Challenges and Design Opportunities for AI-based Tutoring and Assessment Software in Special Education: An Interview Study with Teachers

C.Naomie Williams[1[0009-0001-8987-2811]], Margaret Ellen Seehorn[2[0000-0002-6949-9079]], Haiyi Zhu[1 [0000-0001-7271-9100]], and Vincent Aleven[1 [0000-0002-1581-6657]]

[1] Carnegie Mellon University, Pittsburgh PA 15213, USA
[2] Grinnell College, Grinnell IA 50112, USA
{naomiew, haiyiz}@cmu.edu, seehorne@grinnell.edu, aleven@cs.cmu.edu

Abstract. Approximately 240 million children, 10% of children worldwide, have disabilities. However, few studies have investigated requirements for AI-based tutoring and assessment for this population within a real-world context. It is unclear how special education classrooms experience these emergent tools in practice. We therefore conducted interviews with 18 special education teachers in the United States about the usage, challenges, and perceived benefits of AI in their classrooms. The interviews reveal tensions between special education's need for personalized learning systems and difficulty integrating existing techno-solutions due to: (1) poor interface and curriculum adaptations for students with learning disabilities, (2) under-consideration of special education's additional form of academic assessment when differentiating students, and (3) criteria for subject mastery incompatible with special education students' personalized learning plans. Nevertheless, special education teachers remain optimistic towards AI's promise of flexible instruction for students requiring accessible digital interfaces. From these findings, we conclude with design implications and potential research directions to better synergize emerging techno-solutions with special education classrooms' needs. This work envisions tangible pathways towards inclusive AI for students with disabilities and a more equitable classroom of tomorrow.

Keywords: AI in Special Education, Accessibility, Narrative Interview

1 Introduction

Education promised to be the great equalizer. Technology became the medium to improve access to education and catapult students to better futures. However, the extent this promise has come to fruition for students with learning disabilities is unclear. Literature reviews spanning the last decade emphasize the need for more real-world studies examining the use of AI in special education [1–3]. Prior scholarship for AI in special education has contributed innovative case studies and theoretical approaches to support students with disabilities [1, 3–5]. However, contributions in this domain received critique for lacking the longitudinal data to enhance statistical robustness and real-world relevance” [3]. Notwithstanding limitations, AI-based tutoring and assessment software maintain promise for special education.

AI-based tutoring and assessment technologies such as: iReady, IXL, MATHia, and ALEKS, refer to learning software where the interface and curriculum are personalized for each user based on the student’s behavior and academic performance [6]. Typically, these systems help students practice complex problems, provide adaptive (individualized) step-level guidance within these problems, select problems that involve unmastered material, and support individualized placement. Several meta-analyses show that students learn better when using AI-based tutoring and assessment technologies than without [7]. This suggests that such software may be particularly beneficial for underserved students with disabilities, though the study did not focus on this population [8].

Special education, envisioned and mandated in the US as an inclusive, school-based learning environment for students with disabilities, shares AI-based tutors’ ethos for personalized instruction. In the United States, public schools are required to have an Individualized Education Program, or IEP, for each student receiving special education services. IEPs are living documents co-designed by teachers, parents, school administrators, and related services personnel. Teachers use the IEP as a guide to determine what curriculum modifications, personalized learning goals, and assistive technologies are necessary for an individual student. A mutual ethos of personalization would suggest the possibility of a promising synergy between AI’s capability and goals of special education.

Students in special education have widely differing learning support and technological access needs, often requiring a combination of assistive technologies (e.g. text to speech software and adaptive switch devices as an alternative to mouse and keyboard), personalized learning goals, and curriculum modifications (changes to a student’s educational program). The characteristic properties of AI-based tutoring and assessment software would nominally make them suitable for use in special education, with its high demands for personalization and individualized instruction.

Our goal is to expand our understanding of AI in special education by studying teachers’ perspectives of AI-based tutoring and assessment within classrooms that serve students with multiple, widely differing, and co-occurring disabilities. Prior scholarship highlights a need for real-world data [7, 9], data that clarifies the extent that education, and the emergent technology facilitating, has equalized opportunity for students with disabilities. As scholarship continues to push the boundaries for inclusive education, it’s crucial to understand the ways emerging techno-solutions can reify and expand existing inequities or, if well-designed, limit them. This study's goal is to understand the degree to which that notion is true – and more specifically, the strengths and weaknesses of AI in special education, as seen through the eyes of teachers. We conducted interviews with 18 special education teachers in the United States to gain insights into the usage, challenges, and perceived benefits of AI-based tutoring and assessment software in their classrooms. We present and discuss our findings and conclude with design implications and potential research directions to improve the synergy between AI-based tutoring and assessment software and the needs of special education classrooms.

2 Related Work

2.1 AI-based Tutoring Systems for Personalization and Accessible Learning

There are two general ways AI-based learning technologies currently serve students with disabilities: customized learning support and fundamental accessibility. Accessibility incorporates the needs of people with disabilities into the development and implementation of mainstream technologies to provide equitable access to exclusionary physical and digital environments that otherwise could not accommodate them, including K–12 education [7, 10]. AI-based tutoring systems use artificial intelligence to personalize learning materials and curriculum pacing, two critical components of special education's IEP. AI-based tutors can take the form of intelligent tutoring systems [11], serious games [12], and other emergent technologies such as smart assistants and interactive robots [9]. AI-based tutors can use algorithms that study student behavior, identify strengths and weaknesses in performance, and suggest areas for improvement. For example, a machine learning model was proposed to identify reading, writing, and memory difficulties for students with dyslexia [13]. Although unevaluated, the proposal highlights novel methodology to customize support for students needing additional reading and writing assistance.

AI-based tutoring and assessment technologies could enable pathways to personalize technology-enhanced learning for students with specific disabilities, should special education teachers incorporate AI in their classrooms. Potential avenues include using computer vision to detect facial expressions and determine which activities increase engagement, robots for social-emotional learning, AI-generated recommendations to reduce teacher effort when personalizing instructions, and interactive simulations with multisensory learning scaffolds [7, 13–19]. These contributions present encouraging opportunities and recommendations to tailor educational techno-solutions for specific disabilities.

However, there is limited understanding of how these techno-solutions work in the daily practice of special education, whether in separate classrooms or mixed within regular classrooms. In special education, individual students often have multiple co-occurring disabilities and will be in classrooms with other students who have widely differing disabilities and support needs. This environment differs from prior scholarship, which has primarily studied personalized support for specific disabilities in isolation [9, 36]. We argue this overgeneralization of smaller studies for specific types of disabilities to broader special education could result in decreased software usability, failure to address crucial learning needs, and increase the amount of teacher labor required for instruction within special education classrooms [20].

2.2 Challenges within AI-based Tutoring for Students with Disabilities

Although, as mentioned, there is ample past research on how AI-based tutoring and assessment technologies might be designed to serve students with disabilities, past scholarship has not focused enough on understanding how these technologies would need to function in the daily practice of special education. For example, in 2015, Whyte et al. found that "the existing computer-based interventions have been designed from theoretical models about the science of autism" [2] rather than the real-world experience of teachers and their students with disabilities. Cinquin et al. recommended future research expand beyond case studies to "compare different solutions across a large spectrum of cognitive impairments to assess whether they are able to cope with the diversity of people with disabilities' situations" [1], as existing evaluations of the time were considered too context dependent, with too few learning trials. [2, 21]. In 2019 scholarship for emergent technology for students with disabilities had a limited number of long-term studies where evaluation was modeled in an environment reflective of where the actual e-learning tool would be used [1]. Unfortunately, five years later, "larger sample sizes and longitudinal studies to enhance statistical robustness and real-world relevance" were still needed [3].

Considering the growing adoption of AI-based tutoring and assessment software in classrooms, and limited empirical evaluation, it is imperative we hear from teachers of students with disabilities already utilizing these tools. In this paper we sought answers to the following research questions: **1) How are adaptive tutoring and assessment tools being used in special education classrooms? 2) What factors inform AI-based learning technology's usage, and its perceived benefits, by special education teachers? 3)From the perspective of teachers, have the potential benefits of AI in special education been actualized for students with disabilities?**

3 Methods

To understand how these limitations and unexplored areas of AI-based tutoring and assessment technologies impact students with disabilities, we interviewed 18 special education teachers.

Recruitment: We used Reddit for recruitment, a social media platform of user-organized discussion forums, or subreddits, around specific interests. We reached 63 teachers via education specific subreddits related to our study - r/SpecialEd, r/TeachingResources, r/EdTech, and r/OnlineLearning. Every teacher recruited had experience teaching students with learning disabilities, ranging from 1 year of experience to 30. Selecting 18 special education teachers, representing ~30% of total recruited, we used the following criteria: 1) experience using learning technology in their classrooms, and 2) Be a K-12 educator in a public or private school. By selecting teachers from both public and private schools, we aspired to represent a variety of educational contexts, including specialty schools only serving students with disabilities.

Semi-Structured Narrative Interviews: Our interview protocol inquired into their school's accommodation practices and sought to uncover the frustrations, and successes using educational technology in their classrooms, with an emphasis on tools with adaptive capability. The interview was 60 minutes long and semi-structured, with the majority spent on two retrospective contextual inquiries. First teachers were given the opportunity to share the kind of educational technology they had experience with. Examples questions include: *What kinds of educational software do you use with your students? Do you use any "adaptive" or intelligent" learning software? How does that* *software fit into your typical week?* Then we pivoted to their school's accommodation practices, where teachers were asked to walk us through a recent experience creating learning accommodations for one of their students, asking questions such as: *In the classes you teach, what percentage of students would you say have learning disabilities? Could you describe how students needing accommodations are identified, who is involved in the process, and what documentation you're provided? Of the accommodations mentioned, which ones are used in combination with software?* Afterwards, teachers shared recent stories of learning technology, not limited to adaptive, being used in their classrooms. In the spirit of a semi-structured interview, the questions varied by teacher. Follow-up questions typically depended on answers to previous questions, as the interview tried to steer the interview to points of interest related to the main theme of the study. Finally, we concluded with some speculative questions, allowing the teacher to look towards the future. Example questions include*: Is there anything about the design of the software you would like to change? If you could have anything to make your experience with teaching better for you and your students with accommodations, what would that be? Consider this your wish list in an ideal world.*

Analysis: After finishing all 18 interviews, we completed a thematic analysis via Interpretation Sessions to analyze approximately 20 hours of transcribed audio recordings [22]. 5 coders consisting of the lead author and 4 undergraduate assistants reviewed the transcripts. The first step involved generating codes—labels or short phrases used to assign meaning to pieces of data in the interview transcripts, capturing key observations or insights. Codes represented succinct, essential summaries of utterance in the transcript. The first 3 interviews were qualitatively coded by multiple researchers to establish a baseline and resolve any disagreements in the granularity of the coding [23]. Following the shared coding session, the remaining 15 transcripts were divided among the research team. Each transcript had 2 reviewers and received 2 passes. The first pass examined the data using an inductive lens by deprioritizing "analytic preconceptions" to answer specific research questions [24]. The goal was to understand the experiences of special education teachers through their eyes. The second pass featured a more theoretical thematic analysis aiming to uncover the significance of emerging patterns and broader implications [24]. During this phase, transcripts were analyzed for broader ideas and documenting specific observations related to our research questions. For the second step, we reconvened as a group to analyze the 443 individual codes to complete a bottom-up affinity diagram [22]. The codes were clustered and refined into 3 levels of themes in an iterative process. The first level clustered our 443 codes into 61 themes. These were then grouped into 12 second-level themes. The four top-level themes and 13 second level themes correspond to the level two and three headers in the Results section, respectively.

Ethics and Participant Safety: We assured all teachers participation was voluntary, responses would be anonymized, and consent could be revoked at any time. To ensure participants' identities remain unidentifiable to those not within the study team, participant data is reported at an aggregate level with participant IDs. We acknowledge research participatory in nature can have high emotional and time investment, then fail to produce tangible outcomes for participants [25]. We intend to continue collaborating with the special education teachers who participated in this study. As a demonstration of our commitment to maintain communication post-study, we created an opt-in Google Group mailing list that 13 out of 18 teachers have joined.

Positionality: We acknowledge that our perspectives are shaped by our experiences, which influence our research. Our study team is composed of American citizens working in the United States. The first and second author identified as disabled, with the first author having lived experience as a special education student in the American public school system. Our academic backgrounds span interdisciplinary fields within Human-Computer Interaction, including education, accessibility, social computing, and AI.

4 Results

4.1 Special Education Teacher's Current Practices

Participants reported using multiple types of learning software in special education. These tools can be broadly categorized into three groups: AI-based tutoring and assessment systems, assistive technologies, and learning management systems. Table 1 provides examples for each category. For example, Lexia is an AI-based tutor designed to support students' reading development through personalized, adaptive learning paths, while DreamBox personalizes math instruction using an individual student's performance in lesson and overall progress data. These AI-based tutoring and assessment programs are used to help identify disability based on performance, match students to accommodations, and determine which classroom is the appropriate academic level for the student. Assistive technologies such as Clicker, a word processor with pictorial representations and drag and drop word banks, scaffold sentence building for learners needing additional literacy support. Learning Management systems are used to manage IEP progress and facilitate communication between special education teachers, parents, and other relevant personnel. In contrast to Frontline IEP and YellowFolder, software intended for use within special education, some participants adapted workforce management programs such as ADP Learning and Cornerstone to manage IEPs.

Table 1. Examples of self-reported software from participants

AI-based Tutor & Assessment	Assistive Technology	Learning Management Systems
Lexia	Clicker (Cricksoft)	Frontline IEP
DreamBox	LyriQ Screen Reader	ADP Learning
Mobi Max	Adaptive Switch	CornerStone LMS
TeachTown	Braille Keyboard	Google Classroom
HelpKidzLearn	LiveScribe Pen	YellowFolder

Assessments from AI-based tutoring and assessment software are crucial to identifying students who may need special education services. P3 describes the start of an IEP process, "A student is identified either by parent or a teacher as struggling with whatever...content is being used at that grade level. Maybe the student, hasn't turned in any work, maybe the work that's being turned in is all wrong. Maybe the student's breaking down over the work." Teachers in P1's school refer students who are "far below grade level" to the district psychologist for further evaluation. A student study team is formed; comprised of the general education teacher, special education staff, school administration, district psychologist, and specialists that support specific disabilities such as a speech, occupational, or physical therapist. According to P3, the student study team will "send in an observer to the classroom and run the student through a battery of computer tests. They'll do some reading, math, and may also look at auditory processing. They may ask the nurse to do an auditory test and a vision test to see if the student's having vision problems."

Although crucial, technology-based assessments are often inaccurate, thus teachers use additional assessments to evaluate students. AI-based tutors and other software provide that adapt based on student performance to determine a student's current academic level**.** P2 explains how a technology-based assessment adapts for their sixth-grade students, "The placement test would give them sixth-grade computation problems aligned to the state standards. If they get them right, [the test] will keep going to make sure [the student] knows all the skills in that specific area. Whereas if they start to get them wrong, then it bumps them down to...a [lower] grade level."

However, assessment is not limited to measuring subject mastery. Technology-based assessments are used in the form of IQ tests and adaptive behavior scales, to identify disability and determine a student's accommodations. It is imperative the assessments adapt appropriately and produce accurate results, so students needing special education services can receive accommodations timely. However, these examinations are often time consuming and inaccurate. For example, P5 describes how AI-based psychological testing falls short with a student who is nonverbal, "[students] still have to go through all the psychological testing, but a lot of the times they're not able to get an accurate IQ score, because the kids are nonverbal." Consequently, P5 seeks out additional input from parents and therapists to, "determine where they would fit withing the school system and what accommodations are needed".

Student performance data from AI-based tutoring and assessment software assists in matching students with disabilities to the appropriate special education classroom. P9 and P11 also emphasize the importance of these assessments to "make sure the students end up in the right class for them". Students with disabilities are matched to a classroom before receiving personalized instruction.

After classroom placement, AI-based tutoring and assessment usage shifts towards measuring learning goal progress defined in the IEP creation process, as part of the software's guidance to students in their problem-solving exercises. Learning goals are adaptive and based on resources and a student's needs. P9 uses assessments to "understand student capabilities and create goals." P5 expounds, "A list of goals is created based on how they have done in previous and current assessments and their ability level in general, and these goals are added to their IEP. Once a week, their performance towards accomplishing these goals is evaluated and the average of these values becomes their quarterly grades".

4.2 Special Education Teacher's Current Challenges

AI-based tutoring and assessment software fall short in providing accurate, detailed, and meaningful assessment. Unfortunately, assessment results from software alone provide insufficient information to measure student growth**.** P14 says, "[the data] is a general, shallow hint." P3 believes learning technology can provide "a good baseline" but fails to present the data to teachers in a meaningful way. "Teachers will look at those numbers and think a kid is functioning at grade level when they're not...in specific areas...There's not enough categorization of the sub-context to understand the overall scores." P3 explains that their students have higher scores for comprehension when reading passages are fiction, due to additional contextual clues. The software will not distinguish that, placing the student at a higher level.

AI-based tutoring and assessment software fail to consider the intersecting skills between subjects and the literacy levels of students with limited cognitive abilities, resulting in inappropriate shifts between mastery levels. When the student receives a non-fiction passage, such as an excerpt from a science book or a math word problem, the lack of contextual information reduces comprehension of harder vocabulary words. This misperception of student mastery results in abrupt shifts between levels, instead of gradual change. P1 describes the impact of these oversights in the case of DreamBox, an adaptive math software, "One student who fails a lot of the math questions in DreamBox has a low comprehension level. He has difficulty understanding the math questions. It's not because he has low math skill. It's because his literacy comprehension is bad. DreamBox could not detect this."

AI-based tutoring and assessment software do not consider accessibility barriers on the software and environmental factors when adapting to poor student performance. P18 acknowledges software can show a student is getting poor grades but doesn't consider why. P18 believes, "failure can be because of several things. Maybe they're not interacting with the software. It could be hard for them to use it." For P18, challenges outside the software, within the classroom environment can contribute to low scores. P17 can identify the "local situation" and "trace [it] to this individual [student]". Special education teachers know the issue that is contributing to "low concentration", whereas the software does not. In P7's classroom, "information from the computer is sparse" and teachers are unable to figure out what learning gaps to address solely from that information.

AI-based tutoring and assessment software differentiates students according to grade level, while special education uses performance level and accommodations, adding challenge to progress monitoring. Consequently, students with widely differing needs and disabilities are represented in one classroom. These students may not be of a similar age, adding to the complexity of managing a classroom with students in different developmental stages. According to P4, within the multi-grade classroom, students are further differentiated due to the “notable differences between their levels of need”. Strategies for additional differentiation can vary depending on the teacher. P3 “separates students into different tiers, based on how much support they need: general, extra repetition, and full IEP support without an IEP.” P1 uses academic performance to “split students into three levels” with the “third level split into more levels”. Students within those tiers are “assign[ed] books that are harder or easier to the different levels”.

Variance between accessibility needs and performance level results in multiple AI-based tutoring and assessment in one classroom, increasing complexity of classroom instruction. As P1 puts it, “There isn’t one AI tool for the entire class. All students work on the same subject, but the app differs between individual students based on proficiency.” *However, software differentiation between students is not streamlined.* In contrast to P1, who matches software to students based on performance level, P7 differentiates using accessibility needs. “There is software for students that are visually impaired, for [those] that have a learning disability, or a hearing impairment.”

With multiple AI-based tutoring and assessment software, managing individual progress towards learning goals is slow. In P4’s classroom students are taught as a group. “Students who have understood the concept are not able to move onto the next topic until the students who were not understanding catch up.” *Despite the tiered classroom environment and varying software applications, group learning is crucial*. Likewise, P6 teaches for the slowest learner by “[going] through the classroom activity sessions slowly and at a steady pace to not overwhelm any slower learning students”.

AI-based tutoring and assessments are unaware of the additional forms of progress monitoring and personalized learning goals tracked independently from the software, increasing labor for special education teachers. In P5’s class “students spend an hour to hour and a half daily receiving one-on-one instruction with a teacher or aide.” P2’s students with cognitive disabilities “spend more time being pulled out of classes, and that education focuses on vocational training and life skills.” P12 creates their own “[assessment and evaluation] on the condition or the status of the student. We usually question the student verbally and get more information from the parent.” For P7 there is “not enough time for one-on-one evaluation and discovery so work is done at home.” Special education teachers rely on other forms of data collection to understand student growth and redefine subject mastery outside of the software being used. Moreover, tracking individual student data can be tedious. P1 explains, “despite many teachers standardizing assessments and material using assistive technology, student data and progress is tracked in a non-standardized, pen and paper format.” P3 expounds, “[Each class] I do weekly progress monitoring, which is on a sheet. [Each student] is on a different grade level for that too. I [track] what they're doing, which questions they're missing one-week, which ones they got right next week, and which questions they need more of that way too.” But using this information to assess student progress is difficult when this must be done for each student in every class.

Slow bureaucratic process within special education and discovery of new technology often falls to teachers, making adoption of more effective AI-based tutoring and assessment technologies difficult. P7 further explains that in addition to teaching, teachers identify student needs and discover software solutions if unavailable. “You just have to discover what your students need, and make sure that you get the services provided for them.” Special education teachers advocate for the adoption of new techno-solutions with their school’s administrators but are often limited by slow bureaucracy. P9 laments, “when you give some suggestions on [what] you want introduced to the school, or [say] you want this kind of resource, [administration] doesn't respond immediately with immediate action. That is the lack of support or cooperation from the institution’s management.” Other teachers, such as P6, desire better software, but are forced to use whatever the school decides. “I am just a teacher; I didn't make the decision to use it. This is the system they're using to learn, so I had to start using the software.” When teacher-advocated software is approved and desired for renewal, financial costs are a barrier to continued use. P5 recounts their experience, “Our supervisor wasn't going to pay to have it renewed. Because somehow, the price increased, and our department just freaked out. And... everybody said how much they needed and liked it and then they agreed to continue paying for it.” Without institutional support, teachers absorb the costs, or their students go without. “If we’re not using technology, then we’re just using paper... and it may sound silly, but printers aren't always available, and we can't even print in color. I have to do that at home.”

4.3 Benefits of AI-based tutoring & assessment software in Special Education

Despite these challenges, special education teachers view AI-based tutoring and assessments as beneficial. **AI-based tutoring and assessment software can accommodate assistive technology**. P4 has students who are blind and nonverbal in one class. The class solely communicates digitally using braille keyboards. “Every student has their own braille keyboard which is always connected with the computer. It's easy for them to read and understand what I'm trying to say… it's the only means of communication we can do.” Technology-enhanced learning in P10’s classroom supports LiveScribe, a smart pen designed for students with disabilities. P10 describes how LiveScribe works, “[it] reads aloud any text it physically passes over and a student can then go over the written words or sentences. The pen remembers the text it has read. The playback speed is controlled for maximum comprehension. The student or I can reduce the speed.”

Students with disabilities are more engaged when the software is highly interactive and incorporates gamified elements. When P18 tries to teach traditionally, “it’s simpler using the software. Because I’m teaching verbally, I'm not able to use visual representations. I'm not able to present the kids with games and everythingP12 uses a software that supports customizable multimedia—they can present teaching materials and communicate using videos and images. P16’s classroom enjoys a technology that, “presents mathematics problems in form of games. [I]t also can be used for group activities as well as individual activities." P16 continues to describe the added benefit of technology supporting their multi-lingual classroom environment with the Spanish and English language options. P6 uses learning software on a tablet that has, "stories, [students] can listen to. They also have mathematical examples, physical examples using animations, which is easy for the kids to understand and learn first." Agreeing with P18, "software is simpler, it's easier" for P6. Although P2 had software with progress monitoring, they switched to improve engagement, "I use Prodigy mostly for my sixth and seventh graders because it's a reward for them. They're more willing to use it. And MobyMax... became a chore because it wasn't very fun."

4.4 Special Education Teacher's Aspirations for Future AI-based Tutoring and Assessment Software

Special education teachers want an all-inclusive AI-based tutoring and assessment software that support multiple, widely differing disabilities. P6 explains, "I wish I could have an all-inclusive technological package that.... would serve all my students at once and wouldn't make anyone feel left out, is really going to be a dream come true." P5 wants more options on what their students can do in class, but currently, needs different types of devices to accommodate all students.

AI-based tutoring and assessment technologies should track student progress more holistically, while communicating key learner insights for special education teachers and parents, in an easy-to-understand format. P1 wants a software that, "can track all of a student's progress...including their attendance grade, assignments accuracy, assessments, and performance." Technology Enhanced Learning should facilitate communication between special education teachers and parents. P18 wants a software that, "keeps families informed about their children progress through the practice questions that we...assign them to do at home." *However, progress data needs to be easily digestible with recommendations for achieving a student's specific goals*. P3 feels, "computers don't necessarily give me suggestions on how to reteach this. Obviously, I need to reteach the area because the kid's not getting it."

5 Discussion

Understanding how learning technologies, especially adaptive tools, are used by special education teachers is crucial for iterating on the design and goals for future systems. AI-based tutoring and assessment software specifically, have significant promise to support students with disabilities. In this paper we contribute the first comprehensive window into how teachers use these systems within special education, their perceived benefits, current challenges, and desires for future technologies. AI for education has received increased attention broadly [26]; however, there is still limited empirical data about its practical uses, benefits, and challenges for students with disabilities [3]. Prior scholarship has contributed critical theoretical work that explores ways AI-based tutoring systems and assessments could support specific disabilities individually [1, 3]. Our work affirms that teachers see value in intelligent serious games and multi-sensory learning mediums [7, 12, 17], desire holistic progress tracking that includes social-behavioral goals [15] -- such as those found in students IEPs, comprehensive learning analytics [5], and assistance from AI recommendations [13, 16].

Nevertheless, we uncovered AI-based tutoring and assessment software fail to meet the needs of students already in special education, while simultaneously having a critical role identifying students who may need (but not receive) special education services – resulting in increased labor for special education teachers. AI-based tutoring and assessment technologies provide inadequate progress monitoring, often shifting levels abruptly with large jumps in difficulty. Special education teachers' credit this to AI-based tutoring and assessments' failure to consider the literacy levels of students with limited cognitive abilities and intersecting skills between subjects (Subsection 3.2). Additionally, AI-based tutoring and assessment technologies do not provide streamlined accessibility support for the range of disabilities represented in classrooms. Special education classrooms are not grouped according to grade level, rather performance and accessibility needs. Consequently, teachers find themselves using multiple AI-based tutoring and assessment software in one class to accommodate students' assistive technologies, personalized learning goals, and curriculum modifications defined in each IEP (Subsection 3.2). The variety of AI-based tutors, disabilities, and performance levels in one classroom results in unreliable learning analytics and recommendations from the software, difficulty maintaining curriculum pacing on a class level, and need for additional progress monitoring, data management systems separate from the AI-based tutors. Despite these challenges, teachers continue to dedicate significant time outside of work to find new technology, because the promise of personalization is worth struggling for.

Future technology-enhanced learning studies should **evaluate effectiveness and accessibility across multiple disabilities.** It is difficult for special education teachers to maintain instructional pace that attends to varying support needs, while allowing higher achieving students to reach their full potential (P7, P12, P18). Additionally, onboarding students with disabilities to a new e-learning environment is challenging—teachers report a period of struggle introducing new technologies. Students with disabilities may not adapt in an expected way to technical errors or an interface they find unintuitive. It's imperative to evaluate the system with a sample target population to identify areas of improvement and limit the amount of software updates.

Special education teachers have limited resources and institutional support, making it difficult for teachers to locate and implement resources that could potentially be useful in their classrooms. Teachers are often forced to use software provided by their school or school district. **Although emerging techno-solutions could be helpful, special education teachers may be unable to adopt it.** Therefore, it is imperative that existing techno-solutions have the empirical support in their design and implementation, while maintaining flexibility to support a wide range of widely differing disabilities that can be present in one classroom or possessed by one student. Designing a technology-enhanced learning system for a specific disability or task will be insufficient to support the needs of special education teachers and their students.

Effective techno-solutions facilitate parent-teacher communication. Special education teachers want parents to contribute to defining and evaluating learning goals and accommodations by providing insight into the at-home learning experience (P10, P7, P9, P18). Despite the potential benefit of providing more holistic assessments for student with disabilities' progress, teachers are concerned about overreliance on software outside of the classroom. Many students within special education have few technological resources and minimal parental involvement. Future studies should consider these concerns in the design of new and existing systems, while soliciting teacher and student input throughout the development life cycle. Existing scholarship has highlighted excellent opportunities for emergent technology to help students with disabilities; however, these opportunities are theoretical or from case studies, with limited empirical evidence from intended teachers, students, and parents who could benefit.

Considering the variety of tools special education teachers use with their students, the mixed-ability nature of Special education classrooms, and additional learning goals from the IEP, a priority research direction should **uncover mechanisms to reduce the burden teachers face creating and organizing student personalization data.** We imagine a future system where teachers could provide information from a student's IEP, including but not limited to accessible formats and mastery criteria for learning goals. The AI-based tutoring and assessment system would adjust the default user experience and student progress data would be re-standardized for easy interpretation by teachers. If multiple software systems are needed, a comprehensive dashboard could legibly centralize all student's data, potentially reducing special education teachers' currently numerous forms of data management. To aid in the streamlining of personalization data, we envision a tool that combines special education's tactical in the classroom learning with technology-enhanced practice to track student progress, fulfilling teachers' desire for technology to supplement rather than replace.

Acknowledgements. This study is supported by the National Science Foundation grant awards: #DGE1745016 and #DGE2140739.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Cinquin, P.-A., Guitton, P., Sauzéon, H.: Online e-learning and cognitive disabilities: A systematic review. Comput. Educ. 130, 152–167 (2019). https://doi.org/10.1016/j.compedu.2018.12.004.

2. Whyte, E.M., Smyth, J.M., Scherf, K.S.: Designing Serious Game Interventions for Individuals with Autism. J. Autism Dev. Disord. 45, 3820–3832 (2015). https://doi.org/10.1007/s10803-014-2333-1.
3. Yang, Y., Chen, L., He, W., Sun, D., Salas-Pilco, S.Z.: Artificial Intelligence for Enhancing Special Education for K-12: A Decade of Trends, Themes, and Global Insights (2013–2023). Int. J. Artif. Intell. Educ. (2024). https://doi.org/10.1007/s40593-024-00422-0.
4. Kohnke, S., Zaugg, T.: Artificial Intelligence: An Untapped Opportunity for Equity and Access in STEM Education. Educ. Sci. 15, 68 (2025). https://doi.org/10.3390/educsci15010068.
5. Nguyen, A., Gardner, L.A., Sheridan, D.: A framework for applying learning analytics in serious games for people with intellectual disabilities. Br. J. Educ. Technol. 49, 673–689 (2018). https://doi.org/10.1111/bjet.12625.
6. du Boulay, B.: Recent Meta-reviews and Meta–analyses of AIED Systems. Int. J. Artif. Intell. Educ. 26, 536–537 (2016). https://doi.org/10.1007/s40593-015-0060-1.
7. Chang, C.-H.S.: Design and evaluation of a multi-sensory scaffolding gamification science course with mobile technology for learners with total blindness. Comput. Hum. Behav. (2022).
8. Christle, C.A., Yell, M.L.: Individualized Education Programs: Legal Requirements and Research Findings. Exceptionality. 18, 109–123 (2010). https://doi.org/10.1080/09362835.2010.491740.
9. Panjwani-Charania, S., Zhai, X.: AI for Students with Learning Disabilities: A Systematic Review, https://papers.ssrn.com/abstract=4617715, (2023).
10. García, A.M., De Bra, P., Stash, N., Fletcher, G.H.L., Fabri, M., Pechenizkiy, M.: Adaptive web-based educational application for autistic students. CEUR Workshop Proc. 1628, (2016).
11. Nagashima, T., Britti, J., Wang, X., Zheng, B., Turri, V., Tseng, S., Aleven, V.: Designing Playful Intelligent Tutoring Software to Support Engaging and Effective Algebra Learning. In: Hilliger, I., Muñoz-Merino, P.J., De Laet, T., Ortega-Arranz, A., and Farrell, T. (eds.) Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption. pp. 258–271. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-031-16290-9_19.
12. Flogie, A., Aberšek, B., Aberšek, M.K., Lanyi, C.S., Pesek, I.: Development and Evaluation of Intelligent Serious Games for Children With Learning Difficulties: Observational Study. JMIR Serious Games. 8, e13190 (2020). https://doi.org/10.2196/13190.
13. Sharif, M.S., Elmedany, W.: A Proposed Machine Learning Based Approach to Support Students with Learning Difficulties in The Post-Pandemic Norm. In: 2022 IEEE Global Engineering Education Conference (EDUCON). pp. 1988–1993 (2022). https://doi.org/10.1109/EDUCON52537.2022.9766690.
14. Abdul Hamid, S.S., Admodisastro, N., Manshor, N., Kamaruddin, A., Ghani, A.A.A.: Dyslexia Adaptive Learning Model: Student Engagement Prediction Using Machine Learning Approach. In: Ghazali, R., Deris, M.M., Nawi, N.M., and Abawajy, J.H. (eds.) Recent Advances on Soft Computing and Data Mining. pp. 372–384. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-72550-5_36.

15. Garnier, P., Martel, K., Dachez, J., Audry, P., Bourgoin, P., Stawinski, F.: Educators' perspectives on working with a humanoid robot in a French preschool class for autistic children. J. Res. Spec. Educ. Needs. 23, 163–174 (2023). https://doi.org/10.1111/1471-3802.12588.
16. Hamid, S.S.A., Admodisastro, N., Manshor, N., Ghani, A.A.A., Kamaruddin, A.: Engagement Prediction in the Adaptive Learning Model for Students with Dyslexia. In: Proceedings of the 4th International Conference on Human-Computer Interaction and User Experience in Indonesia, CHIuXiD '18. pp. 66–73. ACM, Yogyakarta Indonesia (2018). https://doi.org/10.1145/3205946.3205956.
17. Khamparia, A., Pandey, B., Mishra, B.P.: Effects of microworld game-based approach on neuromuscular disabled students learning performance in elementary basic science courses. Educ. Inf. Technol. 25, 3881–3896 (2020). https://doi.org/10.1007/s10639-020-10142-2.
18. Ouherrou, N., Elhammoumi, O., Benmarrakchi, F., El Kafi, J.: Comparative study on emotions analysis from facial expressions in children with and without learning disabilities in virtual learning environment. Educ. Inf. Technol. 24, 1777–1792 (2019). https://doi.org/10.1007/s10639-018-09852-5.
19. Zingoni, A., Taborri, J., Panetti, V., Bonechi, S., Aparicio-Martínez, P., Pinzi, S., Calabrò, G.: Investigating Issues and Needs of Dyslexic Students at University: Proof of Concept of an Artificial Intelligence and Virtual Reality-Based Supporting Platform and Preliminary Results. Appl. Sci. 11, 4624 (2021). https://doi.org/10.3390/app11104624.
20. Fryia, G.D., Wachowiak-Smolikova, R., Wachowiak, M.P.: Web accessibility in the development of an e-Learning system for individuals with cognitive and learning disabilities. In: 2009 First International Conference on Networked Digital Technologies. pp. 153–158 (2009). https://doi.org/10.1109/NDT.2009.5272167.
21. Grynszpan, O., Weiss, P.L. (Tamar), Perez-Diaz, F., Gal, E.: Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. Autism. 18, 346–361 (2014). https://doi.org/10.1177/1362361313476767.
22. Beyer, H., Holtzblatt, K.: Contextual design. interactions. 6, 32–42 (1999). https://doi.org/10.1145/291224.291229.
23. McDonald, N., Schoenebeck, S., Forte, A.: Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. Proc ACM Hum-Comput Interact. 3, 72:1-72:23 (2019). https://doi.org/10.1145/3359174.
24. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. 3, 77–101 (2006). https://doi.org/10.1191/1478088706qp063oa.
25. Pierre, J., Crooks, R., Currie, M., Paris, B., Pasquetto, I.: Getting Ourselves Together: Data-centered participatory design research & epistemic burden. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–11. Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3411764.3445103.
26. Allman, B., Kimmons, R., Wang, W., Bao, H., Rosenberg, J.M., Koehler, M.J.: Trends and Topics in Educational Technology, 2024 Edition. TechTrends. 68, 402–410 (2024). https://doi.org/10.1007/s11528-024-00950-5.