EMPOWERING SPECIAL EDUCATION TEACHERS: EXPLORING CHALLENGES AND OPPORTUNITIES WITH SOCIALLY-ASSISTIVE ROBOTS

by

Shane Romero

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Computer Science).

Golden, Colorado

Date _____

Signed: _____

Shane Romero

Signed: _____

Dr. Tom Williams
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____

Dr. Iris Bahar
Department Head
Department of Computer Science

ABSTRACT

This thesis addresses the critical intersection of socially assistive robots (SARs) and special education, with a primary focus on the role of teleoperators— special educators—in shaping the effective implementation of SARs. The three interconnected studies provide a comprehensive perspective on the essential aspects of SAR utilization in special education.

In Study One, session summary and progress reports emerged as multifaceted tools integral to the credibility of SAR interventions. Balancing visualization for a diverse audience while ensuring teleoperators have adequate details for informed judgments, such as engagement and proficiency metrics, proves crucial.

Study Two delves into the nuanced needs of special education teachers, revealing the intricate landscape of the field. The challenge lies in designing for diverse tracking methods and multifaceted dimensions of success, including academic goals, behavioral observations, social interactions, and health metrics. This complexity necessitates careful consideration of collaborative dynamics, presenting a challenge in creating a universally effective design.

Study Three highlights the significance of a structured support system, emphasizing community building among special educators. Practical strategies, including continuous assistance, hands-on sessions, accessible resources, and incentivized professional development, empower educators and enhance their understanding of SARs. The study underscores the specific needs of special educators, providing valuable considerations for institutions to support effective implementation.

In conclusion, this collectively emphasize the dual importance of refining teleoperation interfaces and ensuring effective implementation to ensure SAR success in special education. Technological advancements alone are insufficient; the key lies in strategies empowering and supporting educators operating SARs. This research advocates for a holistic approach, focusing on the perspectives of special educators. Such an approach ensures widespread SAR adoption and positive impacts on the lives of children in special education settings.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Special education aims to provide appropriate educational services and support to students with disabilities to help them achieve academic and personal success [1]. In recent years, there has been a growing recognition of the importance of engaging, motivating, and personalized curriculum to meet the diverse needs of students with disabilities [2]. Engagement and motivation are critical components in special education. Engaging curriculum incorporates multiple modalities, allows for flexible pacing, and promotes active participation that can help overcome the barriers faced by students with disabilities and promote their success [3–5]. Motivating curriculum can be achieved by incorporating student interests, providing choices, and setting achievable goals [5–7]. To address this collection of needs, special education teachers need to identify and deliver personalized curricula. A personalized curriculum recognizes the individual strengths, interests, and needs of each student, providing tailored learning experiences that maximize the students' potential [8, 9]. In special education, personalized curriculum can be particularly important as students often require individualized support and accommodations to access and participate in learning experiences [9]. Assistive technologies, such as socially assistive robots offer an additional tool for personalized curriculum, as they can provide individualized support, feedback, and motivation to students with disabilities [10, 11].

Socially assistive robots (SARs) provide meaningful assistance to humans through primarily social interactions [12]. In education, SARs can act as tutors [13] or peers [14] that motivate students to learn academic content [12] and social skills [15, 16]. Similarly, SARs have been shown to improve the efficacy of therapeutic interactions by increasing engagement [17, 18] and motivation [19, 20]. Within the Human-Robot Interaction (HRI) literature there is a significant focus on fully autonomous SARs but when deployed in real scenarios, therapists and educators are typically the ones controlling or teleoperating SARs because they're the ones creating the content and use SARs as a medium to deliver that content [21]. Recent analysis suggests a need to reevaluate the emphasis on autonomy, aligning it with the level of autonomy (LoA) selection guidelines that recommend teleoperation in certain sensitive domains [21]. Teleoperation, where a human operator guides the robot's actions, is not only a more fitting choice in many SAR domains but is also a more respectful acknowledgment of the educators and therapists already providing assistance [22]. In addition, the prevalent focus on autonomous SARs has left a gap in understanding how to design tools to best support therapists and educators in their use of teleoperated robots [22].

When we shift our focus from fully-autonomous to teleoperated SARs not only do we introduce additional stakeholders, in this case teleoperators, but also other design considerations. A study with teleoperators reveals a dual-cycle model of therapy: the outer cycle spanning months or years involves collaboration, goal-setting, and iterative reviews, while the inner cycle focuses on therapist-child interactions within each session [22]. The model emphasizes extending the analysis beyond the session to include authoring interfaces and post-session analysis dashboards, offering valuable insights for improving robotic technology in therapy [22]. The study identifies six themes related to how therapists examine, evaluate, and prepare within inner and outer therapy cycles, emphasizing the importance of evaluation for teleoperators [22].

Evaluation, assessing the value and success of therapeutic activities, allows teleoperators to monitor client progress, report to stakeholders, and influence future session preparations and interventions [22]. Therapists stress continuous evaluation to monitor improvement, determine specific goals for children, and adapt goals based on progress. Recommendations that address the importance of evaluation include enhancing SAR tools with long-term progress displays, insurance approval reports, and facilitating easy performance comparison across sessions [22].

In addressing the research-to-practice gap, personalized curriculum, and visual support tools like dashboards and reports play a crucial role. These tools offer real-time detection of issues, display learning progress, and provide feedback for future sessions, empowering educators to tailor interventions. The opportunity to improve teleoperation interfaces lies in visual support tools like session summary reports, leveraging data from SARs and teleoperators to provide valuable insights, support future sessions, and personalize content. The effectiveness of SAR systems is contingent on considering the role of teleoperators in the evaluation process.

Therefore, one design opportunity for improving teleoperation interfaces is the incorporation of session summary reports, aiding teleoperators in their evaluation process [23, 24]. These reports are used in two key ways: insurance reporting and preparation [22]. Insurance companies can use these reports to determine the efficacy of treatments; while teleoperators can use the data to determine future interventions to deliver. In addition, these interviews revealed that evaluation is institutional meaning that institutions can play a role in providing tools that can ease the evaluation process for teleoperators [22]. In this thesis, we deepen our understanding of teleoperators' needs around session summary reports and how institutions can support the deployment of SARs with a focus on special education teachers.

We initially explored this topic with the Peerbots [25] interface in mind.

1.1 Peerbots

Peerbots, an open source application, that offers an affordable solution for socially assistive robot teleoperation and has been employed in programs focusing on enhancing social skills for children with autism [23, 26]. Peerbots enable a teleoperator to oversee a robot's movement and verbal expression. Teleoperators can create content that the robot will articulate during a session and can incorporate relevant metadata for each verbalized item. Notably, therapists can define both the goal for the articulated content and the required proficiency level. This feature allows teleoperators to assess a client's performance using the information gathered during a session [23, 26].

In our first experiment, we created an initial session summary report motivated by existing treatment and lesson plans, questionnaires used by a non-profit organization that used Peerbots, real-life session transcripts from Peerbots and a documentation report recommended in Elbeidy's work [23].

Figure 1.1 Peerbots teleoperation interface [18]

Through design iterations and semi-structured interviews, we aimed to answer the following questions in our first study:

- Why do summary reports add value to therapists, educators, and administrators?
- How can the visualizations and summaries displayed in these reports be best designed to meet these stakeholders' needs?

We conducted a series of interviews, in which participants first explained their current workflow for planning, conducting, and reflecting on sessions, and then critiqued our most recent summary report designs. After each interview, we iterated our summary report designs based on insights gleaned from the preceding interview. Our final design included a (1) newly designed main panel as well as panels for (2) student/client performance, (3) activity duration and engagement, (4) time spent per goal, and (5) new

task efficiency. One of the most critical pieces of feedback received during our iterative design process was the need for distinct reports for visualizing (a) the contents of an individual session, and (b) progress over time. We then qualitatively analyzed the resulting collection of interviews to identify overarching themes and contributions. Our analysis produced **three key insights**:

- Performance and engagement metrics guide future therapeutic and educational interventions
- Therapists and educators have limited time to reflect after sessions
- Progress reports provide credibility in therapeutic and educational interventions

In summary, our first study addressed the overarching needs of educators, therapists, and administrators by exploring the requirements for comprehensive summary and progress reports, offering broad recommendations that consider multiple perspectives. Our initial experiment highlighted the importance of session summary reports for both therapists and educators, but we overlooked their differing needs. To provide tailored recommendations for each stakeholder, we must examine their unique requirements and incorporate them into the report.

In our second study, we narrowed our participant pool to only special education teachers to provide more tailored recommendations. We sought to understand their needs and organizational goals related to session summary and progress reports, aiming to enhance SAR adoptability through tailored design. We aimed to answer the following research questions in our second study:

- How can session summary and progress reports be best designed to meet special educators' needs?
- What are other design implications that should be considered when building/ deploying SARs in special education classes?

In our second study, we explored how special education teachers track goals to offer targeted recommendations for session and progress reports. This investigation highlighted the challenging process of designing a solution that addresses diverse needs and workflows, underscoring the collaborative effort among special education teachers, paraprofessionals, and specialists in crafting content for a student's Individualized Education Program (IEP). Employing diverse recruitment methods, we conducted in-depth semi-structured interviews where participants elaborated on their goal-tracking processes, including criteria for progress assessment. Through comprehensive thematic analysis, **two key insights** emerged:

- The tracking of goals in special education is essential for both historical documentation and ongoing assessment, yet it presents a complex challenge due to diverse tracking methods and the multifaceted nature of special education goals;

- Session summary reports should support holistic tracking of goals and customizable interactions to align with the diverse teaching styles and collaborative nature of special education.

In our third study, we conducted in-depth interviews with special educators with experience teleoperating SARs within their classrooms. Our objective was to garner insights into the effective deployment of SARs in special education settings and provide recommendations to institutions, specifically, school districts. To achieve this, we broadened our research questions to concentrate on identifying the challenges inherent in teleoperating SARs and pinpointing opportunities to refine their implementation. Our third study sought to answer:

- What are the challenges special educators face when using socially assistive robots in their classrooms?
- What opportunities exist for improving the implementation of SARs in special education?

We recruited special education teachers who have experience with SARs from one Colorado school district. We conducted semi-structured interviews that not only uncovered the hurdles of teleoperating SARs but also identified potential improvements for their implementation in special education. Our qualitative analysis in the third study revealed **two key findings**:

1. A robust and continuous support system is critical in the effective implementation of socially assistive robots;
2. Offering incentives and various resources can motivate special educators to implement socially assistive robots;

Our third study shed light on the pivotal role of the implementation phase, a critical consideration for seamlessly integrating SARs into the daily routines of users in the field, particularly special educators. Unsurprisingly, our exploration of existing literature unveiled a notable gap in research, especially from the perspective of special educators, regarding the implementation of SARs in special education. This gap could be attributed to the historical emphasis on fully autonomous robots, inadvertently sidelining the needs of other stakeholders involved in the process [21]. By exploring the insights of experienced special educators using SARs, our third study contributes to bridging the gap between adoption and successful implementation. It offers broad recommendations for institutions such as school districts or programs, aiming to guide improved SAR implementation and ensure their full potential in enhancing special education classrooms.

In essence, while SARs have made substantial progress in addressing the labor cost associated with personalization through autonomous systems, recent insights suggest a potential misalignment in

prioritizing autonomy over teleoperation, especially in the context of SARs in special education. Our motivation stems from a keen awareness of the critical role played by special educators as teleoperators, stakeholders that have been overlooked. Furthermore, the hurdles encountered by special educators as they strive to integrate SARs into their classrooms highlight the necessity for effective implementation alongside any advancements in SAR design, while also emphasizing the significant role that institutions can play in supporting such effective implementation. Our research aims to offer valuable insights contributing to the development, adoption, and successful implementation of SARs in special education. By centering the experiences of special educators, our goal is to provide targeted recommendations that enhance the usability and effectiveness of SARs, ultimately fostering positive outcomes for both educators and students in special education settings. In all, these studies mark significant contributions to the socially assistive robot literature, particularly within the often overlooked domain of special education, a sentiment echoed by a special educator in our interviews:

> *Thank you for thinking of special ed teachers and ... technology that [is] better and more accessible... [it's] so encouraging to hear about other people caring about this field. So thank you.*

CHAPTER 2

MOTIVATION

Special education is an educational service that is specially designed to meet the needs of a child with disability [1]. In the 2019-2020 school year, the National Center for Education Statistics reported that there were 7.3 million students, or about 14 percent of all public school students that are receiving special education services [27]. The goal of special education is to help these students overcome their challenges and develop the skills necessary to succeed academically and in life [1, 28]. Special education services may include specialized instruction, assistive technology, therapy, and accommodations and modifications to the curriculum and learning environment [1, 27].

Special education serves students with a wide range of disabilities or special needs. Some categories of disabilities that are eligible for special education services according to the IDEA (Individuals with Disabilities Education Act) [29] are autism, deafness, deaf-blindness, emotional disturbance, intellectual disability, speech, language, visual or hearing impairment. Students who have one or more of these disabilities are eligible for special education services [29]. Additionally, students who have developmental delays or are at risk of developing a disability may also receive special education services [29].

2.1 Engagement and Motivation

Engagement and motivation are critical components of special education as they play a significant role in student's learning and academic success. Students with disabilities often experience lower levels of engagement and motivation in the classroom, which can lead to poor academic outcomes [4, 5]. When students are engaged and motivated, they are more likely to participate in classroom activities, develop a positive attitude towards learning, and achieve their academic goals [6, 7]. Additionally, engagement and motivation can improve students' social skills, self-esteem, and overall well-being [4–7]. Special educators must employ evidence-based strategies to promote engagement and motivation, such as using technology-based tools, providing positive feedback, and incorporating student interests into the curriculum [2–5]. By fostering engagement and motivation in special education, educators can help students with disabilities reach their full potential and become successful learners.

Special education teachers need to design a curriculum that targets student's specific needs, but also motivates and engages the student. One way special education teachers address this is with assistive technology.

2.1.1 Challenges in Special Education

Special education teachers face a variety of challenges in their roles, including high caseloads, complex student needs, and a lack of support and resources [30]. Special education teachers are often responsible for supporting students with a wide range of disabilities and need to have a deep understanding of each student's individual needs, learning styles, and strengths. This requires specialized training and ongoing professional development, which may not be available to all special education teachers [30–32]. Often there might only be one special education teacher throughout the entire school which leaves them to function independently from other teachers [32]. In addition to these challenges, special education teachers may also face challenges related to communication and collaboration with other professionals, such as general education teachers, administrators, and related service providers [30]. These challenges can lead to high levels of stress and burnout among special education teachers, which can negatively impact their job performance and their ability to provide effective support to their students [30, 32].

2.2 Assistive Technology

Assistive technology (AT) refers to tools and devices that are designed to help individuals with disabilities overcome barriers and achieve greater independence [33, 34]. AT can be "low-tech" tools such as pencil grips or adapted scissors, or "high-tech" devices such as speech-to-text software or communication boards for nonverbal students with autism [33]. The goal of AT is to enhance an individual's ability to learn, communicate, and perform daily activities, ultimately improving their quality of life [33, 34]. AT can help students access the curriculum, participate in classroom activities, and communicate with their peers and teachers [33, 34]. Additionally, AT can provide opportunities for students to develop independence and improve their self-esteem [33, 34].

2.2.1 Socially Assistive Robots

Socially assistive robots (SARs), a subset of assistive technologies, assist humans through social interactions [12]. SARs have been used with students in education [35], the elderly [36], and individuals with cognitive and behavioral disabilities [12]. SARs typically engage in social interactions through social modalities such as speech, gestures, or emotional expressions [15, 37]. SARs have been shown to help with tasks such as tutoring [12, 35], physical therapy [12, 38], teaching social interaction skills [12, 39] and delivering or supplementing behavioral interventions [37].

2.2.2 Socially Assistive Robots in Education

SARs fit into a broader category of technologies used in education to better assist students with physical, cognitive, emotional, or social needs [40]. Within this category of technologies, SARs have been shown to be particularly effective in supporting various learning topics, ranging from traditional academic subjects like math [12] and second language acquisition [41], to more general social skills like collaboration and self-confidence [15, 16]. SARs have taken on a variety of roles, including teacher's assistant [41], tutor [13], and peer [14]. In these roles, SARs have been shown to help increase enjoyment in students, which in turn has been shown to lead to more motivation and engagement [15]. Additionally, SARs can be beneficial for students with disabilities who may have difficulty interacting with human teachers and peers, providing a non-judgmental and supportive learning environment [38]. Therefore, SARs are a promising form of AT in special education that can enhance the educational experience and support the social and emotional development of students with disabilities.

2.2.3 Socially Assistive Robots in Therapy

SARs have also been used in therapeutic domains, where they have been shown to increase engagement [17, 42] and encourage communication [43]. Some of this work has been in therapy with adults, such as those with intellectual disabilities, in which SARs have been shown to increase social interaction and engagement [44]. Most SAR research in the context of therapy, however, has been performed in the context of therapy with children, such as autistic children [17, 42], where SARs have been shown to not only increase engagement [17] but also to more broadly improve social [17, 19], behavioral [45] and cognitive [45] skills. As discussed in Chapter **??**, one of the key benefits of SARs across both of these domains, is the ability to effectively deliver personalized content.

2.3 Personalization in Educational and Therapeutic Interventions

Educators tailor content based on students' strengths, needs, and interests, to increase motivation and engagement [46]. This personalized approach is especially crucial in learning interventions for autistic children, considering the vast heterogeneity among children with Autism Spectrum Disorder (ASD) [47, 48]. For example, speech and language therapists, who deal with linguistically, culturally, cognitively, and behaviorally diverse groups of children with ASD, are challenged to navigate this heterogeneity in intervention selection and response measurement [49]. Recognizing that a single intervention is not universally effective for all children with ASD [50], therapists must employ diverse tools and approaches, guided by an understanding of intervention effectiveness based on the child's symptom presentation [51].

2.3.1 Personalization in Special Education

Personalization is a critical component of special education [1], as it allows educators to tailor instruction and support to meet the unique needs of each student with disabilities. Personalization can lead to improved engagement, motivation, and academic outcomes for students with disabilities [47, 52, 53]. By personalizing instruction and support, educators can address the diverse learning needs of students with disabilities and provide them with opportunities to develop their strengths and interests [47, 52, 53]. Additionally, personalization can help students with disabilities develop self-awareness, self-advocacy, and independence, which are essential skills for success in school and career [47, 52, 53].

2.3.1.1 Individualized Education Program

An IEP, or Individualized Education Program, is a legal document that outlines the educational goals, services, and accommodations for students with disabilities [54]. The IEP is developed by a team of professionals, including special education teachers, general education teachers, parents or guardians, and related service providers. It is based on a comprehensive evaluation of the student's strengths and needs and includes specific goals and objectives for academic achievement and functional performance. The purpose of the IEP is to provide students with disabilities with an individualized and appropriate education that enables them to reach their full potential in school and beyond [54]. The IEP is reviewed annually and updated as necessary to reflect the student's progress and any changes in their needs or services [54]. Research into IEPs has shown that districts struggle to fill out certain sections of IEPs such as student's present educational performance, IEP goals, and specialized services provided to the student [54].

Personalization is crucial for achieving engagement [15] and positive learning outcomes [42] in therapy. For example, neurodivergent individuals have different therapeutic needs just as they have different educational needs. If a neurodivergent individual has a physical disability, physical or occupational therapy may be appropriate. In contrast, if they have a speech or language impairment, speech and language therapy may be needed. Moreover, each of these therapies must be carefully customized to the needs of individual clients [55, 56]. Yet personalization may come at a steep cost, requiring significant labor that often falls on the shoulders of educators [57] and therapists [21]. These labor costs may be avoidable, however, through well-designed visual support tools.

2.3.2 Role of Visual Support Tools in Personalization

Education and therapy technologies often make use of visual support tools [53, 58]. These tools use visualizations to help therapists and educators make sense of complex data and more easily spot trends and outliers [59–62]. Visual support tools such as dashboards help augment decision-making processes by

facilitating the filtering, analyzing, and visualization of data [63]. In education, dashboards have been developed to support educators when reflecting on their teaching practice or to help educators find at-risk students [64]. Data visualizations can also help educators gain insights that may positively influence their personalized learning interventions [53]. Similarly, in therapy, data visualizations can help clinicians see trends that could enable them to better customize treatment for their patients [58]. Dashboards, for example, can help clinicians digest large amounts of data in a way that enables them to easily personalize their treatment plans [63]. Clinicians are thus better guided in their decision-making process with the help of data visualizations that make processing information more accessible [63].

In addition, research has found that data collection and analysis are essential for evaluating the effectiveness of AT in practice and informing decisions about which tools and interventions to use for individual students [65]. However, collecting and analyzing data on AT use in special education can be challenging, as there are often numerous stakeholders involved in the process, including educators, service providers, students, and families. Additionally, collecting and analyzing data on AT can be time-consuming and resource-intensive, which may limit its use in practice [33, 34, 65].

Despite the promise of and the need for these tools, little work exists on their use in the domain of socially assistive robotics. We argue that this may be because of tensions between the dominant approaches taken by SAR researchers and the actual needs and practices of SAR operators.

2.4 Research to Practice Gap

Below we will discuss the current gap in SAR research including focus on fully-autonomous robots when in fact teleoperation is best suited and the implication of shifting our focus from fully-autonomous to teleoperated robots.

2.4.1 Level of Autonomy

In the realm of Socially Assistive Robots (SARs) research, the predominant focus has been on alleviating the labor cost of personalization, envisioning a future where autonomous robots dynamically adapt to users for long-term personalized interventions [21, 57]. The goal of autonomous SARs is to reduce the burden of long-term individualized interventions by automatically adjusting to user behavior [47] and learning patterns in this behavior [66]. However, recent perspectives challenge this emphasis, suggesting that task criticality, accountability, and environmental complexity might warrant partial or full teleoperation in many SAR domains [21]. Autonomy, defined as a robot's ability to perceive surroundings, formulate plans, and execute actions without external control, is a crucial consideration for researchers examining appropriate levels of autonomy (LoA) in SARs [21]. Task criticality involves balancing

automation and the risk of task failure, with increased automation introducing risks in highly critical tasks. Task accountability influences LoA by determining who should be held accountable when errors occur [21]. Depending on the environment's complexity and dynamics, the deployment of SARs may vary [21]. Considering these dimensions, robots with less autonomy are recommended in tasks with high criticality, high accountability, or unpredictable environments [21].

In the context of education, SAR research has primarily focused on fully autonomous robots, influenced by challenges in deploying teleoperated SARs, such as training difficulties for non-expert users [12] and perceived limitations in teleoperation interfaces [21]. Following the guidelines for LoA, the task criticality in delivering educational content through robots depends on the subject matter, with quantitative topics being easier to correct than socially-oriented subjects [21]. Task accountability is high in educational scenarios involving children, considered a vulnerable population [21]. The unpredictable nature of a classroom environment suggests SARs may benefit from a human supervisor, typically the teacher or teacher's assistant [21]. Analyzing these dimensions, SARs involved in teaching highly sensitive content are anticipated to have a low LoA [21]. Additionally, human experts in education are already trained to handle unpredictable situations, and SAR designers can leverage their expertise instead of replacing them[21]. In practice, many SARs are typically teleoperated by caregivers (those assisting individuals) already [21].

The shift from fully autonomous to teleoperated SARs underscores the importance of considering teleoperators, particularly therapists and educators, who have traditionally been overlooked [56, 67]. Additionally, it prompts the exploration of neglected dimensions of SAR deployment, such as content authoring and post-session analysis [24].

2.4.2 Evaluation is Important for Teleoperators Using SARs

Interviews with participants experienced in using robots to deliver therapeutic content to children revealed that therapy is cyclic, comprising two cycles: the outer cycle, spanning months or years across multiple sessions, and the inner cycle, occurring within each session on a short timescale of minutes or hours [22]. The outer cycle involves collaboration with stakeholders, goal-setting, intervention delivery, and iterative reviews, while the inner cycle focuses solely on the therapist and child working together during a session [22]. Importantly, the model emphasizes the need to extend the analysis beyond the session itself to include authoring interfaces and post-session analysis dashboards [22]. The significance of the "dual-cycle model lies" in its capacity to reveal valuable insights for improving robotic technology used in therapy [22]. While traditional perspectives often concentrate on the inner cycle and the direct robot-child interaction, Elbeleidy's work advocates for a broader view that gives enhanced emphasis to the outer cycle [22]. Elbeleidy's study explores six themes derived from the interviews, directly linked to how therapists

examine, evaluate, and prepare within the inner and outer cycles of therapy. One finding was the importance of evaluation for teleoperators [22]. Evaluation involves assessing the value and success of therapeutic activities [22]. The process allows teleoperators to monitor client progress within and between sessions and report to stakeholders such as insurance companies and parents [22]. The evaluation process focuses on continuous assessment and long-term changes in a child's behavior [22]. The iterative nature of evaluation influences how teleoperators prepare for future sessions, as their evaluations determine future interventions [22]. Therapists stress the need for continuous evaluation to monitor improvement and ensure appropriate interventions [22]. Therapists periodically determine specific goals for children, which may change over time based on the child's progress [22]. Due to the importance of evaluation, it is recommended for roboticists to enhance SAR tools by displaying long-term progress, offering reports for insurance approval, and facilitating easy comparison of a child's performance across sessions [22].

2.4.3 Addressing the Research to Practice Gap through Session Summary Reports

As mentioned earlier, a personalized curriculum is critical to successful interventions. Visual support tools such as dashboards or visual reports [53, 58] play a crucial role in creating personalized content. These visual reports can present performance information, offering real-time detection of issues by displaying students' learning progress and providing alerts or feedback for future sessions [68, 69]. Visual tools empower educators to gain insight into their clients and students, informing how they tailor interventions [53, 58]. One opportunity for improving the teleoperation interface is a visual support tool, like a session summary report, aiding teleoperators in their evaluation process [23]. This presents an opportunity to leverage data collected from SARs and teleoperators, providing educators with valuable insights to support future sessions and personalize content. Evaluation is crucial; therefore, effective SAR systems must consider teleoperators.

2.4.4 Institutional Support: A Dual Role in Evaluation and Implementation

Institutions play a crucial role in the success of SARs in both evaluation and implementation processes [22, 70]. In terms of evaluation, institutions are instrumental in standardizing the evaluation process, providing valuable tools like Google Forms to collect and summarize data. This not only streamlines the evaluation workflow but also alleviates the workload of preparing for clients and evaluation from teleoperators [22]. Furthermore, institutions contribute significantly to the continuity of service by offering essential documents to assist new therapists in their preparation [22]. In essence, institutions emerge as vital contributors to therapy and evaluation, offering tools that not only standardize processes but also retain crucial long-term documentation about children, ensuring a seamless continuation of service when

therapists change.

Moreover, the significance of institutional support extends beyond evaluation into the realm of implementation. Implementation involves seamlessly integrating technology into the daily routines of educators and therapists, ensuring sustained and effective usage [71]. Khaksar et al.'s longitudinal study spanning 12 months emphasizes the pivotal role of organizational support in championing the successful implementation of socially assistive robots in educational services [72]. Another perspective suggests the incorporation of technology leadership into the responsibilities of principals because they are crucial decision-makers in technology integration [70]. Another work highlights that effective leadership, especially from principals, significantly influences a school's effectiveness when implementing technology-related projects [73]. This multifaceted involvement of institutions underscores their dual responsibility in both evaluating the efficacy of SARs and effectively implementing them within educational frameworks.

To summarize, tension exists between the SAR literature's focus on autonomy and the actual need for enhanced teleoperation faced by therapists and educators in the field. Our first two studies aim to understand how a session summary report can support teleoperators in their evaluation process and what specific features and considerations should be considered when designing such a feature. Motivated by the critical role institutions play in the evaluation of therapeutic interventions and implementation of technology in schools, our third study seeks to delve into the involvement of institutions in the successful implementation of SARs in special education settings.

By understanding how institutions can actively support the integration of SARs into special education practices, we aim to contribute valuable insights that bridge the existing gap between technological innovations and their practical, day-to-day application in educational settings. Our focus on the role of institutions underscores the broader implications of SAR implementation, shedding light on the collaborative efforts necessary for its successful integration and long-term effectiveness. By centering our research on teleoperators' needs, specifically educators within special education, we contribute critical literature in SAR research regarding adoptability and implementation. Our studies emphasize how to improve SAR design for better adoptability through session summary reports, relieving the burden of evaluation from teleoperators, while also addressing how to enhance implementation processes as more classrooms and special education teachers adopt SARs in the future.

CHAPTER 3
STUDY ONE METHODOLOGY

To gain deeper insights into the significance of an evaluation tool, such as a session summary report, for teleoperators, and to understand their specific preferences regarding its design, we employed an iterative design methodology. This approach combined design critiques, interviews, and design iterations, followed by a qualitative analysis of the interview data. Iterative design, a widely embraced approach, involves creating multiple design iterations through frequent prototyping, testing, and refinement, with the flexibility to revisit and circle back to different design stages throughout the process [74, 75].

To facilitate our iterative design approach, we began by creating a preliminary design of a session summary report based on treatment and lesson plans found online, and based on prior research conducted by Elbeleidy et al [26]. We then iteratively improved this design, conducting semi-structured interviews and design critiques with domain experts in between design iterations. The interviews allowed us to learn more about teleoperator workflow pre-, during-, and post- (therapy or class) sessions. These interviews were guided by prepared questions on topics like the types of tools, if any, interviewees used to collect data and analyze sessions, and the types of metrics they used to track and measure progress. These guiding questions allowed us to get a better idea of the kind of reporting that would ease the evaluation process. The design critiques of our intermediary designs (conducted with interviewees immediately after these interviews) helped us to evaluate our choices of content and visualizations based on comprehensibility, usability, and necessity. The data collected during these interviews and critiques were then qualitatively analyzed to help us understand our participants' needs and to help guide our next design iteration.

We will now provide additional details and methodological background for each element of this process.

3.1 Preliminary Design Development

To create our preliminary design, we leveraged five key sources.

1. **Previous Research:** We took inspiration from previous research that recommended reports to support post-session reflection [26] and behavior analysis [19].

2. **Treatment and Lesson Plans:** We collected online sample treatment and lesson plans to get a general sense of the information therapists and educators keep track of.

3. **Program Surveys:** We collected post-session survey questions from an organization that uses Peerbots in their interventions with children to inform the type of information a session summary

report should collect and visualize.

4. **Session Transcripts:** We consulted transcripts of the use of the Peerbots system by therapists and educators during the authoring and execution of real sessions. Our goal was to discern the specific information, if any, that was documented during these sessions, laying the foundation for understanding the essential elements to be collected and visualized in a comprehensive report.

5. **Preliminary Design** We took inspiration from a previous Session Summary Report prototyped specifically for the Peerbots platform described by [23]. We did not build *directly* off of this preliminary design as it was overly constrained to the current capabilities of the Peerbots platform so instead we created an initial design that was more universal so our participants could contribute their expertise without background knowledge of Peerbots.

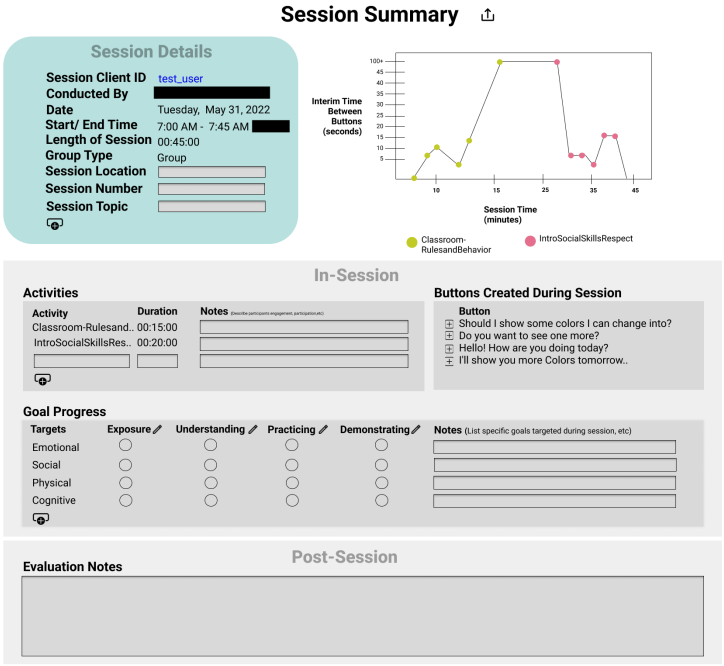


Figure 3.1 Initial Design for session summary report created before conducting interviews. This report is meant to summarize the activity after each session. Critiqued by Participant 1.

The resulting design is shown in Figure 3.1. This initial design has 3 key elements.

1. **Session Details (Top Left):** Here, session identification information is shown, including date, start time, and end time which was inspired by the same session details presented in Elbeiledy's work [23]. Fillable fields for Session Location and Session Topic are also included and were informed by the Program Surveys we reviewed. These sections allow teleoperators to quickly see which session they

are viewing and the topic covered. Finally, there is space for teleoperators to add additional fields as they see fit.

2. **Session Time Chart (Top Right):** Here, a graph breaks down the session length according to the activities conducted with the SAR. This graph was created using metadata collected from a real session with Peerbots. Each dot represents a button pressed on the Peerbots interface that permits the SAR to speak. The color of each dot indicates the activity being conducted. The length between button presses is visualized to help teleoperators identify where the longest pauses occurred. Long pauses could signal an activity being performed outside the context of the SAR interaction or could signal extended communication with the child or children. We wanted to visualize this to help teleoperators remember activities that might have been crucial in the session but for which the SAR was not used – or which the SAR could not track. This graph was inspired by the Session Details section presented in Elbeiledy's recommended post-session report [23, 26] but changed to leverage colors for easy identification of the activities being conducted with the SAR.

3. **In-Session (Middle):** This area captures what happened during a session, and is comprised of three sub-areas:

   (a) **Activities (Upper Left):** Here, the activities performed through Peerbots are automatically tracked, along with the duration of each activity. This component was motivated by the program surveys mentioned above. This subsection also includes fillable fields that allow teleoperators to leave anecdotal feedback and ad-hoc notes. This information is important to show because it informs the utilization of SARs in therapeutic and educational interventions. Longer duration of activity with SAR could mean that it's working out well with a child/ children and is an effective way to engage them.

   (b) **Buttons Created During Session (Top Right):** Here, live "programming" by the teleoperator (in which new dialogue options are created on the fly) is recorded. By keeping track of new buttons, this subsection aims to help teleoperators identify options that should be added permanently to future sessions (an action performed by selecting the "+" icon). This will help them run future sessions more easily because they have responses ready to use versus having to spend time authoring them during a session. This section was motivated by the chart in Elbeiledy's work [23] that shows the buttons created during a session.

   (c) **Goal Progress (Bottom):** Here, teleoperators can record proficiency reached for each goal. Goals and proficiency levels are added by content authors through the Peerbots interface for

each dialogue option. This subsection also includes fillable fields for anecdotal and ad-hoc feedback. The proficiency of each child is important to track because this is the main indicator of progress. By being able to track proficiency for specific goals teleoperators can gauge not just progress but where more intervention is needed. This section was motivated by the program surveys, online treatment, and lesson plans we reviewed prior. Those three sources emphasized the need to measure progress for specific goals.

4. **Post-Session (Bottom):** Here, a text box is provided where teleoperators can reflect on the session as a whole, and provide evaluation notes on the child or children's overall progress. This was inspired by the program surveys, online treatment, and lesson plans we reviewed that enabled therapists and educators to leave as many notes as possible. This is important because there are things that a SAR can't track or can't interpret so it's important to leave space for the trained therapist and educator to summarize the session in their own words.

After creating this initial design, we were ready to begin our iterative design approach.

3.2 Recruitment and Participants

We reached out to forty-five child therapists and K-12 educators via email. Forty-one of these were sourced from online school directories, LinkedIn, and Psychology.com's therapist directory. We specifically targeted child therapists and K-12 educators because research has shown the benefits of using SARs with children in therapy [17] and education [15]. We also reached out to select K-12 instructors with whom we already had a relationship, to identify other educators who might meet our criteria and be interested in participating. Finally, we reached out to SAR experts who had previously participated in research with our lab. From this set of candidates, we recruited five domain experts, including two therapists, two educators, and a non-profit administrator whose therapeutic and educational programs used SARs. This set of interviewees thus covered multiple distinct perspectives on SARs. **??** provides more demographic information such as background and experience with robots or SARs for each of these participants. After scheduling their interviews, participants were sent consent and data authorization forms.

3.3 Study Procedure

Each session was split into a semi-structured interview and a design critique. The semi-structured interview enabled us to gather information about the participants' workflows before, during, and after a session. The design critique was used to get feedback on the content and visualizations of our report(s).

Table 3.1 Interview Participants

Pseudonym	Profession	Experience with robots or SARs	Recruitment Type
Participant 1	Non-profit administrator	Yes; Peerbots	Recruited from our professional network
Participant 2	Child therapist	None	Psychology.com
Participant 3	High-school STEM teacher	Yes, but not as a medium to teach	Colleague of personal connection
Participant 4	High-school STEM teacher	Yes, but not as a medium to teach	LinkedIn
Participant 5	Child occupational therapist	Yes; Peerbots	Recruited from our professional network

3.3.1 Semi-Structured Interviews

We conducted one-on-one interviews to dive deeper into each participant's experience. All interviews were conducted and recorded using Zoom. Semi-structured interviews were used to focus on pre-determined themes while allowing the conversation to flexibly expand into other relevant areas [76]. We prepared a series of questions before the start of interviews, inspired by the treatment and lesson plans we found online, and by prior research focused on teleoperators of SARs [18, 26]. Collected questions were chosen to acquire a better sense of each participant's workflow pre-, within-, and post-session. We also prepared questions to dig into specific areas of experience that participants might have. Our prepared questions are found in Appendix A.

3.3.2 Design Critique

Design critique is a reflective practice used to gain holistic insight into a particular design [77]. This design method encourages information sharing and uses feedback from stakeholders to infuse important values into future designs [78]. A significant part of a design critique includes analyzing a previously existing design to produce descriptions of its design considerations while developing a rich contextual understanding of its effects [79]. In our design critiques, we asked each participant to provide feedback on our session summary reports, with specific attention to their interpretation of the report, what parts of the report they found valuable, and what additions they would like to see in the report. Later in the process, as we will discuss in more detail, we created an additional design for a Progress Report that participants critiqued. Verbal feedback and commentary were recorded during design critiques. As part of this critique, additional questions were asked to further evaluate our designs for usability, comprehensibility, and usefulness. These questions are listed in Appendix B.

3.4 Iterative Design and Qualitative Analysis

After each interview, we iterated on our designs by identifying major problems brought up by the preceding participant. We then updated our designs to address these problems. Our iterative design process enabled us to work closely with real users to ensure their needs were met. Each iteration provided new insights for improving our design. After all of our interviews had been completed, we performed a thematic analysis of those interviews, to identify themes and patterns and develop a deeper understanding of our results [80].

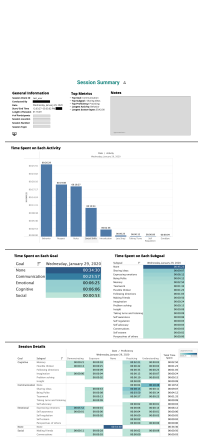
Following the thematic analysis approach, we first open-coded the transcripts from all five interviews. We labeled participant utterances according to seven key factors: (1) Comments or feedback relating to session summary reports (why they're important, what metrics do they want to see, etc), (2) Upsides of SARs in their programs, (3) Downsides of SARs their programs, (4) Role of robots in their programs, (5) Preparation techniques, (6) Educator-specific needs, (7) Therapist-specific needs. After open coding was complete, we grouped the annotations to better understand the connections between our seven factors above and the general themes we identified.

CHAPTER 4

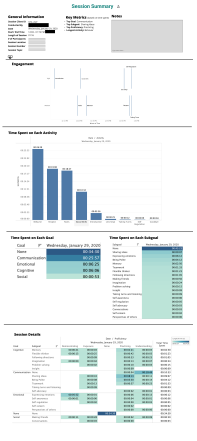
STUDY ONE RESULTS

In this section, we will discuss the results of our iterative design procedure, discussing each component of our final design. Our final design includes a (1) newly designed main panel as well as panels for (2) student/client performance, (3) activity duration and engagement, (4) time spent per goal, and (5) new task efficiency. As we will discuss later on, one of the most critical pieces of feedback received during our iterative design process was the need for distinct reports for visualizing (a) the contents of an individual session, and (b) progress over time. As such, we will show and discuss both the session summary report and progress report versions of each of our visualizations. Intermediate designs are shown in Figure 4.1 (for session summary reports) and Figure 4.2 (for progress reports).

(a) Design 2, critiqued by Participant 2.



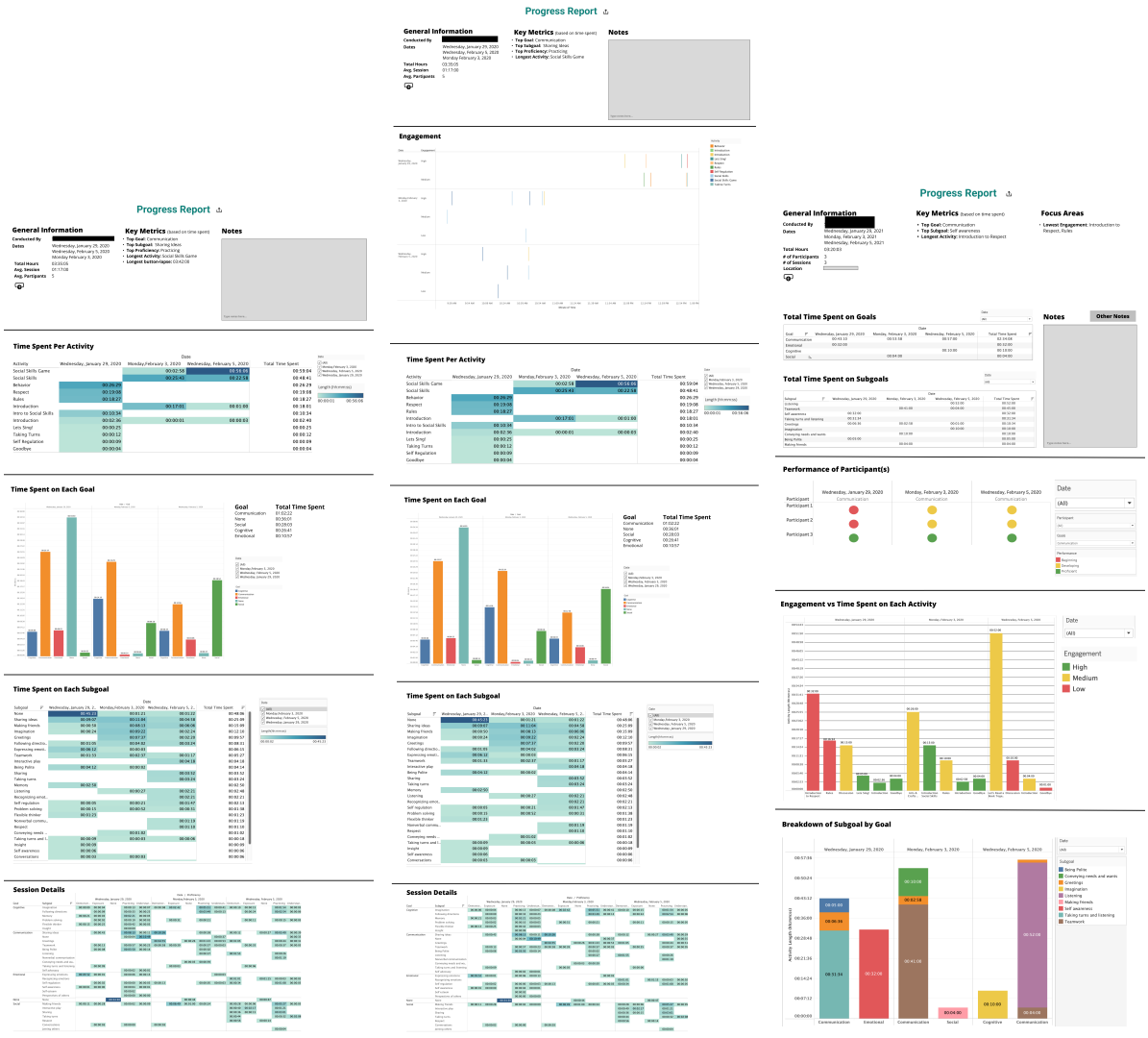
(b) Design 3, critiqued by Participant 3.



(c) Design 4, critiqued by Participant 4.

(d) Design 5, critiqued by Participant 5.

Figure 4.1 Design Iterations for session summary reports



(a) Design 3, critiqued by Participant 3. (b) Design 4, critiqued by Participant 4. (c) Design 5, critiqued by Participant 5.

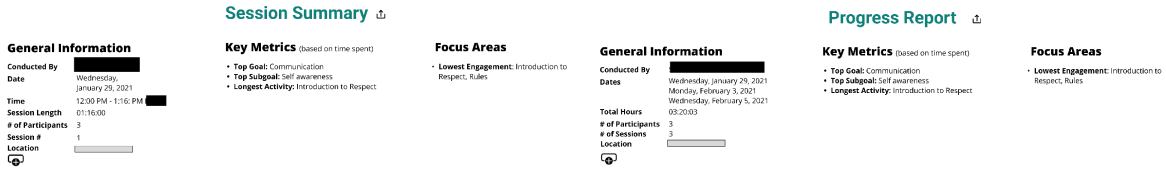
Figure 4.2 Design Iterations for progress reports. Progress reports were not created until after the interview with Participant 2

4.1 Main Panel

The first panel of our final report design shows the information needed for therapists, educators, and administrators to quickly identify the session they are reviewing (Figure 4.3). This part of the main panel remained unchanged throughout the design iterations. However, the "Key Metrics" and "Focus Areas" sections were added below this section to provide a quick summary of the other panels, due to the significant time constraints that therapists and educators face. Previous work has suggested that educators with access to existing technologies with personalization features underutilize them due to these time constraints [48], a finding that was backed up by our interviews.

> *Participant 4 (when asked if it'd be helpful to have metrics tracked):* "It'd be really awesome to have like the para doing this because the teacher's gonna have to do about a billion other things. And so like remembering to rank it, like you could rank it afterward if you had a planning period right afterward. But usually you're like teach, teach, teach, teach, teach [...] when you're managing a class there's so much going on already to then remember to like log some data point, it was really difficult."

Based on this observation, we believe that our initial summary panel is critical for conveying the actionable essence of what has been tracked by a SAR with a concision that is sensitive to these time constraints. As such, the Key Metrics section shows only the goal, subgoal, and activity the therapist or educator spent the most time on, while the Focus Area section shows only the activities that had the lowest engagement.



(a) Main Panel: Session Summary Report

(b) Main Panel: Progress Report

Figure 4.3 Main Panel

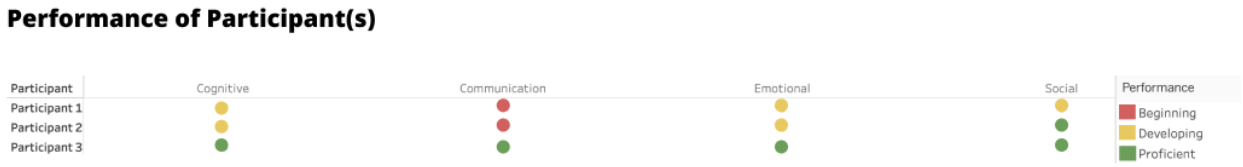
4.2 Student/Client Performance Panel

Participants stressed the importance of showing progress toward goals, to assess how each child is tracking toward their goals.

> *Participant 4:* [School administrators] want to know that you're...not...just subjectively gauging how they're learning, but they want...hard data to show it.

Participant 5: I would usually just give a narrative of the session itself and that's, I guess, sounding a little bit more objective, but then I would say if there was any progress towards certain goals that we were working towards and which supports were used, which adaptations were used and then what the plan was for the next session. And then there was always an education piece for the parent or caregiver in the session.

This need was met by our Participant Performance Panel, which visualizes the proficiency level achieved by each student or client towards each of their goals. The same version of this panel was used in both the Session Summary Report and the Progress Report.



(a) Participant Performance Panel

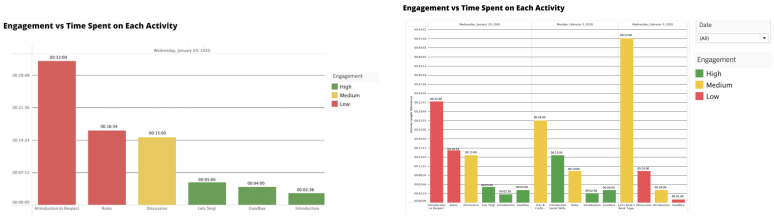
Figure 4.4 Participant Performance Panel

4.3 Activity Duration and Engagement Panel

Participants expressed the need to understand how time was spent during a session for each activity, and how this related to engagement.

Participant 1: Knowing how much the robot was engaged is a really good thing in a classroom situation where you're going to have to report to the administration of a school[...]how much we use the robot in class because they want to know what why they are paying to have this robot brought in.

In particular, participants wanted to see how much time was spent on each activity, and the level of engagement for that activity. Research also shows that more engaged students are more motivated to participate and learn [67]. This need was met by our Activity Duration and Engagement Panel (Figure 4.5), which shows the overall engagement and time taken for each activity. Our fourth design included a separate Engagement table. However, in our fifth design, we decided to consolidate our Engagement table and our Time Spent per Activity table to reduce cognitive load.



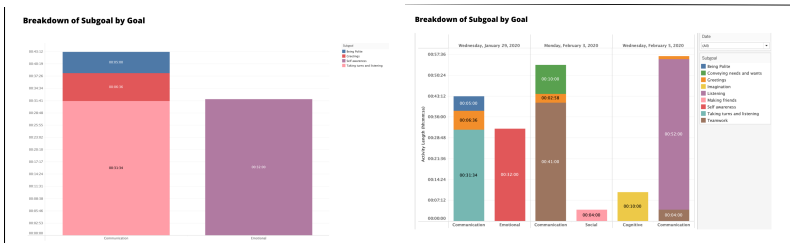
(a) Activity Duration and Engagement: Session Summary Report

(b) Activity Duration and Engagement: Progress Report

Figure 4.5 Activity Duration and Engagement Panel

4.4 Time Taken per Goal Panel

Participants also expressed the need to understand how much time was taken on each goal and subgoal. This need was met by our Time Taken per Goal Panel (Figure 4.6). As before, we initially included separate panels for Time Spent Per Goal and Time Spent Per Subgoal but consolidated them to reduce cognitive load.



(a) Time Taken per Subgoal and Goal: Session Summary Report

(b) Time Taken per Subgoal and Goal: Progress Report

Figure 4.6 Time Taken per Goal Panel

4.5 New Task Efficiency Panel

Participants also expressed the need to understand how much time was spent on tasks unrelated to the session itself, such as activities conducted without a SAR, or the time needed to author new buttons on the fly for the sake of teleoperation.

> *Participant 5:* It would be helpful to know how much time between the last button that I hit and creating a new one. Like I was getting a new one and like, oh my gosh, it takes a long time, like further justifying why there should be some modification to that to ease that process a little bit. I saw that really kind of hinder the social grace and rapport we had already built

because then there would be like kind of this awkwardly long wait before the robot responded. And then that would like call into question the validity of the robot to the child.

Participants reported that knowing how long especially extensive verbalizations took to create would help to determine whether the use of a robot was worth the time cost. The manual tasks that teleoperators have to do when using any socially assistive robot should be tracked, and information about them should be made available to users for them to make informed decisions about whether to use a robot and if so, how to use it efficiently. Figure 4.7 shows the *task efficiency panel* from our progress report. This table shows new button(s) created, which activity they were created for, and how long it took to create them. A Notes field was added to allow for any anecdotal reporting of activities conducted without the SAR, or observations not captured by the SAR.

(a) New Task Efficiency Panel: Session Summary Report (b) New Task Efficiency Panel: Progress Report

Figure 4.7 New Task Efficiency Panel

4.6 Qualitative Analysis

After completing our iterative design process, we performed a thematic of the transcripts collected from our interviews, to allow us to further reflect on our results and identify possible limitations. To do so, we engaged in open coding of our interview transcripts then clustered codes into themes, and reflection and analysis of those resulting themes.

This analysis provided us with additional insights as to how our final designs might best provide value to educators, therapists, and administrators, and revealed three key themes that had arisen throughout our interviews:

- Performance and engagement metrics guide future therapeutic and educational interventions
- Therapists and educators have limited time to reflect after sessions
- Progress reports provide credibility in therapeutic and educational interventions

4.6.1 Performance and engagement metrics guide future therapeutic and educational interventions

Engagement and performance metrics help therapists, educators, and administrators see how well their students or clients are progressing and if the activity is effective. Participants 2 and 5, who are both therapists, explained that tracking emotional responses would be helpful in their work, as this is most often how they measure progress in emotional regulation or social understanding.

> *Participant 5:* I would always take note of the emotional regulation piece of it because I think that comes into the social interaction part a lot. And if you can't regulate your emotions, you're probably not going to be interacting appropriately with somebody [...] I would've liked to see if a child responded appropriately to a facial expression that the robot was using or a tone of voice, just to kind of pick up on those social cues, because that was so often a goal of ours when you really couldn't figure that out based on software.

> *Participant 2 (when asked what they would like summarized in a report that will be helpful when filling out a treatment plan):* Maybe like mood or something like that for kids that struggle with like, you know, irritability or just anger [...] it's all related. I feel like, you know, like the mood affects behavior and how you interact with people and if you follow rules or not.

Previous research into socially assistive robots has also noted their robots' limitations in not being able to adapt to the mood of the learner [13], as this can impact a student's ability to learn and engage in school. Research has also noted the importance of engagement in measuring SAR effectiveness due to its correlation with the effectiveness in educational [47] interventions.

> *Participant 4:* My engagement levels are pretty high. I need to gauge performance and understanding of concepts. I'm doing a lot more assessment of knowledge and skills than I am of behavior and engagement.

Moreover, different assessments are used in therapy and education to track and measure progress. For example, P-scales [45] are used for students with learning difficulties, and Goal Attainment Scales [81] are used in therapy. This suggests a need for different interfaces based on whether an educational or therapeutic perspective is taken.

> *Participant 5 (when asked how they measured success):* goal attainment scale and that had kind of a spectrum of, so like a zero meant that the goal was met. Then you could have a plus one or plus two if they surpassed the goal and then a negative one or a negative two if they were below what the goal was. And so you would usually have to have an objective measure.

Finally, there is no one set of goals that apply across all therapeutic and educational domains. Therapeutic and educational efforts often have different goals even when they share high-level objectives, and goals differ from teleoperator to teleoperator. It's thus important for socially assistive robotic systems to allow teleoperators to flexibly adapt the set of goals tracked within and between sessions. Moreover, as reflected in our ultimate designs, real-world logistical concerns mean that reports need to include information beyond goals and subgoals, such as duration and timing information. Including such information helps teleoperators see where their time is spent, and where they might need to focus to better achieve their goals.

4.6.2 Therapists and educators have limited time to reflect after sessions

Report visuals need to provide actionable insights while being mindful of the limited time available to therapists and educators for reflection. Educators, for example, often teach large back-to-back classes across multiple subjects, severely limiting the time available to peruse and reflect on reports:

> *Participant 4 (when asked to look at the table that provided time spent on each activity, goal, and subgoal):* "This is making me like, not happy to look at this, like so cool to take this data, but no teacher's going to sit down and like pore through this[...] It needs to be pretty simple, quick, as teachers [...] have like five-minute chunks of time to look at something [...] we don't have like an hour and a half to sit down [and look at this].

Due to these time constraints, any visualization that a report automatically creates needs to focus on tangible and actionable insights. Critically, what is actionable may differ depending on the domain:

> *Participant 5:* I can look at it quickly and kind of know what the majority of the time is spent on and that's nice [...]that it's fast [...] So like for me as a therapist, I'd be like, well, it might be tough to progress in a goal area if you're not being consistent.

4.6.3 Progress reports provide credibility in therapeutic and educational interventions

Most interviewees reported engaging in deep reflection on their sessions or classes only every few weeks. For educators, this occurs around progress report season, or when meeting with administrators. Therapists would typically reflect on session outcomes when submitting treatment plans to insurance companies every six weeks or so. Because socially assistive robots are often used to facilitate long-term goals, this reflection and assessment must occur frequently, and not simply at the end of a course of intervention that may have lasted a year or longer [45, 47].

Our participants recognized this. Most reported that they tried to engage in *some* reflection after each session or class. The tension between the time needed for these reflective practices is grounded in the time

constraints described above. Our participants reported trying to navigate this tension in several ways. Participant 1, for example, would manually compile progress reports from the Google Forms they were using, and manually create graphs and charts based on that data. Similarly, Participant 4 maintained a folder that contained their teacher effectiveness goals and reported using this to mark off goals over time.

> *Participant 4:* I make a folder every year with goals in it and mark things that I'm working on then meet with my administrator to go over each one [...] I know that administrators love data [...] like [...] scores on the AP pre-assessment that I gave [...] They want to know that you're [...] not [...] just subjectively gauging how they're learning, but they want like hard data to show it.

These observed end-user programming and design efforts motivate perhaps the largest change we made throughout our iterative design process: separating concise session summary reports from more expansive progress reports that aggregate data from multiple sessions.

These progress reports also serve a critical administrative purpose, as they are critical for providing credibility and proof that progress has been achieved. This proof is often needed for SARs to be justifiably deployed in clinical practice [19], and many therapists are hesitant to use robotic technologies without evidence for their utility [19, 82]. Similarly, Participant 1 (an administrator) spoke of the need to use progress reports as proof that bringing robots into classrooms was worthwhile:

> *Participant 1* It was great for people to be able to see that it's not just anecdotal information. It's not just me saying it works. They had the proof that it did. And so, you know, that was, that was always very, just very helpful and impactful, made the whole program a lot more credible [...]knowing how much the robot was engaged is a really good thing in a classroom situation where you're going to have to report to the administration of a school and you want to show them, look, this is what we do. Here's how much we use the robot in the class because they want to know why are we paying to have this robot brought in.

Finally, therapists talked about their need to provide hard data to insurance companies and clients, to show why their sessions were effective and necessary. Insurance companies like Medicaid require treatment plans to be submitted every 6 months, while others require other documentation; yet all require proof of consistent progress for reimbursement.

> *Participant 2 (when asked when they are required to submit paperwork for their sessions):* We have [...] private clients and then our agency clients. So our agency clients are anyone with Medicaid or funded through a grant program [...] They'll do internal audits [...] and check and

make sure that you have a treatment plan [...]evaluating their symptoms [...] are things getting better? Are they staying the same? Are they getting worse? And they report out on those [...] probably for funding.

Participant 5: You can get audited and you need to show that what you working on is effective [...]if like insurance was fighting me back, if I was recommending like 12 sessions and they only approved six, then I might have to provide supplemental documentation.

For both therapy and education, these stories underscore a critical point: progress reports need to track not only the metrics that therapists and educators use to guide their practice, but also whatever metrics are deemed important by other stakeholders, like parents, employers, and insurance agencies. Moreover, these metrics need to be showcased in a way that is easy to understand and doesn't require a lot of cognitive effort, both for educators, therapists, and administrators as well as for these third parties.

4.7 Discussion

Our iterative design process and subsequent thematic analysis led us to consider several key considerations for the design of future tools like that designed in this work.

4.7.1 Data-tracking

A key tension felt throughout this work was therapists' and educators' desire to track key metrics like engagement, yet the lack of time to enter this data. This could be addressed in several ways. One option common throughout similar research would be to automatically measure engagement through multimodal cues. Therapists and educators alike have expressed a desire for SARs to track social and affective cues in the name of engagement measurement [35]. This solution comes with its challenges, though. First, automated engagement tracking may not be feasible on all SAR hardware platforms, and it presents a substantial algorithmic challenge in contexts like large group settings, especially when a classroom has only a single SAR.

4.7.2 Data Literacy

Data literacy of therapists, educators, or administrators is also important as it can heavily impact the usage of reports, analysis, and decision-making. Inadequate data literacy skills in educators can lead to poor interpretation of analytics and result in decisions that can negatively impact student learning [83]. This is thus a critical challenge for the successful deployment of SAR session summaries. Lack of data literacy skills can also impact therapists', educators', or administrators' motivation to consult and use

reports to improve their performance [84]. Future designs should consider providing operators with access to more complex visualizations only as they gain mastery over their interpretation.

4.7.3 Buy-in from employers and other stakeholders

Finally, our work has demonstrated the importance of gaining buy-in from other types of stakeholders, like employers and insurance companies. These stakeholders can help implement and enforce the usage of reports and can encourage or require the training needed to gain the data literacy needed to interpret those reports. Our participants conveyed that they already kept insurance companies (and what they might request or audit) in mind.

4.7.4 Limitations and future work

While we were able to glean substantial new insight through this work, our research is nonetheless subject to several limitations. First, for logistical reasons, our work is based on the perspectives of only a small number of individuals. Future work should bring in insights from a broader array of stakeholders, such as parents, caregivers, or insurance representatives. Future work could also collect information from more feedback per design iteration to ensure that large changes aren't overly biased toward the desires of individual interviewees. Second, our work focused on iterative design with interface mock-ups. Future work could instead use a functional prototype that can be assessed through usability tests, cognitive walkthroughs, and observations. These efforts would also help understand how users might edit graphs and charts to enhance their utility or engage in other types of end-user programming and design.

4.8 Conclusion

In this work, we used an iterative design approach to develop session summaries and progress reports that meet the needs of therapists, educators, and administrators. These initiatives illuminate the essential content that SAR interfaces must monitor, gather, and summarize. They delve into tracking the performance and engagement of children, correlating it with the time spent on each activity to inform future interventions. Visualization attempts should strike a balance—being easily digestible for a broad audience while offering sufficient details for teleoperators to exercise their expertise and make informed judgments. Our key finding is that both session summary reports and progress reports are key tools that serve multiple purposes for these stakeholders: they not only help clients achieve educational and therapeutic goals but also provide credibility through documentation of client success.

CHAPTER 5

STUDY TWO METHODOLOGY

In our initial experiment, therapists, educators, and administrators underscored the significance of an evaluation tool, such as session and progress reports, enabling us to offer recommendations aligned with their overarching requirements. Customizing these recommendations necessitates a thorough exploration of their distinct needs for an evaluation tool. Given the positive outcomes observed with SARs in children with autism and within educational contexts, our strategic focus is now directed towards meeting the specific needs of special education teachers. Through this, we identified crucial design implications that session summary reports must encompass to be effective in practical settings. Notably, we deliberately abstained from conducting interviews with special education teachers with experience with SARs, as the personalization of content applies to the broader special education context beyond the use of socially assistive robots. It is imperative to comprehend the overall needs and challenges in tracking goals and relevant information outside of socially assistive robot usage. This strategic approach allowed us to gain profound insights into the diverse needs of special educators, seeking a comprehensive understanding by capturing a multitude of opinions and perspectives. We aimed to answer the following research questions in our second study:

- How can visual support tools be best designed to meet special educators' needs?
- What are other design implications that should be considered when building/ deploying SARs in special education classes?

5.1 Recruitment and Participants

We employed three distinct recruitment strategies to engage participants in our study. Firstly, we scoured the websites of local school districts, identifying special education teachers from online directories and adding them to our contact list if their information was available. We reached out to a total of twenty-one special education teachers through this outreach strategy. Secondly, we proactively reached out to schools through calls and emails, seeking their collaboration in sharing our research with their internal team. We only had one institution participate through this channel. Lastly, we leveraged social media platforms, specifically relevant Facebook and sub-Reddit groups dedicated to special education, to share information about our research. This multifaceted approach yielded six initial interviews; however, one participant was deemed ineligible due to a lack of experience in special education.

After scheduling their interviews, participants were sent consent and data authorization forms.

Table 5.1 Interview Participants

Pseudonym	Years of experience as a Special Educator	Source
Participant 1	13	Online school directory
Participant 2	5	Online school directory
Participant 3	14	Cold-outreach to school's administrator
Participant 4	3	Cold-outreach to school's administrator
Participant 5	4	Reddit

5.2 Semi-Structured Interviews

We employed one-on-one interviews to explore each participant's unique experiences, concentrating on their methods of student engagement, motivational strategies, content personalization, goal tracking, and measurement approaches. Utilizing Microsoft Teams, all interviews were conducted and recorded to ensure accurate documentation. Employing a semi-structured interview format allowed us to adhere to pre-determined themes while also permitting the conversation to organically expand into other pertinent areas, aligning with established qualitative research practices [76]. The detailed set of prepared questions can be referenced in Appendix **??**.

5.3 Qualitative Analysis

After all of our interviews had been completed, we performed a thematic analysis of those interviews, to identify, analyze, and define themes within our data [80]. We first generated initial codes labeling interesting features or patterns. Afterward, we looked for broader patterns and connections among codes that may suggest the existence of themes. We reviewed the themes and consolidated them to make them stronger which resulted in two key findings:

1. The tracking of goals in special education is essential for both historical documentation and ongoing assessment, yet it presents a complex challenge due to diverse tracking methods and the multifaceted nature of special education goals.

2. Visual decision support tool should support holistic tracking of goals, customizable interactions to align with diverse teaching styles, and include collaborative features.

CHAPTER 6

STUDY TWO RESULTS

Effective goal tracking in special education is important for historical documentation and continuous assessment, however, crafting a singular visual support tool proved challenging due to the need to accommodate varying needs. Nevertheless, gleaned insights suggest that Socially Assistive Robots (SARs) hold potential for visual support tools by integrating customizable features, fostering collaboration, and aligning with the diverse strategies employed by educators in tracking student progress.

6.1 The Significance and Challenges of Goal Tracking in Special Education

Efficient tracking of goals within the realm of special education is necessary, serving dual purposes of historical documentation and continuous assessment as supported by our first study and Elbeleidy's interview with SAR teleoperators [22]. Participants provided insights into the cyclical process of reevaluation, where a comprehensive review of individual education programs (IEPs) takes place every year or every three years to determine if a student still qualifies for special education services.

> *Participant 1:* "Every year we have a meeting with parents and that's where all the data is reviewed and then we update the IEP goals. (We review) all the data that I've been collecting throughout the year...If it's a reevaluation, which is like every three years where you go ahead and see if the student still qualifies for services, then you can do a record review and look at the history."

> *Participant 3:* "You check it (IEP document) quarterly but we only meet and look at the document yearly."

As mentioned by Participant 3, the evaluation process does not only happen at a specific time but is an ongoing process. Special educators share progress more frequently with multiple stakeholders such as parents.

> *Participant 1:* "Progress reports which are like report cards are sent home each trimester or semester depending on what type of setting you're in and that goes home to the parents. All service providers are on there, so speech, OT (occupational therapist), PT (physical therapist) depending on what services the students have."

Participant 4: "I do updates (to parents) every week. If I miss a week I do a two-week update but I keep it within that window - a week or two weeks. If things are going well, I might just leave a very short two-sentence email. The longer ones are always the things that aren't going well, there's a problem or recurring behavior. I think most of the parents trust my judgment because I've at least earned that repertoire with them. But I think for the sake of peace of mind, not for my own some explainable continuity that's been charted and measured, and you can see that this has been happening for a while."

Parents play a pivotal role in their child's educational journey, ranging from actively contributing suggestions and making final decisions to a more passive role of attending meetings to gain insights into their child's behavior [85]. The intermediary role of parents between the child and the educational system is critical, yet many face challenges with school systems that inadequately support their competence or assist them in effectively navigating the intricacies of their child's education [86, 87]. This type of parental support might be something a teacher wants to do but can't do due to time constraints or limited resources.

Participant 1: "I've had parents ask "We'd love to see a video" and I'm like, I just can't, you know, we're so short. And then I'll have parents who just refuse to use seesaw [software platform for student engagement] or email and they'll only want to talk to me directly. Then you have a parent who wants me to handwrite everything. It's all over."

Another stakeholder in sharing progress reports is the students themselves.

Participant 4: "We keep track over the course of each quarter (and) we do mid-quarter check-ins and then we do end-of-quarter check-ins where we talk about everything from the class to goals. We then talk through strengths and weaknesses based on the parameters that are laid out for them (students). (an official report will be helpful) because it would help communication between teacher and student and in this setting, that's important because most kids can't figure out...They get in trouble a lot. Some of them have no clue why. So something that would communicate back and forth..that they can access and look at and say 'OK these classes I'm more disruptive in'"

However, collecting data and sending these reports are time-consuming.

Participant 1: "I try to (update student's profile on the school's internal site) at least once a week (but) depending on what type of year it is and type of staffing...once a week is great to do... (or) I write (things) down on a sticky (note)."

Participant 2: "I wish I had all the time to be able to have hard data on (my students)."

Participant 3: "We gather current information from each teacher, kind of what the students working on and where they're at. But then we also pull all of our past reports. Our school sends weekly emails to parents on what students are working on academically and what students are working on socially or emotionally."

Participant 4: "I'd have to go through (emails). Just read it. Just unbelievable amounts of text. To come up with some kind of consistent data point telling me how often that happens. And so that's what I think is missing, is being able to see what is the actual consistency with which this behavior shows up."

The importance of goal tracking and the challenges faced in tracking goals due to time constraints or lack of support highlight the need for evaluation tools like session and progress reports, as presented in Study 1. Special education teachers are mandated to explore assistive technology options [88]; therefore, streamlining the evaluation process with session and progress reports not only addresses a pressing need but also makes a compelling case for increased adoption of SARs in educational settings. Additionally, the incorporation of goal-tracking features in SARs not only provides tangible resources but also empowers students and parents to actively participate in and comprehend their child's academic progress. However, designing a comprehensive session summary report presents its own set of challenges, as the subsequent insights will reveal.

6.2 Tracking Goals are Complex and Multi-Dimensional

Navigating the intricacies of goal tracking in special education unveils a spectrum of challenges stemming from diverse tracking approaches. Some educators rely on meticulous documentation, employing methods such as note-taking stored in binders or internal web tools provided by their school district. Conversely, others depend on pure memory when offering feedback or completing paperwork.

Participant 1: "(My school) created this thing called Miller Maps and we broke down goals and created a rating scale. But the coolest thing about these Miller Maps is I scan the QR code and I go to my student's (profile) and there I have the scales and I can select where my students are at and put the data in, but I can also insert videos or pictures then it summarizes it all. So I have that data all the time. I can always go to Miller Maps and look at it...A lot of times I'll just use the iPad and talk to the iPad about what the student did."

Participant 2: "Our school has this thing called Miller Maps that I use to monitor kids. Then I have Data Bingo where we have this big sheet up in the classroom with different note

cards where if a kid meets something we get to put it (on the sheet). So we can collect data before we input it (in Miller Maps) so we can visually see and keep track of where we're at. I (also) videotape my lessons and then I go through and code and pinpoint behaviors that students are doing to demonstrate what I see as engagement and seeing as meaningful participation. And then I code that over time to kind of like quantify a video and quantify data."

Participant 3: "At the end of class or end of the week we kind of fill in a little chart for each student who participates and engages in the lesson. Sometimes I refer back to my Google chats of what I've reported or jotted down in my planner."

Participant 4: "(We) go to (their) goals tab (that) I split up by quarter... I have goals for them to read and then they could check it off in their binders."

Participant 5: "I don't do a lot of keeping of data..."

Participant 5 relied on memory when sending updates via emails to parents, then relied on these emails when filling out formal documentation for their students. Lynn et al. also discovered the same insights that special education teachers predominantly rely on observational information to assess students' development [89]. Moreover, the multifaceted nature of goal tracking transcends academic progress, encompassing the overall health, emotional well-being, and social skills of students. Educators find themselves wrestling with tracking not only academic milestones but also non-verbal cues, alertness, and emotional engagement, illustrating the holistic dimension of goal tracking in special education.

Participant 1: "We have 3 nurses on staff and a lot of my students will have a private duty nurse come to school with them. I have kids that have seizures up to 30 times a day so just kind of monitoring overall health because we're constantly checking 'Do they have a fever?' or even basic bio-metrics, like fever, heart rate."

Participant 2: "I teach all of them...reading, writing, math, science, and social studies. We want kids to be engaged in 50 percent of the interaction (so) I look at if he's leaning forward on his tray, I know he's more engaged (or) if he's looking at his letters or at his writing artifact and if he's spending more time looking at the letters and writing artifact over time to demonstrate that he is increasing his engagement..."

Participant 4: "We have a pyramid of goals that we go from foundational to higher reaching goals, but I set more specific goals based on individual students... (and also track) how many times they got off task or weren't during their paper when they were supposed to."

Participant 5: "Everybody has a different speed of learning, so you have to track each student at their own speed. Then I also (see) how well they interact with their fellow students and they are not being included or they don't keep themselves in one place."

This finding aligns with other research indicating that special education teachers either use unstructured digital documents or large physical files containing vast amounts of information about academic and social progress, typically in the form of unstructured data - responses and observations on how the student is progressing [90]. As shown above, some participants already use existing systems to track goals. Therefore, seamless integration with existing student data systems used by teachers or school districts is imperative for efficient record-keeping. This integration ensures that the data collected through SAR interactions becomes an integral part of the broader educational framework. The findings emphasize the holistic nature of goal tracking, highlighting the need for SARs to accommodate different tracking methods, and preferences, and integrate with existing school district tools for optimal effectiveness.

In the collaborative landscape of education, the necessity for shared insights and data-driven decision-making becomes apparent. Participants emphasized the importance of collaboration with substitute teachers and paraprofessional educators. Paraprofessional educators serve as teacher assistants whose role has expanded to include formal assessments and initial instructions [91]. As participants noted, collaborative efforts extend to meetings where they and other educators collectively analyze challenges across different classes and subjects, emphasizing the scattered yet interconnected nature of the information.

Participant 1: "We (support staff) work as a team, so they come in 30 minutes early and go over what is going on at that time, including learning objectives...(and) I'll have them enter (data) into Miller Maps or they'll just tell me."

Participant 2: "We try to look at the data once a month during our PLCs (professional learning community), like my co-teacher and I will come together with our instructional coach and look at the data and then we have data walks as a classroom staff. We try to have that once a week to just kind of make sure we're on the same page. but that's if we have data."

Participant 3: "We (them and other teachers) meet and talk about our (students) and their foundational skills. And then if we're noticing something academically say, it's like we're noticing something in reading, then we ask all of the teachers at the table. 'Are you noticing the same struggle in your class?', What does it look like in social studies?', 'What does it look like in English?' and then that information is recorded in our meeting notes so it starts to kind of get like scattered."

Participant 4: "We do have team meetings once a week where we kind of talk, but it's not student by student. We have a platform where we can see other messages that were sent. We can see the grade books. There a lot of extra work to do to try and chart where they're doing well"

Educators need to generate effective documentation in real time as they observe and engage with students. Nonetheless, the significance of documentation extends beyond the immediate classroom, as it plays a crucial role in communicating with and gaining insights from other stakeholders as observed in Elbeiledy's work where documentation was found to be important to SARs teleoperators [22].

Furthermore, collaborative efforts extend beyond educators, involving a diverse team of licensed professionals. In special education, professional collaboration is viewed as a beneficial tool for helping teachers serve students with disabilities, with the idea that various professionals working together are tied to long-term success for students [92]. This collaborative ecosystem includes occupational therapists, physical therapists, speech-language pathologists, teachers of the deaf and hard of hearing, and teachers of the visually impaired. This collaborative model shows that tracking goals is not solely the responsibility of teachers; it involves a coordinated effort with various specialists.

Participant 2: "I collaborate with my occupational therapist on math because kids need help with access or speech will have their own goal. OT (occupational therapist) will have their own goal. Academics have reading, writing, and math then PT (physical therapist) has like a gross motor goal. So if we're working on a goal then I'll collect the bulk of the data just because I am with the students the most, but then they'll have their stuff as well."

Participant 5: "(I) won't be the one implementing (speech goals). The speech therapist would take the child away and perform that intervention."

Hence, any functionality incorporated into socially assistive robots for goal tracking should consider incorporating features that allow teachers to collaborate with fellow teachers and specialists when tracking goals for a student. As previously highlighted, a comprehensive session summary and progress reports stand out as essential tools not only in aiding the special educator's evaluation process but also in fostering greater involvement from students and parents in monitoring progress. However, as discussed, crafting a session summary or progress report poses challenges, given the individualized approaches teachers employ in tracking goals and the varied aspects they consider, spanning core curriculum, behavior, and social and health metrics.

6.3 Limitations and Future Work

This study aimed to gain insights into the goal-tracking practices of special educators, exploring their methodologies, measured metrics, and associated challenges. While the study yielded valuable information, it's noteworthy that participants lacked experience with socially assistive robots (SARs). Focusing on educators with SAR experience could potentially offer more nuanced recommendations specific to SAR design. Furthermore, the diversity in tracking methods observed among participants may be influenced by variations between different schools. Future research endeavors could consider developing visual decision support tools tailored to the specific context of individual schools, aiming to provide more targeted and context-specific recommendations.

6.4 Conclusion

Our first study shed light on the overarching needs of special education, therapists and administrators for a comprehensive summary and progress reports, offering broad recommendations encompassing multiple perspectives. This investigation highlighted the essential nature of efficient goal tracking specific to special education, fulfilling purposes of historical documentation and continuous assessment. However, it also brought to light the intricate challenge of designing a summary and progress report tool that could accommodate diverse tracking methods and the multifaceted nature of special education. The experiences of participants underscored the necessity for streamlined methods that enhance historical documentation, facilitate ongoing assessment, and ensure effective communication among all stakeholders - parents, students, teachers, and licensed professionals. The complexities go beyond individual academic goals, encompassing broader behavioral observations, social interactions, and health metrics. Therefore, any design feature targeting goal tracking in SARs must navigate this intricate landscape, recognizing and seamlessly integrating into the collaborative dynamics inherent in the special education environment.

CHAPTER 7

STUDY THREE METHODOLOGY

In our first experiment, we underscored the significance of session and progress reports, providing broad recommendations for a design that encompassed the needs of therapists, educators, and administrators. In our second study, we focused on special education teachers to provide tailored recommendations and better understand their unique needs. This second study reaffirmed the importance of evaluation tools like session and progress reports; but due to the complex and inter-connected nature of special education, coming up with one design would've ignored these intricacies and complexities that are vital in special education. The first two studies allowed us to investigate the evaluation tools shown to be important for SAR teleoperators [22]. It was also shown that evaluation is institutional and institutions have a role in providing evaluation tools. With this in mind, our third study focused on how the implementation of SARs can be improved from the perspective of special educators that can be used by institutions in improving their current process or in future deployments of SARs. While existing SAR literature reflecting the perspectives of special educators tends to underscore concerns about cost, safety, appearance, and the role of SARs, it often neglects their insights into effective implementation [15, 93, 94]. However, a crucial precondition before contemplating SAR enhancements such as session and progress reports is ensuring that special education teachers master the fundamental proficiency of using these technologies. Implementation strategies must harmonize with educators' educational goals, fostering a supportive environment for the sustained integration of technology [95]. Moreover, the significance of institutional support extends beyond evaluation into the realm of implementation.

To bridge the existing research-to-practice gap spotlight the specific needs of special education teachers in SAR research and provide recommendations for institutions such as school districts that can enhance current and future SAR deployments. The primary objective was to engage special education teachers with direct SAR experience, aiming to address the following research questions:

1. What challenges do special educators encounter when integrating socially assistive robots into their classrooms?
2. What opportunities exist for enhancing the implementation of SARs in special education?

7.1 Recruitment and Participants

In collaboration with a Colorado school district actively employing Misty and Peerbots in special education settings, I gained access to a pool of participants. Previous partnerships between the school

district and my research lab facilitated a seamless collaboration. Before initiating the interviews, we navigated through the school's Institutional Review Board (IRB) process to ensure ethical research practices. Initially, a school administrator acted as an intermediary, reaching out to the twelve identified special educators on my behalf. After a brief period, we were able to directly approach the educators. Thus far, six special educators have shared their insights. These educators have experience with students from kindergarten to twelfth grade, offering a diverse perspective across different educational levels. A summary of participant information is presented in the table below:

Table 7.1 Interview Participants

Pseudonym	Length of SAR Usage in Years	Years of Experience as a Special Educator
Participant 1	1	11-20
Participant 2	1	1-10
Participant 3	3	1-10
Participant 4	3	1-10
Participant 5	1	1-10
Participant 6	1	20-30

7.2 Semi-structured Interviews

We facilitated one-on-one interviews to gain comprehensive insights into each participant's experience with Misty and Peerbots.

Peerbots, an open source application, that offers an affordable solution for socially assistive robot teleoperation and has been employed in programs focusing on enhancing social skills for children with autism [23, 26]. Peerbots enable a teleoperator to oversee a robot's movement and verbal expression. Teleoperators can create content that the robot will articulate during a session and can incorporate relevant metadata for each verbalized item. Notably, therapists can define both the goal for the articulated content and the required proficiency level. This feature allows teleoperators to assess a client's performance using the information gathered during a session [23, 26].

7.3 Misty

Misty is a robotics platform, equipped with sensors and customizable features such as eyes, voice capabilities, and movements [96]. In our study, participants leveraged Peerbots to manage Misty's verbalization and eye color.

Utilizing Microsoft Teams, all interviews were systematically conducted and recorded to ensure accuracy and thorough documentation of responses. Employing a semi-structured interview format, the discussions centered around predetermined themes, offering a framework for exploration while allowing

flexibility for participants to delve into other pertinent areas. The specific interview questions guiding the conversations can be referenced in the Appendix (see **??**). Each interview session spanned a duration of 30 to 45 minutes, providing ample time for participants to share nuanced perspectives and contribute to the depth and richness of the gathered data.

7.4 Qualitative Analysis

Following the interviews, the transcripts underwent thorough analysis using Dovetail, a dedicated software for transcription analysis and coding. Employing a thematic analysis methodology, we systematically coded the interviews to unveil underlying patterns. The process initiated with the generation of initial codes, which were subsequently scrutinized to identify overarching themes. The initial analysis uncovered six distinct themes. Upon revisiting these themes, we identified two core themes that succinctly captured the essence of participants' experiences and insights.

1. A support system is critical in the effective implementation of socially assistive robots;
2. Offering incentives and various resources can motivate special educators to implement socially assistive robots.

CHAPTER 8
STUDY THREE RESULTS

Facilitating access to resources, streamlining existing materials, and ensuring ongoing support structures are essential measures to improve the implementation of SARs in special education classrooms. These insights offer valuable guidance for institutions, such as school districts, currently using or considering the deployment of SARs. Moreover, the findings underscore the significance of improving implementation strategies alongside potential enhancements for SARs.

8.1 A support system is critical in effective implementation of socially assistive robots

Research has identified challenges faced by educators in effectively utilizing assistive technology due to a lack of experience and knowledge in the process, emphasizing the need for training or professional development to bridge this gap and boost teacher confidence in implementing these technologies in classrooms [97–100]. Our participants' experiences align with these findings, as those who perceived ample support reported increased confidence in using SARs, and incorporating them into their teaching practices frequently.

One participant, who was part of the initial cohort of special education teachers using SARs, highlighted the resources and onboarding experience they received:

> *Participant 3:* "The first night was just an introduction to the robot. We got to explore it a bit, turn it on, and play around. After that, it was more content creation, and they (Peerbots representatives) helped us out with that. Once a month, they would come to the school, informing us about palettes or content the robots needed to talk, about and demonstrating the capabilities of the robot."

This resulted in them using SARs every week.

> *Participant 3:* "I'm comfortable with (SARs) now and especially having kind of got in on the ground floor when we adopted it (SARs). I (now) use (SARs) once a week ...for social, emotional, and communication goals for my students based on their individualized education plans. So I will (write content) based on the child's goal. So a lot of kids are practicing like answering personal questions or recognizing emotions in others. I have not used it (SARs) with the curriculum. It's all really specific to my kids."

Another participant who started using SARs more recently expressed confidence after just one session and felt assured that support was available if needed.

> *Participant 5:* "I pretty much (use SARs) daily because I use them for one-on-one instruction with students in literacy and math. And then we use (SARs) to have conversations like 'How are you feeling today?'"

.

However, teachers who felt unsupported reported demotivation and infrequent use of SARs compared to their more confident counterparts.

> *Participant 1:* "I didn't use it consistently. I wish I would have known more because we did ask the same questions. That still is beneficial to students who need that repetition. I just didn't get far enough into it where I knew exactly how to customize or add to it. I would use it again or recommend it to one of the teachers that I coach right now but I think there would need to be some sort of support for sure."

> *Participant 2:* "I have the program (Peerbots) on my iPad, and the person who was like my in-house helper no longer works at the school, so I'm kind of on my own and I don't know how to troubleshoot. I don't know how to get the technology to talk to each other. (The school district) had a group that met every Monday or once a month or something like that, that I was going to, which was helpful because then I had troubles I could bring them up and kids would help me work through that. But this year it's all different."

These varied experiences underscore how different onboarding processes impact teachers' confidence in learning and implementing SARs. Structured and consistent support during the initial stages is crucial for building a solid foundation, contributing to teachers' long-term success with SARs. Furthermore, maintaining a continuous support system is essential for sustained confidence and competence. As Silvera-Tawil et al also found, the effectiveness of SARs was dependent on the teacher's ability to program the robot [93].

Another support system that is deemed valuable is community building among special education teachers. A lack of knowledge and confidence about robotic technology is also found to be related to a lack of exposure [101, 102]. Creating a sense of community among special education teachers can help expose them to effective usage of SARs.

> *Participant 3:* "Get teachers in a classroom where it's being used, like the practical application of the robot, and see how they (kids) could respond to it. The first year we did it

(launched SARs), I had a child who had never said a complete sentence and was almost completely nonverbal. By the end of the year working with a robot, he was speaking in complete sentences."

Participant 5: "I knew that my colleague had gotten a lot of positive feedback from her students, and we tend to have a boy-heavy program, and she had a lot of boys and kind of around the same age that I did at that time. So I knew if it had worked for her, it would work here."

Participant 6: "I heard about it and thought it would be a cool thing to try. And then one of the teachers who has been using it a lot more consistently was talking about how she's had students (who started) communicating when they've interacted with the robot... It was (also) nice to collaborate with other teachers who have been using it and kind of share palettes and ideas with them."

Establishing knowledge-sharing platforms that facilitate educators in exchanging experiences, sharing best practices, and accessing valuable resources is crucial. This community-driven approach fosters a supportive network, emphasizing the importance of shared expertise in navigating challenges and optimizing the utilization of SARs in special education.

In conclusion, effective implementation of SARs in special education classrooms requires a comprehensive support system. Teachers who feel supported exhibit greater confidence and are more likely to integrate SARs into their teaching practices consistently. Institutions should prioritize providing a consistent onboarding experience and ongoing support to empower special educators in utilizing SARs successfully. Additionally, fostering a sense of community among special education teachers is essential for exposing them to effective SAR usage and creating a collaborative environment for knowledge exchange.

8.2 Offering incentives and various resources can motivate special educators to implement socially assistive robots

Motivating educators to embrace socially assistive robots (SARs) can be effectively achieved through strategic incentives and tailored professional development programs. Offering professional development workshops, coupled with opportunities to earn credits, proves instrumental in building teachers' knowledge and skills. The positive impact of such workshops on teacher confidence is well-documented in the literature [98, 101]. The development of teachers' knowledge through these workshops positively affected their confidence in teaching engaging robotics-based STEM activities as well [101].

Participant 4: "You could have a (professional learning community) that's like socially assistive robots with Peerbots. And this could be like a little cohort that we put together and we meet twice a month. And if we meet a certain amount of hours, then we get a certain amount of credits towards their salary...You can take professional development classes and they can help move you on the pay scale and they can be used for re-certification credits too."

The incorporation of professional development credits aligns with educators' personal and professional growth goals, enhancing the appeal of these workshops. Furthermore, the potential for creating a professional learning community focused on SARs, where participation is tied to credit incentives, offers a structured and collaborative approach to learning. In a study by Brownell et al. of special education teachers, their satisfaction with professional development opportunities was influenced by content, timing, quality, and incentives for participation [**?**]. In addition, several studies in special education indicate a connection between salary and teacher turnover [**?**]. Henke, Choy, Chen, Geis, and Alt (1997) emphasized the importance of compensation in teachers' decisions to stay in the profession or a specific district salary [103].

Another incentive is to cater training based on their attitudes and personality traits to effectively learn how to plan their curriculum while integrating robotic tools [94]. A study by Conti et al explored the influence of personality factors on the acceptability and intention to use SARs in teaching activities [94]. The study suggests that to effectively incorporate robotics in education, teachers should receive tailored training that considers their attitudes and personality traits, enabling them to integrate robotic tools into their educational activities successfully [94]. Addressing individual learning preferences is crucial, as some of our participants expressed a preference for hands-on, kinesthetic learning experiences. Tailoring training materials to accommodate various learning styles, such as step-by-step guidance and interactive sessions, can further optimize the effectiveness of professional development.

Participant 1: "If all the information wasn't kind of thrown at me, if you will at once, maybe more of a scaffolding like here follow these steps to get started and then to make more pallets or to customize it, try these steps. I just didn't have time to sit down and go through tutorials."

Participant 2: "They sent me some visual and written (documents) and I was like 'no, no, I need you to, I'm a doer, I'm a kinesthetic learner and I need you to do it or show me or verbally walk me through it so I know for the future'"

Participant 6: "It'd be cool to have someone come into the classroom that has the knowledge and help us at the moment because you know, it always works out where things are

working when you're figuring it out with the professional there and they're gone and it's not working... Specific training would be really helpful or some ideas on how to use it in the classroom more specifically or opportunities to see it in action."

This study emphasizes the need for continuous support throughout the implementation process. Creating a robust support system, including community building among special educators, facilitates knowledge-sharing and problem-solving. In addition, professional development workshops coupled with professional development credits not only increase motivation to increase knowledge of SARs but can also support retention by helping increase salaries for special education teachers which research has shown to increase teacher satisfaction and retention. Moreover, the findings suggest that incorporating personality-aware training can enhance the effectiveness of professional development. Recognizing and accommodating individual attitudes and traits contribute to more meaningful and targeted training experiences. This aligns with the idea that effective implementation of SARs requires tailored training that considers educators' unique characteristics.

The positive outcomes reported by participants, such as increased student engagement and communication skills, reinforce the potential benefits of SARs in special education. These outcomes contribute to building a compelling case for more effective implementation of SARs in educational settings.

Participant 2: "I've seen more success with kids who have those kinds of struggles talk to the robot rather than an adult. It's less threatening and they can be a bit more candid."

Participant 3: "The robot has been awesome because the robot doesn't have (a ton) of facial expressions. It doesn't have the nonverbal communication cues that people inherently have. You can try to be as neutral as you possibly can and you still have nonverbal cues like tone of voice or movement of body or eye. And so for students with autism who are developing communication, I find that this is so much information to process, and that hinders their ability to communicate authentically. With the robot, it strips its way down to the basic communication level and they're able to generalize a lot more. And then you can add on nonverbal (cues) and all of my students have been able to generalize those skills to human beings once they've mastered that with the robot."

8.3 Limitations and Future Work

While our interviews with special educators provided valuable insights into the implementation of socially assistive robots, it's essential to acknowledge the inherent limitations of my research. Firstly, the perspectives presented are exclusively those of special educators who have integrated SARs into their

classrooms at various times. To gain a more comprehensive understanding, future research should encompass diverse stakeholders, including administrators school districts, or developers/ designers of SARs. Their insights can shed light on potential gaps between teacher needs and the support provided, as well as effective ways to showcase the features of each technology. Secondly, the scope of our interviews is confined to a single school district. A comparative analysis, considering factors such as budgetary constraints, administrative limitations, and other top-down influences on SAR implementation, across diverse school districts, would provide more nuanced insights. Examining how different financial and administrative structures impact the implementation of SARs can enhance our understanding of the specific challenges and opportunities faced by each district. This approach aims to derive practical recommendations that go beyond mere location-based differences, offering a more comprehensive and insightful perspective on optimizing SAR integration in special education classrooms. Neglecting to take into account the perspectives of all stakeholders may result in unforeseen outcomes in the real-world implementation of SARs [104].

8.4 Conclusion

In summary, this study sheds light on the challenges and opportunities in implementing socially assistive robots (SARs) in special education classrooms. Conducted through insightful semi-structured interviews with special education teachers, the research underscores the value of a structured and ongoing support system. By fostering community building among special educators, this approach not only motivates but also ensures the consistent and effective utilization of SARs.

Our interviews emphasize the significance of continuous assistance, hands-on sessions, and readily available resources. Additionally, introducing incentives like professional development credits emerges as a pragmatic strategy to inspire and engage special educators. This not only fuels motivation but also deepens their understanding of SARs, aligning personal and professional interests.

These practical strategies collectively empower educators, enabling them to feel well-prepared, supported, and motivated to harness the full potential of SARs in real-world applications. By highlighting the pivotal role of special educators as key operators of this technology, the study contributes to SAR literature often overlooking the specific needs of those teleoperating SARs in practice. The findings underscore the critical importance of effective implementation for these technologies to be widely utilized and yield positive outcomes in the lives of children in special education.

CHAPTER 9

CONCLUSION

Socially assistive robots have demonstrated considerable success in therapeutic and educational contexts. In real-world applications, therapists and educators often teleoperate with these robots, yet current trends in Human-Robot Interaction research tend to overlook these teleoperators in favor of fully autonomous robots. Shifting our focus to teleoperated SARs introduces new stakeholders—teleoperators—and necessitates unique design considerations. These teleoperators, comprising therapists and educators conducting interventions, emphasize the importance of evaluation in their processes.

In the culmination of our three studies, a comprehensive perspective emerges on the importance of enhancing teleoperation interfaces while ensuring the effective implementation of socially assistive robots (SARs) in special education contexts.

The iterative design approach employed in Study One illuminated the pivotal role of session summary and progress reports as multifaceted tools for therapists, educators, and administrators. Beyond serving as educational and therapeutic aids, these reports contribute to the credibility of interventions that utilize SARs. The emphasis here is on striking a balance in visualization—making information digestible for a broad audience while providing enough details for teleoperators to make informed judgments such as engagement and proficiency metrics.

Building on this foundation, Study Two delved into the specific needs of special education teachers around evaluation tools such as session and progress reports. The findings underscore the complex landscape of special education making it a challenge to create a universal design. This complexity arises from the diverse tracking methods and multifaceted dimensions of success, including individual academic goals, behavioral observations, social interactions, and health metrics. Design features aimed at facilitating goal tracking in socially assistive robots (SARs) must navigate this intricate landscape and consider the collaborative and complex dynamics inherent in special education.

Study Three focuses on the challenges and opportunities for more effective implementation of SARs in special education classrooms. The study emphasizes the significance of a structured and ongoing support system, such as community building among special educators. The identified practical strategies, including continuous assistance, hands-on sessions, accessible resources, and incentivized professional development, collectively empower educators to enhance their understanding of SARs. This study underscores the specific needs of special educators in effective implementation, providing valuable considerations for institutions.

In summary, the collective findings from these studies emphasize the dual importance of refining teleoperation interfaces and ensuring effective implementation in the context of SARs in special education. The success of SARs is contingent not solely on technological advancements but also on strategies that empower and support the educators operating these technologies. These insights advocate for a holistic approach to the adoptability and integration of SARs in special education, focusing on the perspective of special educators. This approach ensures the widespread utilization of SARs and their positive impact on the lives of children in special education settings.

REFERENCES

[1] Bryan G Cook and Barbara R Schirmer. What is special about special education? overview and analysis. *The Journal of Special Education*, 37(3):200–205, 2003.

[2] David H Rose and Nicole Strangman. Universal design for learning: Meeting the challenge of individual learning differences through a neurocognitive perspective. *Universal access in the information society*, 5:381–391, 2007.

[3] Karrie A Shogren, Anthony J Plotner, Susan B Palmer, Michael L Wehmeyer, and Youngshil Paek. Impact of the self-determined learning model of instruction on teacher perceptions of student capacity and opportunity for self-determination. *Education and Training in Autism and Developmental Disabilities*, pages 440–448, 2014.

[4] Kathy Thompson Whaley and Teresa C Bennett. Promoting engagement in early childhood special education. *Teaching Exceptional Children*, 23(4):51–54, 1991.

[5] Carla Pedrosa, Helena Bilimória, Maria de Jesus Cepa, Maria de Deus Mestre, Maria de Fátima Pires, Marisa Ferreira, Natália Serra, and Paulo AS Moreira. Engagement with school in students with special educational needs. *International Journal of Psychology and Psychological Therapy*, 15(3):361–375, 2015.

[6] Caro C Williams-Pierce. Five key ingredients for improving student motivation. 2011.

[7] Jere E Brophy. *Motivating students to learn*. Routledge, 2013.

[8] Rolf Reber, Elizabeth A Canning, and Judith M Harackiewicz. Personalized education to increase interest. *Current directions in psychological science*, 27(6):449–454, 2018.

[9] James D Basham, Tracey E Hall, Richard A Carter Jr, and William M Stahl. An operationalized understanding of personalized learning. *Journal of Special Education Technology*, 31(3):126–136, 2016.

[10] Maja J Matarić and Brian Scassellati. Socially assistive robotics. *Springer handbook of robotics*, pages 1973–1994, 2016.

[11] Jillian Greczek, Elaine Short, Caitlyn E Clabaugh, Katelyn Swift-Spong, and Maja Mataric. Socially assistive robotics for personalized education for children. In *2014 AAAI Fall Symposium Series*, 2014.

[12] D. Feil-Seifer and M.J. Mataric. Defining socially assistive robotics. In *9th International Conference on Rehabilitation Robotics*, pages 465–468. IEEE, 2005. doi: 10.1109/ICORR.2005.1501143.

[13] Martin Saerbeck, Tom Schut, Christoph Bartneck, and Maddy D. Janse. Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1613–1622. Association for Computing Machinery, 2010.

[14] Goren Gordon and Cynthia Breazeal. Bayesian active learning-based robot tutor for children's word-reading skills. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2015.

[15] Irena Papadopoulos, Runa Lazzarino, Syed Miah, Tim Weaver, Bernadette Thomas, and Christina Koulouglioti. A systematic review of the literature regarding socially assistive robots in pre-tertiary education. *Computers & Education*, 155, 2020.

[16] Eija Kärnä, Kaisa Pihlainen, Erkki Sutinen, and Marjo Virnes. Can robots teach? preliminary results on educational robotics in special education. volume 2006 of *Conference: Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies*, pages 319–321. IEEE, 2006. doi: 10.1109/ICALT.2006.1652433.

[17] Brian Scassellati, Henny Admoni, and Maja Matarić. Robots for use in autism research. *Annual review of biomedical engineering*, 14, 2012.

[18] Saad Elbeleidy, Terran Mott, Dan Liu, and Tom Williams. Practical considerations for deploying robot teleoperation in therapy and telehealth. In *IEEE Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2022.

[19] Laurie A Dickstein-Fischer, Darlene E Crone-Todd, Ian M Chapman, Ayesha T Fathima, and Gregory S Fischer. Socially assistive robots: current status and future prospects for autism interventions. *Innovation and Entrepreneurship in Health*, 5, 2018.

[20] Adriana Tapus and Maja J Mataric. Socially assistive robots: The link between personality, empathy, physiological signals, and task performance. In *AAAI spring symposium: emotion, personality, and social behavior*, pages 133–140, 2008.

[21] Saad Elbeleidy, Terran Mott, and Tom Williams. Practical, ethical, and overlooked: Teleoperated socially assistive robots in the quest for autonomy. In *Companion Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (alt.HRI)*, page 577–587, 2022.

[22] Saad Elbeleidy, Terran Mott, Dan Liu, Ellen Do, Elizabeth Reddy, and Tom Williams. Beyond the session: Centering teleoperators in socially assistive robot-child interactions reveals the bigger picture. *Proc. ACM Hum.-Comput. Interact.*, 7(CSCW2), oct 2023. doi: 10.1145/3610175. URL https://doi.org/10.1145/3610175.

[23] Saad Elbeleidy, Aryaman Jadhav, Dan Liu, and Tom Williams. Robot teleoperation interfaces for customized therapy for autistic children. In *Companion proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, page 750–753, 2022.

[24] Saad Elbeleidy. Improving teleoperation interfaces to support therapists in robot-assisted therapy. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 2022. doi: 10.1145/3491101.3503804.

[25] Home. URL https://www.peerbots.org/.

[26] Saad Elbeleidy, Daniel Rosen, Dan Liu, Aubrey Shick, and Tom Williams. Analyzing teleoperation interface usage of robots in therapy for children with autism. In *ACM Conference on Interaction Design and Children*, pages 112–118, 2021.

[27] James M Kauffman. *Revitalizing special education: Revolution, devolution, and evolution*. Emerald Group Publishing, 2022.

[28] Mary Abbott, Cheryl Walton, Yolanda Tapia, and Charles R Greenwood. Research to practice: A "blueprint" for closing the gap in local schools. *Exceptional children*, 65(3):339–352, 1999.

[29] Paul H Lipkin, Jeffrey Okamoto, Council on Children with Disabilities, Council on School Health, Kenneth W Norwood Jr, Richard C Adams, Timothy J Brei, Robert T Burke, Beth Ellen Davis, Sandra L Friedman, Amy J Houtrow, et al. The individuals with disabilities education act (idea) for children with special educational needs. *Pediatrics*, 136(6):e1650–e1662, 2015.

[30] Marilyn S Kaff. Multitasking is multitaxing: Why special educators are leaving the field. *Preventing school failure*, 48(2):10, 2004.

[31] James McLeskey. Supporting improved practice for special education teachers. *Journal of Special Education Leadership*, 24(1):26–35, 2011.

[32] Vivian I Correa and Jennifer Y Wagner. Principals' roles in supporting the induction of special education teachers. *Journal of Special Education Leadership*, 24(1):17–25, 2011.

[33] Craig A Michaels and Jennifer McDermott. Assistive technology integration in special education teacher preparation: Program coordinators' perceptions of current attainment and importance. *Journal of special Education technology*, 18(3):29–44, 2003.

[34] Sandra Alper and Sahoby Raharinirina. Assistive technology for individuals with disabilities: A review and synthesis of the literature. *Journal of Special Education Technology*, 21(2):47–64, 2006.

[35] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. Social robots for education: A review. *Science Robotics*, 3, 2018.

[36] Ryan Van Patten, Amber V Keller, Jacqueline E Maye, Dilip V Jeste, Colin Depp, Laurel D Riek, and Elizabeth W Twamley. Home-based cognitively assistive robots: maximizing cognitive functioning and maintaining independence in older adults without dementia. *Clinical Interventions in Aging*, 15:1129–1139, 2020.

[37] Arielle AJ Scoglio, Erin D Reilly, Jay A Gorman, and Charles E Drebing. Use of social robots in mental health and well-being research: systematic review. *Journal of medical Internet research*, 21 (7), 2019.

[38] Catherine Plaisant, Allison Druin, Corinna Lathan, Kapil Dakhane, Kris Edwards, Jack Maxwell Vice, and Jaime Montemayor. A storytelling robot for pediatric rehabilitation. In *Proceedings of the Fourth International ACM Conference on Assistive Technologies*, pages 50–55. Association for Computing Machinery, 2000.

[39] Iain Werry, Kerstin Dautenhahn, Bernard Ogden, and William Harwin. Can social interaction skills be taught by a social agent? the role of a robotic mediator in autism therapy. *Cogn. Technol. Instrum. Mind*, pages 57–74, 2001. doi: 10.1007/3-540-44617-6_6.

[40] Eija Kärnä. Technology in finnish special education-toward inclusion and harmonized school days. *Informatics in Education*, 6(1):103–114, 2014.

[41] Zhen-Jia You, Chi-Yuh Shen, Chih-Wei Chang, Baw-Jhiune Liu, and Gwo-Dong Chen. A robot as a teaching assistant in an english class. In *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, pages 87–91, 2006. doi: 10.1109/ICALT.2006.1652373.

[42] Brian Scassellati, Laura Boccanfuso, Chien-Ming Huang, Marilena Mademtzi, Meiying Qin, Nicole Salomons, Pamela Ventola, and Frederick Shic. Improving social skills in children with asd using a long-term, in-home social robot. *Science Robotics*, 3(21), 2018. doi: 10.1126/scirobotics.aat7544.

[43] Kazuyoshi Wada, Takanori Shibata, Tomoko Saito, and Kazuo Tanie. Analysis of factors that bring mental effects to elderly people in robot assisted activity. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1152–1157, 2002.

[44] Saminda Sundeepa Balasuriya, Laurianne Sitbon, Margot Brereton, and Stewart Koplick. How can social robots spark collaboration and engagement among people with intellectual disability? In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, OZCHI'19, page 209–220. Association for Computing Machinery, 2019. doi: 10.1145/3369457.3370915.

[45] Ben Robins, Kerstin Dautenhahn, R Te Boekhorst, and Aude Billard. Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? *Universal access in the information society*, 4(2):105–120, 2005.

[46] Peggy Grant and Dale Basye. *Personalized learning: A guide for engaging students with technology.* International Society for Technology in Education, 2014.

[47] Caitlyn Clabaugh, Kartik Mahajan, Shomik Jain, Roxanna Pakkar, David Becerra, Zhonghao Shi, Eric Deng, Rhianna Lee, Gisele Ragusa, and Maja Matarić. Long-term personalization of an in-home socially assistive robot for children with autism spectrum disorders. *Frontiers in Robotics and AI*, 6, 2019.

[48] Alyssa M Alcorn, Eloise Ainger, Vicky Charisi, Stefania Mantinioti, Sunčica Petrović, Bob R Schadenberg, Teresa Tavassoli, and Elizabeth Pellicano. Educators' views on using humanoid robots with autistic learners in special education settings in england. *Frontiers in Robotics and AI*, 6, 2019.

[49] Connie Kasari, Alexandra Sturm, and Wendy Shih. Smarter approach to personalizing intervention for children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, 61 (11):2629–2640, 2018. doi: 10.1044/2018_JSLHR-L-RSAUT-18-0029. URL https://pubs.asha.org/doi/abs/10.1044/2018_JSLHR-L-RSAUT-18-0029.

[50] Connie Kasari and Tristram Smith. Interventions in schools for children with autism spectrum disorder: Methods and recommendations. *Autism*, 17(3):254–267, 2013.

[51] Aubyn C Stahmer, Sarah Rieth, Ember Lee, Erica M Reisinger, David S Mandell, and James E Connell. Training teachers to use evidence-based practices for autism: Examining procedural implementation fidelity. *Psychology in the Schools*, 52(2):181–195, 2015.

[52] Goren Gordon, Samuel Spaulding, Jacqueline Kory Westlund, Jin Joo Lee, Luke Plummer, Marayna Martinez, Madhurima Das, and Cynthia Breazeal. Affective personalization of a social robot tutor for children's second language skills. In *Proceedings of the AAAI conference on artificial intelligence (AAAI)*, volume 30, 2016.

[53] Danny Yen-Ting Liu, Kathryn Bartimote-Aufflick, Abelardo Pardo, and Adam J Bridgeman. *Data-Driven Personalization of Student Learning Support in Higher Education*, pages 143–169. Springer International Publishing, 2017.

[54] William H Blackwell and Zachary S Rossetti. The development of individualized education programs: Where have we been and where should we go now? *Sage Open*, 4(2):2158244014530411, 2014.

[55] Dorin Mihai Petrescu. *One size doesn't fit all: An inclusive art therapy approach for communication augmentation and emotion control in children with autism.* Ms thesis, OCAD University, 2013.

[56] Saad Elbeleidy. Towards effective robot-teleoperation in therapy for children with autism. In *ACM Conference on Interaction Design and Children (IDC)*, page 633–636. Association for Computing Machinery, 2021. doi: 10.1145/3459990.3463401.

[57] Caitlyn Clabaugh and Maja Matarić. Escaping oz: Autonomy in socially assistive robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, 2(1):33–61, 2019.

[58] Nikolai Alexander Grieg, Svein-Ivar Lillehaug, and Yngve Lamo. A visual analytics dashboard to support icbt therapists. In *Proceedings of the 17th Scandinavian Conference on Health Informatics (SHI)*, number 161, pages 134–140. Linköping University Electronic Press, 2019.

[59] Rahul C. Basole, Jukka Huhtamäki, Kaisa Still, and Martha G. Russell. Visual decision support for business ecosystem analysis. *Expert Systems with Applications*, 65:271–282, 2016.

[60] Edward R Tufte. The visual display of quantitative information. *The Journal for Healthcare Quality (JHQ)*, 7:15, 1985.

[61] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*, pages 364–371. Elsevier, 2003.

[62] C. North. Toward measuring visualization insight. *IEEE Computer Graphics and Applications*, 26(3): 6–9, 2006. doi: 10.1109/MCG.2006.70.

[63] Ketan K. Mane, Chris Bizon, Charles Schmitt, Phillips Owen, Bruce Burchett, Ricardo Pietrobon, and Kenneth Gersing. Visualdecisionlinc: A visual analytics approach for comparative effectiveness-based clinical decision support in psychiatry. *Journal of Biomedical Informatics*, 45(1): 101–106, 2012.

[64] Katrien Verbert, Erik Duval, Joris Klerkx, Sten Govaerts, and José Luis Santos. Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10):1500–1509, 2013.

[65] D.P. Bryant and B.R. Bryant. *Assistive Technology for People with Disabilities*. Allyn and Bacon, 2003. ISBN 9780205327157. URL https://books.google.com/books?id=b384AQAAIAAJ.

[66] Shomik Jain, Balasubramanian Thiagarajan, Zhonghao Shi, Caitlyn Clabaugh, and Maja J Matarić. Modeling engagement in long-term, in-home socially assistive robot interventions for children with autism spectrum disorders. *Science Robotics*, 5(39), 2020.

[67] Florence Martin and Doris U Bolliger. Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning*, 22(1):205–222, 2018.

[68] Yesom Yoo, Hyeyun Lee, Il-Hyun Jo, and Yeonjeong Park. Educational dashboards for smart learning: Review of case studies. In Guang Chen, Vive Kumar, Kinshuk, Ronghuai Huang, and Siu Cheung Kong, editors, *Emerging Issues in Smart Learning*, pages 145–155, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg. ISBN 978-3-662-44188-6.

[69] Vili Podgorelec and Saša Kuhar. Taking advantage of education data: Advanced data analysis and reporting in virtual learning environments. *Elektronika ir Elektrotechnika*, 114(8):111–116, 2011.

[70] Patricia M. Davies. On school educational technology leadership. *Management in Education*, 24(2): 55–61, 2010. doi: 10.1177/0892020610363089. URL https://doi.org/10.1177/0892020610363089.

[71] Mary Abbott, Cheryl Walton, Yolanda Tapia, and Charles R. Greenwood. Research to practice: A "blueprint" for closing the gap in local schools. *Exceptional Children*, 65(3):339–352, 1999. doi: 10.1177/001440299906500305. URL https://doi.org/10.1177/001440299906500305.

[72] Seyed Mohammad Sadegh Khaksar, Bret Slade, Jennifer Wallace, and Kaur Gurinder. Critical success factors for application of social robots in special developmental schools: Development, adoption and implementation. *International Journal of Educational Management*, 34(4):677–696, 2020.

[73] Ronald E Anderson and Sara L Dexter. School technology leadership: Incidence and impact. 2000.

[74] Eric W. Johnson, Luis A. Castillo, and Jay B. Brockman. Application of a markov model to the measurement, simulation, and diagnosis of an iterative design process. In *Proceedings of the 33rd Annual Design Automation Conference (DAC)*, page 185–188. Association for Computing Machinery, 1996.

[75] Jakob Nielsen. Iterative user-interface design. *Computer*, 26(11):32–41, 1993.

[76] Barbara DiCicco-Bloom and Benjamin F Crabtree. The qualitative research interview. *Medical education*, 40(4):314–321, 2006.

[77] Eli Blevis, Youn-kyung Lim, David Roedl, and Erik Stolterman. Using design critique as research to link sustainability and interactive technologies. In *International Conference on Online Communities and Social Computing*, pages 22–31, 2007.

[78] Christopher Frauenberger, Judith Good, Alyssa Alcorn, and Helen Pain. Conversing through and about technologies: Design critique as an opportunity to engage children with autism and broaden research(er) perspectives. *International Journal of Child-Computer Interaction*, page 38–49, 2013.

[79] Kurt Luther, Jari-Lee Tolentino, Wei Wu, Amy Pavel, Brian P. Bailey, Maneesh Agrawala, Björn Hartmann, and Steven P. Dow. Structuring, aggregating, and evaluating crowdsourced design critique. In *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW)*, page 473–485. Association for Computing Machinery, 2015.

[80] Victoria Clarke and Virginia Braun. *Thematic Analysis*, volume 3, pages 1947–1952. 01 2014. ISBN 978-1-4614-5582-0. doi: 10.1037/13620-004.

[81] Lynne Turner-Stokes. Goal attainment scaling (gas) in rehabilitation: a practical guide. *Clinical rehabilitation*, 23(4):362–370, 2009.

[82] Elizabeth Kim, Rhea Paul, Frederick Shic, and Brian Scassellati. Bridging the research gap: Making hri useful to individuals with autism. *Journal of Human-Robot Interaction*, 2012. doi: 10.5898/jhri.v1i1.25.

[83] Ifeanyi Glory Ndukwe and Ben Kei Daniel. Teaching analytics, value and tools for teacher data literacy: A systematic and tripartite approach. *International Journal of Educational Technology in Higher Education*, 17(1):1–31, 2020.

[84] Dawn Dowding, Jacqueline Merrill, and David Russell. Using feedback intervention theory to guide clinical dashboard design. In *AMIA Annual Symposium Proceedings*, volume 2018. American Medical Informatics Association, National Library of Medicine, 2018.

[85] Roland K Yoshida, Kathleen S Fenton, Martin J Kaufman, and James P Maxwell. Parental involvement in the special education pupil planning process: The school's perspective, 1978.

[86] Karen Czapanskiy. Special kids, special parents, special education. *University of Michigan journal of law reform*, (47.3):733–, 2014. ISSN 0363-602X.

[87] Samantha E Goldman and Meghan M Burke. The effectiveness of interventions to increase parent involvement in special education: A systematic literature review and meta-analysis. *Exceptionality*, 25(2):97–115, 2017.

[88] Dave L Edyburn. Rethinking assistive technology. *Special Education Technology Practice*, 5(4):16–23, 2004.

[89] Lynn S Fuchs et al. Special education practice in evaluating student progress toward goals. 1982.

[90] Ryan Thomas Burns. An application of the autism management platform to tracking student progress in the special education environment. 2015.

[91] Karen H. Jones and William N. Bender. Utilization of paraprofessionals in special education: A review of the literature. *Remedial and Special Education*, 14(1):7–14, 1993. doi: 10.1177/074193259301400103. URL https://doi.org/10.1177/074193259301400103.

[92] Stephen J Hernandez. Collaboration in special education: Its history, evolution, and critical factors necessary for successful implementation. *Online submission*, 3(6):480–498, 2013.

[93] David Silvera-Tawil and Christine Roberts-Yates. Socially-assistive robots to enhance learning for secondary students with intellectual disabilities and autism. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 838–843. IEEE, 2018.

[94] Daniela Conti, Elena Commodari, and Serafino Buono. Personality factors and acceptability of socially assistive robotics in teachers with and without specialized training for children with disability. *Life Span and Disability*, 20(2):251–272, 2017.

[95] Khe Foon Hew and Thomas Brush. Integrating technology into k-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational technology research and development*, 55:223–252, 2007.

[96] URL https://mistyrobotics.com/.

[97] Emily C. Bouck Sara Flanagan and Jennifer Richardson. Middle school special education teachers' perceptions and use of assistive technology in literacy instruction. *Assistive Technology*, 25(1):24–30, 2013. doi: 10.1080/10400435.2012.682697. URL https://doi.org/10.1080/10400435.2012.682697. PMID: 23527428.

[98] Laura M Desimone. Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational researcher*, 38(3):181–199, 2009.

[99] Ellen J Rohaan, Ruurd Taconis, and Wim MG Jochems. Analysing teacher knowledge for technology education in primary schools. *International Journal of Technology and Design Education*, 22: 271–280, 2012.

[100] Nico Verloop, Jan Van Driel, and Paulien Meijer. Teacher knowledge and the knowledge base of teaching. *International journal of educational research*, 35(5):441–461, 2001.

[101] Christina Chalmers and Peter Macbeth. Creating synergy in teacher education through robotics-based stem activities. In *Australian Teacher Education Association (ATEA) Conference*, volume 30, 2013.

[102] Campbell J McRobbie, Ian S Ginns, and Sarah J Stein. Preservice primary teachers' thinking about technology and technology education. *International Journal of Technology and Design Education*, 10: 81–101, 2000.

[103] RR Henke, SP Choy, X Chen, S Geis, and N Alt. Teachers in the 1990's: Profile of a profession (nces 97-460). *Berkeley, CA: Office of Educational Research and Improvement*, 1997.

[104] Katie Winkle, Praminda Caleb-Solly, Ailie Turton, and Paul Bremner. Mutual shaping in the design of socially assistive robots: a case study on social robots for therapy. *International Journal of Social Robotics*, 12:847–866, 2020.

APPENDIX A

STUDY ONE INTERVIEW QUESTIONS

- Can you tell me more about your background? Who do you mainly work with?
- Do you work with neuro-diverse children? What types?
- What kind of therapy do you do? Or what grades do you teach?
- How long have you been conducting therapy? Or teaching? With children?
- Do you work with groups of children?
- Are there other software applications you use? Which ones? Do they integrate with each other?

Preparation for a session

- Can you walk me through how you prep for a session?
- What tools (software, etc) do you use?
- What information do you need?
- What kinds of tools do you wish you had to make preparation easier/ faster?
- Do you prep any customization depending on the child / settings? I.e jotting down specific questions to ask?
- Do you use robots to teach?

During a session

- What is the structure of your session? Planned activities or more discussion-based?
- Have you ever used a robot when conducting therapy?
- If yes to the first question, how long? How often? In groups or?
- What kind of SARs?
- Do you use any tools during a session?
- What key information do you jot down during a session?

- What are the types of quantitative and qualitative data you track?
- Are there clinical guidelines you are required to report or follow?

After a session

- Can you walk me through your post-session workflow?
- Do you summarize sessions? If so, why? Why not? What is the goal?
- where do you summarize your notes, do you send any info to parents/ guardians/ etc?
- What do you wish you could do better in your post-session workflow?
- What trends in progress or client behavior do you watch out for in order to provide the right customization for patients?
- What types of feedback would be helpful post-session to
- Do you use an electronic health record? EHR are electronic version of patients medical history, maintained over time including personal info, medications, progress notes, etc
- Do you have reporting requirements? For example, from your employer or insurance in order to be reimbursed, etc?
- What's tracked within a session and tracked during the duration of your therapy with kids? therapists' documentation practices
- What information is useful to summarize?
- Would data visualizations help at all?
- How do these answers change in a 1:1 vs group setting?

If users have experience with socially assistive robots

- What has your experience with Peerbots or other SARs been like?
- Do those SARs have a summary report or data you can see? Do you use it? How?
- Are robots mainly used for engagement vs actual therapy?

APPENDIX B

STUDY ONE CRITIQUE QUESTIONS

- Does this visualization make sense? What do you think is being shown here?
- Do you find this visualization(s) useful when reflecting on a session? How else could we represent this information better?
- What other measurement would you like to be tracked and measured?

APPENDIX C

STUDY TWO AND THREE INTERVIEW QUESTIONS

Background:

- How many years have you been teaching?
- Teaching special education?
- As a special educator teacher, what types of neurodivergent students do you mainly work with?
- Which grades do you teach?
- Do you work with groups of children?

Engagement and motivation in special education classes

- What do you do to engage students? i. How do you measure this?
- What do you do to motivate students? i. How do you measure this?
- Challenges with engagement and motivation?
- How do your approaches differ with the type of neurodiversity?
- Personalization in special education classes

IEP - Individualized Education Program (IEP)

- How do you use IEP in your classes?
- Challenges with creating, filling it out, executing it for individual students?
- How does this differ with the type of neurodiversity?
- In general, how often does the student's neurodiversity type affect your personalization plans?
- What trends in progress or students do you watch out for in order to provide the right customization for students?

Cross-functional work

- Do you share metrics/ reports to other stakeholders - parents, school districts?

- What metrics do they care about?
- How often do you share it?
- Do you have reporting requirements? For example, from your school district?

SARs

- Current level of SARs knowledge?
- How might SARs be useful in supporting special education classes?
- Initial thoughts on the use of robots in special education?
- If SARs was to be deployed in your classroom, what features would it need in order for it to be useful to you?
- When using technology or trying out new activities in class - what is the hardest part about implementation?
- If there was a report that helped you track metrics/ summarize sessions to help with reflection and personalization, would that be beneficial or harmful?
- Would it be helpful in evaluating your clients?
- What would you like to see in such a report?
- Would you use the report when personalizing or filling out IEP?