frontiers in Education

TYPE Original Research
PUBLISHED 01 April 2026
DOI 10.3389/feduc.2026.1727782

Check for updates

OPEN ACCESS

EDITED BY
Geoff Lindsay,
University of Warwick, United Kingdom

REVIEWED BY
Saodat Gulyamova,
Bukhara State University, Uzbekistan
Hadeel Al-Dawsari,
Princess Nourah bint Abdulrahman University, Saudi Arabia

*CORRESPONDENCE
Omar Abdullah Alsamani
o.alsamani@uoh.edu.sa

RECEIVED 21 October 2025
REVISED 14 February 2026
ACCEPTED 12 March 2026
PUBLISHED 01 April 2026

CITATION
Alsamani OA and Alsamiri YA (2026) The effects of AI-based visual instruction on the reading comprehension of students with dyslexia in Saudi Arabia: a single-case experimental study. Front. Educ. 11:1727782. doi: 10.3389/feduc.2026.1727782

COPYRIGHT
© 2026 Alsamani and Alsamiri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The effects of AI-based visual instruction on the reading comprehension of students with dyslexia in Saudi Arabia: a single-case experimental study

Omar Abdullah Alsamani[1]* and Yasir Ayed Alsamiri[2]

[1]Department of Special Education, College of Education, Hai'l University, Hail, Saudi Arabia, [2]Department of Education, Islamic University of Madinah, Madinah, Saudi Arabia

Students with learning disabilities (LD), particularly dyslexia, often face significant challenges in reading comprehension that traditional instruction may not fully address. Generative Artificial Intelligence (GenAI) offers emerging opportunities to provide personalised and visual instructional support to bridge these gaps. This study investigated the effectiveness of a GenAI-based visual instruction to support the reading comprehension skills of students with dyslexia. A single-case experimental design using a multiple-probe across participants approach was employed with three male students aged 9–11 years diagnosed with dyslexia. All three students were attending a public elementary school in Yanbu city, Saudi Arabia. The intervention utilised ChatGPT-generated visual explanations aligned with the Grade 4 Arabic language coursebook. Five-point quizzes were implemented with the students to assess their reading comprehension of paragraphs of text taken from the Arabic language coursebook. All participants started with an initial low and stable baseline (0%–10%). Mastery in the treatment phase was determined by a score of 80% or higher on the assessments in three consecutive sessions. Using this criterion, two participants showed immediate level changes and the third showed a clear accelerating trend. Maintenance probes were conducted several weeks following the conclusion of the intervention phase, during which all three students showed sustained performance at or above mastery. Visual analysis indicated a strong functional relation, minimal or no phase overlap, and durable effects. These findings provide preliminary evidence for the utility of GenAI-enhanced visual instructional tools to support reading comprehension for students with dyslexia.

KEYWORDS
dyslexia, generative artificial intelligence (GenAI), innovation in education, learning disabilities, reading comprehension, single-case design, special education

Introduction

Advances in technology and artificial intelligence (AI) have transformed educational systems worldwide in recent years. AI technologies are considered powerful tools that can enrich instruction, personalize learning experiences, and at the same time respond to the challenges that students face in their education. It seems that the potential for this type of

digital technology might be especially noticeable at the elementary level, where early intervention and evidence-based practices could play an important role in shaping students' achievement over time, particularly for those with disabilities. In many traditional classrooms, students with learning disabilities (dyslexia) struggle not only with acquiring basic skills such as reading comprehension but also with finding the kind of supportive environment that recognizes their individual needs and nurtures their abilities. GenAI apps can make learning more engaging and give students support that fits their specific individualised needs (Kumari, 2025; Sakowicz and Hamidi, 2025). Owing to their adaptability as well as their rapid capacity for data processing, GenAI systems are viewed as offering instruction that is both highly efficient and easily scalable (Chavez and Palaoag, 2024).

Learning disabilities and dyslexia

Simply put, learning disabilities are a group of information processing disorders that are neurodevelopmental in nature (American Psychiatric Association, 2013, 2025). Although the terms "learning disability" and "specific learning disorder" are often used interchangeably, the former is the terminology used in education to determine what types of services a student should receive to support their learning, and the latter is the term used when diagnosing the individual (American Psychiatric Association, 2025). The diagnostic term, as stated in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013), is used to cover three areas of learning disability, which are reading (dyslexia), writing (dysgraphia), and mathematics (dyscalculia). In addition to these three categories, for the purposes of supporting a student in the school setting, consideration is also given to auditory processing disorder, language processing disorder, nonverbal learning disorders, and visual/perceptive motor deficit, even though these four additional terms are not included in the DSM-5 and certain other lists of learning disabilities (Cleveland Clinic, 2024; Mayo Clinic, 2023; Visser et al., 2024).

Dyslexia, as a one area of learning disability, is defined as a specific learning disability of neurobiological origin. A person with dyslexia has difficulties with accurate and fluent word recognition, as well as poor spelling and decoding abilities. These challenges are linked to a deficit in the phonological component of language (Lyon et al., 2003). They may also demonstrate problems in reading comprehension and reduced reading experience (Visser et al., 2024). For students with learning disabilities in language (dyslexia), problems with phonological awareness, vocabulary learning, and also syntax can make it very difficult for them to read and to understand texts in a proper way. Since reading comprehension is not only the basis of language arts but also an essential part of learning in all other subjects, this type of learning disability carries consequences that may extend widely to affect the overall academic success of the student (Al-Dosari and Al-Subaie, 2022; Fletcher et al., 2018).

The manifestation of these reading comprehension difficulties is deeply influenced by the linguistic and orthographic nature of the student's native language. The Arabic language, for instance, is a morphologically rich language that utilizes a complex root-and-pattern system, demanding strong phonological and visual-orthographic processing skills (Al Ghanem and Kearns, 2015). Furthermore, the Arabic script is cursive and ligated. For example, letters may change their shape based on their position within a one single word, which adds a significant visual-perceptual load for readers (Asaad and Eviatar, 2014). For students with dyslexia, these features may intensify difficulties in word recognition and delay the development of reading fluency and comprehension (Shahbari-Kasem et al., 2024).

This challenges of reading compression is particularly evident among Saudi students with learning disabilities (particularly dyslexia), whose primary language is Arabic. While Modern Standard Arabic (MSA) is the formal language used in textbooks and classroom instruction in Saudi Arabia, these students communicate daily using a regional dialect (such as the Najdi dialect). This linguistic distance between the spoken vernacular and the written standard requires students to perform a mental translation while reading, which can add more pressure on those students with dyslexia (Al-Dosari and Al-Subaie, 2022; Shahbari-Kasem et al., 2024).

It seems that when reading difficulties are not identified in time, students may often struggle with low self-esteem, feelings of anxiety, and, in certain cases, even depression (Alsamiri and Alsamani, 2023; Chavez and Palaoag, 2024). Learning disabilities are more likely than others to remain undiagnosed and unaddressed. One possible reason for this higher rate of non-diagnosis is that many individuals develop compensatory strategies to conceal their challenges, such as relying heavily on memorization, leaning on peer assistance, or rehearsing content until it appears mastered (Westmoreland, 2025). Such practices do not erase the difficulty but instead place a veil over it, delaying recognition, intervention, and the hope of proper support. When intervention is delayed, the challenges posed by learning disabilities may persist into adulthood, which can limit success in post-secondary settings and career, as well as cause ongoing psychological stress, including depression and anxiety (Burenkova et al., 2021; de Beer et al., 2014; Salcido, 2025). Additionally, these disabilities frequently interfere with the development of both receptive language (reading and listening) and expressive language (writing and speaking) (Alsamiri and Alshammari, 2025; Wendling, 2019).

It is critical for students with learning disabilities and difficulties to identify and support them at an early stage. Teachers also need instructional support to help them recognize these students' needs and to apply the most effective practices to improve those students' learning (Alsamani et al., 2023; Alsamiri and Alhassani, 2021; Dani et al., 2024; Kocsis, 2016). In this context, enhancing reading comprehension stands as a matter of teaching technique and as a demand of fairness. It reminds us that equity in learning begins with the ability to understand text for all students when achievable.

Artificial intelligence in education

AI is defined, in very general terms, as computer-based systems that try to imitate some of the human thinking processes (Irgashevna, 2025). Within the field of education, AI

shows up in many different forms. For example, sometimes shows as expert systems, sometimes as natural language tools, sometimes in speech recognition or computer vision. It works with each case by interacting with learners and creating adjustable responses based on the interaction (Irgashevna, 2025). Classrooms are increasingly being opened for these capabilities of AI to assist with a variety of instructional tasks, including lesson customization and the delivery of feedback (Alghamdi et al., 2023; Alshehri, 2023).

The current increasing involvement of AI in the educational setting reflects the growing emphasis on such technology to enhance teaching effectiveness and learner engagement. AI-powered systems have the ability to watch the behaviour of students, find the learning gaps, and then change the content automatically to fit with every student's needs. For those students who struggle, such intelligent tools provide both immediate scaffolding and ongoing support, thus decreasing the chance of withdrawal and disengagement from their studies. AI steps can be seen as the quiet assistant that takes away the heavy weight of repetitive duties. This may assist teachers turn their energy toward higher-order teaching and toward building the essential relationships with their students (Samala et al., 2025; World Economic Forum, 2025).

Recent research has confirmed that, when artificial intelligence is integrated in a thoughtful and purposeful way within educational contexts, it contributes to noticeable improvement in learning outcomes. For example, Irgashevna (2025) demonstrated that real-time AI dashboards help teachers adapt instruction more easily and responsively, while Samala et al. (2025) found significant gains in student motivation and achievement following the introduction of adaptive reading platforms. These findings emphasize the high potential of AI for addressing diverse learner needs, particularly in inclusive classroom settings.

Despite the growing enthusiasm for AI-enhanced instruction, there is a lack of research on whether these tools create a sustained or consistent effect on the individual student. For example, many of the located studies highlight general applications of these systems, including technology such as adaptive text, virtual tutors, and chatbots; very few of the articles identified provided empirical evidence of the instructional effectiveness of these tools (He, 2024; Lee and Kwon, 2024). This is particularly true in the area of reading comprehension, where the interventions investigated have tended to be short-term, small-scale, and lacking in data on the effect of the treatment on individual students with learning disabilities (Lee and Kwon, 2024).

Rationale and objectives of the research

Educational researchers in the Gulf region and the larger Arab World have yet to catch up with the pace of global inquiry (Al-Batal and Al-Arafaj, 2021). Although research conducted in Kuwait (Al-Mutairi et al., 2024), Oman (Maghawry, 2024), and Saudi Arabia (Al-Zahrani, 2024) provide promising insights, these studies were limited in terms of duration, rigor of design, and/or sample size. Specifically, very few studies in the region have examined whether AI reshapes reading comprehension over time or how it functions in the real-world school environment with learners with diagnosed learning disabilities, including dyslexia (Al-Mutairi et al., 2024).

Therefore, this study endeavored to address a critical gap in the literature by employing a single-case, multiple-probe design within a Saudi elementary school for boys to examine the effectiveness of an AI-based reading intervention. The study traced the individual patterns of reading comprehension of the three participants, each of whom had not responded to traditional instructional approaches. Grounded in the belief that every learner has unique needs, and that adaptive technology can offer instruction tailored to these differences, the authors sought to examine whether AI-enhanced applications could serve not only as digital assistants but also as transformative instruments for the development of literacy among students with special educational needs.

The researchers aimed to evaluate the functional relation between AI-based visual instruction and measurable reading comprehension gains over time in three students diagnosed with dyslexia. In addition to analyzing overall performance improvements, the design enabled examination of the immediacy of effect following the introduction of the treatment and whether any improvement to comprehension skills was sustained through a subsequent maintenance phase. The following research questions guided the study:

1. Is there a functional relation between GenAI-based visual instruction and reading comprehension performance in students with dyslexia?
2. Does the reading comprehension of students with dyslexia improve immediately and consistently following the introduction of GenAI-based instruction?
3. Are reading comprehension skills of students with dyslexia maintained over time after the withdrawal of the GenAI-based instructional support?

Methods

Research design

The study was conducted over 32 weeks in 2024; it employed a single-case experimental design. Specifically, a multiple-probe design across participants was used to establish a functional relation between an independent variable and a dependent variable. Multiple-probe design is considered particularly well-suited to educational interventions in which continuous measurement during baseline is unnecessary or may lead to participants becoming disinterested or bored. This study design allowed for experimental control to be maintained while implementing the intervention across participants and limiting practice effects (O'Neill et al., 2022).

Setting

In Saudi Arabia, all students with learning disabilities are educated in the general education classroom for most of the day and then receive additional support from a special education

teacher with specialization in learning disabilities in a dedicated resource room as needed or during regularly scheduled sessions. The setting chosen for the study was a public elementary school in the city of Yanbu, which is located in Medina Province in western Saudi Arabia. The participating school was purposively selected because it had facilities to support students with learning disabilities, such as a resource room, and due to the authors' familiarity and existing relationship with the program. The selection of one setting for all participants helps increase experimental control in the study setting (Kazdin, 2021; Ledford and Gast, 2018). At the time of the study, 16 of the 190 students in the school (8.42%) had been identified as having a learning disability; each of these students regularly received additional support in the resource room. The resource room was organised in a way that reduced distractions (e.g., lighting, ventilation, background noise). This controlled setting ensured procedural integrity across all phases of the study and all participants.

Participants

Purposive sampling with predefined inclusion criteria were followed for recruiting participants for this study. Random sampling is typically employed in group designs to enhance external validity; however, SCED relies on replication logic rather than sampling logic to demonstrate functional relations (Horner et al., 2005). While limiting the immediate generalisability to other settings, this approach ensured that the observed effects could be attributed to the intervention rather than variance between school environments. These inclusion criteria established prior to the start of the study were: (a) students enrolled in one of the upper elementary grades (Grades 4–6), (b) who were aged 9–11 at the time of the study, (c) with a documented diagnosis of a learning disability (dyslexia) that affects reading comprehension, and (d) who were not receiving any concurrent intervention targeting reading comprehension during the study period.

These criteria were verified through a review of school records and in consultation with the students' Arabic language teachers and their parents, including through the use of an Interview Protocol that was administered to gather demographic and other information on each student and his family. Initial diagnostic information was supplemented with structured interviews conducted with both teachers and parents to confirm the persistence and level of the students' reading comprehension difficulties and to rule out the presence of additional confounding factors. All three participants were in the fourth grade when conducting this study. Their ages ranged from 9 years/4 months to 9 years/10 months (see Table 1).

Ethical considerations

We obtained the ethical approval from the Institutional Review Board of Hail University [H-2025-808], and administrative permission was granted by the Saudi Ministry of Education before any data collection. Participation in the study was voluntary. Once the three participants were confirmed as meeting the inclusion criteria, written consent was obtained from their parents and verbal consent was obtained from each student prior to the start of the study. To protect the privacy of the students and their families, all data were anonymised and pseudonyms were assigned to each child; the pseudonyms used were "Ahmed," "Badir," and "Ali." The researchers informed the participants and their families that they could withdraw at any time from the experiment.

Materials and instruments

A structured form for participant selection was used to write demographic information and to check if the inclusion rules were followed.

The primary intervention tool was ChatGPT (specifically the GPT-4 model). This generative AI application was selected for its advanced multimodal capabilities, which was applied to generate visual explanations and illustrative materials carefully adapted to the content of each reading passage. Unlike other tools, this integration ensured high consistency between the Arabic learning material and the visual explanations.

The visual representations used in this study were generated through ChatGPT based on a standardized set of instructional criteria rather than a single fixed prompt. This technique helps us to ensure that the AI-generated output was consistently aligned with the specific linguistic and contextual requirements of each Arabic text segment. The general rules we applied for these prompts were: (a) maintaining a literal and concrete depiction of the narrative's key events to minimize abstract interpretation, (b) ensuring the visual elements directly reflected the vocabulary of the reading passage, and (c) excluding any extraneous visual information that might distract the student from the primary textual meaning. By focusing on these standardized parameters, we were able to create a consistent instructional experience across all participants while allowing the GenAI to adapt to the unique content of each story. These digital visuals were accessed by participants using a tablet device with internet connectivity.

In addition to digital tools, researcher-developed quizzes were employed to monitor and evaluate reading comprehension, support the implementation, the monitoring, and the evaluation of the study. These included a Baseline Assessment Test that

TABLE 1 Participant demographics.

Participant (Pseudonym)	Gender	Age (Years; Months)	Grade Level	Diagnosis
Ahmed	Male	9;4	4th Grade	Specific Learning Disability (Dyslexia)
Badir	Male	9;7	4th Grade	Specific Learning Disability (Dyslexia)
Ali	Male	9;10	4th Grade	Specific Learning Disability (Dyslexia)

was administered to establish each student's initial reading comprehension ability, which served a dual purpose: it confirmed the persistence of comprehension difficulties and helped define the instructional focus for each participant.

To construct these, stories from four units of the Grade 4 Arabic textbook *My Language*, which included "Rain," "The Clever Rabbit," "The Clay Jar," and "My City." Each text was divided up into short, age-appropriate paragraphs and five questions were developed for each. The five-question quizzes served both instructional and evaluative functions, forming the core of each training session and each post-session performance assessment. Participant responses were scored with a "1" for the correct answer and a "0" for the incorrect answer. We converted the scores to a percentage on each assessment to come up with a consistent, easy-to-read measure to compare comprehension across the three phases (baseline, intervention, and maintenance). This also helped us enable precise measurement of each participant's change in performance over time, which is essential to establishing a functional relationship between the independent and dependent variables in such a type of research design. The use of session-level scores also supported visual analysis of changes in level, trend, immediacy of effect, and consistency over time, which can be observed in the final results' visual presentation.

Procedures

The multiple-probe design employed in this study integrated the AI-visual supported instructional approach as the independent variable and reading comprehension was as the dependent variable. All sessions were conducted in a one-to-one format within a dedicated resource room to ensure a controlled environment with minimal distractions. Each session lasted 45 min, consisting of 30 min of guided instruction and 15 min for the delivery of the assessment.

The sequence of each session began with a structured presentation of the assigned reading passage, followed by the introduction of the AI-generated visual explanation. During the instructional phase, the instructor guided the student to connect visual elements to the text without providing answers to the assessment questions. Subsequently, the student was asked to answer five written reading comprehension questions, each targeting a distinct sub-skill (e.g., making inferences, finding the main idea, or synthesizing information).

The sequence of each session began with a structured presentation of one of the assigned reading passages and was followed by a visual explanation generated through the AI application. Participants were then asked to answer five written reading comprehension questions, each of which targeted a distinct reading comprehension sub-skill (i.e., making inferences, finding the main idea, or synthesizing information). The same structure was maintained throughout the study for each question on the assessments to ensure procedural consistency. The sessions occurred in a one-to-one format within a controlled setting, with no additional teacher prompting or external instructional support provided during the baseline, treatment, or subsequent maintenance phases. The application of the AI tool for generating visual instructions followed the standardized parameters and remained consistent across all participants, while the variation was limited to the particular paragraphs employed in each session.

Structure and phases of the study

All sessions throughout the study were conducted individually in a quiet, controlled environment with consistent lighting, ventilation, and minimal ambient noise. Reading passages were adapted from four units of the Grade 4 Arabic textbook (My Language) previously described in the materials section. Since each of the students had been diagnosed as having a learning disability (dyslexia), they usually receive daily educational support in a designated resource room with a special education teacher with a specialization in LDs. To accommodate the research, part of these sessions were devoted to conducting the study. The interventionist was the students' regular special education teacher who worked with them in the resource room.

The dependent variable in the study was reading comprehension performance, operationalised as the percentage of correct responses on a 5-point probe, which assessed a range of core comprehension subskills, including literal understanding, inferential reasoning, identification of main ideas, recognition of key details, and integration of information. The independent variable was the instructional intervention using ChatGPT, which was designed to improve the reading comprehension skills of the study participants.

For the initial baseline phase, probes were administered intermittently to verify stability, which was defined as no upward trend and minimal variability, before the treatment was introduced. The intervention was introduced in a staggered sequence to establish experimental control. All sessions were one-to-one in the same resource room; no instructional feedback was provided during probes or post-reading assessments, and the participants received no concurrent comprehension interventions. Mastery on the probes was defined *a priori* as ≥80% accuracy on three consecutive sessions. After mastery was achieved, maintenance probes were scheduled without instructional support at preset post-criterion sessions to test retention. Procedural fidelity was checked in one-third of sessions (98%–100%; M = 99%), and interobserver agreement averaged 99%, supporting internal validity. The specific structure of each phase is presented in the following sections.

Baseline phase

During the baseline phase, probes were administered intermittently to verify stability, which was defined as no upward trend and minimal variability, before the treatment was introduced. This was done to determine each participant's initial level of reading comprehension as well as to confirm their suitability for the study. To preserve objectivity, no instructional feedback, prompting, or reinforcement was provided. In this phase, after reading the text, participants were verbally asked five reading comprehension questions targeting essential subskills, including literal understanding, inference, main idea identification, detail recognition, and information synthesis. Students either read independently or received minimal assistance to sustain engagement. Responses were recorded in real time using standardised observation forms. For each 5-point

probe, correct responses were scored "1" and incorrect responses "0," after which a session-level percentage was calculated. Baseline data followed multiple-probe logic (spaced/staggered sessions) to reduce fatigue. Data collection continued until a stable pattern of responding was documented for each participant (no upward trend, minimal variability), ensuring that subsequent changes could be attributed to the intervention rather than chance or external influences.

Intervention phase

As noted, the intervention phase was commenced in a staggered fashion. Therefore, it was begun for Ahmed at Session 5, Badir at Session 13, and Ali at Session 20. These start points were determined based on the stability of each participant's baseline data. To illustrate, the intervention was introduced to the first participant only after a stable baseline was established at Session 5, and subsequent participants entered the intervention phase in a staggered fashion once the previous participant demonstrated a clear therapeutic effect (Session 13). The instructional intervention was delivered one-to-one with each participant and involved a number of structured sessions, each of which focused on one reading skill (e.g., making inferences, finding the main idea, or synthesizing information). In this phase, AI application 'ChatGPT' served as the primary tool for producing contextualised illustrations and explanatory aids to enhance the reading comprehension of the participants. Again, the materials utilised were drawn from the My Language series. AI application 'ChatGPT' was used to generate tailored visual representations of the content, which were displayed to each participant on a tablet with internet connectivity.

Each session in the intervention phase lasted approximately 45 min; the first 30 min involved guided reading supported by GenAI-generated visuals, and the remaining 15 min were used to present the assessment, which comprised a new paragraph and a corresponding five-question quiz that was scored using the same 0- or 1-point standard that had been employed in the baseline. Percentage scores were calculated for each session. The environment was kept the same in all treatment sessions to reduce the effect from outside things and to make a place that helps students focus and stay engaged. During the assessment parts of the treatment sessions, no feedback and no reinforcement were provided, in order to guarantee that any change in performance could be linked only to the instructional content itself. The intervention continued for each participant until one of two termination criteria was met: (a) attainment of a comprehension accuracy rate of four to five correct answers (80%–100%), or (b) demonstration of stable performance across three consecutive sessions at or above criterion (80%–100%).

Maintenance phase

Following the conclusion of the treatment phase, post-intervention maintenance probes were conducted at scheduled intervals to assess the durability of each participant's skills acquisition. These sessions occurred for (a) Ahmed in Sessions 17, 23, and 25 (+5, +11, and +13 sessions after Session 12); (b) Badir in Sessions 23 and 31 (+5 and +13 sessions after Session 18); and (c) Ali in Session 31 (+3 sessions after Session 28). The variation in these session numbers reflects the individualised progress of each participant through the study phases. This structure enabled visual analysis of level, trend, immediacy of effect, overlap, and maintenance while avoiding withdrawal of a beneficial intervention. The maintenance sessions involved no additional instruction or visual support and followed the same assessment procedures as the prior phases.

Ensuring reliability and validity

To ensure procedural fidelity and data integrity, several oversight tools were employed. Procedural fidelity was checked in 33.33% (five out of fifteen) intervention sessions by using a checklist made by the researcher and filled by another trained person (see Appendix 1). This observer was a PhD candidate with twelve years of teaching experience with children who have learning disabilities, and he was also employed at the same school where the study was conducted. The observer ensured the sessions followed the structure, the data were recorded correctly, and the protocol was applied. Across the sessions that were watched, fidelity never faltered—rising and holding between 98% and 100%, with an average of 99%, like a rhythm that almost never missed a beat. Interobserver agreement for both procedural fidelity and students' response scores was calculated using the smaller/larger × 100 method and averaged 99%, indicating high consistency between observers. The combination of fidelity checks, consistent environmental controls, and visual data analysis provided a rigorous basis for evaluating the effectiveness of the intervention.

Results

This section presents the individual findings for each participant using visual analysis of the key indicators of changes in level, trend, variability, immediacy of effect, data overlap, and consistency across phases (see Figure 1). These indicators aligned with the research questions, which focused on the functional relationship between the intervention and reading comprehension outcomes, immediacy and consistency of performance change, and maintenance of skills following the intervention.

Findings related to Ahmed

Ahmed's scores in his baseline (Sessions 1–4) were uniformly 0%, with flat level, no upward trend, and negligible variability, establishing a stable baseline and confirming the absence of the target skill. With intervention onset in Session 5, performance increased immediately to 30% (strong immediacy of effect), then rose to 50% in Sessions 6–7, 60% in Session 8, and 80% in Session 9. He exceeded criterion in Sessions 10–12, scoring 90% on each, yielding a clear change in level from a baseline mean of 0% to an intervention-phase mean of 67.5% (median shift: 0% → 70%). Data patterns show an accelerating positive trend during early intervention, followed by a stable high plateau at mastery. Non-overlap with baseline was complete (all intervention and maintenance points exceeded the 0% baseline range), and variability decreased as performance approached criterion, indicating strengthened control by the independent variable.

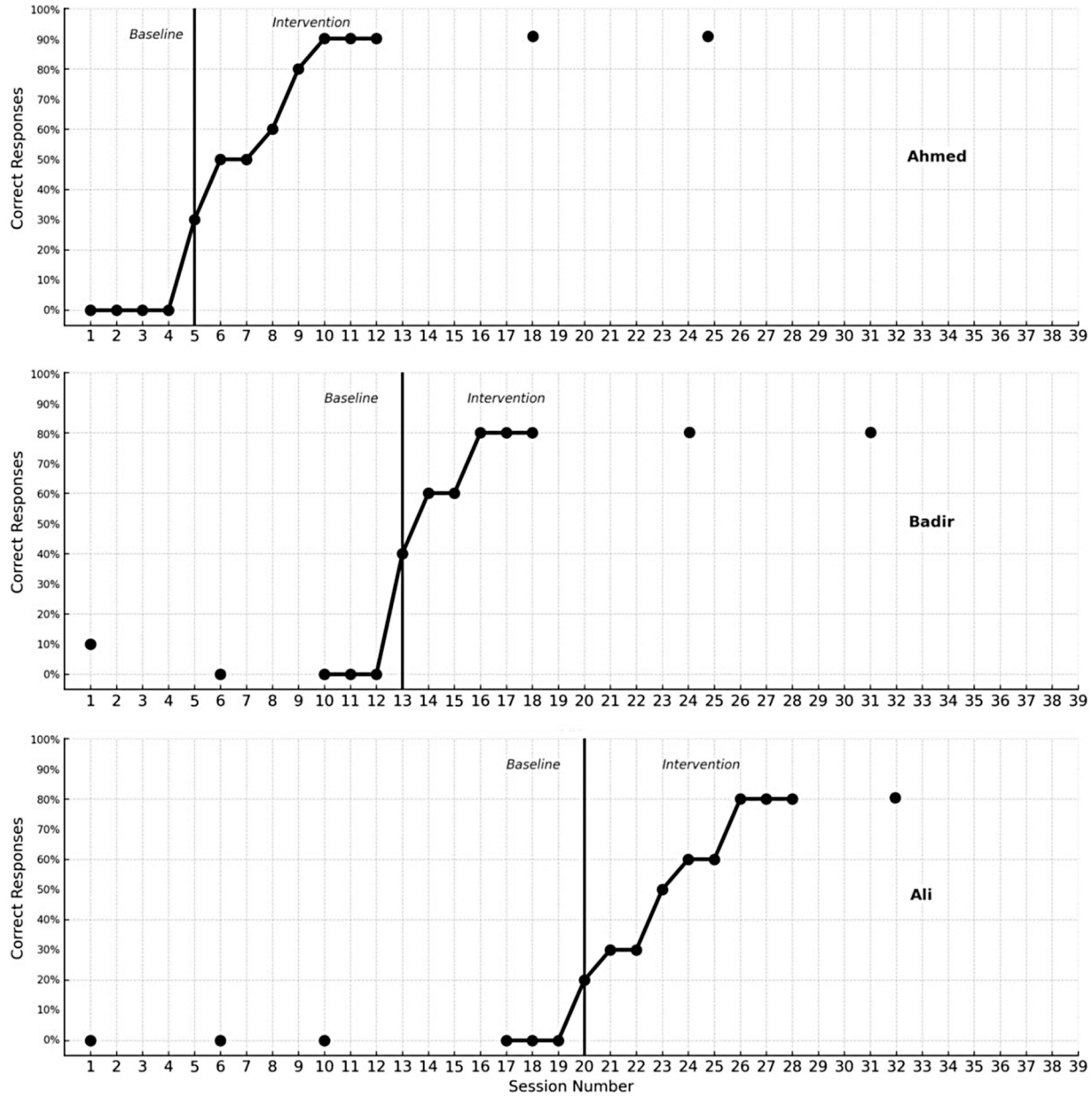

FIGURE 1
Scores across phases and sessions for each participant.

On maintenance probes administered without instructional support in Sessions 17, 23, and 25, Ahmed scored on each 90%, demonstrating durable retention with no reversion toward baseline across spaced follow-ups (+5, +11, +13 sessions post-intervention). Taken together, these findings of immediacy, level change, zero overlap, consistent trend toward mastery, and stable maintenance provide strong evidence of a functional relation between the GenAI-generated visual instruction and Ahmed's reading comprehension performance.

Findings related to Badir

Badir got 10% in the first session, but then he scored 0% in Sessions 6, 10, 11, and 12. These data indicate low and variable performance in the baseline that was consistently below the mastery threshold, which confirmed he lacked the targeted comprehension skills and established a stable baseline for evaluating the effects of the intervention. The treatment for Badir started in Session 13, and right away he showed better result with 40% in that first session. This was then followed by a steady rise in his performance, as he achieved 60% in both Session 14 and 15. By Session 16, he had reached the mastery level of 80% and kept it in Sessions 17 and 18. It seems that Badir's maintenance checks, which were done in Sessions 23 and 31, tend to show that he was able to keep the skills after the intervention. To clarify, he stood firm in both probes at or above the 80% mark. This shows the gains carved during intervention had settled into lasting memory even without new teaching delivered to Badir.

Findings related to Ali

Ali's starting performance was low and stayed the same that there was no sign that it got better by itself. He obtained a score of 0% across six nonconsecutive baseline sessions (sessions 1, 6, 10, 17, 18, and 19) that reflects a flat trend without any variability between sessions. His progress line lay silent and unbroken, showing no spark of spontaneous change. This confirmed the absence of the target reading comprehension skill and provided a solid foundation for evaluating the intervention effect provided afterward. We started the intervention in session 20, and Ali exhibited gradual, consistent improvement; scoring 20% in the first session, then 30% in sessions 21 and 22, and 50% in session 23. His performance kept going up step by step, with 60% in sessions 24 and 25, and then he got 80% in session 26 and stayed there in sessions 27 and 28. In session 31, the maintenance test showed he was still 80%, and this was without any extra teaching or support. During the maintenance probe in session 31, Ali's performance stood unmoved at 80%, as if the skill had taken root and refused to fade even in the absence of fresh instruction.

Discussion

This study aimed to examine the effectiveness of an GenAI-enhanced visual instructional tool for supporting the development of reading comprehension skills among students with learning disability (dyslexia) using a single-case experimental design with a multiple-probe across participants approach. The intervention showed marked performance improvement, with each participant exceeding the 80% mark across three consecutive sessions in the treatment and through to the maintenance. This suggests a functional relationship between the intervention and the gains in reading comprehension.

Participant Ahmed showed a rapid and consistent change in level and trend once the treatment began, with no overlap between baseline and intervention phases. The pattern of Ahmed's results also provides evidence of durable skills acquisition consistent with answering the study's third research question, which asked whether the improvements to the reading comprehension skills of the participants were maintained after the treatment was withdrawn. Furthermore, visual analysis of the students' scores confirms a distinct and immediate change following the intervention, sustained performance above criterion, and stability during maintenance. Similarly, in Badir's results, the immediate change in level, lack of data overlap, and consistency of response following initiation of the intervention provide strong evidence of a functional relation between the intervention and performance gains. These results are relevant to both the first and second research questions, which asked whether there was a functional relation between the treatment and reading comprehension performance and whether such skills improved "immediately and consistently" after the intervention was introduced. The maintenance of the intervention's effect is also demonstrated by Badir's scores in the maintenance phase. Taken together, the data for Badir show an immediate and stable improvement in reading comprehension following the treatment, with improved performance maintained across post-intervention sessions. Visual analysis of the data shows a strong functional relation between the use of the GenAI 'ChatGPT'-generated visual instruction and improvement in comprehension performance.

Although the initial change in Ali's case was less immediate than those observed in the other participants, the consistent upward trend and eventual attainment of the criterion of his scores also support a functional relation between the intervention and his comprehension gains. Ali's results show that the skills he learned in the intervention stayed with him for a long period afterward. Taken as a whole, Ali's data reveal a gradual yet distinct enhancement in his reading comprehension following the implementation of the treatment. He kept the skills over time and showed steady gains in reading after the treatment.

The overall outcomes of the study demonstrate strong empirical evidence indicating a functional relation between the GenAI-based visual instructional procedures and improvement in reading comprehension among all participants. In relation to the third research question, data from the post-intervention maintenance probes indicated the persistence of comprehension gains over extended periods in the absence of supplemental instructional input, thereby providing empirical confirmation of the durability and stability of treatment effects. Multiple features of the intervention protocol are likely to have contributed to the observed outcomes. Foremost among these was the main systematic integration of visual representations tied to the target texts, which functioned to establish a multimodal instructional environment demonstrably favorable to both enhanced comprehension and long-term retention of material. It seems that when the AI tool converted complex or abstract content into visuals, students were able to stay engaged with the material for longer. This might be similar to the findings of He (2024), who suggested that adaptive AI texts may help adjust instruction to learner needs and pace.

Beyond visual scaffolding, the intervention also incorporated real-time responsiveness. Specifically, the AI system enabled immediate clarification and adjusted support as students progressed, helping to close comprehension gaps as soon as they arose. This circle of feedback may have helped the skills to grow and made students learn by themselves more. The present study's results converge with the empirical findings of Fatmi et al. (2025), who suggested that AI-based personalised environments have the potential to improve reading comprehension. In addition to cognitive and academic gains, the current study also revealed important motivational benefits. Observational data indicated that the participants became increasingly engaged and confident as their performance improved. These findings echo those of Miller (2010) and Rosalina and Nasrullah (2019), which indicated that enhanced comprehension fosters higher self-esteem and greater classroom participation among students with learning disabilities. Fletcher et al. (2018) likewise identified the cultivation of reading competence as a pivotal factor in enhancing the academic integration and self-perception of this student population.

Regional studies also support the introduction of AI in inclusive classrooms in many Gulf contexts, similar to this study. Al-Mutairi et al. (2024), for instance, demonstrated that GenAI-based simulations supported the enhancement of science

comprehension for students with learning disabilities in Kuwait. Likewise, Al-Zahrani (2024) reported that chatbot-assisted instruction significantly increased reading engagement and achievement among Saudi elementary learners. Moreover, Behforouz and Al Ghaithi (2024) achieved similar findings in their study involving interactive chatbots with English as a Foreign Language (EFL) students in Oman. These examples reinforce our results and highlight the potential for AI to improve both academic achievement and learner motivation in Arabic-speaking contexts.

Another important implication of the study is that the adaptability of AI technologies makes them well-suited to support different learning styles and abilites. This finding is similar to those of Maghawry (2024), who found comprehension gains among deaf and hard of hearing students using AI-enhanced instructional tools in Oman, and Ogunlade et al. (2023), who documented improved spelling and reading fluency among Nigerian students with dyslexia following GenAI-enhanced instruction. Such findings affirm that AI can offer flexible, scalable support for diverse learners across cultural and educational contexts.

Conclusion

It appears that while the application of GenAI tools in teaching students with learning disabilities (dyslexia) is still in the early stages, this study provides an important contribution and perspective to ongoing work in the field of education for such groups of learners with special needs. AI applications such as ChatGPT, which was employed in the current study, are quickly evolving in a way that creates tremendous opportunities for supporting students with learning disabilities (dyslexia) to develop their skills. In this study, the researchers employed a single-case experimental design, the intervention produced immediate, consistent, and sustained improvements in all student performance, which highlight a functional relationship between the GenAI-supported visual instruction and targeted reading gains. The evidence supports generative AI as a potential aid for personalizing learning and raising engagement for students with dyslexia.

Implications for practice and policy

The findings of this study have a number of practical implications for teachers and other educational stakeholders who work with students who have learning disabilities that affect reading comprehension skills, including dyslexia. The first recommendation involves embedding GenAI-generated visual supports (including adaptive illustrations and simplified textual scaffolds) into reading comprehension instruction, thereby enabling individualised support mechanisms that promote both comprehension and engagement. The second centers on the incorporation of GenAI applications, particularly ChatGPT-like, into structured pedagogical routines in one-to-one or small-group contexts. As demonstrated in this study, even brief but targeted sessions employing GenAI-generated materials yielded clear and sustained improvements in student performance. The third recommendation highlights the necessity of expanding professional development frameworks designed to train educators in the effective use of GenAI-based instructional tools within special and inclusive educational environments, with the overarching goal of advancing personalised learning and enhancing reading comprehension outcomes for students with learning disabilities (dyslexia). Beyond teacher training, educational authorities and the Ministry of Education should develop national policy frameworks that recognize GenAI-generated materials as valid assistive technologies within special education. This should include integrating GenAI-supported protocols into Individualized Education Programs (IEPs) and establishing policies that provide the necessary technical infrastructure, such as high-speed internet and high-performance tablets. Finally, establishing clear ethical guidelines for the use of AI in classrooms will ensure that these tools are used responsibly to enhance personalized learning outcomes for students with disabilities.

The educational authorities and the Ministry of Education should consider introducing GenAI-supported instructional frameworks into the special education learning system. Policies should also be developed to provide teachers with ethical guidelines and technical infrastructure (e.g., high-speed internet and tablets) to facilitate the use of such tools for students with disabilities.

Limitations and recommendations for future research

While the findings of this study provide strong preliminary support for the use of GenAI-based visual instruction in improving reading comprehension among students with learning disabilities (dyslexia), several limitations must be acknowledged. This study included only three students, which is deemed appropriate with this single-case design allowing for detailed individual analysis; however, the limited number of participants and specific educational context limit the generalizability of the findings. In addition, the study involved only male participants due to the gender segregated nature of Saudi education (Alsamani et al., 2026). Future research should be conducted to investigate the use of such an intervention with female students.

In addition, while immediate and sustained improvements were observed, the study did not assess long-term retention or generalization of comprehension skills across subjects or instructional settings. Future studies with extended follow-up would provide a clearer picture of the intervention's lasting effects.

Furthermore, although the AI app used in this study integrated multiple instructional features, such as visual scaffolding, simplified explanations, and adaptive pacing, the authors did not isolate the specific contribution of each component. It remains unclear whether one feature of the intervention was particularly influential to the effect of this study or whether the combined use of supports was necessary for achieving gains. Ultimately, while methodological safeguards were implemented to reduce the effect of extraneous variables, the possibility cannot be excluded that contextual influences, such as differential levels of teacher input, parental engagement, or other environmental supports, exerted an impact on the

students' performance and, by extension, on the overall pattern of results. Further research with more environmental controls, or integrating parallel qualitative artifacts, should be conducted to help to reduce the influence of external variables. Experimental designs comparing individual features, such as visuals alone vs. visuals plus explanation, could help refine best practices for these types of GenAI-enhanced tools.

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 Research Ethics Standing Committee (REC) at UoH. 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

OA: Conceptualization, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. YA: Conceptualization, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. The authors extend their appreciation to the Deanship of Scientific Research, Islamic University of Madinah, Saudi Arabia, for funding this research work.

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.

References

Al Ghanem, R., and Kearns, D. M. (2015). Orthographic, phonological, and morphological skills and children's word reading in Arabic: a literature review. *Read. Res. Q.* 50 (1), 83–109. doi: 10.1002/rrq.84

Al-Batal, Z. M., and Al-Arafaj, N. A. (2021). Challenges faced by parents of primary school girls with learning disabilities. *Saudi J. Spec. Educ.* 18 (1), 49–83.

Al-Dosari, H. B. H., and Al-Subaie, A. M. (2022). The reality of the practices of Arabic language teachers in addressing spelling errors among primary students in Al-Aflaj governorate. *J. Fac. Educ.* 38 (1.2), 151–210. Available online at: https://mfes.journals.ekb.eg/article_222260.html

Al-Mutairi, S. F. M. F. F. S., Jalil, A. S. M. A., and Mohsen, A. S. A. (2024). The use of simulation through artificial intelligence to develop scientific concepts among middle school students with learning difficulties. *Stud. Psychol. Educ. Couns.* 7 (1), 147–166. doi: 10.21608/dapt.2024.345433

Al-Zahrani, A. M. (2024). The impact of using chatbots in mobile learning on developing reading skills among students with learning disabilities. *Acad. J. Res. Sci. Publ.* 66 (1), 124–158.

Alghamdi, E. A., Alghamdi, S. H., and Alfarani, L. A. (2023). The effect of using artificial intelligence applications in improving selective attention of students with learning disabilities in the primary stage. *J. Educ. Psychol. Sci.* 7 (43), 64–82. doi: 10.26389/AJSRP.N240823

Alsamani, O. A., Alsamiri, Y. A., and Alabdulwahab, R. A. (2026). Twice-exceptionality in primary schools: teachers' perceived barriers to identifying and supporting 2e students. *Acta Psychol. (Amst)* 262, 106031. doi: 10.1016/j.actpsy.2025.106031

Alsamani, O. A., Alsamiri, Y. A., and Alfaidi, S. D. (2023). Elementary school teachers' perceptions of the characteristics of twice-exceptional students. *Front. Educ.* 8, 1150274. doi: 10.3389/feduc.2023.1150274

Alsamiri, Y. A., and Alhassani, S. A. (2021). Teachers' perceptions about primary school students with learning disabilities problems with receptive and expressive language in Al-Madinah region. *Int. J. Res. Educ.* 45 (2), 100–132. doi: 10.36771/ijre.45.2.21-pp100-132

Alsamiri, Y. A., and Alsamani, O. A. (2023). From self-doubt to self-efficacy: Saudi elementary teachers' reflections on their experiences and challenges of teaching 2E students. *J. Educ. Soc.* 2 (16), 367–388. doi: 10.36046/2162-000-016-019

Alsamiri, Y., and Alshammari, A. (2025). Language learning difficulties among gifted students with learning disabilities from their teachers' perspectives. *J. Educ. Psychol. Sci. Qassim Univ.* 18 (2), 535–563.

Alshehri, B. (2023). Pedagogical paradigms in the AI era: insights from Saudi educators on the long-term implications of AI integration in classroom teaching. *Int. J. Educ. Sci. Arts* 2 (8), 159–180. doi: 10.59992/IJESA.2023.v2n8p7

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

American Psychiatric Association. (2025). What Is Specific Learning Disorder? Available online at: https://www.psychiatry.org/patients-families/specific-learning-disorder/what-is-specific-learning-disorder (Accessed May 07, 2025).

Asaad, H., and Eviatar, Z. (2014). Learning to read in Arabic: the long and winding road. *Read. Writ.* 27 (4), 649–664. doi: 10.1007/s11145-013-9469-9

Behforouz, B., and Al Ghaithi, A. (2024). The impact of using interactive chatbots on self-directed learning. *Stud. Self-Access Learn. J.* 15 (3), 317–344. doi: 10.37237/150302

Burenkova, O. V., Naumova, O. Y., and Grigorenko, E. L. (2021). Stress in the onset and aggravation of learning disabilities. *Dev. Rev.* 61, 100968. doi: 10.1016/j.dr.2021.100968

Chavez, O. J., and Palaoag, T. (2024). AI-driven mobile application: unraveling students' motivational feature preferences for reading comprehension. *J. Res. Innov. Teach. Learn.* 17 (2), 226–242. doi: 10.1108/JRIT-02-2024-0045

Cleveland Clinic. (2024, January 16). Learning Disabilities and Disorders. Available online at: https://my.clevelandclinic.org/health/diseases/4865-learning-disabilities-what-you-need-to-know (Accessed 07, 2025).

Dani, A. P., Pusdekar, Y. V., Dagdiya, K. R., and Deshmukh, V. R. (2024). Evaluating the impact of training teachers to identify learning disabilities: a pre-experimental study on knowledge enhancement. *Cureus.* 16 (3), e55685. doi: 10.7759/cureus.55685

de Beer, J., Engels, J., Heerkens, Y., and van der Kirk, J. (2014). Factors influencing work participation of adults with developmental dyslexia. *BMC Public Health* 14, 77. doi: 10.1186/1471-2458-14-77

Fatmi, K., Saad, Q., Shoukat, B. S., and Yousaf, M. (2025). The effectiveness of AI-based Reading intersections for students with learning disabilities: a psychological evaluation. *Crit. Rev. Soc. Sci. Stud.* 3 (3), 750–767. doi: 10.59075/eaxsqp94

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., and Barnes, M. A. (2018). *Learning Disabilities: From Identification to Intervention*, 2nd Edn. New York, NY: Guilford Press. Available online at: https://colegiopspchubut.com.ar/storage/2023/02/Learning-disabilities.pdf

He, X. (2024). Enhancing reading comprehension with AI-generated adaptive texts. *Int. J. New Dev. Educ.* 6 (7), 46–52. doi: 10.25236/IJNDE.2024.060708

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., and Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Except. Child.* 71 (2), 165–179. doi: 10.1177/001440290507100203

Irgashevna, I. M. (2025). Improving Reading comprehension in higher education: investigating the impact of artificial intelligence in teaching reading. *Eur. Int. J. Multidiscip. Res. Manag. Stud.* 5 (3), 24–26. doi: 10.55640/eijmrms-05-03-05

Kazdin, A. E. (2021). *Single-case Research Designs: Methods for Clinical and Applied Settings*, 3rd Edn. New York, NY: Oxford University Press.

Kocsis, J. (2016). *Primary teachers' knowledge about learning disabilities* (Master's thesis). Nipissing University. ProQuest Dissertations & Theses Global. Available online at: https://www.proquest.com/docview/1787822215

Kumari, R. K. (2025). "Artificial intelligence in special education: enhancing learning outcomes for students with disabilities," in *Transforming Special Education Through Artificial Intelligence*, ed. A. G. Walters (Hershey, PA: IGI Global), 79–112. doi: 10.4018/979-8-3693-5538-1.ch003

Ledford, J. R., and Gat, D. L. (2018). *Single Case Research Methodology: Applications in Special Education and Behavioral Sciences*, 3rd Edn. (New York, NY: Routledge). doi: 10.4324/9781315150666

Lee, S. J., and Kwon, K. (2024). A systematic review of AI education in K-12 classrooms from 2018 to 2023: topics, strategies, and learning outcomes. *Comput. Educ.: Artif. Intell.* 6, 100211. doi: 10.1016/j.caeai.2024.100211

Lyon, G. R., Shaywitz, S. E., and Shaywitz, B. A. (2003). A definition of dyslexia. *Ann. Dyslexia.* 53 (1), 1–14. doi: 10.1007/s11881-003-0001-9

Maghawry, A. A. F. (2024). The effectiveness of a program to improve reading comprehension skills in children with hearing disabilities using artificial intelligence. *J. Arab Stud. Educ. Psychol.* 151 (2), 23–60. doi: 10.21608/saep.2024.363936

Mayo Clinic. (2023, September 30). Auditory Processing Disorder (APD). Available online at: https://www.mayoclinic.org/diseases-conditions/auditory-processing-disorder/symptoms-causes/syc-20555261 (Accessed July 6, 2025).

Miller, D. (2010). *The Book Whisperer: Awakening the Inner Reader in Every Child*. San Francisco, CA: Jossey-Bass.

O'Neill, R., McDonnell, J., Jenson, W., and Billingsley, F. (2022). in *Single-case Designs in Educational and Community Settings*, Trans. B. N. Al-Otaibi (New York, NY: International Publisher for Publishing and Distribution).

Ogunlade, B. O., Babatunde, R. A., and Fakuade, O. V. (2023). Using conversational artificial intelligence tool in rebuilding Reading and spelling on dyslexic learners in Nigerian schools. *Int. J. Engl. Spec. Purp.* 1 (1), 1–11. Available online at: https://dergipark.org.tr/en/pub/joinesp/issue/79780/1342594

Rosalina, E., and Nasrullah. (2019). The correlation between self-esteem and student's reading comprehension. *Engl. Lang. Teach. Educ. J.* 2 (2), 70–78. Available online at: https://eric.ed.gov/?id=EJ1283078

Sakowicz, M., and Hamidi, F. (2025). Generative AI for inclusive education: supporting students with disabilities through universal design for learning. *Am. J. STEM Educ.* 15, 1–22. doi: 10.32674/yeyaaq64

Salcido, G. (2025, March 4). *Undiagnosed Learning Disabilities: A Barrier to Career Success and Personal Well-being*. AZ Big Media. Available online at: https://azbigmedia.com/business/undiagnosed-learning-disabilities-a-barrier-to-career-success-and-personal-well-being/

Samala, A. D., Rawas, S., Wang, T., Reed, J. M., Kim, J., Howard, N.-J., et al. (2025). Unveiling the landscape of generative artificial intelligence in education: a comprehensive taxonomy of applications, challenges, and future prospects. *Educ. Inf. Technol.* 30 (3), 3239–3278. doi: 10.1007/s10639-024-12936-0

Shahbari-Kasem, A., Schiff, R., and Saiegh-Haddad, E. (2024). Development of morphological awareness in arabic: the role of morphological system and morphological distance. *Read. Writ.* 38, 2235–2267. doi: 10.1007/s11145-024-10581-0

Visser, C., Shrivastava, R., and Sohal, A. (2024, August 1). Types of Learning Disabilities: A Comprehensive Guide. Accessibility Checker. Available online at: https://www.accessibilitychecker.org/blog/types-of-learning-disabilities/ (Accessed August 9, 2025).

Wendling, B. J. (2019). *Essentials of Evidence-based academic Interventions*. Hoboken, NJ: John Wiley & Sons.

Westmoreland, S. (2025, February 13). *Revolutionizing Learning Disability Identification Through Process Data Analysis*. Washington, DC: National Center for Learning Disabilities. Available online at: https://ncld.org/ld-identification-through-process-data-analysis/

World Economic Forum. (2025, January 9). How AI and Human Teachers Can Collaborate To Transform Education. Available online at: https://www.weforum.org/stories/2025/01/how-ai-and-human-teachers-can-collaborate-to-transform-education/ (Accessed June 11, 2025).

Appendix A: Procedural fidelity checklist

Observer Name:_____ Date: _____

Participant: _____ Session No: _____

Phase: Baseline Intervention Maintenance

Instructions for Observer: Please observe the entire session. For each step listed below, mark "YES" if the instructor performed the step correctly according to the protocol, or "NO" if the step was missed, modified, or implemented incorrectly.

Total Observed Indicators: _____/11

Fidelity Score: _____ %

Item	Fidelity Indicator	Observed (Yes/No)
1	The session was conducted in the designated resource room under controlled environmental conditions (lighting, seating, and minimal noise).	
2	The assigned printed reading passage was presented to the student prior to visual support.	
3	The student was given the opportunity to read the text independently before the introduction of AI-generated visuals.	
4	The instructor entered the designated text segment into ChatGPT using the standardised intervention prompt.	
5	The AI-generated visual representation accurately reflected the content of the reading passage.	
6	The visual representation was clearly displayed to the student using the tablet device.	
7	The instructor guided the student to connect visual elements to textual meaning without providing answers to assessment questions.	
8	The instructional phase followed the structured format (approximately 30 min of guided reading with visual support).	
9	No additional instructional prompting, feedback, or reinforcement was provided during the assessment phase.	
10	The five-item reading comprehension probe was administered immediately following the instructional phase.	
11	Student responses were recorded accurately using the standardised scoring form (0 = incorrect, 1 = correct).	