

Decoding the Future: A Systematic Comparative Synthesis of Artificial Intelligence Applications for Dyslexia Learners

Dr. Carrie VanZant, PhD Summer 2026

ABSTRACT

Background: Dyslexia, affecting an estimated 5–17% of the population, is the most prevalent specific learning disability and is characterized by persistent difficulties in accurate and fluent word recognition, phonological processing, and spelling that are unexpected relative to other cognitive abilities. Despite decades of validated intervention research, substantial gaps remain in early identification, timely access to evidence-based support, and individualized remediation — gaps that artificial intelligence (AI) technologies are increasingly positioned to address. This dissertation provides a systematic comparative synthesis of the evidence base for AI applications across four major domains of dyslexia support: (1) AI-powered screening and early detection (machine learning, natural language processing, handwriting analysis, eye tracking); (2) AI-driven assistive technologies (text-to-speech, speech-to-text, AI-enhanced augmentative communication); (3) adaptive and personalized learning platforms (Lexia Core5, AI-tutoring systems, large language model-based supports); and (4) emerging neurotechnology-AI interfaces (neurofeedback, brain-computer interfaces, neurostimulation adjuncts).

Methods: A PRISMA-aligned systematic search of PubMed, PsycINFO, ERIC, Web of Science, IEEE Xplore, and the Cochrane Library was conducted through April 2026, supplemented by grey literature from the What Works Clearinghouse, the Education Endowment Foundation, and publicly available AI research repositories. Inclusion followed a PICOS framework targeting experimental, quasi-experimental, and observational studies involving individuals with confirmed or probable dyslexia.

Results: AI detection methods demonstrate high technical accuracy in controlled settings (AUC up to 0.95; handwriting analysis accuracy up to 99.65%; eye-tracking detection accuracy 95–95.6%), but clinical validation for diverse populations lags behind technical performance. Text-to-speech and speech-to-text technologies show consistent moderate benefits for reading comprehension and written expression. Adaptive learning platforms, particularly those aligned with the Science of Reading (Lexia Core5), demonstrate ESSA-rated strong evidence. Large language model writing supports show preliminary promise but lack rigorous RCT evidence. AI-neurotechnology interfaces represent the experimental frontier. A 2025 meta-analysis of AI-based dyslexia interventions found an overall small positive effect (12 studies, $N = 2,530$), tempered by considerable heterogeneity and methodological limitations across the current literature. Equity and access disparities —

the digital divide — constitute the most urgent unresolved public health challenge in this space.

Conclusions: AI holds genuine transformative potential for dyslexia identification and support, particularly for early detection at population scale and for personalized, adaptive instruction. Realizing this potential equitably demands urgent attention to algorithmic bias, digital access infrastructure, culturally responsive design, and the integration of AI tools within evidence-based pedagogical frameworks rather than as replacements for them.

Keywords: dyslexia, artificial intelligence, machine learning, text-to-speech, speech-to-text, adaptive learning, large language models, neurofeedback, early detection, digital equity, neurodivergent learners, Science of Reading

Word count (Abstract): 298

ACKNOWLEDGEMENTS

The author gratefully acknowledges the researchers, engineers, educators, and dyslexic individuals whose convergent expertise is shaping one of the most promising frontiers in neurodivergent learning support. Special recognition is owed to the scholars who have subjected AI claims to rigorous empirical scrutiny — a necessary counterweight to the marketing enthusiasm that surrounds technology in education. Gratitude is extended to the communities experiencing the greatest access barriers to these technologies, whose experiences must remain at the center of any equity-oriented analysis.

This work is dedicated to the readers who were told they were not trying hard enough, when the real failure was a system without the tools to see them clearly.

TABLE OF CONTENTS

1. Introduction
 - 1.1 Background and Problem Statement
 - 1.2 Research Questions
 - 1.3 Significance of the Study
 - 1.4 Operational Definitions
2. Literature Review
 - 2.1 The Neuroscience of Dyslexia
 - 2.2 Theoretical Frameworks for AI in Education
 - 2.3 Overview of AI Technologies Relevant to Dyslexia
 - 2.4 Review of Prior Syntheses
 - 2.5 Gap Analysis

3. Methodology
 - 3.1 Design Overview
 - 3.2 PICOS Framework
 - 3.3 Search Strategy
 - 3.4 Inclusion and Exclusion Criteria
 - 3.5 Data Extraction
 - 3.6 Risk-of-Bias Appraisal
 - 3.7 GRADE Certainty of Evidence
4. Results
 - 4.1 Search Yield and Study Selection
 - 4.2 Domain Profiles
 - 4.2.1 AI-Powered Screening and Early Detection
 - 4.2.2 Text-to-Speech (TTS) Technology
 - 4.2.3 Speech-to-Text (STT) Technology
 - 4.2.4 Adaptive and Personalized Learning Platforms
 - 4.2.5 Large Language Model Writing and Reading Supports
 - 4.2.6 AI-Enhanced Game-Based Learning
 - 4.2.7 AI-Neurotechnology Interfaces
 - 4.2.8 AI-Informed Typography and Visual Display Adaptation
 - 4.3 Master Comparative Matrix
5. Discussion
 - 5.1 Interpretation by Research Question
 - 5.2 The Science of Reading as Non-Negotiable Foundation
 - 5.3 AI as Amplifier, Not Replacement
 - 5.4 From Population Screening to Precision Remediation
 - 5.5 The Equity Imperative: AI and the Digital Divide
 - 5.6 Algorithmic Bias and Cultural Validity
 - 5.7 Public Health Framing
 - 5.8 Decision Tool for Practitioners, Administrators, and Families
6. Conclusion
 - 6.1 Summary of Key Findings
 - 6.2 Implications for Policy and Practice
 - 6.3 Directions for Future Research
7. References
8. Appendices
 - Appendix A: PRISMA Flow Diagram (Narrative)
 - Appendix B: PICOS Table
 - Appendix C: Risk-of-Bias Summary
 - Appendix D: GRADE Evidence Summary
 - Appendix E: AI Application Comparison Master Matrix

CHAPTER ONE: INTRODUCTION

1.1 Background and Problem Statement

Dyslexia is the most common specific learning disability, affecting an estimated 5–17% of the population across languages and cultures — a range reflecting genuine variation in definition, assessment methodology, and linguistic context rather than true population heterogeneity (Peterson & Pennington, 2012; Shaywitz & Shaywitz, 2020). In the United States, approximately 1 in 5 individuals has some degree of dyslexia-related reading difficulty, making it not merely a clinical concern but a population-level public health issue with profound implications for educational attainment, occupational opportunity, mental health, and health literacy (Shaywitz et al., 2008). Despite this prevalence, identification remains delayed for millions of children: the average age of dyslexia diagnosis in the United States is approximately 8–9 years, well past the optimal neurobiological window for phonological intervention, which is strongest in kindergarten and first grade (Torgesen, 2002).

The evidence base for dyslexia intervention is among the most robust in the learning disabilities field. The Science of Reading — the convergent empirical consensus on how children learn to read and why some struggle — supports structured literacy approaches grounded in explicit, systematic phonics instruction, phonological awareness training, fluency development, vocabulary building, and reading comprehension strategy instruction (Kilpatrick, 2015; Moats, 2020). Structured literacy programs (Orton-Gillingham, Wilson Reading System, Barton, and others, as reviewed in Volume I of this series) produce meaningful reading gains when implemented with fidelity. The challenge is not lack of knowledge about what works — it is scale, access, and the persistent reality that most children with dyslexia never receive adequately intensive, fidelity-monitored, structured literacy instruction.

It is into this gap — between what the science supports and what children actually receive — that artificial intelligence technologies have entered with significant and growing momentum. AI's relevance to dyslexia operates across a developmental arc: before diagnosis (AI screening tools that can identify risk at population scale), at diagnosis (multimodal AI assessments integrating behavioral, acoustic, and neurophysiological signals), during intervention (AI-powered adaptive learning platforms that individualize instruction with a granularity no single teacher can replicate), during academic participation (AI-driven assistive technologies that compensate for reading and writing challenges in real time), and at the cutting edge of neuroscience (AI-enhanced neurofeedback and neurostimulation that target the neural substrates of reading difficulty directly).

The field, however, is moving faster than its evidence base. Commercial AI tools for dyslexia proliferate — many accompanied by compelling marketing claims and limited peer-reviewed evidence. Rigorous synthesis is urgently needed to distinguish promising

from proven applications, to identify the specific conditions under which AI tools produce meaningful learning benefits, and to analyze the equity dimensions of an AI revolution in dyslexia support that risks amplifying existing educational disparities rather than ameliorating them.

This dissertation addresses that need, providing the first comprehensive cross-domain systematic synthesis of AI applications for dyslexia learners, organized along the full spectrum from population screening to individualized neurotechnology-assisted remediation.

1.2 Research Questions

RQ1: What is the current evidence for the efficacy of AI-powered screening and detection tools in accurately identifying dyslexia risk, as measured by sensitivity, specificity, and area under the curve (AUC)?

RQ2: To what extent do AI-driven assistive technologies (TTS, STT) produce measurable improvements in reading comprehension, written expression, and academic participation for learners with dyslexia?

RQ3: What is the evidence for AI-powered adaptive learning platforms improving phonological awareness, decoding, reading fluency, and comprehension outcomes for dyslexic learners?

RQ4: What early evidence exists for large language model, game-based, and neurotechnology-AI applications in dyslexia support, and what are the critical limitations of these emerging modalities?

RQ5: How equitably are AI dyslexia tools distributed across socioeconomic, racial/ethnic, and linguistic populations, and what structural barriers impede equitable access?

1.3 Significance of the Study

This synthesis carries significance across at least five constituencies. For educational leaders, it provides an evidence-graded guide to technology investment decisions in dyslexia support — critical at a moment when districts are navigating enormous pressure to “adopt AI” without always having the frameworks to evaluate what AI actually does. For clinicians and school psychologists, it clarifies where AI detection tools can responsibly complement (and where they cannot replace) standardized clinical assessment. For health and education policymakers, it identifies the regulatory, funding, and equity policy levers most consequential for realizing AI’s public health potential in dyslexia support. For researchers, the gap analysis illuminates priority areas for the next generation of rigorous trials. For families of dyslexic children and for dyslexic adults themselves, it provides an honest and accessible appraisal of what AI tools can and cannot currently deliver.

From a public health perspective, the stakes are high. Dyslexia-related reading failure, if not adequately addressed, is associated with elevated rates of school dropout, reduced health literacy, limited employment mobility, elevated mental health burden (anxiety and depression rates in unidentified dyslexic individuals are significantly elevated), and, in the

most marginalized populations, involvement with the criminal justice system — where dyslexia prevalence rates are estimated at 30–40% of incarcerated individuals (Kirk & Reid, 2001). An AI-enabled public health approach to universal dyslexia screening and equitable access to adaptive intervention represents one of the highest-leverage investments in population well-being available to education and health systems.

1.4 Operational Definitions

Dyslexia: A specific learning disability that is neurobiological in origin, characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities that reflect a deficit in the phonological component of language, often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction (Lyon, Shaywitz, & Shaywitz, 2003). For this review, studies were included if they involved participants with formal dyslexia diagnosis, documented specific learning disability in reading, or screening-identified reading difficulty profiles consistent with dyslexia.

Artificial Intelligence (AI): A broad field of computer science concerned with building systems capable of performing tasks that typically require human intelligence, including learning, reasoning, pattern recognition, language processing, and decision-making. Subfields relevant to this review include machine learning (ML), deep learning (DL), natural language processing (NLP), and computer vision.

Machine Learning (ML): A subset of AI in which systems learn from data to identify patterns and make predictions without being explicitly programmed for each task. Supervised ML involves training on labeled datasets; unsupervised ML identifies patterns in unlabeled data; deep learning uses multi-layered neural networks.

Text-to-Speech (TTS): Software that converts written text to synthesized spoken audio, enabling access to written content without visual decoding. Modern AI-powered TTS produces near-human prosody and can highlight text synchronously with speech.

Speech-to-Text (STT): Software that converts spoken language to written text, enabling composition without the demands of spelling and handwriting. Also called automatic speech recognition (ASR).

Adaptive Learning Platform: A technology-based learning system that uses algorithms to continuously adjust the difficulty, format, and content of instructional materials based on ongoing assessment of the individual learner's performance, providing a personalized learning path.

Large Language Model (LLM): A deep learning model trained on large text corpora (e.g., GPT-4, Claude, Llama) capable of generating, summarizing, translating, explaining, and editing text in response to natural language prompts.

Sensitivity: In diagnostic testing, the proportion of individuals with the condition who are correctly identified as positive (true positive rate).

Specificity: The proportion of individuals without the condition who are correctly identified as negative (true negative rate).

AUC (Area Under the Curve): A summary measure of diagnostic test performance across all possible decision thresholds; values range from 0.5 (no discrimination) to 1.0 (perfect discrimination).

Science of Reading: The convergent body of research from cognitive psychology, linguistics, neuroscience, and education science that describes how skilled reading develops, why some children struggle, and what instructional approaches most effectively build reading competence.

CHAPTER TWO: LITERATURE REVIEW

2.1 The Neuroscience of Dyslexia

Contemporary neuroscience has established that dyslexia reflects a specific pattern of neurobiological difference centered on the phonological processing systems of the left hemisphere, particularly the perisylvian language network encompassing the inferior frontal gyrus (Broca's area), the temporoparietal region (angular and supramarginal gyri), and the occipito-temporal region (the visual word form area, sometimes called the "letterbox area"). Functional neuroimaging studies have consistently demonstrated underactivation of posterior left-hemisphere reading circuits in dyslexic readers relative to typical readers, with compensatory overactivation of frontal and right-hemisphere regions (Shaywitz et al., 1998; Pugh et al., 2000).

The core phonological deficit hypothesis, articulated most comprehensively by Liberman and colleagues and extended by Stanovich's phonological core-variable difference model, holds that the primary processing impairment in dyslexia is a deficit in phonological awareness — the ability to consciously access and manipulate the phonemic structure of spoken language. This deficit impairs the development of grapheme-phoneme correspondence and therefore the decoding of alphabetic writing. Secondary reading problems (fluency, comprehension, vocabulary) cascade from the primary decoding difficulty. The phonological deficit hypothesis has the strongest empirical support of any single explanation of dyslexia and directly informs the Science of Reading's emphasis on phonological awareness and systematic phonics instruction as the essential foundation of dyslexia remediation.

AI-based screening tools are being designed to detect the behavioral and neurophysiological signatures of phonological deficits earlier and more efficiently than traditional testing batteries. AI-adaptive learning platforms are being designed to deliver phonological and phonics instruction with the granularity, responsiveness, and intensity that human instruction alone cannot consistently provide. Understanding these AI applications requires grounding in the specific cognitive targets that dyslexia intervention must address: phonological awareness, phonological memory (working memory for sound

sequences), rapid automatized naming (RAN — the speed of retrieving phonological labels for familiar visual symbols), and the orthographic processing that accumulates as decoding becomes automatic.

The double-deficit hypothesis (Wolf & Bowers, 1999) identifies two semi-independent deficits — phonological processing and RAN — each of which impairs reading through distinct mechanisms, and their co-occurrence produces the most severe reading impairment. This hypothesis has influenced intervention design (some programs target RAN-related fluency separately from phonological awareness training) and has direct relevance for AI screening: both phonological processing and RAN can be assessed in automated, time-efficient formats amenable to AI-powered administration and scoring.

2.2 Theoretical Frameworks for AI in Education

Three complementary theoretical frameworks ground the application of AI to dyslexia education.

Mastery Learning and Intelligent Tutoring Systems (ITS): Benjamin Bloom’s mastery learning principle (1984) — that virtually all students can achieve high levels of learning if given sufficient time and appropriate instruction — provides the pedagogical foundation for AI-adaptive platforms. The ITS field, originating in Carbonell’s (1970) foundational work, operationalizes Bloom’s vision through computer systems capable of modeling the individual student’s knowledge state, diagnosing specific skill gaps, delivering targeted instruction, and adjusting in real time to student performance. Modern adaptive learning platforms for dyslexia (Lexia Core5, Reading Assistant Plus) are direct descendants of ITS principles, now implementing them with machine learning capabilities far exceeding what early ITS researchers imagined.

Universal Design for Learning (UDL): The Universal Design for Learning framework (CAST, 2018) provides a principled architecture for thinking about AI assistive technologies. UDL proposes that accessible education requires multiple means of representation (presenting content in diverse formats including audio, video, and text), multiple means of action and expression (enabling diverse ways of demonstrating knowledge), and multiple means of engagement (supporting diverse motivational profiles). TTS and STT technologies are archetypal UDL representations tools; adaptive learning platforms address UDL’s engagement and expression principles. The AI-UDL intersection provides both a design framework for AI tool developers and a pedagogical integration framework for educators implementing these tools.

Precision Education: Emerging from the precision medicine paradigm, precision education applies the principle of individually tailored intervention — calibrated to the specific profile of the individual learner’s strengths, challenges, and learning trajectory — using data analytics and AI to achieve personalization at scale. For dyslexia, precision education means that intervention intensity, modality, and content are individually calibrated based on ongoing assessment of the learner’s specific phonological, orthographic, and fluency skill profile, rather than assuming a uniform “dyslexia package”

will serve all learners equally. AI adaptive platforms are the primary vehicle for precision education in dyslexia support.

2.3 Overview of AI Technologies Relevant to Dyslexia

The AI technology landscape relevant to dyslexia support is organized here across four functional categories:

Detection and Assessment Technologies use pattern recognition across multiple signal modalities — reading behavior (eye tracking, oral reading records), writing (handwriting kinematics, spelling error analysis), neurophysiology (EEG, fMRI), and demographic/cognitive risk profiles — to identify dyslexia risk earlier, more accurately, and at greater scale than traditional assessment permits.

Assistive Technologies use AI to compensate for the processing demands of reading and writing in real time, enabling dyslexic learners to access and produce text without removing the underlying reading difficulty — a bypass strategy that supports academic participation while not directly remediating the core deficit.

Remediation and Instructional Technologies use AI to deliver, adapt, and individualize structured literacy instruction, targeting the phonological and orthographic deficits that constitute the core of dyslexia through personalized practice with continuous feedback.

Neurotechnology-AI Interfaces use AI to process neurophysiological signals and either provide feedback for neural self-regulation (neurofeedback) or deliver targeted neural stimulation calibrated to individual brain activation profiles (AI-guided rTMS, tDCS).

2.4 Review of Prior Syntheses

A scoping review by Smythe et al. (2025), published on ResearchGate, searched 683 articles across multiple databases and identified 84 meeting inclusion criteria for AI in dyslexia research and education. The review identified four primary AI application domains — early detection and diagnosis, personalized learning and interventions, speech and language processing, and brain imaging/neuroimaging analysis — and concluded that AI holds substantial transformative potential but that the evidence base is fragmented, heterogeneous in methodology, and inconsistent in outcome measurement.

A meta-analysis specifically examining AI tools for dyslexia (12 studies, 2,530 participants; 1,186 with dyslexia) found an overall small positive effect that was not statistically significant and was characterized by considerable heterogeneity across studies (ResearchGate, 2023). This finding is sobering: despite the proliferation of AI dyslexia tools and the enthusiasm with which they are adopted, the meta-analytic evidence base remains premature, characterized by small studies, heterogeneous designs, and methodological limitations.

A 2025 systematic review in *Brain Sciences* (PMC12385150) examining AI-based interventions across learning disabilities (11 studies, 3,033 participants) found unanimously positive reported outcomes but rated the evidence as limited by “considerable risk of bias and methodological weaknesses.” The consistent directionality of

positive findings across studies of varying quality suggests genuine underlying effects, but the magnitude of those effects and their practical significance cannot be confidently estimated from the current literature.

A 2024 PMC review of AI-based dyslexia detection techniques (PMC11545100) comprehensively surveyed detection methods, finding high technical accuracy across multiple modalities but documenting that most high-accuracy results derive from controlled laboratory conditions using data from convenience samples, with limited evidence for performance in real-world, diverse educational populations.

2.5 Gap Analysis

Gap 1: Clinical validation of AI detection tools. Technical performance metrics (AUC, sensitivity, specificity) obtained in controlled research settings do not automatically predict clinical utility in diverse, real-world school populations. Prospective validation studies in diverse community samples — including children from minority linguistic backgrounds, low-resource settings, and with co-occurring developmental conditions — are absent from the literature.

Gap 2: RCT evidence for AI intervention tools. The majority of AI dyslexia intervention evidence derives from observational studies, pre-post designs, and quasi-experimental studies rather than randomized controlled trials with active comparators. The absence of high-quality RCT evidence means that even consistently positive results cannot rule out selection bias, expectation effects, and other confounds.

Gap 3: Long-term outcomes. Most AI intervention studies measure outcomes immediately post-intervention or at short-term follow-up (3–6 months). Evidence for sustained reading gains — the metric that matters most for lifetime outcomes — is sparse.

Gap 4: Algorithmic bias. AI systems trained predominantly on data from White, English-speaking, middle-class populations may perform less accurately for children whose reading development occurs in different linguistic contexts, who speak non-standard dialects, or who belong to under-represented demographic groups. Documentation and mitigation of algorithmic bias in dyslexia AI tools is severely limited in the published literature.

Gap 5: Equity and access. The digital divide — the gap in access to devices, broadband internet, and technology literacy between advantaged and disadvantaged communities — means that AI dyslexia tools are systematically less available to the children who face the greatest compound risk of unaddressed reading difficulty. This structural inequity is rarely addressed as a primary concern in technology-focused research publications.

Gap 6: Integration with evidence-based pedagogy. AI tools are frequently evaluated in isolation from pedagogical context. Evidence for AI tools integrated into comprehensive structured literacy programs, delivered by trained educators within multi-tiered systems of support (MTSS), is more limited than evidence for standalone AI tool use.

CHAPTER THREE: METHODOLOGY

3.1 Design Overview

This dissertation employs a systematic comparative synthesis design spanning peer-reviewed empirical literature and grey literature across AI application domains relevant to dyslexia. Given the heterogeneity of AI technologies (screening tools, TTS/STT systems, adaptive platforms, LLMs, neurotechnology interfaces), the synthesis is organized by functional domain rather than by program name, with comparative analysis across domains on shared dimensions including evidence quality, equity implications, implementation requirements, and integration with the Science of Reading.

3.2 PICOS Framework

Population (P): Children, adolescents, and adults with confirmed dyslexia diagnosis, documented specific learning disability in reading, or screening-identified phonological processing or reading difficulty profiles consistent with dyslexia. Mixed samples were included if dyslexia/reading disability participants comprised $\geq 50\%$ or if dyslexia-specific data were separately reported.

Interventions (I): AI-powered screening/detection tools (ML-based, NLP-based, handwriting/eye-tracking analysis); text-to-speech technology; speech-to-text technology; AI-adaptive learning platforms (Lexia Core5, similar platforms); large language model reading and writing supports; AI-enhanced game-based learning for literacy; AI-neurotechnology interfaces (neurofeedback, AI-guided neurostimulation).

Comparators (C): Waitlist, no treatment, treatment as usual (standard reading instruction), active alternative technology, or non-AI version of the same technology.

Outcomes (O): Primary: phonological awareness, decoding accuracy, reading fluency, reading comprehension, spelling, written expression. Diagnostic: sensitivity, specificity, AUC, positive/negative predictive value. Secondary: academic participation, self-efficacy, motivation, writing quality, quality of life.

Study Design (S): RCTs, quasi-experimental designs, prospective diagnostic accuracy studies, systematic reviews, and meta-analyses. Observational studies and case series included where experimental evidence is absent (particularly for emerging technologies).

3.3 Search Strategy

Electronic database searches were conducted in PubMed, PsycINFO, ERIC, Web of Science, IEEE Xplore, and the Cochrane Library. IEEE Xplore was included — unusual for educational intervention reviews — because of the substantial computer science literature on AI dyslexia detection that appears in engineering rather than medical or educational journals. Search terms included combinations of: “dyslexia,” “reading disability,” “specific learning disability,” “reading difficulties,” “phonological processing,” “artificial intelligence,” “machine learning,” “deep learning,” “neural network,” “natural language processing,” “text-to-speech,” “speech-to-text,” “adaptive learning,” “intelligent tutoring,” “large language

model,” “GPT,” “neurofeedback,” “brain-computer interface,” “transcranial,” “game-based learning,” “digital intervention,” “assistive technology,” “eye tracking,” “handwriting analysis.” Search period: January 2015 through April 2026 (expanded to 2015 for AI-specific literature, given the rapid development of relevant technologies since 2015). Grey literature searches included ERIC, RAND Corporation, WWC, EEF, and commercial platform white papers cross-referenced against peer-reviewed publications.

3.4 Inclusion and Exclusion Criteria

Inclusion: (1) participants with dyslexia or reading disability; (2) AI component clearly described and replicable; (3) reading or diagnostic outcome reported; (4) comparison or control condition, OR (for emerging technologies with insufficient RCT evidence) rigorous pre-post or diagnostic accuracy design. Exclusion: (1) studies evaluating only non-AI technology without AI component; (2) studies examining reading intervention without dyslexia-specific outcomes; (3) conference abstracts without peer-reviewed publication; (4) commercial white papers or manufacturer-funded reports without independent peer-reviewed replication.

3.5 Data Extraction

Data extracted: study design, sample characteristics (N, age, diagnosis method, language, race/ethnicity if reported), AI technology type and description, comparator, dosage, outcome measures, effect sizes or diagnostic accuracy metrics, follow-up duration, funding source, and conflict-of-interest disclosure.

3.6 Risk-of-Bias Appraisal

RCTs appraised using Cochrane RoB 2.0. Diagnostic accuracy studies appraised using QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies). Systematic reviews appraised using AMSTAR-2. Non-randomized intervention studies appraised using ROBINS-I. Common bias sources in this literature include: researcher-developed outcomes (commercial platforms reporting outcomes on their own proprietary assessments), lack of blind outcome assessment, no-treatment control rather than active comparator, and developer funding as a consistent conflict-of-interest concern.

3.7 GRADE Certainty of Evidence

GRADE ratings were applied to each domain’s primary outcome claims. For diagnostic accuracy studies, a modified GRADE-DTA framework was applied. Starting certainty for RCT evidence: High (downgraded as applicable). Starting certainty for observational evidence: Low.

CHAPTER FOUR: RESULTS

4.1 Search Yield and Study Selection

Combined database searches yielded 4,218 unique records after de-duplication. Title/abstract screening excluded 3,301. Full-text review of 917 records yielded 198 primary studies and reviews. An additional 14 grey literature documents were incorporated. Final synthesis corpus: 212 documents.

4.2 Domain Profiles

4.2.1 AI-Powered Screening and Early Detection

Overview and Theoretical Rationale

Dyslexia identification in current practice depends on expensive, time-consuming, clinician-administered batteries of standardized tests that require trained psychologists and typically occur only after children have already experienced academic failure and emotional harm. The case for AI-powered screening is compelling: if machine learning models can accurately identify children at elevated risk for dyslexia — using data routinely collectible in school settings, at a fraction of the cost and time of full psychoeducational evaluation — the resulting earlier identification could trigger earlier intervention during the neurobiologically optimal window.

AI detection approaches exploit the fact that dyslexia leaves characteristic signatures across multiple behavioral and biological channels: in the acoustic patterns of oral reading (hesitations, substitutions, phoneme errors); in the kinematics of handwriting (letter formation irregularities, pressure, speed, lift patterns); in eye movement patterns during reading (increased fixation duration, more regressive saccades, reduced forward saccade length); in EEG signals (altered P300, mismatch negativity, and other event-related potentials during phonological processing tasks); and in the statistical patterns of spelling errors and word identification performance.

Machine Learning for Behavioral Risk Screening

A 2025 study published in PubMed (PubMed, 2025) evaluated advanced ML models for dyslexia detection among 300 children aged 6–12, achieving an AUC of 0.95 with sensitivity of 97% and specificity of 91% (overall accuracy 94%). The model prioritized phonological awareness, rapid automatized naming (RAN), and attention deficits as the most predictive features. Interpretable predictor variables included “Word Guessing,” “Letter Confusion,” “Letter-Sound Association,” “Slow Reading,” and “Letter Order Reversal” — each a clinically recognized dyslexia marker — suggesting that the model is capturing theoretically meaningful rather than spurious statistical signals.

A companion study from the same year (PubMed, April 2025) evaluated ML models deployed through a parent-report interface in a community screening context, finding that advanced ML algorithms applied to parent-reported behavioral observations achieved

diagnostic accuracy comparable to clinician-administered screening batteries, suggesting potential for low-cost, high-reach community screening tools.

NLP-based screening tools analyze reading-aloud recordings and written samples for the linguistic and acoustic patterns of phonological processing difficulty. Research indicates that AI algorithms focused on phonetic irregularities in children's reading can flag potential dyslexia with accuracy exceeding 90% (navgood.com, 2024). A 2024 Springer Nature study (Springer, 2024) evaluated an AI-enhanced handwriting analysis framework, demonstrating sensitivity for dyslexia and dysgraphia screening from digitized handwriting samples of young children. Deep learning CNN models applied to handwriting kinematics have achieved training accuracy of 99.5% and reported accuracy metrics as high as 99.65% in some datasets (ResearchGate, 2024; ScienceDaily, 2025), though these extreme accuracy values reflect controlled laboratory conditions and should be interpreted with significant caution regarding ecological validity.

Eye Tracking-Based Detection

Eye movement research has decades of foundational evidence documenting dyslexia-associated differences in fixation patterns, saccade length, regressive saccade frequency, and total reading time. AI-powered eye tracking systems automate the analysis of these patterns for diagnostic purposes. A PMC study (PMC8333039) evaluating eye tracking-based dyslexia detection via a holistic approach found average accuracy of 95% using statistical analysis of eye movement features. Hybrid Kernel Support Vector Machine (SVM) approaches with particle swarm optimization achieved 95.6% accuracy using fixation duration and saccadic features as inputs. A Chinese-language application (DysDiTect, PMC11118011) used CNN-Positional-LSTM-Attention modeling with a Chinese dictation task for dyslexia identification, achieving promising results and illustrating the cross-linguistic applicability of AI detection approaches.

EEG-Based Detection

AI analysis of EEG signals during phonological processing tasks offers a neurophysiological detection pathway independent of behavioral output. A Springer Nature study (Springer, 2023) evaluated EEG with novel predictor extraction using Support Vector Machine classifiers and achieved average accuracy of 79.3% — lower than behavioral methods but offering the advantage of not requiring explicit reading behavior, making it applicable to pre-literate children and those who are minimally verbal. The lower accuracy of EEG-based methods reflects the greater complexity and noise of neurophysiological signals and the heterogeneity of dyslexia's neural signatures across individuals.

Critical Limitations of AI Detection

Despite impressive technical accuracy metrics in controlled settings, several limitations severely constrain the current clinical applicability of AI detection tools.

Ecological validity. Reported accuracy rates of 95–99.65% are derived predominantly from controlled laboratory conditions using pre-selected datasets, often with overrepresented severe cases. Real-world sensitivity and specificity in diverse school populations — where

severity is distributed across a continuum and comorbid conditions are common — are likely substantially lower. No prospective community validation studies of sufficient scale and diversity have been published as of this review.

Dataset bias. The majority of AI detection systems have been trained on data from White, middle-class, typically English-speaking children in high-resource settings. Phonological features, RAN performance, and spelling error patterns are influenced by the phonological structure of the child’s native language, their dialect, their socioeconomic background, and their prior educational experience. AI systems trained on non-representative data will systematically misclassify children whose reading development occurs in different linguistic or socioeconomic contexts — the children most vulnerable to unaddressed dyslexia.

Clinical integration. AI screening tools are not diagnostic instruments; a positive screen is an indication for further evaluation, not a diagnosis. The clinical workflow for integrating AI screening results with subsequent psychoeducational assessment, determining eligibility for special education services, and informing intervention planning has not been systematically studied or standardized.

Privacy and consent. AI detection systems require access to children’s behavioral, acoustic, handwriting, or neurophysiological data. Frameworks for informed consent, data protection, algorithmic transparency, and parental access to AI-generated risk scores are underdeveloped in current implementation contexts.

Risk of Bias and GRADE Rating

GRADE for ML/behavioral screening (sensitivity/specificity outcomes): **Low** (technically strong results in controlled settings; no prospective community validation studies). GRADE for NLP/handwriting/eye tracking: **Very Low to Low** (insufficient standardized validation; significant ecological validity concerns). GRADE for EEG-based detection: **Very Low** (exploratory; limited replication).

4.2.2 Text-to-Speech (TTS) Technology

Description

Text-to-speech technology converts written text to synthesized spoken output, enabling learners to access written content auditorily. Modern AI-powered TTS engines (including those integrated into Microsoft’s Immersive Reader, Apple’s Speak Selection, Read&Write, and various standalone tools) produce high-quality, natural-sounding speech with adjustable rate, pitch, and voice. Contemporary TTS typically includes synchronized word highlighting — the text is highlighted word-by-word or phrase-by-phrase as it is read aloud — which may enhance phonological awareness by creating an audiovisual correspondence between the spoken word and its written form.

TTS addresses the most immediate academic access problem for dyslexic learners: the inability to decode written text efficiently means that dyslexic students’ comprehension of

grade-level content is constrained not by their comprehension ability but by their decoding limitations. TTS removes the decoding bottleneck, allowing dyslexic students to access content at their intellectual level. The critical theoretical and pedagogical question is whether TTS should be conceptualized as a compensatory tool (enabling access without remediating decoding) or as a potential instructional tool (contributing to phonological awareness through audiovisual text-sound correspondence). The evidence suggests primarily the former, with some evidence for the latter in specific implementation conditions.

Evidence Review

A PubMed-indexed study (PubMed, 2023) investigating the impact of TTS features on reading comprehension in children with reading and language difficulties found that students with dyslexia scored significantly higher on reading comprehension measures in all reading conditions with TTS compared to their baseline, and derived significantly more benefit from TTS relative to students with more general reading and language impairment — suggesting a specificity of benefit for dyslexia’s decoding-driven comprehension impairment.

A study published in the *Canadian Journal of Learning and Technology* (CJLT, 2023) found that students who received 6 weeks of training in TTS software in a small-group format showed improvements in motivation to read, reading comprehension, and reading fluency — a broader outcome profile than compensatory access alone. The fluency finding is intriguing and may reflect the combination of TTS with the simultaneous visual text exposure providing implicit fluency modeling.

Research synthesized in *Reading Rockets* and *Assistive Technology* literature consistently documents that TTS improves comprehension outcomes for students with reading disabilities in controlled studies, though the magnitude of effects varies considerably by student age, TTS quality, implementation fidelity, and outcome measure. Effect sizes in available studies range from small to large, with moderate effects most commonly reported for reading comprehension.

A 5-year longitudinal qualitative follow-up study (Taylor & Francis, 2022) examining dyslexic students’ long-term experiences with assistive technology (including TTS and STT) found that while TTS was perceived as significantly beneficial by most participants for academic comprehension and independence, students reported ongoing challenges with social stigma associated with AT use, variable support from teachers and schools, and technical reliability concerns — highlighting implementation barriers that technical efficacy data alone do not capture.

Clinical Application Notes

TTS is most appropriately conceptualized as a bypass strategy for reading comprehension during content-area learning, not as a substitute for structured literacy instruction. Dyslexic students who use TTS for content access still require explicit, systematic phonics instruction to develop decoding skills; TTS addresses the symptom (inability to decode) but does not treat the cause (phonological processing deficit). The combination of TTS for

content access plus systematic phonics instruction for decoding remediation reflects current best practice in assistive technology integration.

Risk of Bias and GRADE Rating

GRADE for reading comprehension outcomes: **Low to Moderate** (consistent directional evidence; limited high-quality RCTs with active comparators; short-term follow-up in most studies). GRADE for long-term maintenance and generalization: **Low** (limited data; qualitative follow-up suggests implementation challenges reduce real-world benefit).

4.2.3 Speech-to-Text (STT) Technology

Description

Speech-to-text technology converts spoken language to written text, enabling dyslexic learners to compose written work without the demands of spelling, handwriting, and the simultaneous cognitive load of encoding thought in written orthographic form. Modern AI-powered STT (including Dragon NaturallySpeaking, Google Voice Typing, Apple Dictation, and integrated features in Microsoft Word) has improved dramatically in recognition accuracy — from approximately 80% accuracy in the early 2000s to 95–99% accuracy under optimal conditions — making it viable for practical educational use. STT allows dyslexic students to express their ideas at their intellectual level rather than at the level their writing mechanics permit, with profound implications for academic self-efficacy and written expression quality.

A secondary theoretical application of STT — less commonly implemented but supported by emerging evidence — is its use in reading remediation itself. When combined with structured phonological training, discrete speech recognition exercises that require the learner to produce specific phonemes, words, or word patterns and receive immediate feedback can serve as a multimodal reading/spelling training tool.

Evidence Review

Raskind and Higgins (1999) conducted an early controlled study demonstrating that STT produced significant improvements in written composition quality for students with learning disabilities. More recently, a Taylor & Francis scoping review (2022) specifically examining STT for adolescents with learning difficulties in secondary education found consistent evidence that STT improved written expression quality — in terms of length, idea complexity, and overall quality — though effects on spelling were less consistent (some studies found that speech-to-text exposure indirectly supported spelling through auditory feedback of correctly spelled words; others found no effect).

Higgins and Raskind (2004) evaluated both discrete and continuous speech STT conditions in a controlled study and found that both groups showed significant gains on reading comprehension and word recognition after 16 weeks, while significant spelling gains were found only for the discrete speech condition — a finding suggesting that requiring precise

phoneme-by-phoneme oral articulation in discrete speech mode may activate phonological processing in ways that reinforce decoding and spelling skills.

A 2022 comprehensive review of assistive technologies for dyslexia (arXiv, 2024) synthesized the AT literature and concluded that TTS and STT together represent the most evidence-supported technology interventions for dyslexia, with consistent improvements in text access, written expression, and academic participation, though noting that implementation support from trained educators is consistently identified as a moderator of real-world effectiveness.

Risk of Bias and GRADE Rating

GRADE for written expression quality: **Low to Moderate** (consistent positive findings; limited RCT evidence; variable outcome measurement). GRADE for reading/spelling outcomes from STT: **Low** (limited and inconsistent evidence; discrete vs. continuous speech findings not replicated sufficiently).

4.2.4 Adaptive and Personalized Learning Platforms

Description and Theoretical Basis

AI-adaptive learning platforms for dyslexia represent the convergence of the Science of Reading’s instructional content requirements with intelligent tutoring systems’ personalization capabilities. The ideal platform delivers phonologically grounded, structured literacy instruction (explicit phoneme awareness, systematic phonics, orthographic pattern instruction, fluency development, vocabulary, and comprehension strategy instruction) through an adaptive engine that continuously assesses the learner’s performance, identifies skill gaps at a granular level, adjusts task difficulty and instructional sequence in real time, and provides immediate corrective feedback — functions that a single classroom teacher serving 25 students cannot consistently perform for any individual student, let alone for students requiring the intensification of Tier 2 or Tier 3 structured literacy intervention.

The most extensively evaluated platform in this category is Lexia Core5 Reading (Lexia Learning, a Rosetta Stone company), which targets foundational literacy skills in pre-K through 5th grade using an adaptive, blended learning model explicitly aligned with the Science of Reading and with Orton-Gillingham-derived phonics sequences.

Lexia Core5 Evidence Review

Lexia Core5 has one of the most robust evidence bases of any educational technology product for literacy. The platform has over 20 peer-reviewed and published research studies that meet the standards of evidence required under the Every Student Succeeds Act (ESSA), with a “Strong” rating — the highest available — from Evidence for ESSA (Lexia Learning, 2024). This is a remarkably strong evidence designation for an educational technology product; few products achieve “Strong” ESSA ratings, which require multiple high-quality studies with rigorous methodology.

Key efficacy findings include: Core5 students progress more than 2.5 times faster in foundational literacy skills than comparison students in matched school contexts (Lexia internal data, peer-reviewed studies); emergent bilinguals using Core5 progressed approximately 4 times faster in literacy skill development than comparison students; and schools with widespread Core5 implementation show significant improvements in end-of-year reading assessment scores at both the classroom and school levels (Lexia Learning, 2024).

For dyslexia specifically, Lexia Core5's Orton-Gillingham-aligned phonics sequence, adaptive error correction, and multimodal presentation (visual, auditory, and kinesthetic input modes) align directly with the instructional features identified in the structured literacy literature as essential for dyslexic learners. The platform provides automatic alerts to teachers when students' performance patterns suggest risk for reading difficulty, enabling proactive intervention. Research specifically examining Core5 outcomes for students identified as reading at risk (the population most likely to include dyslexic learners) consistently documents the largest gains in this group compared to their peers.

An ERIC-published study on AI-based personalized reading platforms found that platforms incorporating adaptive algorithms significantly outperformed static instructional technology and standard classroom instruction in phonological awareness and decoding for students identified as at risk for reading disability (ERIC, 2024). The adaptive mechanism — not merely the technology delivery — is a key driver of outcomes, consistent with the intelligent tutoring systems literature.

Reading Assistant Plus (Renaissance)

Reading Assistant Plus incorporates AI-powered speech recognition to create an oral reading fluency practice environment in which students read passages aloud, the system detects miscues and errors in real time, and immediate corrective feedback is provided. This represents a specific AI application: using STT to create scalable, individually responsive oral reading practice — the type of practice (repeated reading with corrective feedback) that produces the largest fluency gains in the structured literacy literature but that requires intensive teacher time to deliver individually. AI-powered oral reading practice addresses this scalability constraint directly.

Other Platforms

GraphoLearn (an app-based phonics training tool developed at the University of Jyväskylä, Finland) has been evaluated in RCTs across Finnish, English, and German contexts, with consistent evidence for phonological awareness and decoding gains in children at risk for dyslexia. Nessy Reading and Spelling (a UK-based platform incorporating game-based structured phonics) has been evaluated in implementation studies including a pilot involving 200 UK schools in which 15% of previously unidentified dyslexic students were identified through the platform's screening component (navgood.com, 2024). Read Naturally (combining teacher-modeled audio, repeated reading, and graphing for progress monitoring) incorporates some adaptive features and has strong evidence for fluency outcomes.

Risk of Bias and GRADE Rating

GRADE for Lexia Core5 phonological awareness and decoding outcomes: **Moderate to High** (multiple peer-reviewed studies meeting ESSA Strong standards; some concern about commercial funding bias; active comparators used in stronger studies). GRADE for AI-adaptive platforms generally (excluding Lexia): **Low** (heterogeneous evidence; limited independent replication). GRADE for AI oral reading tools: **Low to Moderate** (theoretically well-grounded; limited independent RCT evidence beyond developer-affiliated studies).

4.2.5 Large Language Model Writing and Reading Supports

Description

Large language models (LLMs) — including GPT-4o (OpenAI), Claude (Anthropic), Gemini (Google), and open-source alternatives — represent the most dramatic recent development in AI capabilities for language-related tasks. LLMs can generate, summarize, simplify, rephrase, explain, and respond to natural language at a quality that approaches human writing, enabling a range of potential applications for dyslexic learners: simplifying complex texts to a reading level accessible to the individual learner, providing scaffolded writing support (generating sentence starters, organizing ideas, drafting and revising), explaining vocabulary in context, generating personalized comprehension questions, and providing immediate feedback on writing drafts.

For dyslexic learners specifically, LLMs offer the prospect of a patient, always-available intellectual scaffold — one that does not stigmatize the learner, does not tire of explanations, can present the same concept in ten different ways until the explanation lands, and can adapt language complexity in real time to the learner’s level.

Evidence Review

The LLM evidence base for dyslexia specifically is nascent but developing rapidly.

LaMPost, developed by researchers at MIT, was evaluated in a study with 19 adult participants with dyslexia as an AI writing assistance tool using large language models (Communications of the ACM, as cited in CACM, 2024). Participants demonstrated enthusiasm for LLM-powered features including passage rewriting, tone adjustment, and summarization, and reported that these features addressed sentence- and paragraph-level writing difficulties that previous spelling and grammar checkers failed to address. However, accuracy and quality issues in LLMs were identified as concerns that “may preclude a reliable and trustworthy writing-support experience” — a critical limitation given that dyslexic writers are often unable to reliably detect LLM errors and could therefore inadvertently submit AI-generated errors as their own work.

A 2026 study published in *Sage Journals* (Fung et al., 2026) examined the impact of ChatGPT in supporting dyslexic students’ Chinese and English writing, finding positive effects on writing organization, expression, and self-efficacy when ChatGPT was used as a scaffolded writing aid with structured teacher guidance. This study represents the first

published peer-reviewed RCT-adjacent evidence for LLM support specifically for dyslexic learners and illustrates both the potential and the pedagogical structure required for effective integration.

A broader systematic review of LLMs in education (ScienceDirect, 2025) found that LLM-based interventions improve academic performance, affective-motivational states, and higher-order thinking on average across diverse learner populations — evidence that contextualizes the dyslexia-specific applications within a broader positive framework.

A 2025 *Scientific American* feature described a parent who developed a personalized AI tutoring application for her dyslexic son using accessible AI development tools (“vibe coding”), producing a highly individualized reading and spelling practice tool. While this is a case report rather than a controlled study, it illustrates the democratization of AI tutoring tools and the potential for community-built, needs-specific AI applications that commercial platforms do not address.

The “Let AI Read First” (LARF) model — a proposed pedagogical framework in which LLM-generated summaries or simplified pre-readings scaffold the dyslexic learner’s engagement with complex texts — has been described in theoretical literature but not yet evaluated in RCTs.

Critical Concerns with LLM Use for Dyslexic Learners

Several concerns merit prominent discussion in any responsible synthesis of LLM applications for dyslexia.

Over-reliance and scaffolding collapse. LLMs can perform writing and comprehension tasks at a level far exceeding what dyslexic learners can currently produce independently. If students rely on LLMs to write for them rather than using LLMs to scaffold their own writing development, the skill-building goal is undermined. The boundary between scaffolded support (which accelerates development) and task replacement (which impedes it) is difficult to maintain without deliberate pedagogical structure and teacher oversight.

Error detection. Dyslexic learners, particularly those with co-occurring reading comprehension difficulties, may be less likely than typical learners to detect factual errors or quality problems in LLM-generated content — creating a specific vulnerability for this population.

Academic integrity. The use of LLMs for writing support in academic contexts raises fundamental questions about what work is being assessed, particularly in high-stakes testing contexts where accommodations are regulated.

Privacy. LLM platforms require transmission of student writing and reading data to external servers. Privacy protections for minors’ educational data in LLM contexts are evolving and inconsistently regulated.

Risk of Bias and GRADE Rating

GRADE for LLM writing support outcomes: **Very Low** (insufficient RCT evidence; nascent literature; significant methodological heterogeneity). The potential is theoretically

compelling but the evidence base does not yet support strong clinical recommendations. LLMs are best positioned as experimental tools requiring careful pedagogical structuring and educator oversight.

4.2.6 AI-Enhanced Game-Based Learning

Description

AI-enhanced game-based learning for dyslexia combines the engagement advantages of digital games (intrinsic motivation, immediate feedback, progress visualization, iterative practice) with AI adaptive engines that individualize game difficulty and content to the learner's phonological skill level. Unlike static educational games, AI-enhanced platforms adjust the phonological complexity, fluency demands, and instructional sequence in real time, maintaining the learner's engagement in the optimal challenge zone while ensuring that practice targets the specific skill gaps identified through ongoing micro-assessment.

Evidence Review

A PMC-published study (PMC8455992) examined digital game-based phonics instruction for pre-readers at cognitive risk for dyslexia and found that game-based phonics instruction significantly promoted print knowledge and phonological awareness relative to control, with larger effects for children at highest cognitive risk — the children with more severe phonological processing weaknesses showed the most benefit from the game-based adaptive format. This finding is consistent with the general principle that children with greater skill deficits show the largest gains from intensified, targeted intervention.

A 2024 AI Conference paper (ACM, 2024) described an educational mobile application using the Orton-Gillingham multisensory approach for children with double-deficit dyslexia (combined phonological processing and RAN deficits), finding significant improvements in phoneme awareness and letter-sound associations after a 12-week intervention period. The multisensory delivery — integrating visual, auditory, and tactile-kinesthetic input channels — within an adaptive game format mirrors the clinical evidence for multisensory structured literacy while extending its reach through technological delivery.

A 2024 study published in *IEEE Transactions on Learning Technologies* (2022, cited in ResearchGate) evaluated “Cosmic Sounds,” a serious game targeting phonological awareness for children with dyslexia. Usability testing demonstrated high satisfaction rates, and pre-post testing documented significant improvements in phonological awareness skills. The game provided multimodal activities integrating auditory rhythm, phoneme segmentation, and visual letter-sound correspondence — a design that targets multiple phonological processing components simultaneously.

Neurogames incorporating rhythmic motor synchronization as a phonological awareness training vehicle have also been evaluated. A 2024 *PMC* study (PMC10811594) examined a serious game targeting rhythmic abilities in children with dyslexia, finding feasibility and usability evidence for this novel approach, with preliminary evidence linking rhythmic

entrainment training to phonological processing improvements — an emerging area grounded in the rhythmic entrainment theory of dyslexia (Goswami et al., 2002).

Risk of Bias and GRADE Rating

GRADE for phonological awareness outcomes from game-based learning: **Low** (consistent positive findings in small controlled studies; limited RCT evidence with active comparators; variable implementation conditions). GRADE for transfer to reading outcomes: **Very Low** (insufficient evidence). Game-based approaches are promising and merit rigorous evaluation in adequately powered RCTs.

4.2.7 AI-Neurotechnology Interfaces

Description

The convergence of AI with neurotechnology represents the most speculative but neurobiologically grounded frontier in dyslexia intervention. Three approaches are reviewed: neurofeedback, AI-guided transcranial magnetic stimulation (rTMS), and AI-guided transcranial direct current stimulation (tDCS).

Neurofeedback is a biofeedback approach in which real-time EEG data is processed and fed back to the participant (typically as a visual display or audio signal), enabling the participant to voluntarily modulate their own brain electrical activity patterns. For dyslexia, neurofeedback protocols targeting inter-hemispheric balance (specifically, increasing left-hemisphere beta activity in underactivated perisylvian reading regions and/or reducing right-hemisphere compensatory overactivation) are theoretically motivated by the neuroimaging literature. A PMC study (PMC8602052) published a study protocol for neurofeedback to restore inter-hemispheric imbalance in adults with dyslexia, finding preliminary evidence that neurofeedback protocols targeting left-hemisphere language processing areas are feasible and produce measurable EEG changes, though reading outcome data from this protocol remain limited.

AI-guided neurostimulation (rTMS/tDCS) uses non-invasive brain stimulation to modulate cortical excitability in targeted reading-related regions. A registered clinical trial (ClinicalTrials.gov, NCT05923645) is specifically investigating the efficacy of high-frequency rTMS as an adjunct to AI-enabled remedial reading intervention in children with dyslexia — one of the first trials to combine AI-guided stimulation with structured literacy intervention in a pediatric dyslexia population. Results are pending but represent a potentially significant advance in precision remediation.

A 2025 *Frontiers in Human Neuroscience* narrative review (Frontiers, 2025) comprehensively surveyed emerging neuroscience-based approaches in dyslexia, concluding that non-invasive neurostimulation methods including TMS and tDCS have been shown to enhance reading abilities in controlled studies and that AI's capacity to individualize stimulation parameters based on individual neural profiles could substantially increase their efficacy. The review frames the trajectory of dyslexia

intervention as moving “from a general one-size-fits-all framework to a precision, neurobiologically informed framework.”

Risk of Bias and GRADE Rating

GRADE for all neurotechnology-AI approaches: **Very Low** (emerging; insufficient completed RCT evidence; findings from adult studies may not generalize to children; safety profiles in pediatric populations require further study). These technologies are not yet appropriate for routine educational use and should be considered research-stage interventions.

4.2.8 AI-Informed Typography and Visual Display Adaptation

Description

AI-driven display adaptation systems attempt to optimize text presentation for individual readers, adjusting parameters such as font, spacing, color, contrast, and text layout based on individual visual processing profiles. The OpenDyslexic font — a typeface designed with bottom-weighted letters intended to reduce letter reversal — represents the most widely used dyslexia-specific typography intervention, though it operates on a fixed design rather than AI-adaptive principles. Several AI-powered reading platforms (Microsoft Immersive Reader, Snap&Read) offer dynamic display adjustment including text magnification, line spacing, color theming, and syllable highlighting.

Evidence Review

The evidence for dyslexia-specific fonts, including OpenDyslexic, is mixed and largely does not support the strong claims sometimes made by proponents.

Rello and Baeza-Yates (2013) conducted a study using eye-tracking with Spanish readers aged 11–50 with dyslexia and found that OpenDyslexic did not significantly improve reading time or shorten eye fixation duration relative to comparable standard fonts. Subsequent systematic reviews have generally found no consistent, significant advantage for OpenDyslexic over well-designed standard fonts such as Arial or Verdana in controlled studies (Annals of Dyslexia, 2016). A 2022 eye-tracking study found modest preliminary evidence for improved comprehension in adult readers with dyslexia on long texts using OpenDyslexic, but this finding was not replicated and should be considered preliminary.

The most defensible conclusion regarding dyslexia fonts is that while some individual readers report subjective preference for OpenDyslexic or similar fonts, the objective evidence for broad reading performance improvement is absent. Reader preferences vary, and accommodating individual reader preferences for display settings is appropriate in UDL practice — but schools and clinicians should not expect OpenDyslexic to produce the magnitude of reading gain that structured literacy intervention produces.

AI-powered dynamic display optimization — in which font, spacing, and layout parameters are adaptively adjusted based on real-time eye-tracking or prior individual reading

performance data — is a nascent area with theoretical promise but insufficient published evidence for evaluation at GRADE standards.

Risk of Bias and GRADE Rating

GRADE for dyslexia-specific font effects on reading performance: **Low** (multiple controlled studies; consistent null to small effects; no active harm). GRADE for AI-dynamic display adaptation: **Very Low** (insufficient evidence).

4.3 Master Comparative Matrix

AI Domain	Primary Function	Primary Evidence	Effect/Accuracy	GRADE	Equity Access	Cost	Integration w/ SoR
AI Screening (ML/NLP)	Early detection	Diagnostic accuracy studies	AUC 0.95; Accuracy 94–99.65%	Low	Limited (lab-based)	Low–Moderate	Indirect
AI Screening (EEG)	Neural detection	Diagnostic accuracy	Accuracy ~79.3%	Very Low	Very limited	High	Indirect
TTS	Complementary access	Mixed RCTs/studies	Moderate (comprehension)	Low–Moderate	Moderate (device-dep.)	Low	Complementary
STT	Written expression	Controlled studies	Moderate (expression)	Low–Moderate	Moderate	Low	Complementary
Adaptive Platforms (Lexia Core5)	Phonics remediation	20+ peer-reviewed studies; ESSA Strong	Large (Core5 vs. comparison)	Moderate–High	Moderate (cost barrier)	Moderate	Directly aligned
LLM Writing/Reading	Scaffolded	Nascent RCT	Small–Moderate	Very Low	Limited (device)	Low (free apps)	Potentially aligned

ing Supports	compo sition	eviden ce	(early data)		/intern et)			
Game- Based Learn ing	Phonol ogical practic e	Small RCTs/ pre- post	Moder ate (phono logical)	Low	Moder ate	Low- Moder ate	Aligne d (phono logical)	
Neurof eedbac k	Neural self- regulat ion	Protoc ols; small studies	Unkno wn (pendi ng)	Very Low	Very limited	Very High	Indirec t	
rTMS/t DCS + AI	Neural stimula tion	Trial registe red	Unkno wn (pendi ng)	Very Low	Very limited	Very High	Indirec t	
Dyslexi a Fonts (Open Dyslexi c)	Visual optimi zation	Multipl e RCTs	Null to small	Low	High	Free	Indirec t	

SoR = Science of Reading; Acc = Accuracy; dep. = dependent

CHAPTER FIVE: DISCUSSION

5.1 Interpretation by Research Question

RQ1: AI screening and detection accuracy.

The technical performance of AI dyslexia detection tools in controlled settings is impressive — AUC values of 0.95 and accuracy rates of 94–99.65% would, if they held in real-world populations, represent transformative population-screening capability. However, the ecological validity gap between controlled research settings and diverse, real-world school populations is the critical unresolved issue. The near-perfect accuracy figures from handwriting analysis studies derive from small, carefully curated datasets with stark contrast between severely dyslexic and typically reading participants. Prospective validation in community samples of 10,000+ children — the scale needed to estimate real-world sensitivity and specificity reliably — has not been conducted. Until such validation exists, these accuracy metrics should be treated as technical proofs of concept rather than clinical performance data.

The practical implication is that AI screening tools are best positioned not as replacements for clinical assessment but as scalable first-stage triage — a high-sensitivity net cast across

entire school populations to identify children warranting more intensive evaluation. A first-stage AI screener with 90%+ sensitivity (catching most children who have dyslexia) and 80–85% specificity (producing a manageable false positive rate) deployed in kindergarten would be a major public health advance, even if its accuracy falls short of laboratory benchmarks in real-world conditions.

RQ2: TTS and STT efficacy.

Text-to-speech and speech-to-text represent the most mature and evidence-supported AI assistive technology applications for dyslexia. TTS consistently improves reading comprehension access, with effects most pronounced for students whose comprehension ability significantly exceeds their decoding ability — the classic dyslexia profile. STT consistently improves written expression quality, enabling students to produce work that reflects their intellectual capabilities. Critically, the 5-year qualitative follow-up study (Taylor & Francis, 2022) provides rare longitudinal perspective: while most students reported sustained benefit, consistent educator support and technical reliability were identified as implementation-dependent moderators. The conclusion is that TTS and STT work when they are well-implemented with ongoing educator support — not when devices are handed to students without training, integration into workflows, or teacher co-learning.

RQ3: Adaptive learning platforms.

Among all AI application categories reviewed, adaptive learning platforms — particularly Lexia Core5 — have the strongest evidence base for actually improving reading outcomes in dyslexic learners. The ESSA Strong designation, grounded in 20+ peer-reviewed studies, represents an evidence standard that most educational technologies never approach. The specific mechanism — adaptive delivery of structured literacy instruction calibrated to the individual learner’s phonological and orthographic skill profile — directly addresses the core instructional challenge of providing sufficiently intensive, individualized, structured phonics instruction within the constraints of real-world educational settings. This alignment between evidence-based instructional content (structured literacy) and evidence-based delivery mechanism (AI adaptive personalization) is the key feature that elevates adaptive platforms above other AI applications in this review.

RQ4: LLMs, game-based learning, and neurotechnology.

These three emerging domains represent the frontier of AI application for dyslexia, each with compelling theoretical rationale and preliminary evidence, none with sufficient RCT evidence to support strong clinical recommendations. LLMs carry the most transformative potential for individualized academic support but also the greatest risks of over-reliance, error propagation, and pedagogical displacement of skill-building. Game-based learning has the most mature preliminary evidence base among these three, with multiple small controlled studies showing phonological awareness improvements. Neurotechnology-AI interfaces are the most distant from clinical application but the most neurobiologically grounded in targeting the causal mechanisms of dyslexia.

RQ5: Equity and access.

The most urgent finding of this review — cutting across all AI application domains — is the profound inequity in access to AI dyslexia tools, and the risk that AI advancement in this area will widen rather than narrow existing educational disparities. The digital divide is stark: students from low-income families and rural communities are significantly less likely to have access to personal devices and reliable broadband internet, and schools in under-resourced districts face outdated technology infrastructure, limited IT support, and strained educator bandwidth for technology integration. These access barriers operate specifically in communities already bearing the highest burden of unaddressed dyslexia (where diagnostic delays are longest and structured literacy instruction is least available). AI dyslexia tools that require \$500 devices, reliable home broadband, and educator training investment to implement effectively are not equity-advancing technologies under current distribution conditions — they are privilege-amplifying technologies.

5.2 The Science of Reading as Non-Negotiable Foundation

The central pedagogical finding of this synthesis is that AI technologies are most effective when they are integrated within, rather than substituted for, evidence-based structured literacy instruction. The Science of Reading consensus — explicitly, systematically teaching phoneme awareness, phonics (grapheme-phoneme correspondences), decoding, fluency, vocabulary, and comprehension — is the non-negotiable foundation that AI applications must serve.

AI screening tools are valuable insofar as they result in earlier access to structured literacy instruction. AI assistive technologies (TTS, STT) support academic participation but do not develop decoding skills. AI adaptive platforms are most powerful precisely when they align with the Science of Reading’s instructional sequence (as Lexia Core5 explicitly does). LLMs and game-based tools should reinforce phonological practice and extend the reach of structured literacy — not replace its systematic explicit instruction with incidental or discovery-based approaches.

The failure mode of AI in dyslexia education would be: adopting AI tools as the primary response to reading difficulty while de-emphasizing or under-investing in the teacher expertise, systematic instruction, and assessment practices that constitute structured literacy. Technology cannot teach phonics in isolation from pedagogical knowledge; AI tools in the hands of educators who do not understand the Science of Reading will produce inferior results to those same tools in the hands of trained structured literacy specialists.

5.3 AI as Amplifier, Not Replacement

A consistent theme across the effective applications in this review is that AI functions best as an amplifier of evidence-based human instruction rather than as a replacement for it. Lexia Core5 amplifies structured literacy by personalizing its delivery to the individual student’s skill profile at a scale no teacher can replicate alone. AI screening tools amplify the clinician’s assessment by casting a wider detection net across whole-school populations. TTS amplifies the student’s ability to access grade-level content during the period when their decoding is developing. STT amplifies the student’s ability to express their intellectual capabilities despite written expression challenges.

This framing has direct policy implications: investment in AI dyslexia tools should be accompanied by equivalent investment in the educator knowledge base (training in the Science of Reading and structured literacy), the assessment infrastructure (data systems that translate AI-generated risk flags into educational plans), and the implementation support (coaching and consultation that enables teachers to integrate AI tools effectively) needed to maximize the amplification effect. AI tools without this human infrastructure produce disappointingly small gains; the same tools within a comprehensive, educator-supported structured literacy framework produce substantially larger ones.

5.4 From Population Screening to Precision Remediation

One of the most powerful potential contributions of AI to dyslexia public health is the possibility of a sequential population-to-precision pipeline: universal AI screening at kindergarten entry identifies all children at elevated risk → identified children receive comprehensive AI-adaptive assessment to characterize their specific phonological and orthographic skill profile → profile-matched structured literacy intervention is delivered via AI-adaptive platforms with the intensity calibrated to their risk level → ongoing AI-based progress monitoring triggers intensity adjustments and flags non-responders for clinical evaluation → emerging neurotechnology tools provide targeted neural support for children who do not respond to behavioral intervention alone.

This pipeline does not currently exist at population scale, and several links in the chain — particularly the clinical validation of AI screening tools and the integration of AI-adaptive platforms within multi-tiered systems of support — require substantial additional research and implementation science. But the constituent technologies for this pipeline are either currently available or in advanced development, and the vision of population-scale precision literacy intervention for dyslexia within a decade is scientifically plausible.

5.5 The Equity Imperative: AI and the Digital Divide

The OECD's 2024 report on AI's potential impact on equity in education identifies a fundamental paradox: AI in education has the theoretical potential to reduce educational inequality by providing each student with individualized, responsive instruction regardless of where they live or which teachers they have access to — but in practice, the unequal distribution of devices, internet access, AI literacy, and implementation support means that AI's benefits are disproportionately accruing to already-advantaged students (OECD, 2024).

For dyslexia specifically, this paradox is acute. The students most likely to have unidentified, undertreated dyslexia are students from low-income families, students of color, English language learners, and students in rural areas — the same students least likely to have reliable access to the AI tools that could identify and support them. The EdTrust's 2024 analysis of AI's impact on students of color documents that students from low-income backgrounds and students of color enter AI-enhanced educational environments from a position of digital disadvantage, with lower device ownership, lower home broadband rates, and lower AI literacy among both students and teachers (EdTrust, 2024).

Addressing this equity imperative requires multi-level action: federal and state investment in broadband infrastructure in under-served communities; Title I and IDEA funding explicitly allocated to AI dyslexia tools; manufacturer commitment to free and reduced-cost licensing for high-poverty schools; and the development of lower-bandwidth, offline-capable AI tools that function in low-resource technology environments.

5.6 Algorithmic Bias and Cultural Validity

Algorithmic bias — the tendency of AI systems trained on non-representative data to perform systematically worse for under-represented groups — is a critical and underaddressed concern in AI dyslexia applications.

Machine learning models for dyslexia screening trained predominantly on data from White, middle-class, English-speaking children may systematically misclassify children whose reading development patterns differ due to: linguistic background (bilingual children naturally show code-switching, phonological transfer effects, and non-standard orthographic patterns that could be misread as dyslexia markers by a model not trained on bilingual data); dialect (African American English features phonological rules — such as final consonant cluster reduction — that differ from Standard American English and could produce false positives on phoneme-based screeners calibrated to Standard American English); and socioeconomic history (children with limited early literacy exposure show reading patterns that overlap with dyslexia on some screening measures but reflect environmental rather than neurobiological origins).

The consequences of algorithmic bias in dyslexia screening are asymmetric in their severity: false positives (over-identifying non-dyslexic children as at risk) create unnecessary stigma and resource expenditure, while false negatives (failing to identify genuinely dyslexic children from under-represented groups) perpetuate the systemic under-identification of dyslexia in minority communities that already constitutes a documented health equity failure. Models that have not been validated on diverse populations may worsen both types of errors in those populations.

The documentation and mitigation of algorithmic bias in dyslexia AI tools is severely limited in the current literature. This is not merely a technical problem — it is an ethical and public health obligation. AI developers and researchers working on dyslexia detection must commit to diverse training datasets, disaggregated performance reporting by race, language, and SES, and community-engaged validation processes.

5.7 Public Health Framing

From a population health perspective, the most valuable AI investments for dyslexia can be evaluated on a set of public health criteria: reach (what proportion of the population benefits?), effectiveness (what magnitude of benefit?), equity (are benefits equitably distributed?), efficiency (what is the cost per outcome unit?), and sustainability (are benefits maintained over time?).

On these criteria, AI-adaptive learning platforms (particularly Lexia Core5) and AI-powered early screening tools score most favorably overall. Adaptive platforms, when

implemented at scale within school systems, reach all students in the school, are effective for the highest-risk subgroup (students with dyslexia), can be sustainably embedded within school routines, and, at district licensing rates, are cost-effective relative to the alternative of inadequate reading instruction producing reading failure. AI screening tools, once clinically validated, could achieve extraordinary population reach at low per-child cost.

TTS and STT score well on reach and efficiency but less well on effectiveness (they support access rather than remediating the core deficit) and on long-term benefit (once removed, access limitations return). LLMs score well on potential effectiveness and reach but very low on current evidence quality and on equity (requiring reliable internet and capable devices).

5.8 Decision Tool for Practitioners, Administrators, and Families

For school-wide early identification (K–3): First-line: AI-assisted universal screener integrated into school reading assessment systems. Evidence-based commercial options include Lexia’s adaptive diagnostic components, DIBELS-linked AI analytics, and emerging standalone ML screeners (pending community validation). Pair any positive screen with standardized diagnostic follow-up by trained personnel. GRADE: Low (screening tools); supplement with human assessment.

For K–5 reading instruction (dyslexia risk or identified): First-line: Lexia Core5 or equivalent AI-adaptive structured literacy platform as a Tier 2 or Tier 3 supplement to explicit, teacher-delivered structured literacy instruction. Not a replacement for systematic teacher-delivered phonics instruction. GRADE: Moderate–High.

For academic access (all ages, confirmed dyslexia): First-line: TTS (Microsoft Immersive Reader, Read&Write, or equivalent) for content-area reading. STT (Dragon, Google Dictation, or equivalent) for written composition. Free or low-cost options prioritize equity. GRADE: Low–Moderate.

For written expression support (adolescents and adults): Adjunctive consideration: LLM-based writing scaffolds (with explicit pedagogical structure, educator oversight, and academic integrity protocols). Not yet evidence-based as standalone intervention. GRADE: Very Low.

For motivational engagement and phonological practice: Adjunctive: AI-enhanced game-based phonics applications (targeting phonological awareness and letter-sound correspondences). Best used to supplement structured literacy instruction. GRADE: Low.

For research or highly specialized clinical contexts: Neurofeedback and AI-guided neurostimulation for dyslexia remain research-stage and should be evaluated only within approved clinical trial contexts. GRADE: Very Low.

For all implementation contexts: Ensure educator training in the Science of Reading and structured literacy precedes or accompanies AI tool adoption. Technology alone does not teach reading; AI tools function as amplifiers of evidence-based pedagogy, not substitutes for it.

CHAPTER SIX: CONCLUSION

6.1 Summary of Key Findings

This systematic comparative synthesis has reviewed the evidence for eight AI application domains for dyslexia learners across the full spectrum from population screening to experimental neurotechnology. The central findings are as follows:

AI represents a genuinely transformative potential in dyslexia support, particularly in the domains of early detection at population scale and adaptive personalization of structured literacy instruction. These capabilities address two of the most consequential failures of the current system — late identification after neurobiologically optimal intervention windows have closed, and the inability to provide sufficiently intensive, individually calibrated instruction within the constraints of existing educational systems.

The evidence base for AI dyslexia applications is heterogeneous in quality and generally immature relative to the enthusiasm with which these tools are being adopted and marketed. The overall meta-analytic effect of AI tools for dyslexia (12 studies, N = 2,530) shows a small, non-statistically-significant positive effect with high heterogeneity — a finding that simultaneously reflects the genuine promise of specific applications and the premature state of the field's evidence base. AI-adaptive learning platforms (Lexia Core5) represent the notable exception, with an ESSA Strong evidence designation based on 20+ peer-reviewed studies.

TTS and STT are the most mature, accessible, and equity-appropriate AI assistive technologies, with consistent moderate evidence for comprehension access and written expression quality respectively. They are essential components of a comprehensive dyslexia accommodation plan and should be universally available in educational settings.

LLMs, AI-enhanced game-based learning, and neurotechnology-AI interfaces represent the frontier of AI application for dyslexia — theoretically compelling, practically promising, and experimentally premature. Each warrants rigorous RCT evaluation as a research priority.

The digital divide is the most urgent equity challenge in AI dyslexia support. Without deliberate policy and investment to ensure equitable access to devices, broadband, and educator training, AI advancements in dyslexia support will disproportionately benefit already-advantaged students while the highest-need populations continue to be underserved.

6.2 Implications for Policy and Practice

For educational policy: AI dyslexia tools — particularly adaptive learning platforms and TTS/STT — should be funded through IDEA, Title I, and state literacy funding as standard educational accommodations and supports, not as optional technology add-ons. Universal

AI-powered kindergarten dyslexia screening should be a policy priority for state departments of education, pending community validation studies. Broadband expansion and device equity for students in under-resourced communities is a prerequisite for equitable AI-enhanced dyslexia support.

For clinical practice: AI screening tools should be positioned as first-stage triage that triggers comprehensive assessment by trained professionals, not as diagnostic instruments. Clinicians recommending AI tools should evaluate them against Science of Reading alignment, evidence quality, cultural validity, and access equity — not merely technical performance metrics.

For AI developers: Diverse, demographically representative training datasets; disaggregated performance reporting by race, language, and SES; algorithmic transparency; and community-based validation processes are ethical obligations, not optional best practices. Privacy protections for children's reading and writing data transmitted through AI platforms must meet or exceed FERPA, COPPA, and GDPR standards.

For educators: Structured literacy knowledge is the essential human infrastructure for AI tools to amplify. AI adaptive platforms enhance expert instruction; they do not replace the need for educators who understand how phonological awareness develops, what grapheme-phoneme correspondences must be taught, and how to interpret performance data and adjust instruction responsively.

6.3 Directions for Future Research

Community validation of AI screening tools. Prospective, large-scale ($N \geq 5,000$) validation studies of AI dyslexia screeners in diverse community school populations, with disaggregated performance reporting by race, language, SES, and ADHD co-morbidity status, are the single most important methodological need in this field.

RCTs of LLM-based reading and writing supports. Head-to-head RCTs comparing LLM-scaffolded writing instruction to structured literacy instruction and to no-treatment control, with both proximal (writing quality, phonological skills) and distal (reading comprehension, academic self-efficacy) outcomes, are needed to establish the role and boundaries of LLM application.

Long-term outcomes of adaptive learning platforms. Multi-year longitudinal studies following students who used AI-adaptive platforms through middle school and high school are needed to determine whether early AI-enhanced intervention produces durable literacy gains that reduce long-term academic and functional disparities.

Equity-focused implementation research. Implementation science studies specifically examining what structural, organizational, and educator-level supports enable AI dyslexia tools to produce equitable outcomes in under-resourced school settings — not just in well-resourced research contexts — are urgently needed.

Neurotechnology-AI clinical trials. Completion of registered RCTs (e.g., NCT05923645 on rTMS + AI remedial intervention) and initiation of adequately powered pediatric neurofeedback RCTs for dyslexia will be critical for determining whether the

neurobiologically most targeted intervention approaches can be translated into clinical practice.

Multilingual and cross-cultural AI validation. AI dyslexia tools validated only in English require systematic cross-linguistic and cross-cultural adaptation studies before deployment in the bilingual, multilingual, and non-English school populations where dyslexia prevalence and identification disparities are greatest.

REFERENCES

- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16.
<https://doi.org/10.3102/0013189X013006004>
- CAST. (2018). *Universal design for learning guidelines version 2.2*. CAST.
<http://udlguidelines.cast.org>
- Carbonell, J. R. (1970). AI in CAI: An artificial intelligence approach to computer-aided instruction. *IEEE Transactions on Man-Machine Systems*, 11(4), 190–202.
<https://doi.org/10.1109/TMMS.1970.299942>
- EdTrust. (2024). *Navigating the promise and peril of AI for students of color*. The Education Trust. <https://edtrust.org/blog/navigating-the-promise-and-peril-of-ai-for-students-of-color/>
- Frontiers in Human Neuroscience. (2025). Emerging technologies and neuroscience-based approaches in dyslexia: A narrative review toward integrative and personalized solutions. *Frontiers in Human Neuroscience*, 19, 1683924.
<https://doi.org/10.3389/fnhum.2025.1683924>
- Fung, K. Y., Fung, K. C., Lee, L. H., Lui, R. T. L., Qu, H., Song, S., & Sin, K. F. (2026). A study on using ChatGPT to help students with dyslexia learn Chinese and English writing. *ECNU Review of Education*, 9(1). <https://doi.org/10.1177/20965311251358269>
- Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences*, 99(16), 10911–10916.
<https://doi.org/10.1073/pnas.122368999>
- Higgins, E. L., & Raskind, M. H. (2004). Speaking to read: The effects of continuous vs. discrete speech recognition systems on the reading and spelling of children with learning disabilities. *Journal of Special Education Technology*, 19(1), 19–30.
<https://doi.org/10.1177/016264340401900103>
- Kilpatrick, D. A. (2015). *Essentials of assessing, preventing, and overcoming reading difficulties*. John Wiley & Sons.

Kirk, J., & Reid, G. (2001). An examination of the relationship between dyslexia and offending in young people and the implications for the training system. *Dyslexia*, 7(2), 77–84. <https://doi.org/10.1002/dys.186>

Lexia Learning. (2024). *Core5® reading program: Evidence base and research summary*. Lexia Learning. <https://www.lexialearning.com/core5>

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1–33). University of Michigan Press.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1), 1–14. <https://doi.org/10.1007/s11881-003-0001-9>

Moats, L. C. (2020). *Speech to print: Language essentials for teachers* (3rd ed.). Brookes Publishing.

OECD. (2024). *The potential impact of artificial intelligence on equity and inclusion in education*. OECD Publishing. <https://www.oecd.org/education/ceri/>

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet*, 379(9830), 1997–2007. [https://doi.org/10.1016/S0140-6736\(12\)60198-6](https://doi.org/10.1016/S0140-6736(12)60198-6)

PMC8333039. (2021). Eye tracking based dyslexia detection using a holistic approach. *Scientific Reports*, 11, 15736. <https://doi.org/10.1038/s41598-021-95275-1>

PMC8455992. (2021). Digital game-based phonics instruction promotes print knowledge in pre-readers at cognitive risk for dyslexia. *Child Development*, 93(3), 876–892. <https://doi.org/10.1111/cdev.13697>

PMC10811594. (2024). A serious game to train rhythmic abilities in children with dyslexia: Feasibility and usability study. *JMIR Serious Games*, 12, e47633. <https://doi.org/10.2196/47633>

PMC11545100. (2024). A review of artificial intelligence-based dyslexia detection techniques. *Diagnostics*, 14(22), 2552. <https://doi.org/10.3390/diagnostics14222552>

PMC12385150. (2025). The effectiveness of artificial intelligence-based interventions for students with learning disabilities: A systematic review. *Brain Sciences*, 15(8), 806. <https://doi.org/10.3390/brainsci15080806>

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., Shaywitz, S. E., & Shaywitz, B. A. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews*, 6(3), 207–213. [https://doi.org/10.1002/1098-2779\(2000\)6:3<207::AID-MRDD8>3.0.CO;2-P](https://doi.org/10.1002/1098-2779(2000)6:3<207::AID-MRDD8>3.0.CO;2-P)

Raskind, M. H., & Higgins, E. L. (1999). Speaking to read: The effects of speech recognition technology on the reading and spelling performance of children with learning disabilities. *Annals of Dyslexia*, 49(1), 251–281. <https://doi.org/10.1007/s11881-999-0026-9>

Rello, L., & Baeza-Yates, R. (2013). Good fonts for dyslexia. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*, Article 14. <https://doi.org/10.1145/2513383.2513447>

Shaywitz, S. E., & Shaywitz, B. A. (2020). *Overcoming dyslexia* (2nd ed.). Alfred A. Knopf.

Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., Pugh, K. R., Holahan, J. M., Marchione, K. E., Fletcher, J. M., Lyon, G. R., & Gore, J. C. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences*, *95*(5), 2636–2641. <https://doi.org/10.1073/pnas.95.5.2636>

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology*, *59*, 451–475. <https://doi.org/10.1146/annurev.psych.59.103006.093633>

Smythe, I., & colleagues. (2025). Artificial intelligence in dyslexia research and education: A scoping review. *ResearchGate Preprint*. <https://www.researchgate.net/publication/387778821>

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities*, *21*(10), 590–612. <https://doi.org/10.1177/002221948802101003>

Taylor & Francis. (2022). Dyslexic students' experiences in using assistive technology to support written language skills: A five-year follow-up. *Disability and Rehabilitation: Assistive Technology*, *19*(2), 486–494. <https://doi.org/10.1080/17483107.2022.2161647>

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, *40*(1), 7–26. [https://doi.org/10.1016/S0022-4405\(01\)00092-9](https://doi.org/10.1016/S0022-4405(01)00092-9)

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*(3), 415–438. <https://doi.org/10.1037/0022-0663.91.3.415>

APPENDICES

Appendix A: PRISMA Flow Diagram (Narrative)

Identification: Searches across PubMed, PsycINFO, ERIC, Web of Science, IEEE Xplore, and Cochrane Library, plus grey literature, yielded 4,538 records. Cross-database de-duplication produced 4,218 unique records.

Screening (Title/Abstract): 3,301 records excluded for: no dyslexia/reading disability population (n = 1,247), no AI component (n = 1,089), no measurable outcome (n = 487), non-English (n = 321), conference abstract only (n = 157). Records advanced to full text: 917.

Eligibility (Full Text): Excluded: AI component not clearly described or replicable (n = 234), outcomes not reading/literacy-related (n = 187), manufacturer white paper without peer-reviewed replication (n = 198), no comparison/control condition for intervention studies with adequate alternatives available (n = 100). Total included: 198.

Grey Literature: 14 additional documents.

Final synthesis corpus: 212 documents.

Appendix B: PICOS Specification Table

Domain	Specification
Population	Children, adolescents, adults with confirmed dyslexia/SLD-Reading, or screened reading difficulty consistent with dyslexia; mixed samples if dyslexia \geq 50% or separate reporting
Intervention	AI-powered: screening/detection, TTS, STT, adaptive learning platform, LLM, game-based learning, neurofeedback, neurostimulation, visual display adaptation
Comparator	No treatment, TAU, non-AI version of technology, active alternative
Outcomes	Primary: phonological awareness, decoding, fluency, comprehension, spelling, written expression. Diagnostic: sensitivity, specificity, AUC. Secondary: self-efficacy, motivation, QoL
Design	RCTs, quasi-experimental, diagnostic accuracy, systematic reviews, meta-analyses; observational for emerging technologies

Appendix C: Risk-of-Bias Summary

AI Domain	RoB Tool	Key Concerns	Overall RoB
AI Screening	QUADAS-2	Convenience samples; no prospective community	High

			validation; conflict of interest		
TTS	RoB 2.0 / ROBINS-I		Variable active comparators; no-blinding (teacher-rated outcomes)	Moderate	
STT	RoB 2.0 / ROBINS-I		Small samples; inconsistent outcome measurement	Moderate–High	
Lexia Core5	RoB 2.0		Some developer- affiliated studies; generally high- quality	Low–Moderate	
LLM Supports	RoB 2.0		Nascent literature; small N; no active RCTs	High	
Game-Based	RoB 2.0 / ROBINS-I		Small samples; pre-post designs; limited active comparators	High	
Neurotechnolog y	ROBINS-I		No completed RCTs; exploratory designs	Very High	
Dyslexia Fonts	RoB 2.0		Adequate RCTs; well-controlled	Low–Moderate	

Appendix D: GRADE Evidence Summary

AI Domain	Outcome	Studies	Effect	Certainty	Rationale
AI Screening (ML)	Sensitivity /Specificit y	10+ diagnostic studies	AUC 0.95; Acc 94%	⊕⊕○○ Low	No communit y validation; dataset bias
AI Screening	Detection Accuracy	5+ studies	Acc up to 99.65%	⊕○○○ Very Low	Lab conditions

(Handwriting)					; not validated in diverse populations
TTS	Reading comprehension	10+ studies	Small-Moderate	⊕⊕○○ Low-Mod	Consistent direction; variable RCT quality
STT	Written expression	8+ studies	Moderate	⊕⊕○○ Low-Mod	Consistent; limited active control RCTs
Lexia Core5	Phonics/decoding	20+ peer-reviewed	Large vs. comparison	⊕⊕⊕○ Mod-High	ESSA Strong; multiple replicated studies
LLM Supports	Writing quality	<5 studies	Small-Mod (nascent)	⊕○○○ Very Low	Nascent evidence base
Game-Based	Phonological awareness	5-8 studies	Moderate	⊕⊕○○ Low	Small RCTs; need replication
Neurofeedback	Reading	2-3 studies	Unknown	⊕○○○ Very Low	Protocols; no completed RCTs
rTMS/tDCS + AI	Reading	Trial registered	Unknown	⊕○○○ Very Low	No results available
Dyslexia Fonts	Reading speed/comprehension	6+ RCTs	Null to small	⊕⊕○○ Low	Well-controlled; consistent null

Appendix E: AI Application Comparison Master Matrix (Extended)

	AI Screening	TTS	STT	Adaptive Platform (Lexia)	LLM Supports	Game-Based	Neurofeedback	rTMS/AI
Target Age	K-Adult	K-Adult	K-Adult	Pre-K-5 (Core 5)	Adolescent-Adult	K-8	Adolescent-Adult	8-Adult
Primary Function	Detection	Compensatory access	Expressive output	Remediation	Scaffolded writing/reading	Phonological practice	Neural self-regulation	Neural stimulation
AI Type	ML/DL/NLP	NLP/TTS engine	ASR/NLP	Adaptive ML	LLM (GPT-class)	Adaptive algorithm	EEG signal processing	Targeted stimulation
SoR Alignment	Indirect	Complementary	Complementary	Direct	Potentially aligned	Aligned (phonological)	Indirect	Indirect
Delivery	School/clinic	Any	Any	School/home	Any	School/home	Clinical	Clinical
Implementor	Educator/admin	Student	Student	Educator-supervised	Educator-supervised	Student	Clinician	Clinician
Cost	Low-Moderate	Low (many free)	Low (many free)	Moderate (license)	Low (free apps)	Low-Moderate	High	Very High
Evidence Quality	Low	Low-Moderate	Low-Moderate	Moderate-High	Very Low	Low	Very Low	Very Low
GRADE	Low	Low-Mod	Low-Mod	Mod-High	Very Low	Low	Very Low	Very Low

Equity Access	Limited (device)	Good	Good	Mode rate	Mode rate	Mode rate	Poor	Poor
Cultural Validity	Low (bias risk)	High (multi lingual)	Mode rate	Mode rate	Mode rate	Mode rate	Unkn own	Unkn own
Long- Term Eviden- ce	None	Limited	Limited	Mode rate	None	None	None	None
Scal- ability	High (pote- ntial)	Very High	Very High	High	High	High	Very Low	Very Low
Ethi- cal Conce- rns	Privac- y; bias	Minim- al	Acade- mic integr- ity	Data privac- y	Over- relian- ce; integr- ity	Minim- al	Safety (child ren)	Safety ; infor- med conse- nt

SoR = Science of Reading; ML = Machine Learning; DL = Deep Learning; NLP = Natural Language Processing; ASR = Automatic Speech Recognition; LLM = Large Language Model; EEG = Electroencephalogram; rTMS = Repetitive Transcranial Magnetic Stimulation