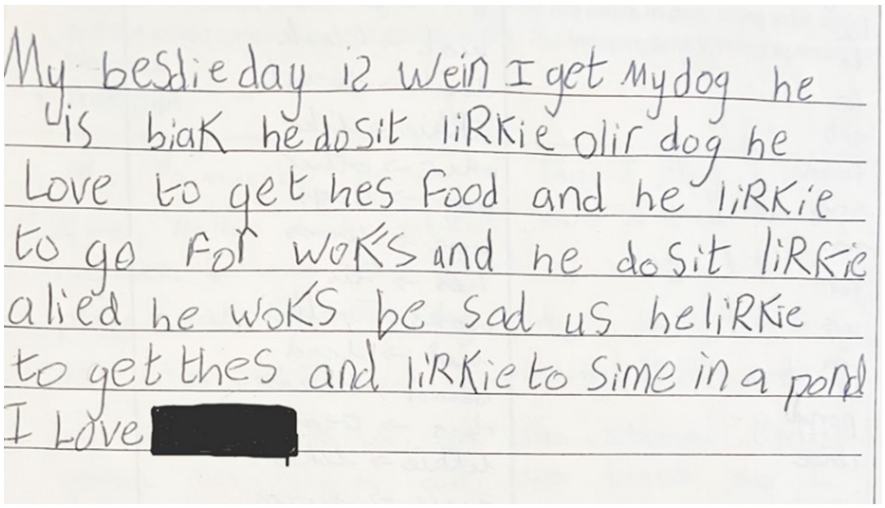


FIGURE 6
Category 2 Readable with the transcript (127 writing samples: 47.56% of 267). Transcript: My best day is when I got my dog. He is black. He doesn't like other dogs. He loves to get his food and he likes to go for walks, and he doesn't like a lead. He walks beside us. He likes to get treats and likes to swim in a pond. I love xxxxx (total words 56).

The incorrect spelling of like ("liRkie") is more difficult to explain as the capital R has been added consistently in all four attempts to spell the word. This unusual spelling, "liRkie," is not close to how the word sounds or looks and unlike "olir" and "besdie," it is not a word that is commonly mispronounced in NI, but may be a phonological issue specific to this child which a classroom teacher would be able to identify. Because the letters 'l' and 'c' are omitted from the spelling of the word black, the transcript is needed to identify this word. The word "treats" also requires the transcript to identify the word (thes).

The correct spelling of the word "love" provides evidence of the correct selection of letters for the vowel sound. A phonologically plausible spelling would be "luv." Evidence of orthographic knowledge is limited to the small bank of high frequency words and the correct spelling of the words "love" and "food." The vowel sound "oo" in "food" could be spelled different ways as in, "rude, blue, knew, fruit, feud, soup, through" but "oo" is the most common spelling for this phoneme and therefore, "food" may have been spelled phonemically. This was the only sample in this category that had two words that demonstrated orthographic awareness. The other samples had one or no words demonstrating the selection of the correct letters for words with sounds with multiple mappings.

The percentage of correctly spelled words that included orthographic choices in category 2 is 13.84% of the total number of words spelled correctly. The word "the" is spelled correctly in 66 of the 127 samples. If the word "the" is not included in the orthographic group then the percentage would be 8.39%.

3.3 Category 3

In Figure 7 the text is easily read because the reader can draw on contextual clues made possible because of the greater number of correctly spelled words: 'upon, was, to, sleeping, in, the, car, and, brother, mummy, at, find, so, we, out, of, car, on, apple, won, pink, puppy.' Correct spellings are predominantly phonemically regular words: "sleeping, pink, brother, time" (the word time is also a frequently displayed word in classrooms), or high frequency words: "a, I, to, was, in, the, and, at."

There is evidence of orthographic knowledge in the words, "apple, puppy, mummy (double letters), won (a phonologically plausible spelling is wun), car (kar). The misspelling of the word "bule" (blue) may be an attempt to spell the word drawing on orthographic knowledge with the letters being placed in the wrong order.

The phonemic knowledge demonstrated in these phonologically plausible mspellings allows incorrectly spelled words to be read as they sound, even though they may not look like the word. For example, the word "used" is misspelt "youst." "Once" is misspelled "ones."

With is misspelled "wif" because some children pronounce the sound at the end as an "f" sound (RQ5). "W" has been added to going ("gowing") because a "w" sound is often pronounced in the word (RQ5). The silent "h" in ghost is left out (gost). The misspellings for the words parking space "prekeing spas" also demonstrate the child's attempt to spell the words by sound.

The wrong letters for the vowel sound have been selected for the word train (trean), the word ride (rad) and the word went (want, wante).

The transcript is needed to decipher a small number of words that are not identifiable because the letters used do not represent the sounds in the word sufficiently well or because the reader feels a check is needed to ensure the correct interpretation of what is written, blake (blanket), tlel (tail), kudt (couldn't), chot (coat).

The percentage of correctly spelled words that demonstrate orthographic knowledge in category 3 is13.87% of the total number of words spelled correctly. The word "the" is spelled correctly in 61 of the 97 samples. If the word "the" is not included in the orthographically correct spellings then the percentage would be 9.74%.

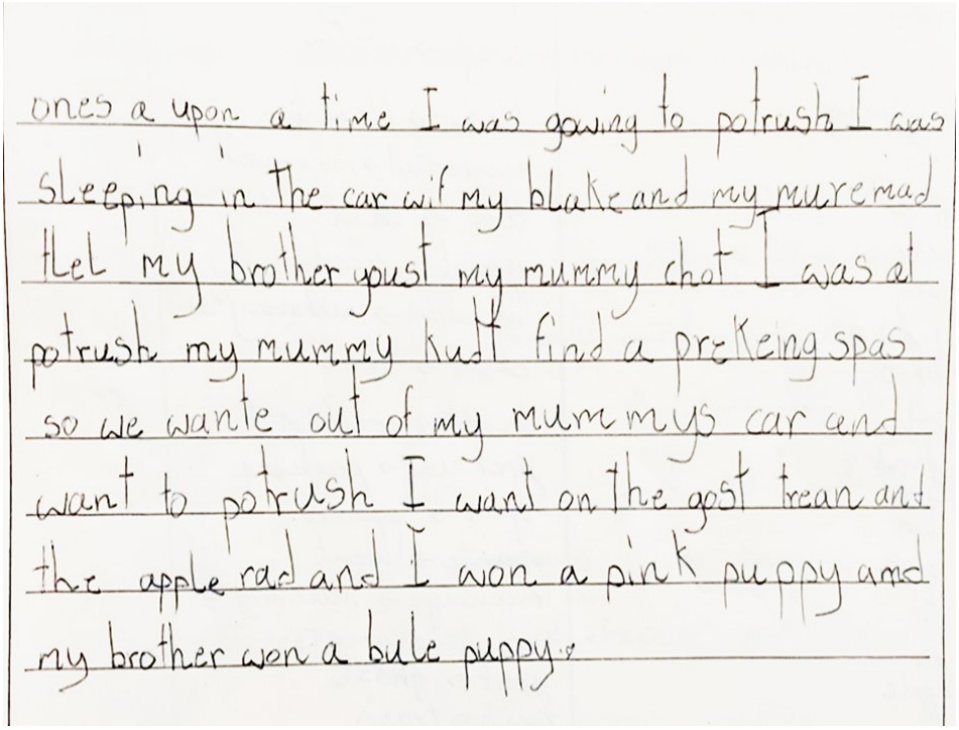


FIGURE 7
Category 3 Transcript needed to read some words (97 Writing samples: 36.3% of 267). Transcript: Once upon a time I was going to Portrush. I was sleeping in the car with my blanket and my mermaid tail. My brother used my mummy's coat. I was at Portrush. My mummy couldn't find a parking space, so we went out of my mummy's car and went to Portrush. I went on the ghost train and the apple ride, and I won a pink puppy, and my brother won a blue puppy (total words 75).

3.4 Category 4

There were 23 samples in this category, and they could all be read without the transcript, but children did not write at length. This means that there was not a wide range of vocabulary used. Spelling errors were still evident but the words being attempted were easily identifiable. Most of the words used in the category 4 samples were regular in their spelling, raising the possibility that children in this category wrote so little because they wanted to limit what they were writing to words they knew they could spell.

In Figure 8 there were three incorrect spellings: "aksed" (asked), "biled" (build), "finisht" (finished). Whilst there is some evidence of orthographic knowledge in this sample, for example, "home" evidence of orthographic knowledge is limited. Although "k" and "s" are in the wrong order in the word "asked," this incorrect spelling provides some evidence of orthographic knowledge. It has all of the correct letters needed to spell the word, and "ed" has been correctly added to make the past tense. When the word asked is spelled by sound the final letter is often written as a "t" because the end sound is often heard as a "t" sound, as seen in the misspelling of "ed" in the word "finisht" (finished). The misspelling of the bound morpheme "ed" in the word finished provides evidence that the child is spelling by sound and has not been taught or is not applying the orthographic rule consistently for adding "ed" to make the past tense of regular verbs.

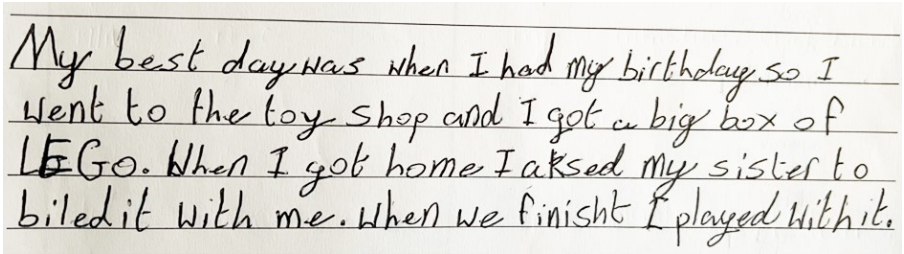


FIGURE 8
Category 4 Transcript not needed (23 writing samples: 8.61 of 267). Transcript: My best day was when I had my birthday, so I went to the toy shop, and I got a big box of lego. When I got home, I asked my sister to build it with me. When we finished, I played with it (total words 44).

The percentage of correctly spelled words that included orthographic choices was 23% of the total number of words spelled correctly. If the word "the" is not included in the orthographically correct spellings then the percentage would be 20%.

4 Discussion

Bell and McLean (2012) claimed that some dyslexic children have severe phonological difficulties that teaching phoneme-to-grapheme correspondence cannot address. The 17 samples in category 1 support this claim (RQ1). There is insufficient evidence of a small bank of high frequency words for this group of children as the only correctly spelled high frequency words were I, in, to, the, is, a (RQ2). The majority of spelling attempts appear as jumbled letters, bearing no resemblance to the actual words, except for two regular words big and dog in two of the samples. There is limited or no evidence of phonologically plausible spellings indicating that there is insufficient knowledge of phoneme-to-grapheme correspondence despite 4 years of schooling to develop this knowledge. This raises issues regarding the ethics of persisting with a phoneme-to-grapheme only approach such as SSP, when it is beyond the child's processing ability. To ensure a focus and reliance on phonemic knowledge the DfE (2023) core phonics criteria discourages the teaching of orthographic units (initial consonant clusters and onset and rime). It is, therefore, reasonable to suggest that these children have not been taught alternative strategies. Taking into account Treiman et al. (2019) findings that orthographic correctness at 5–6 years of age predicts later spelling ability, it is a concern that these children's spelling difficulties are entrenched to the extent seen in these samples. This suggests difficulty with orthographic processing and the development of orthographic knowledge in addition to the phonological difficulties. It is important to note that this group represents only 6.36% of all of the 267, 8–9 year old children in this study. These were children with the severest difficulties in literacy development in 143 schools.

Orthographic knowledge of what a word looks like is necessary for spelling accuracy but other factors can impact the acquisition of orthographic knowledge. Spelling errors (Category 2: Figure 6) such as "olir" for other (mispronunciation), besdie for best (colloquial) and dosit for doesn't (poor discrimination of phonemes) indicate that spelling errors can be a result of more than an orthographic deficit or a lack of knowledge of orthographic rules, and that a combination of factors should be considered including the child's pronunciation and/or auditory discrimination difficulties.

The phonologically plausible incorrect spellings used by the 224 children in categories 2 and 3 indicate that their primary strategy is phoneme-to-grapheme correspondence (alphabetic mapping) without the orthographic knowledge needed to select the correct letters for phonemes with multiple mappings. Orthographic knowledge was observed in less than 14% of correctly spelled words in the samples in Categories 2 and 3 suggesting an orthographic deficit (RQ3). The teaching of phoneme-to-grapheme correspondence did not enable the development of orthographic knowledge as observed by the prevalence of phonologically plausible spelling errors (RQ4). Because the spelling errors demonstrated knowledge of phoneme-to-grapheme correspondences, and a lack of awareness of what the words they were trying to spell should look like, these findings are consistent with Georgiou et al.'s (2021) argument that the orthographic deficit is as significant as the phonological deficit. The samples in category 2 and 3, do not support RQ2 as some orthographic knowledge had developed beyond a small bank of high frequency words. However, many of these were high exposure words, for example, the, was, mummy, daddy, love, football.

As seen in Table 1, in the 23 samples in category 4 there was a higher proportion of orthographic knowledge observed, and children in this group could spell more than a small bank of high frequency words. These children did not write at length, and the range of different words used may indicate that they were limiting what they could write to what they could spell correctly.

Are the issues identified in 2019 still current?

Six specialist teachers (ST), each with a Master degree in teaching and assessing children with dyslexia, were asked to consider the difficulties exhibited by children experiencing severe literacy difficulties that they were working with in the school year 2024/2025, to establish whether the authors' analysis of the 267 samples in this study, issues with reading rate (baseline assessment of reading rate (YARC) in October 2019), and the critique of the DfE (2023) core phonics criteria are representative of ongoing issues. The STs independently considered the statements in Tables 2–6 below. Only statements agreed by the STs were included in the tables and taken forward to a focus group meeting.

Table 2 shows the number and age groups of the children considered by each of the STs. The teachers were working with a total of 84 primary aged children in January 2025. The largest groups were children aged 7–8 years ($n = 23$) and 8–9 years ($n = 25$).

Findings from the independently completed questionnaire found that 83 children decode phoneme by phoneme with only one

TABLE 1 Average numbers per sample for: total words written, range of words used, correct and incorrect spellings.

Category	Total number of words written	Range of different words used	Correct spellings	Incorrect spellings
2	31	20	9	11
3	39	24	16	8
4	41	23	19	4

TABLE 2 Number and age groups of the 84 children considered by specialist teachers (ST) in January 2025.

AGE	5–6	6–7	7–8	8–9	9–10	10–11	Total 1
Year group	Y2	Y3	Y4	Y5	Y6	Y7	
ST1		1	4	2	1	2	10
ST2	2	1	6	5	2	4	20
ST3	1	1	4	9	6	1	22
ST4		2	1	2	4	2	11
ST5			1	7	5	1	14
ST6			7				7
Total 2	3	5	23	25	18	10	84

TABLE 3 READING DECODING AND BLENDING - Statements arising from concerns regarding DfE (2023) core phonics criteria and baseline assessment of 267 participants.

Question number	Age of children in each year group	5–6	6–7	7–8	8–9	9–10	10–11	Total
	Total number of children in each year group	3	5	23	25	18	10	84
1 A	Children decode words phoneme by phoneme when reading.	3	5	22	25	18	10	83
1 B	Children decode words phoneme by phoneme when reading. Because this is their "default" position, it is very hard to improve automatic sight word recognition using the same decoding phoneme by phoneme strategy.	3	5	22	24	16	10	80
1 C	Because decoding is their "default" position, it is very hard to establish new/different strategies.	3	5	15	22	16	10	71
2	Children don't recognise patterns in words because they have not been trained to look for these.	3	5	22	25	18	10	83
3	Children are entrenched in the habit of "sounding out" every letter in every word, even when they recognise the whole word.	3	5	13	22	11	5	59
4	Children are taught to decode words phoneme by phoneme and apply this to irregular words because they don't know when a word is regular or irregular.	3	5	20	21	13	9	71
BLENDING								
5	Children take each sound in turn all through a word by saying each sound individually, breaking the speech stream between each sound.	3	5	23	25	18	10	84
6	Children run each sound into the next sound ensuring that there is no break between each sound in the speech stream.	0	0	0	0	0	0	0

Specialist teachers' responses to these statements in relation to children they worked with in the school year 2024/2025.

TABLE 4 SPELLING - Statements arising from the analysis of the 267 samples of writing, and the issues noted in this paper regarding DfE (2013) see page 5, DfE (2014) see page 4 and DfE (2023) see pages, 5, 7 and 8.

Question number	Age of children in each year group	5–6	6–7	7–8	8–9	9–10	10–11	Total
	Total number of children in each year group	3	5	23	25	18	10	84
1	Children spell by identifying the sounds they hear in words	3	5	23	25	18	10	84
2	Children can "stretch out" a word orally, to segment the word into individual phonemes, without distorting the sounds in the word.	0	0	0	0	0	1	1
3	Children "stretch out" a word orally, to segment the word into individual phonemes, and often distort the sounds in the word, even at the level of regular CVC words	3	5	15	18	13	8	62
4	Children never "stretch out" the sounds in a word to segment the word into individual phonemes/sounds	0	0	8	7	5	1	21
5	Children observed spelling aloud don't name the letters, they say the sound	3	5	21	20	11	3	63
6	Children don't name the letters because they don't know the letter names?	3	5	9	10	7	2	44
7	Children observed spelling aloud who know the letter names and say the sounds, revert back to saying the sound even when encouraged to say the letter name	1	3	9	10	7	2	32
8	When spelling, the letters chosen for the phonemes represent the way the child pronounces the word. Poor articulation and pronunciation impact spelling by sound alone, because the child doesn't know what the word should look like.	3	5	20	25	18	10	81

Specialist teachers' responses to these statements in relation to children they worked with in the school year 2024/2025.

TABLE 5 SPELLING – Statements resulting from the use of phonologically plausible but incorrect choice of letters for phonemes with multiple mappings in the 267 samples of writing.

Question number	Age of children in each year group	5–6	6–7	7–8	8–9	9–10	10–11	Total
	Total number of children in each year group	3	5	23	25	18	10	84
1	Children can link the consonant phonemes to the letter when it is one to one mapping but mix up sounds represented by more than one letter e.g., spelling cat with a k - kat	1	0	16	13	15	7	52
2	Children have difficulty with consonant letters that have more than one sound, for example, y as in (yellow and happy), g (as in girl, giraffe) choosing the j sounds for g; c (as in circle and cat) choosing the s sound for c or vice versa	3	5	21	22	17	7	75
3	Children cannot choose the correct spelling for long vowel sounds and vowel digraphs because of the range of different spellings for these sounds	3	5	23	25	18	10	84
4	Children aged 8–11, recognise few, if any, of the vowel digraphs when presented with these in isolation. For example, "ee" may be the only digraph that is recognised				24	16	9	49/53
5	Children use phoneme-to-grapheme correspondence to spell, and choose letters that represent the phoneme, but it is the wrong choice of letters for phonemes with multiple mappings	3	5	23	25	18	10	84

Specialist teachers' responses to these statements in relation to children they worked with in the school year 2024/2025.

TABLE 6 Auditory discrimination- Statements arising from spelling errors in the 267 samples of writing that indicate that children may have auditory discrimination difficulties.

Question number	Age of children in each year group	5–6	6–7	7–8	8–9	9–10	10–11	Total
	Total number of children in each year group	3	5	23	25	18	10	84
1	Children cannot hear the difference between the "w" sound as in went and "wh" sound as in when and mix them up when spelling	1	4	21	23	16	7	72
2	Children have great difficulty choosing the correct letter for the short vowel sounds a, e, i, o, u even though the short vowel sounds do not have multiple mappings	3	5	22	22	16	10	78

Specialist teachers' responses to these statements in relation to children they worked with in the school year 2024/2025.

child able to recognize initial consonant blends as an orthographic unit, without the necessity to blend the individual phonemes. For 80 of the 84 children it was very hard to achieve sight word recognition using this decoding strategy because they were entrenched in the habit of "sounding out" every letter and could not recognize words by sight. Furthermore, for 71 children it was very hard to establish a new strategy because decoding phoneme by phoneme had been overlearned to such an extent. These children (71) applied this decoding strategy to attempt to read irregular words because they didn't know when a word was regular or irregular. 59 of the 84 children were entrenched in the habit of "sounding out" every letter in every word, even when they recognized the whole word. Only one child demonstrated the emergence of recognizing patterns in words.

In the focus group meeting the specialist teachers agreed that children had been taught to focus on phonemes only and had not been taught to look for letter patterns at onset and rime, syllable and even whole word level. One ST provided the example of a child reading the irregular word "the" making a sound for each letter "t", "huh" "eh." The method for teaching phonics was SSP for all 84 children.

All 84 children "sounded out" each phoneme in turn, breaking the speech stream. In the focus group meeting the STs agreed that this does not constitute blending, it was simply evidence of recognition of phonemes. One teacher expressed the view that taking each sound in turn does not develop word recognition, at the rate needed for reading fluency. There is often a delay after saying the sounds as the child tries to remember the word. Consequently, reading is very labored and pedantic. All of the STs were in agreement and each ST could identify a child or children who had to be told what the word said after the child had been successful in sounding out the phonemes. The sounding out did not facilitate word recognition for children with the severest orthographic difficulties.

All 84 children spell by identifying the sounds they hear in words. 62 children "stretch out" words to segment them and often distort the sounds, even for phonemically regular CVC words. Only one of the 63 children who stretch out words to segment them could do so without distorting the sounds in the word and 21 of the 84 children never stretched words out to segment them.

63 children say the sounds in the words not the letter names. 44 of these 63 children did not know all of the letter names. Letter sounds had been taught as labels for letters rather than letter names

because this is a requirement in SSP programmes. Children who knew the letter names persisted in saying the sounds even when encouraged to say the letter names. All of the STs agreed that letter names should be taught before letter sounds to avoid the confusion when letters have more than one sound, and to avoid the resulting disadvantage when spelling.

Spelling for 81 of the children was impacted by how the child pronounces words.

The STs were in agreement that the multiple mappings presented significant difficulties for all of the children. 52 of the 84 children could link the consonant phoneme to the letter when it was 1: 1 mapping. 75 children were reported to have difficulty with consonant letters that represent more than one sound. All 84 children use phoneme-to-grapheme correspondence to spell, and choose letters that represent the phonemes, but it is the wrong choice of letters for phonemes with multiple mappings. All 84 children could not choose the correct spelling for long vowel sounds and vowel digraphs because of the range of different spellings for these sounds. 49 of 53 children (8–11 years) recognized few, if any, of the vowel digraphs when presented with these in isolation. In the focus group meeting all of the STs agreed that "ee" may be the only digraph that is commonly recognized.

All of the specialist teachers agreed that teaching onset and rime patterns produced positive results when teaching spelling, but that greater benefits would have been achieved if this teaching had occurred earlier in their literacy development as they were now dealing with entrenched difficulties.

In the focus group meeting the STs agreed that difficulty discriminating sounds impacted spelling accuracy. 72 of the 84 children could not hear the difference between the "w" sound in the word went and the "wh" sound in the word when, and these are frequently mixed up when spelling. 78 children had great difficulty discriminating, and therefore choosing, the correct letter for the short vowel sounds a, e. i, o, u, even though the short vowel sounds do not have multiple mappings.

5 Conclusion

The writing samples in this study demonstrate over-reliance on phoneme-to-grapheme correspondence when spelling resulting in varying degrees of difficulty acquiring orthographic knowledge. This research does not support Treiman et al.'s (2019) conjecture that phonological plausibility may lead to orthographic correctness for poor spellers. When considered in conjunction with McMurray's (2004, 2020) claim that the optimum period for the development of orthographic knowledge is 5–8 years of age (years 1–4 NI), these Year 5 children (8–9 years of age) have passed this optimum period and may have been disadvantaged because they were taught to spell by sound, drawing on phoneme-to-grapheme correspondence only. Phonologically plausible spellings should not be considered progress and teachers should be alert to the possibility that a child of 5–6 years of age may have orthographic processing difficulties which will become more entrenched if there is not explicit structured teaching to develop orthographic knowledge.

It is important now, that a clear distinction is made between alphabetic mapping (phoneme-to-grapheme links) and orthographic mapping (letter-sound links representing onsets, rimes, syllables and irregular words). Children with the underlying cognitive processing ability to process orthographic knowledge (i.e., orthographic units) implicitly, learn to read easily and fluently, and abstract the orthographic knowledge needed for spelling accuracy, from their reading experience. Learning phoneme-to-grapheme correspondences is supported by a child's orthographic processing ability, and whilst there is no doubt that phoneme-to-grapheme correspondence should be taught first, observations from the 267 spelling samples suggest earlier intervention to develop orthographic knowledge may benefit this group. The DfE (2023) core phonics criteria advises against the inclusion of initial consonant blends and clusters within synthetic phonics programmes, that is, two or three phonemes recognized as an orthographic unit, but these could be included in SSP programmes to initiate first steps toward orthographic learning for children with working memory and/or orthographic processing difficulties. When phoneme-to-grapheme correspondences have been taught, teaching should move quickly to develop recognition of these early orthographic units within synthetic phonics. This is a lead in to the systematic and structured teaching of spelling.

Children who learn to read easily but are unable to spell correctly are children whose orthographic processing is sufficient for recognition of orthographic patterns when reading, but insufficient for recall when spelling. Children with strong orthographic processing read fluently and their spelling demonstrates orthographic correctness from a very young age as demonstrated by Treiman et al. (2019). As teaching phoneme-to-grapheme correspondence was insufficient for the development of spelling skills for the 267 children in this study and reading fluency is impacted by the over reliance on decoding when reading, it is reasonable to suggest that orthographic processing difficulties may be a cause of reading failure alongside the phonological deficit, and not a consequence of it.

It has been a weakness in the interpretation of the phonological deficit that account was not taken of the orthographic issues that arise due to multiple mappings between phonemes and graphemes, the complex syllable structure in English, and the impact these two factors have on memory storage and retrieval. The findings from this study support McMurray's (2020) claim that the acquisition of orthographic knowledge has been left to chance, with the expectation that children will acquire this knowledge through implicit learning when reading. It is clear that the children in this study with severe specific literacy difficulties of a dyslexic-type could not acquire orthographic knowledge from reading experience and that teaching phoneme-to-grapheme correspondences only is insufficient for its development.

6 Recommendations

Further research should consider whether early intervention to support the development of orthographic knowledge together with, and not instead of, teaching phoneme-to-grapheme correspondences can positively impact the spelling outcomes for children with severe literacy difficulties as seen in the

267 children in this research, and whether the observational findings in this research apply to the wider group of children with literacy difficulties in mainstream primary classrooms. McMurray (2004, 2006, 2020) provides robust evidence that learning to spell is the vehicle that can provide the systematic sequence and structure needed to significantly impact spelling accuracy through the development of vocabulary understanding, orthographic knowledge, and the associated links with phonology and morphology for all ability groups (impact factor 1.19). However, there were only two experimental and two matched control schools in this study. A larger sample is needed to ensure that there are sufficient numbers of children with severe literacy difficulties to determine to what extent it would proactively address their needs. It may, therefore, be beneficial to replicate McMurray's Ph.D research in a large number of mainstream primary schools with children 5–8 years of age, to ensure that the phoneme-to-grapheme knowledge that has been developed with children 4–5 years of age using SSP, is supported and utilized in the development of orthographic units that represent larger units of sound to enhance spelling accuracy and reading fluency. The average spelling score was almost one standard deviation above the mean (113) with all children in the experimental schools increasing their standardized score over the period of the research (McMurray, 2004, 2006, 2020). For this reason, future research should be within a whole class approach to improve the vocabulary knowledge and spelling performance of all ability groups within mainstream classes as greater success is achieved when intervention is proactive rather than reactive. Given the high impact factor in McMurray's Ph.D research the difficulties experienced by children with orthographic processing difficulties (up to 25% as evidenced by DfE statistics), who are unable to make satisfactory progress with systematic synthetic phonics only, should be addressed if this research is replicated with fidelity.

However, a large-scale study could be problematic in the UK given current DfE policy. There has been widespread adoption of SSP for reading and spelling throughout primary education following the Rose report (Rose, 2006) and the publication of the DfE's core phonics criteria (DfE, 2010, 2023). The Education Endowment Foundation (EEF) (2025) note that "*most schools use an SSP approach.*" This makes it extremely challenging to undertake robust research into the effectiveness of additional approaches because of the strict requirements set by DfE (2023) for phonics teaching and intervention for those who struggle.

Data availability statement

The datasets presented in this article are not readily available because access to the writing samples dataset is restricted to original researchers. Requests to access the datasets should be directed to SM, sharon@literacydifficulties.com.

Ethics statement

The studies involving humans were approved by Stranmillis University College. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

SM: Conceptualization, Writing – review & editing, Data curation, Methodology, Writing – original draft, Formal analysis, Investigation. MM: Formal analysis, Writing – review & editing, Methodology, Investigation, Data curation. PB: Formal analysis, Writing – review & editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Acknowledgments

This research would not have been possible without the co-operation and contribution of the Education Authority literacy service teachers and administrators.

Conflict of interest

The author (s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Alloway, T. P. (2011). *Improving working memory supporting Children's learning*. London: SAGE.

Bell, S., and McLean, B. (2012). "Good practice in training specialist teachers and assessors for people with dyslexia," in *Special educational needs: A guide for inclusive practice*, eds L. Peer and G. Reid (London: SAGE Publications), 127–140.

Bowers, J. S. (2020). Reconsidering the evidence that systematic phonics is more effective than alternative methods of reading. *Educ. Psychol. Rev.* 32, 681–705. doi: 10.1007/s10648-019-09515-y

Carroll, J. M., Holden, C., Kirby, P., Thompson, P. A., Snowling, M. J., The Dyslexia, et al. (2025). Toward a consensus on dyslexia: Findings from a Delphi study. *J. Child Psychol. Psychiatry* 66, 1065–1076. doi: 10.1111/jcpp.14123

Cassar, M., and Treiman, R., (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *J. Educ. Psychol.* 89, 631–644.

Daffern, T. (2017). Linguistic skills involved in learning to spell: An Australian study. *Lang. Educ.* 31, 307–329. doi: 10.1080/09500782.2017.1296855

Daffern, T., and Critten, S. (2019). Student and teacher perspectives on spelling. *Aust. J. Lang. Literacy* 42, 40–57. doi: 10.1007/BF03652025

Deheane, S. (2009). *Reading in the Brain*. Penguin Group (USA).

DfE. (2010). *Phonics teaching materials: core criteria and self-assessment*. London: Crown copyright. Available online at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.excaliburtsa.org.uk/excalfiles/phonics-partner-links/Training%20Modules/Module1/DfE%20phonics_core_criteria_and_the_self-assessment_process.pdf#:~:text=The%20self%2Dassessment%20provides%20information%20about,core%20criteria%20for%20a%20high&text=%C2%A9%20Crown%20copyright%202010 (accessed December 23, 2025).

DfE. (2013). *English programmes of study: Key stages 1 and 2 National curriculum in England*. New Zealand: Crown copyright.

DfE. (2014). *Statutory guidance National curriculum in England: English programmes of study*. London: DfE.

DfE. (2016). *Phonics screening check and Key Stage 1 assessments in England 2016*. Available online at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://assets.publishing.service.gov.uk/media/5a7f2099ed915d74e62288dd/SFR42_Phonics_KS1_2016.pdf (accessed January 16, 2026).

DfE. (2019a). *Phonics screening check and key stage 1 assessments in England, 2019*. London: DfE.

DfE. (2019b). *National curriculum assessments at key stage 2 in England, 2019 (revised)*. London: DfE.

DfE. (2023). *Essential core criteria for the validation of SSP programmes*. Online 1st April 2021. Available online at: https://www.gov.uk/government/publications/phonics-teaching-materials-core-criteria-and-self-assessment/validation-of-systematic-synthetic-phonics-programmes-supporting-documentation (accessed June 15, 2021 and January 15, 2026).

DfE. (2024a). *Phonics screening test attainment: England 2024*. London: DfE.

DfE. (2024b). *Key Stage 2 attainment*. London: DfE.

Education Endowment Foundation (EEF). (2025). *Phonics: Teaching & learning toolkit*. Available online at: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/phonics (accessed November 21, 2025).

Education Policy Institute [EPI]. (2024). *What can quantitative analyses tell us about the national impact of the phonics screening check?*. Available online at: https://epi.org.uk/publications-and-research/what-can-quantitative-analyses-tell-us-about-the-national-impact-of-the-phonics-screening-check/ (accessed November 5, 2024).

Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Read. Res. Quart.* 55, S45–S60. doi: 10.1002/rrq.334

Ehri, L. C. (2024). Clarifying the role of orthographic mapping in sight word reading. *Read. League J.* 4–13.

Fernandez, G., Vaquero, Jimenez, L., and Defior, S. (2011). Dyslexic children show deficits in implicit sequence learning, but not in explicit sequence learning or contextual cues. *Ann. Dyslexia* 61, 85–110. doi: 10.1007/s11881-010-0048-3

Georgiou, G. K., Martinez, D., Vieira, A. P. A., and Guo, K. (2021). Is orthographic knowledge a strength or a weakness in individuals with dyslexia? Evidence from a meta-analysis. *Ann. Dyslexia* 71, 5–27. doi: 10.1007/s11881-021-00220-6

Holden, C., Kirby, P., Snowling, M. J., Thompson, J. M., and Carroll, J. M. (2025). Towards a consensus for dyslexia practice: Findings of a delphi study on assessment and identification. *Dyslexia* 31:e1800. doi: 10.1002/dys.1800

Kilpatrick, D. A. (2020). *How the Phonology of speech is foundational for instant word recognition. Perspectives on language and literacy*. Pikesville, MA: International Dyslexia Association.

Koutsouris, G., Norwich, B., and Bessudnov, A. (2021). Interpreting RCT, process evaluation and casestudy evidence in evaluating the Integrated Group Reading (IGR). programme: A teacher-led, classroom-based intervention for Year 2 and 3 pupils struggling to read. *Educ. Rev.* 73, 346–373. doi: 10.1080/00131911.2018.1557597

Lavan, G., and Talcott, J. B. (2020). *Brook's what works for literacy difficulties*, 6th Edn. Available online at: http://www.thedyslexia-spldtrust.org.uk/7/news/391/6th-edition-of-what-works-now-available/ (accessed July 2, 2025).

Lennox, C., and Siegel, L. S. (1994). "The role of phonological and orthographic processes in learning to spell," in *Handbook of spelling: Theory, process and intervention*, eds G. D. A. Brown and N. C. Ellis (Chichester: John Wiley & Sons Ltd).

Mather, N. (2024). It's time for a multiple view of dyslexia. *Communique* 23, 22–24.

Mather, N., and Jaffe, L. (2021). Orthographic knowledge is essential for reading and spelling. *Reading League J.* 2, 15–25.

Mather, N., McCallum, R. S., Mee Bell, S., and Wendling, B. J. (2024). *'Tests of Dyslexia'*. California: Western Psychological Services.

Martinet, C., Valdois, S., and Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition. *Cognition* 91, B11–B22.

McMurray, S. (2004). *Learning to spell*. Ph.D thesis. Belfast: Queen's University.

McMurray, S. (2006). Learning to spell: Raising standards in spelling and independent writing. *Support Learn.* 21, 100–107.

McMurray, S. (2020). Learning to spell for children 5-8 years of age: The importance of an integrated approach to ensure the development of phonic, orthographic and morphemic knowledge at compatible levels. *Dyslexia* 4, 442–458. doi: 10.1002/dys.1663

McMurray, S. (2022). Why systematic synthetic phonics (SSP). Alone cannot meet the needs of all children. *Dyslexia Handb.* 2022, 124–126.

Miles, K. P., and Ehri, L. C. (2019). "Orthographic mapping facilitates sight word memory and vocabulary learning," in *Reading development and difficulties bridging the gap between research and practice*, eds D. S. Kilpatrick, R. M. Joshi, and R. K. Wagner (Switzerland: Springer).

PIRLS. (2021). *International results in reading*. PIRLS: Chestnut Hill, MA.

Rose, J. (2006). *Independent Review of the Teaching of Reading*. London: DfE.

Standards and Testing Agency. (2018). Available online at: https://assets.publishing.service.gov.uk/media/5c0e570d40f0b60bacefd1be/2018_key_stage_1_teacher_assessment_exemplification_expected_standard.pdf (accessed October 25, 2025).

Stothard, S. E., Snowling, M. J., and Hulme, C. (1996). Deficits in phonology but not dyslexic? *Cogn. Neuropsychol.* 13, 641-672.

Treiman, R., and Wolter, S. (2020). Use of letter names benefits young Children's spelling. *Psychol. Sci.* 31, 43–50. doi: 10.1177/0956797619888837

Treiman, R., Kessler, B., and Caravolas, M. (2019). What methods of scoring young children's spelling best predict later spelling performance? *J. Res. Read.* 42, 80–96. doi: 10.1111/1467-9817.12241

Wechsler, D. (2016). *Intelligence scale for children*, 5th Edn: Pearson.

Wechsler, D. (2017). *Wechsler Individual Achievement Test - Third UK Edition (WIAT-111 UK)*. Pearson Clinical.

Wright, D.-M., and Ehri, L. C. (2007). Beginner's remember orthography when they learn to read words. The case of doubled letters. *Appl. Psycholinguist.* 28, 115–133.

Wyse, D., and Goswami, U. (2008). Synthetic phonics and the teaching of reading. *BERJ* 34, 691–710. doi: 10.1080/01411920802268912

Wyse, D., and Hacking, C. (2024). *The Balancing Act: An Evidence -Based Approach to Teaching Phonics, Reading and Writing*. Abingdon: Routledge.

YARC. (2011). *York assessment of reading for comprehension*, 2nd Edn. London: GL Assessment.

Zarić, J., Hasselhorn, M., and Nagler, T. (2021). Orthographic knowledge predicts reading and spelling skills over and above general intelligence and phonological awareness. *Eur. J. Psychol. Educ.* 36, 21–43. doi: 10.1007/s10212-020-00464-7

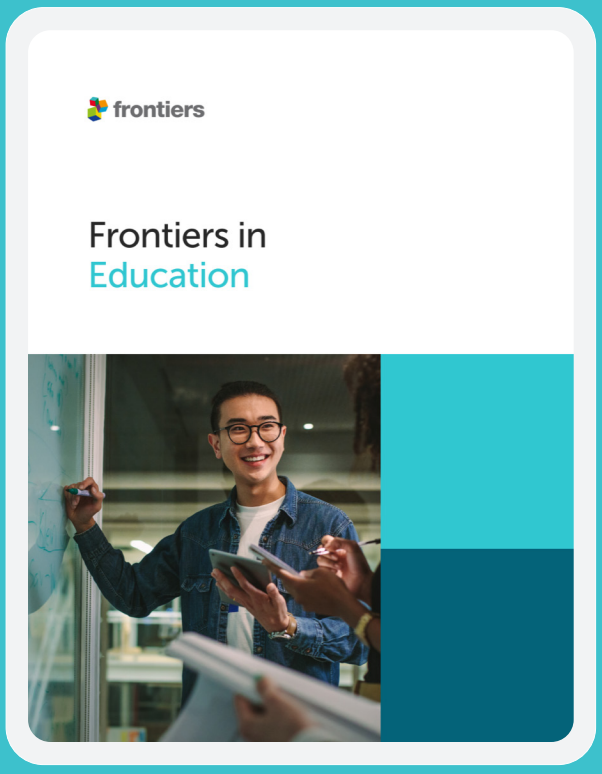
Frontiers in Education

Explores education and its importance for individuals and society

A multidisciplinary journal that explores research-based approaches to education for human development. It focuses on the global challenges and opportunities education faces, ultimately aiming to improve educational outcomes.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact