RAISE: Robotics & AI to improve STEM and social skills for elementary school students

Charles E. Hughes[1,2,3]*, Lisa A. Dieker[3], Eileen M. Glavey[4],
Rebecca A. Hines[3], Ilene Wilkins[5], Kathleen Ingraham[3],
Caitlyn A. Bukaty[3], Kamran Ali[1,2], Sachin Shah[1,2], John Murphy[1,2]
and Matthew S. Taylor[6]

[1]Synthetic Reality Laboratory, School of Modeling, Simulation, and Training, College of Graduate
Studies, University of Central Florida, Orlando, FL, United States, [2]Department of Computer Science,
College of Engineering and Computer Science, University of Central Florida, Orlando, FL,
United States, [3]Center for Research in Education Simulation Technology, College of Community
Innovation and Education, University of Central Florida, Orlando, FL, United States, [4]Department of
Language, Literacy, Ed.D, Exceptional Education, and Physical Education, University of South Florida,
Tampa, FL, United States, [5]UCP of Central Florida, Orlando, FL, United States, [6]Department of Special
Education, Salva Regina University, Newport, RI, United States

The authors present the design and implementation of an exploratory virtual
learning environment that assists children with autism (ASD) in learning
science, technology, engineering, and mathematics (STEM) skills along
with improving social-emotional and communication skills. The primary
contribution of this exploratory research is how educational research
informs technological advances in triggering a virtual AI companion (AIC)
for children in need of social-emotional and communication skills
development. The AIC adapts to students' varying levels of needed
support. This project began by using puppetry control (human-in-the-
loop) of the AIC, assisting students with ASD in learning basic coding,
practicing their social skills with the AIC, and attaining emotional
recognition and regulation skills for effective communication and
learning. The student is given the challenge to program a robot, Dash™,
to move in a square. Based on observed behaviors, the puppeteer controls
the virtual agent's actions to support the student in coding the robot. The
virtual agent's actions that inform the development of the AIC include
speech, facial expressions, gestures, respiration, and heart color changes
coded to indicate emotional state. The paper provides exploratory findings
of the first 2 years of this 5-year scaling-up research study. The outcomes
discussed align with a common approach of research design used for
students with disabilities, called single case study research. This type of
design does not involve random control trial research; instead, the student
acts as her or his own control subject. Students with ASD have substantial
individual differences in their social skill deficits, behaviors,
communications, and learning needs, which vary greatly from the norm
and from other individuals identified with this disability. Therefore, findings
are reported as changes within subjects instead of across subjects. While
these exploratory observations serve as a basis for longer term research on a
larger population, this paper focuses less on student learning and more on

evolving technology in AIC and supporting students with ASD in STEM environments.

1 Introduction

The Project RAISE toolkit focuses on supporting students with autism spectrum disorder (ASD) in learning coding, science, engineering, and mathematics (STEM) content, while developing social-emotional communication skills. The project, funded by the U.S. Department of Education Office of Special Education Program's Stepping Up initiatives, is structured into three phases. Each phase provides opportunities for children with ASD to learn new skills in coding along with support in self-regulation skills for use in individual instruction, peer-to-peer activities, and classroom-based instruction. In the first 2 years of a 5-year project, our efforts focused on single case study research, common in the field of special education when working with students with ASD (Hammond and Gast, 2010; Leford and Gast, 2018).

Phase 1: Students learn coding skills and receive communication coaching while building a relationship with
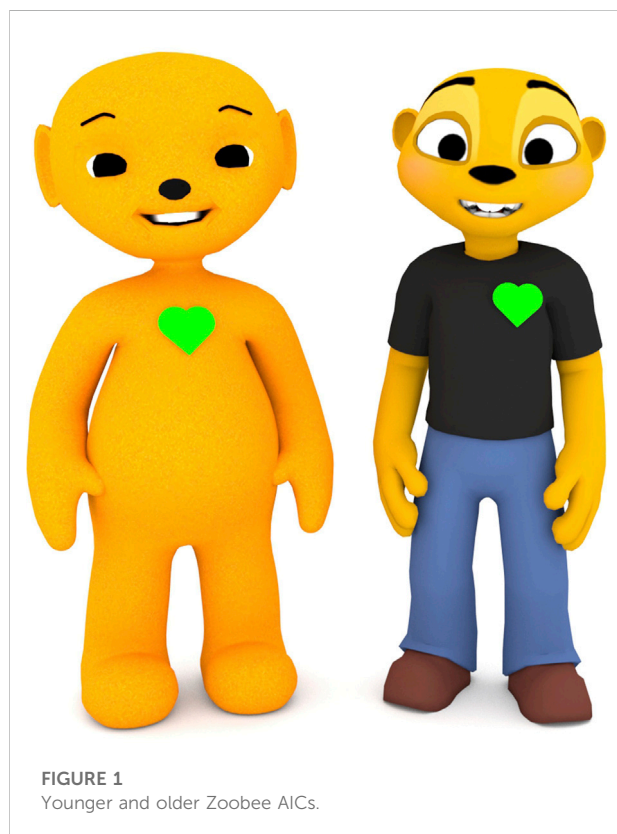
FIGURE 1
Younger and older Zoobee AICs.

a socially assistive AI companion (AIC) called Zoobee™. This phase improves the student's STEM skills as well as provides an opportunity to communicate and create a relationship with the AIC character who has simple pure indicators of emotion (facial expressions, heart color, breathing patterns, and heart rate). Zoobee was initially created as one character, but as the project progressed the need for older and younger versions emerged. Therefore, Zoobee currently exists in two varieties. The first version is designed for children in the early elementary grades (Pre-K - Grade 2). The second is designed for upper elementary students (Grades 3–5) and was co-designed by students, including those with ASD, who voted on several variations of this character. This process of student-driven AI development is one that we believe is valuable for the field to consider in the development of an AIC that is relatable and acceptable to the population targeted for use. Note, as in (Belpaeme et al., 2018) our subjects have made it clear that anthropomorphic characteristics are important in creating an effective relationship between the child and the companion. See Figure 1 for both versions of the AIC.

Phase 2: Students with ASD are given a chance to select a peer (with teacher guidance to ensure they are a positive peer) to share their new skills in being able to code the Dash robot. During this time of teaching a peer to code, the student with ASD receives support from their new AIC friend, Zoobee. The AIC assists in remembering steps and provides challenges (now let's try to make a corner) if the student with ASD and their peer quickly master the first task. The purpose of this phase, and the use of the AIC, is to further strengthen the positive support provided by the AIC and to introduce Zoobee to peers so when the AIC comes into the general education setting other students are aware of this avatar's positive use, and develop stronger social skills with a "real" peer while increasing the sense of accomplishment in a STEM area for the student with ASD (Humphrey and Symes, 2011).

Phase 3: Zoobee affords personalized interactions in the classroom (STEM-related instruction) by providing appropriate support in the general education setting. The use of the AIC in this phase is to build on the student's developing relationship with the virtual companion providing them with positive feedback for their attitudes towards learning, accomplishments in the academic tasks, and focus on completing tasks. The AIC in this phase provides very generic positive praise ("great work staying on task") typically at a rate of

about every 30–45 s or more frequently if biometric data (increased stress) indicates a need. The point of the AIC in this phase is to help the student with self-regulation by providing a "virtual friend" to support them instead of needing the teacher or instructional assistant (IA; sometimes referred to as paraprofessional) to constantly monitor the student's regulation. The goal is to create an environment for students that allows them to develop their own ability to self-regulate with the support of the AIC that is informed by gesture, activity, and engagement recognition algorithms.

The underlying technology used to deliver the experiences with Zoobee is TeachLivE™, a virtual learning environment originally designed for providing professional development to secondary school teachers (Dieker et al., 2019). The TeachLivE software is implemented in Unity™ and provides real-time interactions with virtual avatars supported by digital puppeteering or AI (see Section 3). The human-in-the-loop can supplement the AI behaviors (e.g., Zoobee's verbal responses) to create patterns of language, facial, or gesture responses to rapidly prototype needed AI components to move from human to a complete AI system. The sensing of a user's emotional response occurs through facial expressions and biosensors but in the development phase emotion can be further identified and responded to by the most complex tool available, the human brain of the interactor controlling Zoobee to create clear and easily delivered patterns of AI. Captured data from both the AI already built into each phase and the human responses are analyzed after each phase of the study to drive the AI-enabled interactions, and off-line to support analysis of the system's effectiveness and find correlations that might improve the real-time actions of the AIC. Collectively, the AI behaviors and those occurring as standardized patterns in the AIC are blended and scaffolded to create what eventually will be a completely AI system across all three phases of this study.

2 Target users—Children with autism spectrum disorder

According to the Diagnostic Statistical Manual of Mental Disorders, 5th Edition (DSM-5), one of the cornerstone criteria to receive a diagnosis for ASD includes "persistent deficits in social communication and social interaction across multiple contexts," and may include deficits in social-emotional reciprocity, nonverbal communication, and developing and maintaining relationships (DSM-5; American Psychiatric Association, 2013). This same deficit in the ability to adapt is a component for students identified with intellectual disability and social-emotional gaps and is at the core of the definition of emotional behavioral disabilities (Individuals with Disabilities, 2004). Therefore, long-term deficits in social emotional learning (SEL) may impede relationships with

colleagues and stable employment as well as cause psychological difficulties for students with a range of disabilities (Laugeson et al., 2015).

Many students with disabilities have demonstrated a rapport with technology (Valencia et al., 2019), engaging with its use to learn both academic and life skills (Miller and Bugnariu, 2016; Odom et al., 2015). Miller and Bugnariu's (2016) literature review on the impacts of virtual environments on the social skills of students with ASD demonstrated positive results in many social areas, including behavior and eye contact (Belpaeme et al., 2018). These authors suggest additional research is needed in comparing virtual environment skill development versus human-to-human skill development. This comparison is important not just for students with ASD but others with social and communication skill deficits as well. The Project RAISE team's approach aims to improve the social skills of students with disabilities through engaging technologies (i.e., socially assistive virtual companions) while also teaching participants skills in communication combined with increasing students' skills in STEM fields. Aspects of this work (communications skills, social-emotional learning, and automating companion behaviors) are informed by similar research on socially assistive robots (Belpaeme et al., 2018; Donnermann et al., 2022; Scassellati et al., 2018; et al., 2022; Syriopoulou-Delli and Gkiolnta, 2022).

Learning to code is a component of the STEM fields and leads to life skills (e.g., Bers, 2010; Lye and Koh, 2014; Taylor, 2018). Physical robots and online coding software developed for young learners have led to an increase in research on early elementary students developing STEM skills at a young age (Works, 2014; Strawhacker and Bers, 2015; Sullivan and Heffernan, 2016) and possibly sparking an early interest in STEM careers further down the line. Taylor (2018) and Taylor et al. (2017) published research on teaching young students (PreK-Grade 2) with intellectual disability basic coding skills using the Dash robot developed by Wonder Workshop. The researchers found participants learned to code and build on their learning over a series of explicitly taught steps. Of greater importance, in both studies, students with disabilities succeeded in learning both robotics and coding when given the proper universal design for learning (UDL; Rose and Meyer, 2008) support and guidance. Robotics and coding could lead to strengthening problem-solving skills, collaboration, communication, and critical thinking, areas students with disabilities need to learn for success in school and life (Geist, 2016; Taylor, 2018).

All participants in our studies have been clinically diagnosed with ASD in their school setting. We did not ask for further documentation beyond that they are receiving services under the Individualized Education Program of the label ASD. The project was developed with this population in mind under a federal earmark by the U.S. Department of Education to create socially assistive robots for students with ASD.
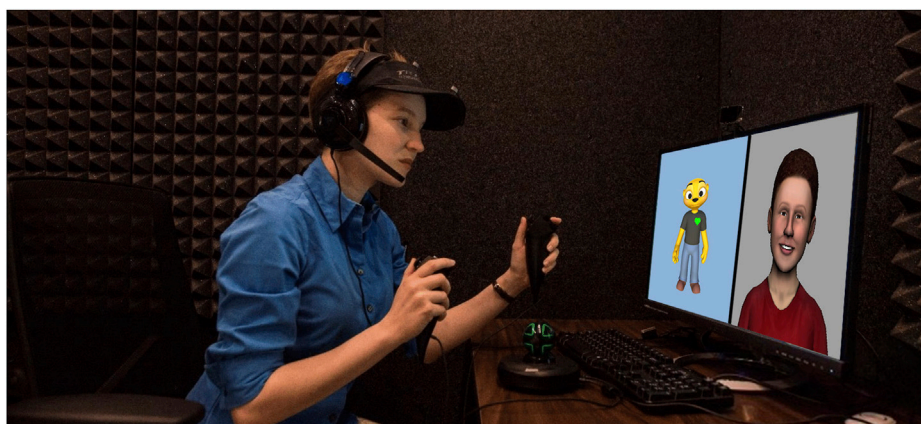
FIGURE 2
Interactor at work puppeteering Zoobee and viewing participant.

3 Virtual learning environments and single case study design

Initially, we used a "Wizard of Oz" approach, called puppeteering, to control the AIC's behaviors. This approach starts as a direct application of our TeachLivE interactor paradigm (Hughes et al., 2015; Ingraham et al., 2021) where a single, highly trained person can puppeteer up to six characters in an interactive virtual world. The current model keeps the cognitive load on the interactor to a reasonable level by allowing all controls to be based on a personalized set of gestures, where each gesture translates to a character pose. Controls like changing which character is being inhabited, altering facial poses, and triggering animated behaviors are mapped onto the buttons, touchpads, triggers, and joysticks of a pair of 6DOF controllers that provide these features (e.g., the HTC Vive or Hydra controllers). This mapping and puppeteering approach allows the interactor to observe an audio/video feed of the participant and adapt performance appropriately. Figure 2 shows a split screen with Zoobee on the left and a young man on the right. We used a virtual stand-in for this figure to avoid using the image of an actual subject. The puppeteer also has another monitor (not shown here) to observe the robot's movements.

The Project RAISE team used the TeachLivE environment, created and patented at the University of Central Florida, to assist the development of the AI by first using the complexity of the human brain to form the patterns the AIC would exhibit. The team combined the puppeteer with using single case study design to examine the nuances and individual changes within a range of individuals with ASD instead of a traditional study looking at group-design outcomes. In this AI-driven project, the purpose of the use of TeachLivE combined with individualized research design models is to rapidly prototype the virtual character

moving from puppetry into a system where we capture and analyze the interactor's performance in the context of the participant's behaviors, verbal/nonverbal, and initiating/responding. These data, along with data captured from sensors on the participant, provide us with sufficient information to train algorithms using multimodal data to learn and mimic the way the interactor both initiates behaviors and responds to the participants. The goal is not perfection in the interactions, as we have no hopes of being able to carry on a complex dialogue as can be provided by a human interactor; rather our intent is to create a level of comfort and personalization for an AIC to support self-regulation in children to achieve educational goals without adult prompting. To support an automated approach, while creating personalization, we have intentionally limited the verbal interactions by our interactor to ensure a pathway that is realistic for machine learning, and yet, responsive to the individualized and wide range of needs of students with ASD in the classroom. Limiting dialogue is important to reach a realistic state of AI at this point in the development phases of this project. During the robot programming phase, the advice provided to the student with ASD from the AIC is simple and aligns with successful ways to program Dash in a square and then suggest other tasks once the primary goal is accomplished. When the AIC moves into the educational environment, the goal is not to distract the student during STEM learning activities but align the use with affirming messages synchronized to stress indicators of the student and prompting positive on-task behavior and social communication.

We are training a deep network that contains multiple branches corresponding to multi-modal data input streams, such as biosensors, poses, eye gazes, vocalizations, and facial expressions to drive the behavior of the AIC, Zoobee. The bottleneck layer of each branch of multi-modal data

corresponds to the feature representation of that input stream. To integrate features from these multiple branches, we are employing three different architectures—an architecture where we fuse these features' representations in the initial layers of our deep network, known as early fusion; another architecture, where we fuse these features' representations in the latter layers of our deep network, known as late fusion; and a third compromise option, known as hybrid fusion, that maintains early fused and unfused stacks, fusing these multiple feature representations at a late stage (Nojavanasghari et al., 2016). We will select our final architecture based on the performance of our multi-modal network. To train the network, we are collecting the training data along with the ground-truth annotations of the behavior of a virtual character provided by the interactor during our initial "Wizard of Oz" approach. Similar to PoseGAN (Ginosar et al., 2019), our goal is to predict gestures. However, in our case, the input audio signal is augmented by other modalities such as biosensor data (Torrado et al., 2017) and facial expressions (Bai et al., 2019; Ali and Hughes, 2020, 2021; Keating and Cook, 2020; Ola and Gullon-Scott, 2020). Moreover, our approach uses a variety of audio technologies, such as an array of directional microphones manufactured by Sony integrated with Amazon web services to identify the speaker, and their location, and to analyze the content (Anderson et al., 2019) and emotions (Abualsamid and Hughes, 2018; Ciftci and Yin, 2019) of the auditory landscape.

3.1 RAISE user-friendly visual programming environment

The RAISE User-friendly Programming environment (RAISE-UP) is designed to be accessible to young learners, including those who are easily distracted, have difficulty reading, or have trouble sequencing (Scassellati et al., 2018). The Blockly programming application designated for use with the Dash robot is designed for students as young as 6 years old (Wonder Workshop,), but relies on users to have a minimum first or second grade reading ability, as well as strong fine motor functioning. The Project RAISE team through the single-case study designed found some participants with ASD stalled in their progress and become confused or frustrated with several aspects of the application including the number of categories and blocks of code (widgets) offered in Blockly; movement of the code from the categories to the programming workspace; sequencing the code appropriately; and deleting unintended widgets by using the interactive trash can.

RAISE-UP was designed to use the Blockly programming environment (Seraj et al., 2019) with a custom subset of the widgets normally available for controlling the Dash robot (Ben-Ari and Mondada, 2018). In effect, RAISE-UP is both a limited (you only need to see the relevant widgets) and an expanded (we have combined some widgets into new ones that are always used

together, for example, a count block with a default value is embedded in the repeat, forward, and right widgets) version of the Wonder Workshop comprehensive learning environment. As such, the environment created for this project is not a replacement but, rather, a UDL-informed version (Rose and Meyer, 2008; Meyer et al., 2014) providing scaffolding to support individual students. This helps in mastering Phase 1 and 2 while gaining the skills and confidence to complete tasks without direct adult intervention. The AIC and RAISE-UP combined allow for a completely AI environment for student learning to code. This was observed in Year 2 in our initial single subject research.

Another positive of the RAISE-UP environment is that, because we developed the environment from the ground up, we can add useful features to assist understanding, e.g., we alter the robot's color to indicate when it is responding to program commands and light up widgets on the screen as each is being executed. Finally, focusing on our specific use, RAISE-UP can communicate with our AIC so Zoobee "knows" when a child is on target or needs support. In Figure 3, we see a finished product for having the robot traverse a right angle; changing the repeat value to four makes a square. Using the configuration tool (see below), this code can be preassembled, partially assembled, or completely left to the creativity and needs for support of the individual child. This UDL-informed version of the RAISE-UP application is critical for the population served on this project and has served a dual purpose of helping to inform the triggering of the AIC when the student introduces coding errors.

To provide universally designed assistive features, RAISE-UP uses a configuration tool allowing the teacher, IA, or researcher to choose appropriate widgets or even add composite widgets for the student's programming task. This tool can be used to create an uncluttered environment with no irrelevant widgets to distract the student or to tempt them to complete other irrelevant tasks. Prior to the RAISE-UP adaptations, we observed children fixating on changing Dash's colors, having it make fun noises, or attempting to reconnect to the robot even when there was an existing live connection. To allow the student to see the functionality of the Dash robot, the robot displays green while in motion and red when completed. Of course, providing support for too long a time can limit students' independence and creativity, so the ability to modulate distractors through a more UDL approach was critical for the individual student, but also for using the AIC. We ensure the website itself is not a distraction in terms of layout and design. More to the point, we can start with a limited set of widgets and then slowly incorporate more as the student's skills advance. This scaffolding of support alongside more structure to the coding interaction was a critical intersection for creating AI prompts within the scope of the project, for example, creating a square instead of having random abstract tasks. We learned early in the project that creating an AIC is possible within confined and UDL-designed tasks.

FIGURE 3
Traversing a right angle.

3.2 Settings

This project took place in schools that are 50% neurotypical and 50% students with disabilities The potential of a comorbidity of students with ASD also having a compromised IQ was addressed by the exclusion of those with less than third grade reading skills and those with severe communication challenges. Moreover, as found in Mutleur et al. (2022), we observed that fewer than a quarter of the students with ASD in this study were identified with intellectual disabilitiy.

We obtained both university and school district IRB approval for the study with the first 2 years focused on case study and single subject design research. Upon receiving permission from the two participating districts, the research began. This funding from the U.S. Department of Education (84.327S) includes very specific requirements for inclusion of traditional, charter, and rural schools to ensure any technology developed is applicable for use across a range of settings. Also, unique to this project is that, at the completion of our work, all components are to be provided as open educational resources (OER). In Year 1 the project was required to work with one pilot site (five students) and each year the AIC and research is to be scaled up into more school sites and more advanced research designs (final two-year randomized control [RCT] studies). In Year 1 all outcomes were based on case studies of students with ASD and collecting exploratory data to inform the development of the AIC and study procedures. In Year 2 (ending in 2022) a small-scale single subject study began across the Year 1 site and two additional schools. The following are the demographics of the schools involved in the project at this time.

The Year 1 and 2 school site in this study (Site A) was a Title I charter school located in a community in Orange County, Florida in a high-need setting. During Year 1, the 2020/2021 school year, 180 children toddler-fifth grade attended the school. All students (100%) were free/reduced lunch-eligible and 57% percent of the population were students with disabilities. The student population was 70% African American, 25% Caucasian, and 5% Hispanic and Asian/Pacific Islander. The site's needs and priorities aligned with the needs and requirements for the project based on the school's identified interest in increasing robotics skills; and social, communication, and self-regulation (SEL) skills of the student population. Robotics was being integrated into the curriculum along with a new SEL curriculum. Additionally, students had a wide range of disabilities at this site making it ideal for including our target population of students with ASD. During Year 2 this same site continued to work with us on the project. In the 2021/2022 school year, 187 children toddler-through fifth grade attended the school, an inclusive education program in which 51% of the population were students with disabilities. The student population was 78% African American, and the other 22% were Caucasian, Hispanic, and Asian/Other. One hundred percent of the students received free lunch.

In Year 2 we continued to work with Site A and two additional sites were added. Site B was from the same charter school local education agency (LEA) as Site A and Site C was recruited from a rural district, required in this project to

demonstrate the usability of the technology in a range of environments.

Site B served 191 students ages 6 weeks to fifth grade, with an average class size of 15 students during the 2021/2022 school year. The school serves students with and without disabilities with 72% of the students receiving special education services and including low-ratio classrooms for students with the most significant disabilities. Most classrooms had a teacher and two IAs. The school serves a diverse population of students in an urban setting—53% Hispanic, 22% African American, 20% Caucasian, and 5% Asian/Other. All elementary students at the site are free/reduced lunch eligible and the school is a designated Title I school.

Site C is an elementary school in a rural district in Florida. This was a Title I school serving 576 students Pre-K through Grade 5 during the 2021/2022 school year. Students with disabilities comprised 16% of the population and the school had both academically inclusive and self-contained special education classrooms. The school served a diverse population of students in this rural setting 37% Hispanic, 29% Caucasian, 26% Black/African America, and 7% combined multiracial and other subgroups. All students at the school (100%) were considered economically disadvantaged and received free lunch.

3.3 Subjects

In Year 1 at Site A, we conducted exploratory case studies with five students who identified as African American in Grades 3 through 5. These case studies allowed the team to examine individual students' reactions to each phase of the study and provided a testbed for the AIC development by the team. As required by the research study inclusion criteria, all five students were receiving special education services based on their Individualized Education Programs (IEPs) with goals in the Social, Emotional, and Behavioral domain and the Communication domain. All students in Year 1 were identified with ASD, the priority disability category for the research. Additionally, all students were screened by their classroom teacher using the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS), which can be used to evaluate a student's overall behavior or subsets of social behavior, academic behavior, and emotional behavior (Severson et al., 2007). Based on results of the SAEBRS, all students were identified by their teachers as "at risk" in terms of their overall behavioral functioning and in the subsets of social and emotional behavior. Signed parental consent and student assent were obtained prior to participation. Three of five students completed all three phases of the study. One student only completed Phase 1 sessions, as he continued to require verbal redirection and physical prompting (hand-over-hand) to code Dash to move in a square at his ninth session. One student who previously provided verbal assent, declined to continue

participation during Phase 1 sessions. We are uncertain if this was due to a lack of interest in the AIC, coding, or just the change of environment. Therefore, we asked for feedback from the classroom teacher, who said this student does not like change or new environments. This type of adaptive behavior deficit is typical for many students with ASD as it is part of the evaluation system for receiving this diagnosis, but we did worry if this type of issue would continue. Also, note this study was occurring during the COVID-19 pandemic and students had returned to school, were required to wear a mask, and to meet social distancing, so this study had to take place in a lobby, a noisy environment. These types of learnings are why the project started with a single case study design to determine the right level of skills needed and inclusion criteria as the project scales up to more schools using a RCT design in Years 4 and 5.

In 2021/2022, Year 2, two students from Site A, three students from Site B, and two from Site C were consented and completed all three phases of the Project RAISE intervention using a single-subject research design. The single-subject research design is common in the field of special education when working with small and unique low incidence populations where finding a large number of individuals to conduct RCT studies is difficult (Horner et al., 2005). This research design allows the team to look at an individual and their change within each phase and to further tweak phases of a study before moving into a pre-post RCT study as is intended in Years 4 and 5 of this project. Year 2 students, like Year 1, were identified and recommended by their teachers and building administrators as students with ASD in Grades 3–5 who could benefit from activities aimed to increase their social-emotional and communication skills. These students also met the same criteria for ASD and having deficits identified on the SAEBRS by their teachers required to participate in the study. The findings from this year showed positive gains across Phases 1 and 2 for all students. Three additional Year 2 students (one from Site A and two from Site B) are expected to engage with the Project RAISE toolkit in the summer of 2022 in a modified capacity due to time constraints during summer school programming.

The findings in Year 2 to this point showed students involved in Year 2 of this study did not decline participation, but one student did decline to work with our team in Phase 3. Therefore, we continue to reflect upon and wonder about when, how, where, and which students are best (with and without disabilities) to benefit and work with an AIC to help increase social and communication skills as well as to learn basic coding.

3.4 Professional development of teachers and instructional assistants

Another requirement of this project is to impact teacher practice and in this study the team is working to support both

teachers and IAs. The team completed professional development (PD) for teachers and IAs at each of the three sites focused on clarifying specific roles and procedures during each stage of the project and in supporting us and the students with ASD. Information regarding the study itself, including the use of the Dash robot and Zoobee the AIC, were not included in the PD to avoid interfering with the data, but instead best practice instructional strategies for including students with disabilities to increase communication and on-task behavior were provided. Specifically, the PD focused on using a component of cooperative learning, think-pair-share (Rahayu and Suningsih, 2018; Baskoro, 2021), an evidence-based practice (EBP) in the field used during mathematics instruction to increase the opportunity for students with ASD to talk with and practice social communication skills with their peers. Due to this study occurring during the pandemic and to allow for scaling-up of the project across multiple school sites, the PD was delivered *via* Zoom.

Fidelity of implementation was a primary topic of the PD to ensure all teachers and IAs refrained from correcting or conversing with the students during the intervention to avoid interrupting participants' interactions with Zoobee, the AIC. For staff in the initial development phase of Year 1, this PD was provided synchronously *via* Zoom and feedback from school staff was encouraged. Based on feedback and questions, the PD was updated in Year 2 to reflect teacher and IA input. A final version of the PD was prepared and recorded to allow for review by Year 1 school-based staff and for use with new staff in the following years of the study. The PD was broken into four short sections with a question at the conclusion of each segment to check for understanding by the respondent. A concluding question asked respondents to indicate their confidence in understanding their respective role on the project and their ability to carry out the activities in the project with fidelity. Once submitted, the form was automatically sent to the Project Coordinator for documentation and to follow up with any respondents who indicated a lack of understanding of expectations or were not confident in their ability to maintain fidelity of implementation.

3.5 Data collection tools

Throughout the study, we created new data collection tools, developed the AIC, began to understand student responses, and evaluated individual student learning gains through case studies and single subject research. The following tool is the primary instrument being used to observe students pre-post in the general education setting to determine the impact of the three phases on students social and communication interactions as well as time-on-task (ToT). This tool in the current study is being used consistently as a measure in all interactions through the case studies and single-subject design. This same form will be used in

future research in an RCT model for pre-post behaviors as well as a data point at the end of each of the phases.

The team created additional student support tools to ensure positive changes within each phase of the research based on lessons learned from the Year 1 case studies. We created a picture-based rating form *via* a Qualtrics survey for individual students with ASD to reflect on how much they did or did not enjoy the sessions during Phases 1 and 2 of the study. We also created social narratives or stories (Odom et al., 2013; Acar et al., 2016) for each phase of the study, noted by the National Clearinghouse on Autism Evidence and Practice (NCAEP) as one of 28 EBPs to support students with ASD. The stories were based upon introducing students to Zoobee and the features found in the heart changing throughout the process to indicate the emotional state of the AIC. This direct instruction to help students understand the role of the AIC emerged from lessons learned in the case studies in Year 1.

All sessions were observed by the research team *via* Zoom and Zoobee interacted with the students through this platform. We recorded all Zoom sessions to capture the social communication and interactions between participants, Zoobee, and peers. Engaging in conversation, for the purpose of this project, was defined as two or more verbal exchanges. Videos were coded by a speech/language specialist using an instrument developed by us (Table 1). In addition, observation transcripts from each session were analyzed to determine trends in language and conversations. These language patterns were further analyzed using AI driven technology to identify common words and phrases for the further development of the AIC.

Following the three phases the research team again observed students in the natural classroom environment using the ToT tool noting every 30 s if a student was or was not on-task across a 10-minute segment. Figure 4 demonstrates that the engagement level of participants was high. The data collection tools and data points gathered within each subject assisted the team in both shaping future research and diving deeper into the response or lack of response of the AIC for individuals with ASD.

3.6 AI tools use and findings

We provide four primary outcomes achieved to this point in this exploratory research, related to AI tools and findings. First, we learned that the use of the "Wizard of Oz" model was essential to development of the project objectives as students with ASD and students in general do not interact in predictable patterns to an AIC. However, with the humanin-he-loop and through voice AI analyses, patterns emerged. Second, the use of triggers to adapt AI interactions is still emerging (Donnermann et al., 2022). For example, the creation of the Blockly-based system, RAISE-UP, based on UDL principles was essential to trigger the AI; yet, without additional information, such as emotional recognition, the programming of the AI did not have enough data to increase

TABLE 1 Verbal reciprocity coding.

Date												Time					
Location												**Duration**					
Phase	Baseline			1-Code w/Zoobee								2-Peer Coding w/Zoobee				3-Independent w/Zoobee	
O = occurrence X = missed																Freq	Ratio
Engages in conversation (>2 turns)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		

FIGURE 4
Time on task during year 1.
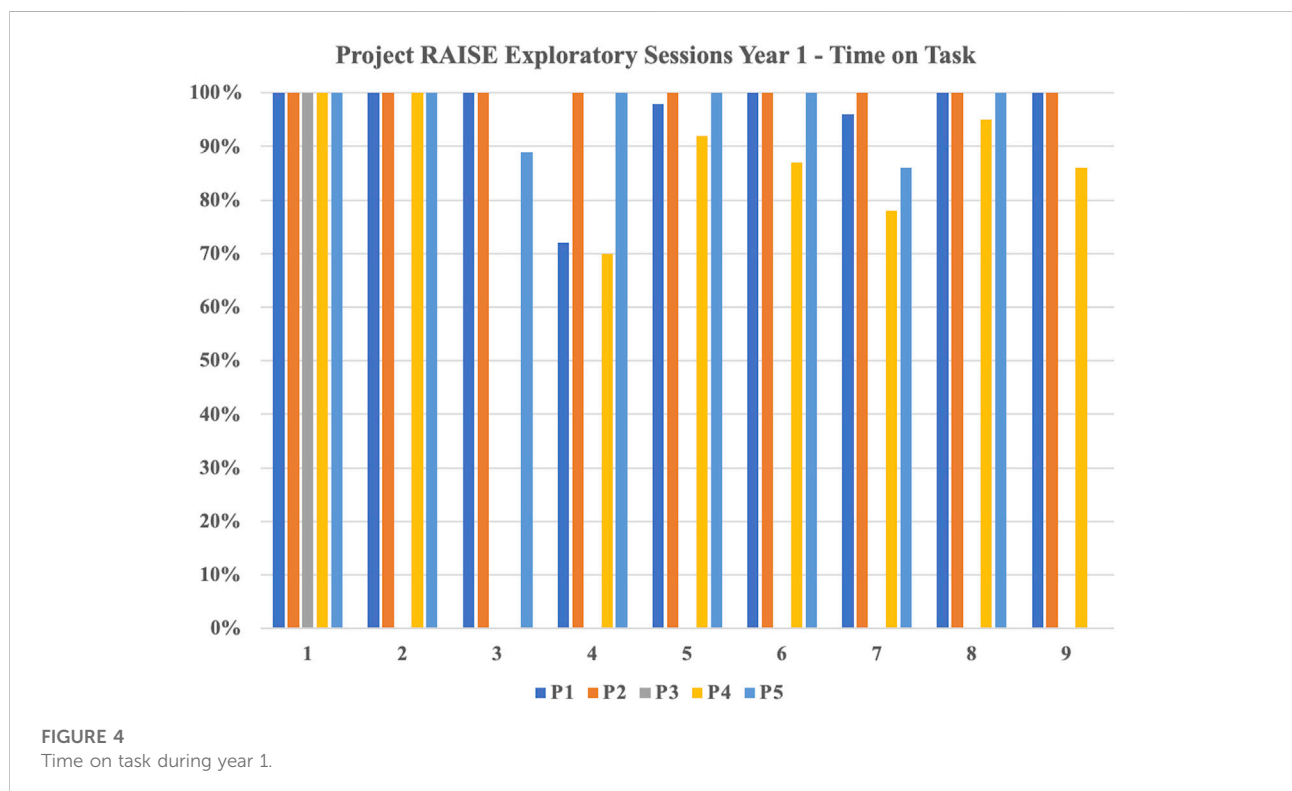
social and communication skills. Third, we encountered the challenges of acquiring neuro or biometric data on students who do not present typical patterns of behavior (e.g., students who run around the room, or go on self-directed behaviors on the iPad, or look away from the iPad constantly). The unique designs of case study and single-subject research allowed the team to conduct an iterative development cycle to inform and improve the outcomes of each phase using a rapid-development model in response to nuances identified within each individual student. This type of design may provide a better platform for future research in creating AI tools for educational environments across diverse learners than creating a tool and using a more traditional RCT model that potentially shows a lack of impact due to not capturing the nuances of individual learners.

3.6.1 Emotion recognition from the video

Facial expression recognition (FER) is widely employed to recognize affective states using face images/videos. In this project, we have employed our FER algorithm (Ali and Hughes, 2020, 2021) to detect the participant's presence and extract expression features from the face region to perform emotion recognition. While FER is very effective in determining if the student is at the tablet and extracting emotions from that view, our method was inaccurate when a participant occludes their face or moves away from the tablet. To overcome this problem, we employed a multi-modal emotion recognition approach by using physiological signals such as heart rate (HR) and heart rate variability (HRV). These aggregated tools provided some level of support for the AIC when combined with the adapted version of the

coding software. This change solved most issues in Phases 1 and 2 of this study but, as noted below, not in Phase 3.

3.6.2 RAISE-Pulse

To collect and process the physiological data of the participants we developed a smartwatch application, which we call RAISE-Pulse. RAISE-Pulse measures the HR and HRV data from a smartwatch sensor, which transmits the data in real-time in raw format to a web server using a wi-fi or Bluetooth connection. The HR and HRV data are then input to a machine learning algorithm for stress and emotion analysis. We used the Samsung Galaxy Watch 3 to measure these rates and developed our app using the Open Source Tizen OS. RAISE-Pulse measures and transmits the data once every second after connecting to the remote web server. We developed a Python-based web server to receive the data in real-time. This information is helpful to trigger the AI in a predictable activity, such as in Phases 1 and 2, but the use of this information in the natural environment is noisy as it is hard to differentiate between when a student is excited by a STEM task versus being frustrated—both resulting in a raise of HR, lowering of HRV, and increase in breathing patterns. Again with the exploratory nature of this funding and project we pose more questions than answers related to the use of an AIC, especially when working with students with neurodiversity.

3.6.3 Stress and emotion recognition using heart rate

Previous research shows that HR and HRV are viable signals for the classification of different affective states (Levenson et al., 1990; Zhu et al., 2019). For example, HR increases during fear, anger, and sadness and is relatively lower during happiness, disgust, and surprise (Levenson et al., 1990). Therefore, emotion recognition using HR has been investigated by the research community in the past due to its benefits over emotion recognition using facial videos/images, especially since we are sensitive to the fact that facial and emotional recognition software is not currently representative of neurodiversity and there is a lack of understanding about the facial patterns unique to children with ASD (Kahlil et al., 2020). Also, since HR is controlled by the human nervous and endocrine systems, it cannot be hidden or occluded, unlike facial expressions. Therefore, with a wide range of commercially available heart rate sensors, measuring emotions using HR is a common practice (Shu et al., 2020; Bulagang et al., 2021). To recognize emotions using HR, we compiled our dataset collected during our initial "Wizard of Oz" approach. We annotated the HR data by initially annotating the video frames with the participant visible by employing our facial recognition algorithm (Ali and Hughes, 2020, 2021). Because the videos were synchronized with the HR signal, annotating the video frames provided the emotion labels of the HR signal. The expression annotation provided by the initial step is considered a

suggestion and based on this suggestion, a human annotator makes the final annotation decision, which is considered the ground truth emotion of the HR signal and the video frames. After compiling our HR dataset, we trained our machine learning algorithm to recognize emotions using the HR/HRV. In the future, we also will explore other modalities such as skin temperature (SKT), skin conductance (SC), and respiration to explore how these data and FER can be fused to better identify the state for triggering AI, or in our case an AIC, for individuals who are neurodiverse.

3.6.4 Gesture, activity, and engagement recognition

In the general education setting (Phase 3), the AIC provides positive feedback by automatically detecting the participant's hand-raising gesture (Lin et al., 2018), writing activity (Beddiar et al., 2020), and engagement (Monkaresi et al., 2016; Nezami et al., 2019). However, the ability to trigger the AIC to reinforce these activities is a challenge with which we still struggle. For example, if a student with ASD has a repetitive behavior of raising their arms, how will the AIC differentiate this from the child wanting to participate. Unlike in Phases 1 and 2 where the RAISE-UP application can provide the AIC information about the appropriateness of the task, we worry about "negative training" from the AIC in the natural classroom environment as might occur from positive reinforcement of an undesirable behavior.

4 Lessons learned

Throughout the first year of the project the team worked on all three phases of intervention to build protocols, identify equipment, and pilot technologies for school use in the project. We began initial investigations and acquired exploratory outcomes using individuals as their own control for changes. These exploratory findings are from 2 years of developmental research involving case studies and single-subject research. The following are exploratory findings from the research in each phase with indications of how they further the development of the AIC.

4.1 Year 1 exploratory results

4.1.1 Pre-Intervention: Driving School, social story, and classroom observation

During the initial development it was determined that participants needed an introduction to coding the robot used in the study. An IA working with participants used a social story about coding with the Dash robot and students then completed *Driving School*. Driving School is a puzzle within Wonder Workshop's Blockly application designed to teach students the

TABLE 2 Year 1, phase 1: participants' completion.

Participant	Number of sessions	Level of mastery	Portion completed independently
P1	5	Full square	Full square
P2	5	Full square	Straight line
P3	1*	Driving school	N/A
P4	10**	Straight line, turn	N/A
P5	6	Full square	N/A

basics of using block code (e.g., robot movement, sequence of code) and introduce students to the robot and coding features. Driving School provides directions for the user to follow, as well as hints if there are difficulties. The goal of incorporating Driving School in Project RAISE was to allow the participants time to feel comfortable with the robot and Blockly before entering the instruction portion of Phase 1 with the AIC.

During the pre-intervention phase, we also observed the students' social, communication, and ToT skills in the natural environment. In Years 4 and 5 we will be conducting a RCT study, and this is the pre-post observation prior to the intervention package of all three phases being introduced to the student. Therefore, we, along with the case studies and single-subject design, note the current level of performance in the natural STEM classroom environment.

4.1.2 Year 1, phase 1: Coding robot

The data collected for individual case studies of five students in Phase 1 included students' communication skills with the AIC, Zoobee, along with the overall use of the AIC for the students, and students' completion/mastery of coding goals (i.e., coding Dash to move in a square; see Table 2). Four out of five of the students completed this phase when they coded Dash to move in a square successfully either independently or with help from Zoobee. If sessions progressed without success, students received hand-over-hand support from the IA to ensure they met the criteria of coding a robot in a square during Sessions 9 or 10. Only one student to this point, across both years, needed as many as nine sessions to complete the task. The average was five sessions with the range being three to nine. The reason for the scaffolding of support was to ensure students were not removed from their general classroom activities for more than the total time noted in the IRB process. Also, this type of learning helped us to understand how robust the AIC needed to be as the project scales up to more schools, sites, and participants.

Each session of Phase 1 (four sessions were designed and scripted to be repeated as necessary) was introduced to the participants by Zoobee to teach the coding steps to make Dash move in a square. Students participated in four to nine Phase 1 sessions depending on the number of sessions needed to complete a full square. Each session began with Zoobee greeting

the participant and stating the goal for the day. The session then built on the participant's accomplishments of the previous session. In Session 1, the participant was taught how to move Dash in a quarter of a square (i.e., a straight line). The focus of Session 2 was to teach the participant to move Dash in half a square (i.e., forward, turn, forward), which is where one student decided he no longer wanted to participate in the study. Finally, the goal of Session 3 was to move Dash in a full square (i.e., forward, turn, forward, turn, forward, turn, forward). If the student was successful at moving Dash in a full square during any session, Zoobee asked the participant if they would like to try a new challenge (e.g., using repeat, adding sounds).

Since this work was a case study design, we provide an example below of one student involved in Year 1 of the study to demonstrate the applicability of this type of design in developing AICs for unique populations of students, such as those identified as ASD. We also share how this case study informed the overall design of the research and the AIC capabilities for Year 2 of the project.

An example of one case study participant was a fourth grade student who identified as an African American male and was identified with ASD and intellectual disability. Based on his academic records and his IEP, the student was performing below grade level in reading and mathematics at the time of the study. Based on teacher input and his IEP, he was a considerate student and respectful toward his teachers and peers. He demonstrated excitement when interacting with peers during small group mathematics and physical outdoor activities, but demonstrated difficulty initiating peer interactions, responding to peer interactions, maintaining attention, and making eye contact during whole group activities, negatively impacting his peer relationships and task completion. The student's IEP included special education services of direct instruction in social skills to engage, cooperate, and interact productively with adults and peers in his learning environment. The student responded well to Zoobee and was conversant from the initial meeting and in pre-classroom observations.

Pre- and post-intervention and during each phase of the study, sessions were recorded and analyzed to determine ToT and participant engagement in conversations with Zoobee, peers, and in the classroom setting. Time-on-task was defined as 1)

TABLE 3 Word frequency query results—participant 1 sessions.

Word	Word frequency
Make	83
Square	55
See	51
Okay	48
Get	39
Turn	35
Now	28
Yeah	25
Dash	23
Well	27

student verbally responding to Zoobee or in the classroom responding to peer or teacher, 2) student appearing to physically respond to Zoobee or teacher whether or not verbal response is noted, 3) student appearing to engage with Dash or teacher in a manner consistent with instructions from the teacher or Zoobee, 4) student engaging independently with Dash if instructed to work independently by Zoobee ("You try it," etc.) or the teacher. Time-on-task data were collected using duration recording during which a timer was started and stopped throughout an entire session to obtain the total amount of time the student was engaged in the targeted behavior.

Additionally, recorded sessions were transcribed through a free, commercially available software, Otter.ai. Transcripts were then imported into NVivo software for coding and analysis. This analysis identified all keywords and phrases spoken by Zoobee, which were not part of the original coding script and represented the natural language occurring between the participant and Zoobee. The word frequency query function of NVivo was used to list the most frequently occurring words by Zoobee in the participant transcripts. For example, Table 3 shows the results of the word frequency query for Participant 1's sessions. Based on the results of the word frequency query, the word "make" was run as a text search query to gain a greater understanding of the context of how the word was used and determine if it should be included in the programmed behaviors. Findings of the text search query for "make" are seen in the word tree displayed in Figure 5. Visual analysis of the word tree revealed that the word "make" most often occurred with the word "Dash" in the context of making Dash "go" or "move" revealing the phrase "make Dash go" as a potential phrase for the AIC.

Transcripts for each of the four remaining participants across all three intervention phases were analyzed to support the reliability of the observation instrument as well as to gather context and build the script. This process along with patterns noted by us and the human-in-the-loop created a standardized script for the AI to be triggered by progress with the RAISE-UP

app, HR/HRV data, and facial tracking. How this work will be triggered in what our team terms the "wild, wild west" of a natural classroom environment of daily changing tasks, demands, communication, and directions is still emerging at this time, but this work has been well-served by using non-statistical approaches during these iterative development cycles.

4.1.3 Year 1, phase 2: Coding robot with peer

Phase 2 sessions followed the same pattern as Phase 1 with the addition of a peer learning alongside of the participant. Phase 2 was designed to provide the participant opportunities to interact with a peer and teach the peer how to code the Dash robot to move in a square to further reinforce the STEM skill of basic coding for students with ASD. This phase also served the purpose of further developing rapport between the student and the AIC and to have the friend (neurotypical) who would socially accept and support Zoobee as a virtual friend for use in the natural classroom environment. Study participants and their peers completed three to five Phase 2 sessions depending on the number of sessions needed for the student to teach the peer to successfully move the Dash robot in a full square. Like Phase 1, Zoobee provided goals for the session and monitored the programming needs of the students (e.g., questioning, difficulty remembering next block of code). The AIC also acted as a support for the participant in practicing social-emotional and communication skills with the peer, e.g., turn taking and discussions, both critical skills that can be reinforced in interaction with socially assistive companions such as virtual characters and robots (Syriopoulou-Delli and Gkiolnta, 2022).

Most participants were excited to bring a peer to meet Zoobee and show them how to code the Dash robot. Zoobee provided suggestions and cues as needed and often reminded participants to let their peer have a turn. This reminder from the AIC often led to the peers interacting with the AIC and the participant observing, rather than teaching (or even communicating with the peer). For instance, one participant remarked "I can't do this" when trying to work with Zoobee and walked away from the programming and left the peer to try. The peer was then able to work through the code with support from Zoobee and self-corrected if the code did not run as expected. In the following session, the participant was more vocal to the peer (i.e., "Do you want to try?"), but walked away as the peer began interactions. The balance between the use of the AIC to support the student with ASD and taking over when the student is unsuccessful is an area the team is further reflecting upon in future years of this research.

4.1.4 Year 1, phase 3: Zoobee (AIC) companion in the general education setting

For Phase 3, the AIC joined the participating student with ASD in the classroom during math class. The AIC appeared on a tablet with a headset connected, so the student could hear and speak to the AIC. Throughout mathematics instruction, the AIC

FIGURE 5
Text query word tree for "make".
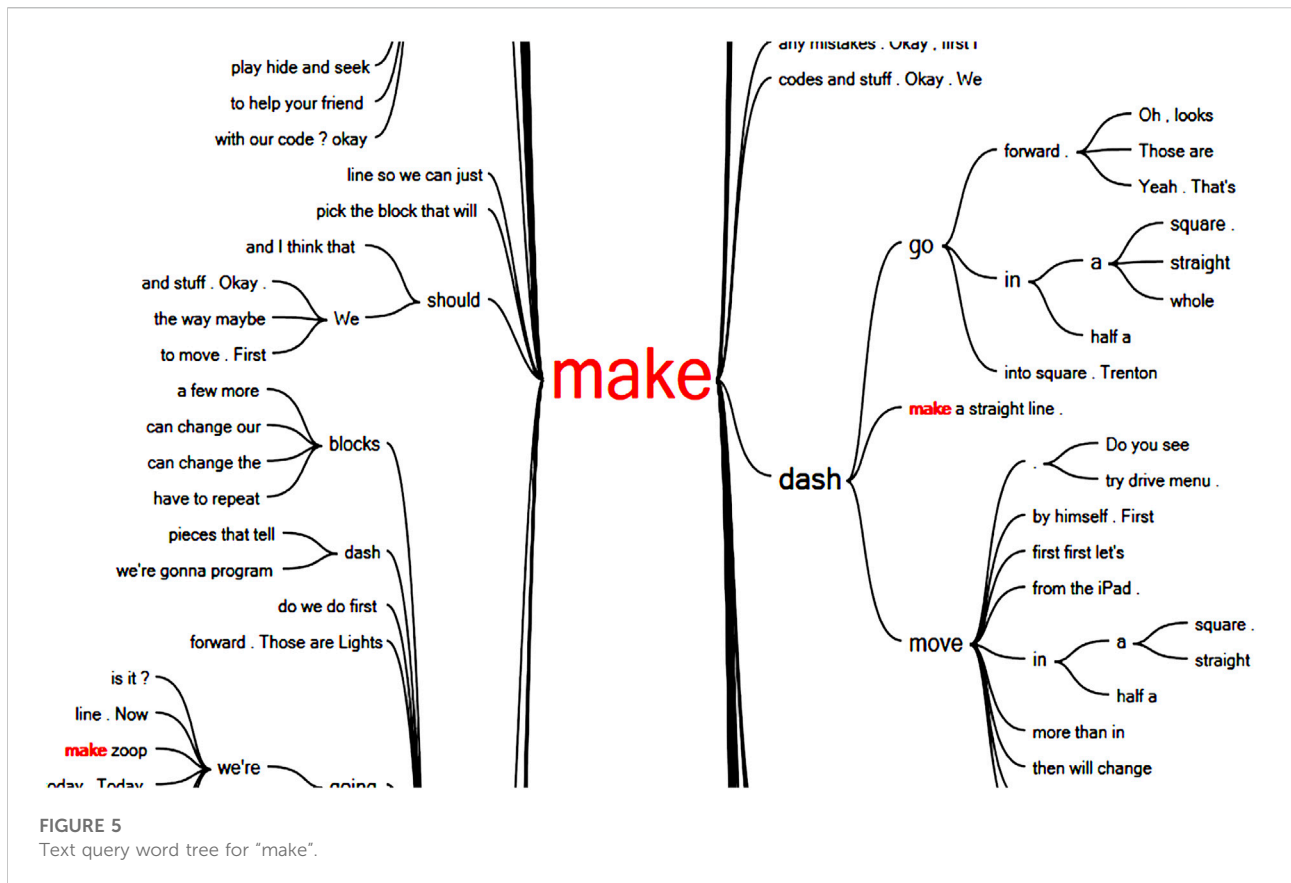
supported and reinforced the student's social-emotional behavior, self-regulation, and communication skills. For Year 1, Phase 3 observations were limited and conducted only to establish procedures for use in future years of the project. Preliminary findings in this phase of the study did lead to our realizing both the potential power of the use of the AIC in the classroom but elevated the potential challenges of triggering AI behaviors in such a noisy and complex environment. Findings from Year 1 suggest participants were accepting of the presence of the AIC in the classroom environment. We also found a need for a pattern to be established to provide "just the right dosage" of support without being distracting. The human-in-the-loop was able to establish some patterns of statements to use on an AI loop, but we continued to struggle with "what" data would trigger the AI correctly in the complex classroom environment.

Overall, the findings from Year 1 showed us that participants with ASD in the study could follow the AIC's directions and guidance during Phase 1 to complete a square or at least a portion of a square. During this initial investigation, we found the need to change aspects of the programming sessions, including when to present or teach Driving School, key phrases to redirect participants' attention and focus (e.g., "return to the iPad"; positive reinforcement), and IA's support. Over the course of each phase, participants also increased their language production. Figure 6 shows Participant 1's increase in time talking across nine sessions as he gained comfort with Zoobee and confidence in his coding skills.

4.2 Year 2 findings and insights about the use of the AIC within each phase

In Year 2 the project team expanded to work in an additional school in an urban setting and added a school in a rural setting. In this year we used a single-subject case study design (Horner et al., 2013) pre- and post-intervention as well as across all three phases of the study. In a single subject design, students stay in a phase until they reach a level of stability and their data is compared only to the changes observed within themselves and no cross-student data occurs (i.e., the individuals serve as their own control for change). Data collection for Year 2 just concluded at the time of writing this article, so we are including findings to provide a reflection on the overall lessons learned aligned with the purpose of this article on how student data informed the use of the AIC within each phase, not on student social and communication changes. These findings about the use of the AIC and our insights emerged from work with seven students with ASD involved in Year 2 of the project. All students completed or are in the process

FIGURE 6
Participant 1 percentage of talking.
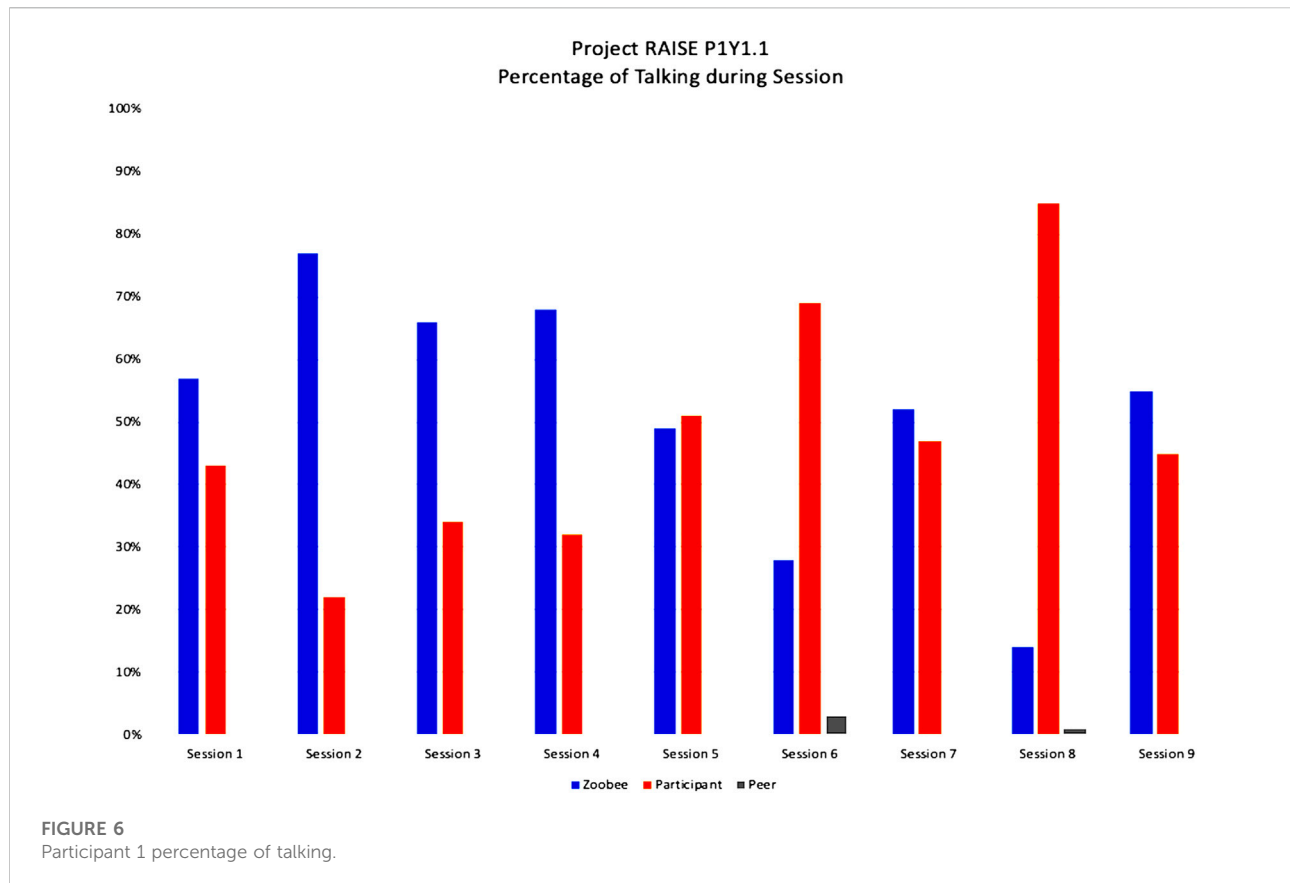
of completing all three phases of the research except one student who adamantly did not enjoy the presence of the AIC. That situation is described to inform others using AICs for neurodiverse learners so they might consider our lessons-learned. Described below are the current findings about the use of the AIC and our reflections at this point in the project.

4.2.1 Year 2, phase 1: Coding robot

All seven participants were observed in the classroom setting for ToT and social communication skills. All seven participants did code Dash the robot to move in a square within a range of one to five sessions (see Table 4). A notable difference between Year 1 and Year 2 is in session completion.

An interesting situation we continue to reflect upon is one student in Year 2 who, though initially afraid of the AIC, Zoobee, became comfortable working through support from the IA. The IA helped the student code the robot Dash in the first session and then in Sessions 2 and 3 Zoobee joined the IA and by Sessions four and five the student accepted the support from the AIC only. This student's acceptance of Zoobee was instant in Phase 2 and the student even described the AIC as a "friend" as they began working with a general education peer. Once the peer accepted Zoobee, the previously resistant student also referred to Zoobee as their "new friend". This

outcome further validated for us the importance of Phase 2. Beyond the data collected, the research team now feels they have gathered enough language patterns to move Phase 1 into AI only, which we anticipate will be triggered by progress or lack of progress by the student observed by the AI agent from the RAISE-UP platform.

4.2.2 Year 2, phase 2: Coding robot with peer

All peers of the participants successfully coded Dash to move in a square and all seven even completed an advanced task (e.g., made a right triangle, made Dash light up). At times, Zoobee had to work with the students with ASD to ensure turn-taking or allowing their peers time to code. In all but one instance, the rapport with the AIC advanced for both the participating student and their peer. A new issue emerged and is being used as a future consideration from one participant. Right before the session was to occur, this participant was involved in a bullying situation. A teacher quickly intervened, but the student entered the session very agitated and began to use abusive statements toward the AIC: "Zoobee you are fat, you are stupid, you are ugly." Despite an increase in verbal communication, this was not the communication we were looking for as a targeted goal. The student did allow his peer to work with the AIC, but the participant continued to verbally

TABLE 4 Year 2, Phase 1: Participants' completion.

Participant	Number of sessions	Level of mastery	Portion completed independently
P1	3	Full square	½ square
P2	4	Full square	N/A
P3	3	Full square	Straight line
P4	4	Full square	½ square
P5	3	Full square	Full square
P6	3 (completed square in first session)	Full square	Full square
P7	3	Full square	Full square

insult Zoobee for the duration of the session. When the session ended, we decided to give the student an opportunity to interact with Zoobee in hopes of repairing the relationship without the peer. Our experts in ASD felt the observed behavior was an attempt to "show-off" and gain credibility with the peer. This one-on-one attempt to build back rapport between the participant and the AI failed. The participant from this point forward continued to not enjoy the interactions with the AIC, and ultimately chose to "turn off" Zoobee, as was their right as a participant in this voluntary project. This finding was interesting to us, and it is noted as an exception and not the norm but did provide us with a new protocol for students with limited communication to have an option to turn off the AIC in Phase 3.

In Phase 2, besides the one incident noted, the students communicated with the AIC, with the one student afraid of the AIC in Phase 1 now calling Zoobee their "friend." Also, all peers did code Dash to go in a square within two to four sessions. As we analyze the individual data in this phase, we continue to struggle with the best way to make this Phase 2 AI, due to the complex nature of the interactions between the peers, coding, and other complexities of engaging with multiple humans and tasks simultaneously. At this time, we are contemplating whether to continue to keep the human-in-the-loop or to design the AI for Phase 2 where Zoobee becomes more of a "help button"-triggered assistant using the AI to provide prompts when errors are noted in the RAISE-UP tool for Dash, or when silence is observed with no movement.

4.2.3 Year 2, phase 3: Zoobee (AIC) companion in the general education setting

Phase 3 sessions were conducted fully in Year 2 (2021/2022) with five sessions per participant and again using a single-subject design model for which data are still being analyzed and two participants completing their last sessions. Collectively, however, all but one participant made positive comments to the AIC when they joined the classroom. In this phase the AIC took a less prominent role, providing prompts and reinforcement related to on-task academic behavior such as "good response" and "great job staying focused on the task". Here, we decided to make the statements less frequent (about every 45 s) and short in duration to ensure the AIC did not distract the learner in the classroom environment.

The pattern of feedback was not meant to provide specific feedback but rather to allow the student to show their AIC friend Zoobee their ability to perform in a STEM classroom and to receive praise for their on-task behavior or redirection for off-task behavior. Currently, Phase 3 still uses a human-in-the-loop. Clear language patterns have emerged, but have yet to be analyzed, across the seven participants to allow us to transition this stage to be only AI behaviors. We realize the statements from the AIC will need to be neutral and on-task behavior-focused but we are looking for ways for teachers or IAs to potentially use a button to trigger reinforcement from Zoobee. For example, Phases 1 and 2 typically end with what we call the "Zoobee celebration dance." This dance is paired with a melody. We decided to not trigger this behavior in Phase 3 as the concern was this type of behavior could be distracting in the classroom setting. However, one student, after they got two answers correct, started singing the celebration dance and making the same moves Zoobee made in earlier phases. We are contemplating introducing options in the AI environment as a future direction that would allow a teacher to trigger the celebratory dance in Phase 3 if they felt it was appropriate and useful. At this time, we are exploring how HR/HRV, emotion tracking, adult button options, and patterns of language can be used in this "in the wild" phase to provide an AIC that promotes an increase in social, communication, and learning outcomes. We are still unclear about the right level of AIC support here, as we believe AI is not yet ready for cases such as answering arbitrary math problem with which the student is struggling. We are encouraged by the fact that six of this year's seven student participants seemed to enjoy the presence of an already established AIC relationship supporting them in the natural environment of their classrooms.

FIGURE 7
Sharing knowledge on coding.

5 Discussion and future directions

The primary purpose of this article was to discuss the development of an AIC that supports unique populations and the interactive, personalized, and individualized process needed when working with students with ASD. We have learned that, despite the challenges also seen with assistive robots (Belpaeme et al., 2018), some automated behaviors are easy to provide even to a population with a wide range of behavioral (moving around the room, refusing to look at the screen), social (telling jokes all of the time, refusing to look at Zoobee's heart), and communication skills (using repetitive statements) based on our individualized observations with the human-in-the-loop preliminary studies. Two responses that are highly effective, despite their simplicity, are the observation of whether students are at the tablet (in our case an iPad) and whether they are making progress on the programming tasks. The presence attribute is easily checked by using a vision-based algorithm for a face appearing in the camera's field of view. We are comfortable in using facial tracking as a sign of "presence" or "not presence" despite our hesitation and sensitivity of using emotion tracking software with its potential bias due to a lack of neurodiversity, specifically ASD, currently present in databases used for training. At present, if a student's face (peer or participant) does not appear on the screen, the AIC is sent a message about this condition and the virtual character encourages the participant to come back to the iPad to continue work on programming the robot. The programming activity also is easily analyzed by the RAISE-UP environment. If the participant is making adequate progress in the programming task but just needs a nudge, for instance, to add a widget to create a repetition block, the Zoobee AIC can either suggest the addition or even send a message back to the programming environment to cause it to automatically drag the repeat widget onto the workspace or even to add it to the evolving program. The choice of what support to provide is informed by prior knowledge about how the participant is progressing.

As the full system evolves and is tested in multiple contexts—schools and geographical settings such as urban and rural—we will be making these tools available to the community through the project website. Specifically, six different aspects of Project RAISE will be available as OER at the end of the project. Teachers and schools will be able to use the tools and resources in these key areas individually or as a collection, free of charge. The areas of resources available for educators or future development and use by the AI community are:

1) Students with autism spectrum disorder (ASD) learning to code with the aid of an AI Companion (AIC) and a programming environment designed around UDL principles (RAISE-UP).
2) Students with ASD teaching a peer to code (see Figure 7 for a virtual image of that environment).

3) AIC supporting students with ASD in the natural classroom environment.
4) Professional development (PD) on cooperative learning in mathematics to increase communication opportunities for students.
5) Instructional assistants and teacher PD on various aspects of the project.
6) Research validated tools for measuring social interaction and communication of students with ASD.

Ultimately, we desire to understand how to impact student learning, teacher behavior, and future outcomes for students with disabilities. Individuals with ASD have the highest rate of unemployment of any disability group (Krzeminska and Hawse, 2020) with an employment rate of only 19.3% (U.S. Department of Labor, 2019), and a noted lack of presence in STEM related fields. Yet students with ASD often have the disposition to work with technology (Valencia et al., 2019), and Krzeminska and Hawse (2020) note employers are recognizing characteristics of individuals with ASD are often of value in STEM related disciplines. We know this work will not magically address the shortage of people who are neurodiverse being represented in STEM related fields, but we do believe creating student-driven AI tools, such as Zoobee, that students can choose to use to support their self-regulation in any environment, could potentially help. We invite others in the VR and AI communities to join us in this journey to positively address the needs of a diverse range of learners.

Data availability statement

The datasets presented in this article are not readily available because of ethical and privacy restrictions. Requests to access the datasets should be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by University of Central Florida Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

References

Abualsamid, A., and Hughes, C. E. (2018). "Modeling augmentative communication with Amazon lex and polly," in *Advances in usability, user experience and assistive technology. AHFE 2018. Advances in intelligent systems and computing.* Editors T. Ahram and C. Falcão (Cham: Springer), 794, 871–879.

Author contributions

All authors contributed to the design and/or implementation of the system and to the authoring of the article. LD, EG, RH, IW, CB, KI, and MT were primarily responsible for the study designs and data analysis. CH, SS, JM, and MT were the designers and developers of the RAISE-UP system. CH, KA, and SS were responsible for the AI components. KI developed and puppeteered the behaviors of Zoobee.

Funding

Acknowledgments

Conflict of interest

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

Publisher's note

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

Acar, C., Tekin-İftar, E., and Yıkmış, A. (2016). Effects of mother-delivered social stories and video modeling in teaching social skills to children with autism spectrum disorders. *J. Spec. Educ.* 50 (4), 215–226. doi:10.1177/0022466916649164

Ali, K., and Hughes, C. E. (2020). "Face reenactment based facial expression recognition," in *Advances in visual computing. ISVC 2020. Lecture Notes in computer science 12509*. Editors (Cham: Springer), 501–513.

Ali, K., and Hughes, C. E. (2021). Facial expression recognition by using a disentangled identity-invariant expression representation. *25th Int. Conf. Pattern Recognit. (ICPR 2020)*, 9460–9467. doi:10.1109/ICPR48806.2021.941217

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders*. Available at: https://doi-org.ezproxy.frederick.edu/10.1176/appi.books.9780890425596.

Anderson, K., Dubiel, T., Tanaka, K., and Worsley, M. (2019). *Proceedings of the 2019 international conference on multimodal interaction*, 506–507. doi:10.1145/3340555.3358662 Chemistry pods: A mutlimodal real time and retrospective tool for the classroom.

Bai, M., Xie, W., and Shen, L. (2019). *Proceedings of 2019 IEEE international conference on image processing (ICIP)*, 31–35.Disentangled feature based adversarial learning for facial expression recognition.

Baskoro, R. A. (2021). The comparison of numbered head together learning models and think pair share in terms of elementary school mathematics learning outcomes. International. *J. Elem. Educ.* 4 (4), 549–557. doi:10.23887/ijee.v4i4.32568

Beddiar, D. R., Nini, B., Sabokrou, M., and Hadid, A. (2020). Vision-based human activity recognition: A survey. *Multimed. Tools Appl.* 79 (41), 30509–30555. doi:10.1007/s11042-020-09004-3

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., and Tanaka, F. (2018). Social robots for education: A review. *Sci. Robot.* 3 (21), eaat5954. doi:10.1126/scirobotics.aat5954

Ben-Ari, M., and Mondada, F. (2018). "Robots and their applications," in *Elements of robotics* (Cham: Springer). doi:10.1007/978-3-319-62533-1_1

Bers, M. U. (2010). Issue editor's notes. New media and technology: Youth as content creators. *New Dir. Youth Dev.* 12 (2), 1–4. doi:10.1002/yd.369

Bulagang, A. F., Mountstephens, J., and Teo, J. (2021). Multiclass emotion prediction using heart rate and virtual reality stimuli. *J. Big Data* 8 (1), 12. doi:10.1186/s40537-020-00401-x

Ciftci, U., and Yin, L. (2019). "Heart rate based face synthesis for pulse estimation," in Advances in visual computing. ISVC 2019. Lecture Notes in computer science *11844*. Editors G. Bebis, R. Boyle, B. Parvin, D. Koracin, A. Ushizima, S. Chai, et al. (Cham: Springer), 540–551.

Dieker, L. A., Straub, C., Hynes, M., Hughes, C. E., Bukaty, C., Bousfield, T., et al. (2019). Using virtual rehearsal in a simulator to impact the performance of science teachers. *Int. J. Gaming Comput. Mediat. Simul.* 11 (4), 1–20. doi:10.4018/IJGCMS.2019100101

Donnermann, M., Schaper, P., and Lugrin, B. (2022). Social robots in applied settings: A long-term study on adaptive robotic tutors in higher education. *Front. Robot. AI* 9, 831633. doi:10.3389/frobt.2022.831633

Geist, E. (2016). Robots, programming and coding, oh my. *Child. Educ.* 92, 298–304. doi:10.1080/00094056.2016.1208008

Ginosar, S., Bar, A., Kohavi, G., Chan, C., Owens, A., and Malik, J. (2019). Learning individual styles of conversational gesture. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 3497–3506. doi:10.1109/CVPR.2019.00361

Hughes, C. E., Nagendran, A., Dieker, L., Hynes, M., and Welch, G. (2015). "Applications of avatar-mediated interaction to teaching, training, job skills and wellness," in *Virtual realities – dagstuhl seminar 2013*. Editors G. Burnett, S. Coquillard, R. VanLiere, and G. Welch (Springer LNCS), 133–146.

Humphrey, N., and Symes, W. (2011). Peer interaction patterns among adolescents with autistic spectrum disorders (ASDs) in mainstream school settings. *Autism* 15 (4), 397–419. doi:10.1177/1362361310387804

Individuals with Disabilities (2004). Individuals with disabilities education improvement act. *Stat* 118.

Ingraham, K. M., Romualdo, A., Fulchini Scruggs, A., Imperiale, E., Dieker, L. A., and Hughes, C. E. (2021). "Developing an immersive virtual classroom: TeachLivE – a case study," in *Current and prospective applications of virtual reality in higher education*. Editors D. H. Choi, A. Dailey-Hebert, and J. S. Estes (Hershey, PA, United States: IGI Global), 118–144.

Keating, C. T., and Cook, J. L. (2020). Facial expression production and recognition in autism spectrum disorders: A shifting landscape. *Child Adolesc. Psychiatric Clin. N. Am.* 29 (3), 557–571. doi:10.1016/j.chc.2020.02.006

Khalil, A., Ahmed, S. G., Khattak, A. M., and Al-Qirim, N. (2020). Investigating bias in facial analysis systems: A systematic review. *IEEE Access* 8, 130751–130761. doi:10.1109/access.2020.3006051

Krzeminska, A., and Hawse, S. (2020). "Mainstreaming neurodiversity for an inclusive and sustainable future workforce: Autism-spectrum employees," in

Industry and higher Education: Case studies for sustainable futures (Springer), 229–261.

Laugeson, E., Gantman, A., Kapp, S., Orenski, K., and Ellingsen, R. (2015). A randomized controlled trial to improve social skills in young adults with autism spectrum disorder: The UCLA peers program. *J. Autism Dev. Disord.* 45 (12), 3978–3989. doi:10.1007/s10803-015-2504-8

Levenson, R. W., Ekman, P., and Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system Activity. *Psychophysiology* 27, 363–384. doi:10.1111/j.1469-8986.1990.tb02330.x

Lin, J., Jiang, F., and Shen, R. (2018). *IEEE international conference on acoustics*. New York, NY, United States: Speech & Signal Processing ICASSP, 6453–6457.

Lye, S. Y., and Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Comput. Hum. Behav.* 41, 51–61. doi:10.1016/j.chb.2014.09.012

Meyer, A., Rose, D. H., and Gordon, D. (2014). *Universal design for learning: Theory and practice*. Wakefield, MA, United States: CAST Professional Publishing. Available at: http://udltheorypractice.cast.org/.

Monkaresi, H., Bosch, N., Calvo, R., and D'Mello, S. (2016). Automated detection of engagement using video-based estimation of facial expressions and heart rate. *IEEE Trans. Affect. Comput.* 8 (1), 15–28. doi:10.1109/taffc.2016.2515084

Mutluer, T., Aslan Genç, H., Özcan Morey, A., Yapici Eser, H., Ertinmaz, B., Can, M., et al. (2022).Population-based psychiatric comorbidity in children and adolescents with autism spectrum disorder: A meta-analysis, *Front. Psychiatry* , PMC9186340. PMID: 35693977. doi:10.3389/fpsyt.2022.856208

Nezami, O. M., Dras, M., Hamey, L., Richards, D., Wan, S., and Paris, C. (2019). *Joint European conference on machine learning and knowledge discovery in databases*. Springer, 273–289.Automatic recognition of student engagement using deep learning and facial expression.

Nojavanasghari, B., Baltrusaitis, T., Hughes, C. E., and Morency, L-P. (2016). EmoReact: A multimodal approach and dataset for recognizing emotional responses in children. *Proc. Int. Conf. Multimodal Interact. (ICMI 2016)*, 137–144. doi:10.1145/2993148.2993168

Odom, S. L., Cox, A. W., and Brock, M. E. (2013). Implementation science, professional development, and autism spectrum disorders. *Except. Child.* 79 (2), 233–251. doi:10.1177/001440291307900207

Odom, S., Thompson, J., Hedges, S., Boyd, B., Dykstra, J., Duda, M., et al. (2015). Technology-aided interventions and instruction for adolescents with autism spectrum disorder. *J. Autism Dev. Disord.* 45 (12), 3805–3819. doi:10.1007/s10803-014-2320-6

Ola, L., and Gullon-Scott, F. (2020). Facial emotion recognition in autistic adult females correlates with alexithymia, not autism. *Autism* 24 (8), 2021–2034. doi:10.1177/1362361320932727

Rahayu, S., and Suningsih, A. (2018). The effects of type learning model numbered head together and think pair share. *Int. J. Trends Math. Edu. Res.* 1 (1), 19–21. doi:10.33122/ijtmer.v1i1.27

Rose, D. H., and Meyer, A. (2008). *A practical reader in universal design for learning*. Cambridge, MA: Harvard Education Press.

Scassellati, B., Boccanfuso, L., Huang, C. M., Mademtzi, M., Qin, M., Salomons, N., et al. (2018). Improving social skills in children with ASD using a long-term, in-home social robot. *Sci. Robot.* 3 (21), eaat7544. doi:10.1126/scirobotics.aat7544

Seraj, M., Katterfeldt, E-S., Bub, K., Autexier, S., and Drechsler, R. (2019). Scratch and google blockly: How girls' programming skills and attitudes are influenced, *Proc. 19th Koli Call. Int. Conf. Comput. Educ.* 19, 1–10. doi:10.1145/3364510.3364515

Severson, H. H., Hill, M. W., Hope-Doolittle, J., Kratochwill, T. R., and Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *J. Sch. Psychol.* 45 (2), 193–223. doi:10.1016/j.jsp.2006.11.003

Shu, L., Yu, Y., Chen, W., Hua, H., Li, Q., Jin, J., et al. (2020). Wearable emotion recognition using heart rate data from a smart bracelet. *Sensors* 20 (3), 718. doi:10.3390/s20030718

Strawhacker, A., and Bers, M. (2015). "I want my robot to look for food": Comparing Kindergartner's programming comprehension using tangible, graphic, and hybrid user interfaces. *Int. J. Technol. Des. Educ.* 25 (3), 293–319. doi:10.1007/s10798-014-9287-7

Sullivan, F. R., and Heffernan, J. (2016). Robotic construction kits as computational manipulatives for learning in the STEM disciplines. *J. Res. Technol. Educ.* 48 (2), 105–128. doi:10.1080/15391523.2016.1146563

Syriopoulou-Delli, C. K., and Gkiolnta, E. (2022). Review of assistive technology in the training of children with autism spectrum disorders, International. *J. Dev. Disabil.* 68 (2), 73–85. doi:10.1080/20473869.2019.1706333

Taylor, M. S. (2018). Computer programming with pre-K through first-grade students with intellectual disabilities. *J. Spec. Educ.* 52 (2), 78–88. doi:10.1177/0022466918761120

Taylor, M. S., Vasquez, E., and Donehower, C. (2017). Computer programming with early elementary students with Down syndrome. *J. Spec. Educ. Technol.* 32 (3), 149–159. doi:10.1177/0162643417704439

Torrado, J. C., Gomez, J., and Montoro, G. (2017). Emotional self-regulation of individuals with autism spectrum disorders: Smartwatches for monitoring and interaction. *Sensors* 17 (6), 1359. doi:10.3390/s17061359

U.S. Department of Labor (2019). *Occupational outlook handbook.* Available at: https://www.bls.gov/ooh/.

Valencia, K., Rusu, C., Quiñones, D., and Jamet, E. (2019). The impact of technology on people with autism spectrum disorder: A systematic literature review. *Sensors* 19 (20), 4485. doi:10.3390/s19204485

Wonder Workshop Wonder Workshop (n.d.) *blockly.* Retrieved from. Available at: https://home.makewonder.com/apps/blockly.

Works, Teaching (2014). *High leverage practices.* Available at: http://www.teachingworks.org/work-ofteaching/high-leverage-practices#sthash.I4xK7DG4.dpuf.

Zhu, J., Ji, L., and Liu, C. (2019). Heart rate variability monitoring for emotion and disorders of emotion. *Physiol. Meas.* 40, 064004. doi:10.1088/1361-6579/ab1887