Taylor & Francis
Taylor & Francis Group

International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

A Literature Review of Conversational Agents for People with Intellectual and Developmental Disabilities

Madhuka Nadeeshani, Jacqueline Johnstone, Kirsten Ellis & Swamy Ananthanarayan

To cite this article: Madhuka Nadeeshani, Jacqueline Johnstone, Kirsten Ellis & Swamy Ananthanarayan (09 Mar 2026): A Literature Review of Conversational Agents for People with Intellectual and Developmental Disabilities, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2026.2634975

To link to this article: https://doi.org/10.1080/10447318.2026.2634975

Published online: 09 Mar 2026.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2026.2634975

A Literature Review of Conversational Agents for People with Intellectual and Developmental Disabilities

Madhuka Nadeeshani, Jacqueline Johnstone, Kirsten Ellis and Swamy Ananthanarayan

Department of Human-Centred Computing, Monash University, Australia

ABSTRACT
Voice-based conversational agents (CAs) are increasingly pervasive in daily life, offering hands-free interaction with technology. Recent advancements have expanded their adoption among underserved populations, including older adults, blind individuals, and those with intellectual and developmental disabilities (IDD). Despite growing interest, the IDD literature on CAs remains scattered and lacks a systematic synthesis. In this paper, we present a review of 37 studies selected from 1,151 records retrieved across five databases. We analyse application areas, design and evaluation approaches, reported limitations, and key design considerations that influence CA usability for people with IDD. Our findings reveal the complexity of designing for a population with varied support needs, and highlight the value of competency-based design methodologies. We also identify gaps in underexplored domains such as health, education, and employment, and emphasize the need to strengthen safeguards that enable safe, transparent, and trustworthy CA use.

KEYWORDS
Conversational agents; intellectual and developmental disabilities; assistive technology; Human-Computer interaction; accessibility

1. Introduction

Conversational Agents (CAs) enable users to interact with smart devices through natural language, providing seamless access to information and services (Hobert & Von Wolff, 2019). These agents exist in various forms, including stand-alone smart speakers (e.g., Amazon Echo,[1] Google Home[2]), smartphone-embedded assistants (e.g., Siri,[3] Google Assistant[4]) and different customized agents tailored for specific applications (Even et al., 2022). They are widely used by people to perform daily activities such as setting timers, establishing routines, streaming multimedia, and home automation tasks (Brill et al., 2019). Increasingly, CAs have been designed to accommodate the needs of underserved populations such as older adults (Huang et al., 2025; Simpson et al., 2020; Zubatiy et al., 2021), blind and low vision people (Choi et al., 2020; Pradhan et al., 2018; Reinders et al., 2023), and people with Intellectual and Developmental Disabilities (IDD) (Balasuriya et al., 2018; Lewis & Vellino, 2021; Mower, Black, et al., 2011). CAs have shown great potential in supporting people with IDD by offering an accessible and intuitive way to interact with technology. They provide hands-free access, facilitate natural and socially engaging conversations while improving safety by reducing communication barriers (Balasuriya et al., 2018; Cho, 2019; Fogg, 1998). Beyond facilitating communication, these agents have the potential to enhance independence and overall quality of life for people with IDD (Volochtchuk et al., 2023). Given that 30% of people with IDD rely on care partners (Bruce et al., 2008), there is a growing need for technologies that foster autonomy and reduce dependency. Voice user interfaces provided by CAs offer an alternative means of accessing technology, fostering greater autonomy and enhancing overall well-being for people with IDD.

CONTACT Madhuka Nadeeshani madhuka.koralalagedona@monash.edu Department of Human-Centred Computing, Monash University, Melbourne, Victoria, Australia

According to the Diagnostic and Statistical Manual of Mental Disorders – 5 (DSM-5) (American Psychiatric Association, 2013), IDD is defined as a group of disorders characterized by intellectual limitations and deficits in adaptive behavior. These limitations impact daily functioning, including communication, cognitive functioning, education, and social interactions (Tassé et al., n.d). IDD encompasses a variety of conditions, such as intellectual disabilities, autism spectrum disorder, cognitive disorders, and developmental disorders, which can manifest in different ways. It is further classified into four severity levels: mild, moderate, severe, and profound, where the severity determines the type and extent of support required, with higher levels necessitating greater assistance (Carulla et al., 2011). Exploring the intersection of CAs and IDD is both timely and crucial, given the growing interest in how these technologies can support people with IDD. Several studies have investigated the use of CAs for people with IDD, yet a significant gap remains in understanding the comprehensive impact of these technologies. Specifically, there is a lack of systematic analysis regarding how CAs are used by people with IDD in different contexts and the unique challenges they encounter.

To the best of our knowledge, only two existing systematic reviews (Syed Mahmudul et al., 2022; Volochtchuk et al., 2023) have indirectly explored the interaction between CAs and individuals with IDD. Volochtchuk et al. (2023) reviewed the use of voice-assisted technology among individuals with different disabilities, including developmental disabilities, brain injuries, physical disabilities, and visual impairments. Their analysis of 68 studies identified seven key themes related to the benefits and challenges of these technologies. However, only seven studies focused on individuals with developmental disabilities, while the majority addressed other types of disabilities. Additionally, their review was restricted to peer reviewed journal publications, which may have excluded relevant studies from conferences. Another review (Syed Mahmudul et al., 2022), examined the use of CAs in healthcare for people with cognitive disabilities, including dementia, Parkinson's disease and elderly people. This review examined both text based and voice based CAs, though most studies primarily focused on text based interactions. Out of the 124 articles analyzed, only 12 studies specifically examined the use of voice based CAs for individuals with cognitive disabilities. Our review focuses specifically on the use of voice based CAs for individuals with IDD, offering a more targeted scope compared to previous reviews that considered a range of disabilities. Additionally, we examine the diverse applications of CAs and methodologies from a Human-Computer Interaction (HCI) perspective. Our contributions include: (i) Analysis of the diverse application areas of CAs for people with IDD across different domains. (ii) Analysis of contribution types design and evaluation methods when working with IDD population (iii) Identification of design considerations to improve the usability of CAs based on insights from existing research studies. (iv) Synthesis of reported limitations across the included studies.

2. Methodology

2.1. Defining terms and database search

To identify relevant literature, we developed a search strategy focused on two core concepts IDD and CAs. For IDD, we included terms such as "intellectual disabilit*", "cognitive disabilit*", "developmental disabilit*", and "intellectual and developmental disabilit*", using wildcard operators to capture variations in phrasing. These terms reflect the spectrum of conditions commonly classified under the IDD umbrella and align with terminology used in the assistive technology. For CAs, our search incorporated both general descriptors and commercial product names, including "virtual assistant*", "voice assistant*", "smart speaker*", "voice-assisted technolog*", "Siri", "Alexa", "Google Home", and "Amazon Echo". A complete list of search terms for both core concepts is presented in Table 1. Terms within each concept were joined using "OR", and the two concept groups were connected using "AND" operator.

The literature search was conducted across five major databases: Scopus, IEEE Xplore, ACM Digital Library, Web of Science, and PubMed. The initial search was conducted in the last quarter of 2023 and yielded 1,083 records. No restrictions were placed on publication year to ensure comprehensive coverage. A follow-up search was conducted at the end of the first quarter of 2025 to capture more recent publications, identifying an additional 68 records. In total, the combined search returned 1,151 results: 703

Table 1. Search terms used in review.

Concept	Search terms
Intellectual and Developmental Disability	"Intellectual disabilit*" OR "intellectual disorder*" OR "cognitive disabilit*" OR "cognitive disorder*" OR "developmental disabilit*" OR "developmental disorder*" OR "intellectual and developmental disabilit*" OR "intellectual and developmental disorder* "
Conversational Agent	"Virtual assistant*" OR "voice assistant*" OR "smart speaker*" OR "voice assisted technolog*" OR "siri" OR "alexa" OR "google home" OR "home pod" OR "amazon echo" OR "google assistant" OR "conversational agent*"

Table 2. Inclusion and exclusion criteria for study selection.

Dimension	Inclusion criteria	Exclusion criteria
Population	Participants with IDD	Participants without IDD (e.g., dementia, learning disabilities, physical disabilities)
Interaction Modality	Voice based CAs (spoken input and/or output)	Text based CAs
Study Design	Design, development, or evaluation of CAs involving IDD participants	No participant involvement or purely technical implementation
Publication Type	Peer reviewed journal articles, conference papers, short papers, and review articles	Abstracts, books, posters, or tutorials
Language	English	Other languages

from Scopus, 308 from the ACM Digital Library, 96 from IEEE Xplore, 24 from Web of Science, and 20 from PubMed.

2.2. Inclusion and exclusion criteria

The research team collaboratively established a set of inclusion and exclusion criteria based on five key dimensions: target population, interaction modality, study design, publication type, and language (Table 2). The search initially captured papers referencing IDD, autism spectrum disorder, Down syndrome, learning disabilities (e.g., dyslexia, dysgraphia, dyscalculia), dementia, cognitive disabilities, and neurodevelopmental disorders. However, we excluded papers on dementia and learning disabilities because they do not align with standard definitions of IDD (CDC Child Development, 2024; Eunice Keneddy Shriver, n.d.). Since IDD commonly co-occurs with motor, linguistic, or sensory conditions, studies involving these conditions were included only when they were explicitly described as part of the users' IDD profile. Only studies that examined voice based modalities were included, while those focusing exclusively on text-based interfaces were excluded. Papers were also required to involve individuals with IDD at any stage of the design, development, testing, or evaluation of CAs, reflecting an emphasis on participatory and user-centered research. Only peer reviewed journal articles, conference papers, short papers, and review articles written in English were considered, while abstracts, books, posters, tutorials, and non-English publications were omitted.

2.3. Screening and study selection

The screening process was conducted in Covidence,[5] a systematic literature review management platform, which identified 63 duplicate records and yielded 1,088 unique papers. Title and abstract screening was then performed independently by two authors using the inclusion and exclusion criteria described above, categorizing each paper as "yes", "no", or "maybe". Papers marked as "maybe", along with any conflicts, were discussed during weekly meetings and resolved collaboratively among all authors. Following the title and abstract screening, 98 papers were selected for full-text review. The same two authors independently assessed these full texts using the same criteria and procedure. Final disagreements were resolved in consultation with all authors, resulting in 33 papers being included in the final review. The screening process followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Sarkis-Onofre et al., 2021) method, and the overall selection flow is illustrated in Figure 1.

To expand the initial set of selected papers, we applied the snowballing method (Wohlin, 2014), conducting both forward and backward searches. Specifically, we reviewed the reference lists of the included papers (backward snowballing) and identified newer studies that cited them (forward

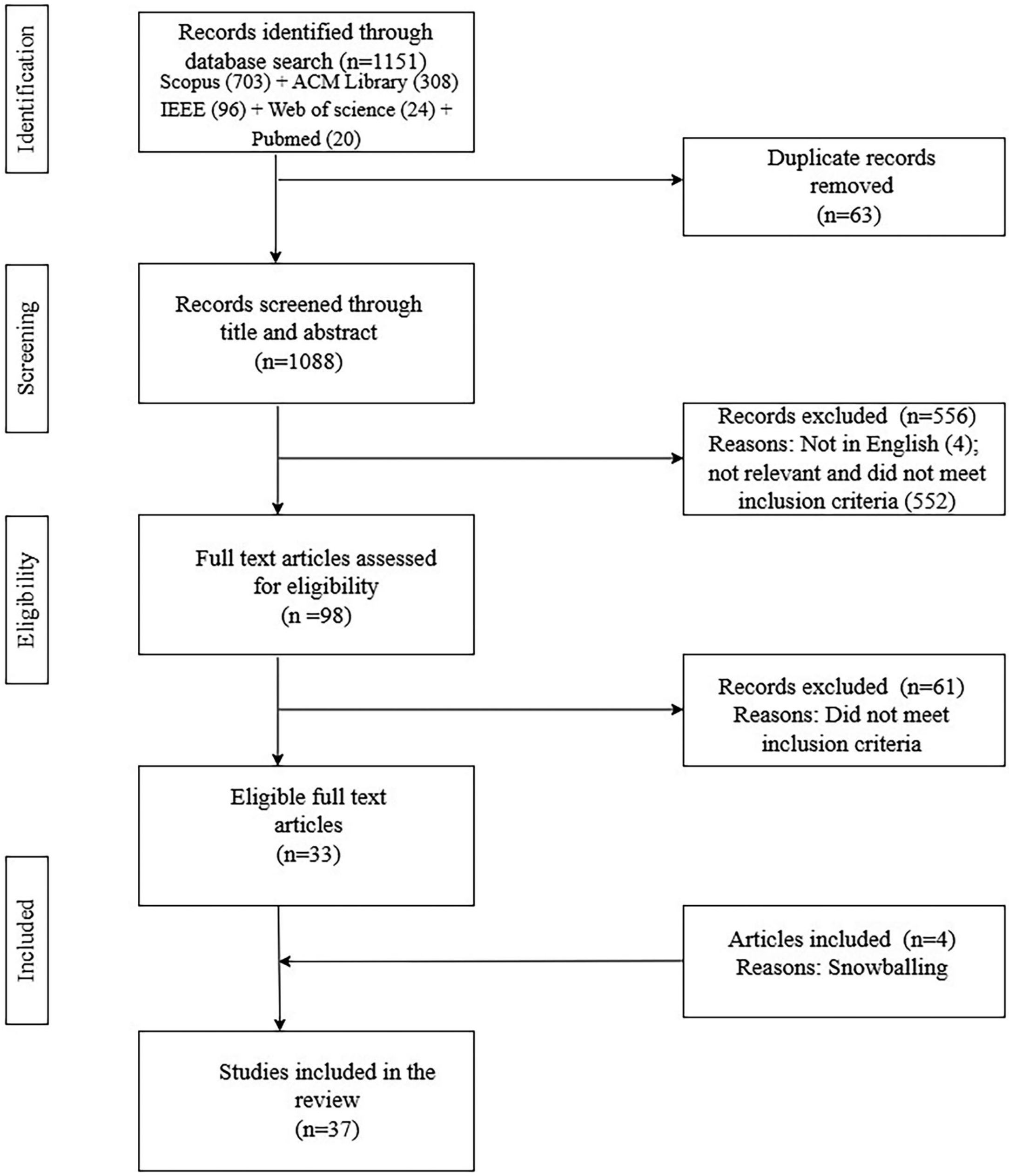

Figure 1. Search process and outcome flowchart.

snowballing). The inclusion and exclusion criteria described above were then reapplied to assess the relevance of these additional records. This process yielded four more papers that met the criteria, resulting in a final corpus of 37 papers.

2.4. Data extraction and verification

To support systematic and consistent data extraction, we developed a structured spreadsheet with predefined column headers, refined collaboratively through team meetings. The spreadsheet was structured to capture key attributes including article title, publication venue, year, design methods, evaluation methods, number of participants, study setting, key findings, application areas, limitations, design considerations, contribution types, and identified challenges. One author performed the initial data extraction for all included studies, and the entries were subsequently reviewed and discussed with the research team during weekly meetings. Any discrepancies or ambiguities were resolved in consultation with all authors, and adjustments were made to maintain accuracy and consistency. The finalized spreadsheet formed the basis for the analysis and synthesis presented in the following sections.

3. Descriptive analysis of the corpus

This section is based on the final set of 37 analyzed articles and outlines key categories, such as publication trends and corpus characteristics, types of CAs, the spectrum and severity levels of IDD, age group distributions, and the different categories of research studies.

3.1. Publication trends and corpus characteristics

Analysis of publication trends revealed a steady increase in research on CAs for individuals with IDD over the past decade (Figure 2). A clear upward trajectory is evident from 2019 onward, reflecting broader advances in CAs and artificial intelligence since the release of Amazon Alexa (2014) and Google Assistant (2016). The number of publications continued to grow through 2023, peaking at nine studies, before a slight decline in 2024. This decline may reflect ongoing publication cycles rather than a decrease in research activity, given that the final search was conducted in the first quarter of 2025. Among the included studies, 24 (65%) were published in peer reviewed journals and 13 (35%) in conference proceedings. The most frequent journals were Disability and Rehabilitation: Assistive Technology and the Journal of Medical Internet Research, while major conference venues included ACM Conference on Human Factors in Computing Systems and International Conference on Human Computer interaction.

Of the 37 papers, six studies (16%) were systematic or literature reviews, while the remaining 31 (84%) focused on the design, development, or evaluation of CAs. The majority of studies (n = 30) reported on unique CA systems described in a single publication, while a small subset of systems appeared across multiple papers, reflecting continued development or longitudinal evaluation efforts. For example, Emoty system (Catania et al., 2019, 2020; Catania & Garzotto, 2023) was investigated in three papers (8%), the Rachel system (Mower, Black, et al., 2011; Mower, Lee, et al., 2011) in two papers (5%), and a transport application developed by Garzotto et al. (Gianotti et al., 2023a, 2023b) in two papers (5%). These recurring systems allowed exploration across multiple stages of design, development, and evaluation, offering longitudinal insights into how CAs evolve through successive iterations. This overall pattern reflects a maturing research landscape and provides the foundation for examining the different types of CAs and their design characteristics, discussed in the following subsection.

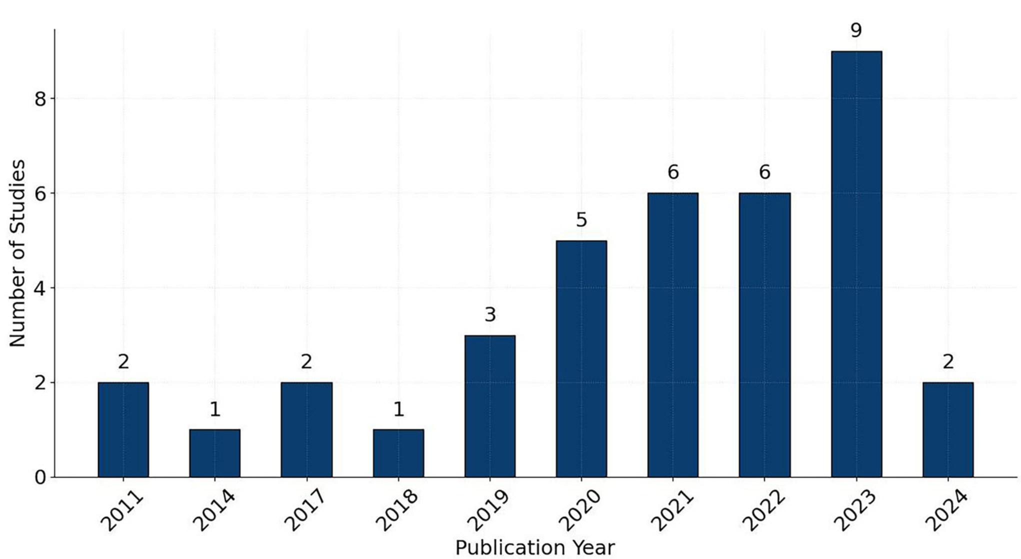

Figure 2. Distribution of reviewed studies by publication year.

3.2. Types of CAs

Based on their functionality and interaction design, the CAs identified in the reviewed studies were grouped into four categories: off-the-shelf CAs (e.g., Google Home, Alexa, Siri), customized CAs developed for specific user needs, virtual agents featuring on-screen animated characters, and voice based chatbots designed for task-specific interactions. All of these CAs enable users to communicate through voice user interfaces. Off-the-shelf CAs are commercially available devices that offer general purpose features, allowing straightforward voice interactions. In contrast, customized CAs are purpose built to address particular user requirements, offering personalized functionality for users. Virtual agents incorporate on-screen animated characters to simulate human like interaction, often combining voice and visual elements to create a more engaging and immersive experience. Voice based chatbots, on the other hand, are typically designed for specific tasks or domains and rely primarily on spoken exchanges.

We did not analyze communication modalities, such as one-way or two-way exchanges, as our primary focus was on exploring voice-based CAs. A total of 13 studies focused on customized CAs (Allen et al., 2018; Arya et al., 2023; Bakhai et al., 2020; Catania et al., 2019, 2020; Catania & Garzotto, 2023; El Rhatassi et al., 2023; Gianotti et al., 2023a; Greuter et al., 2022; Papadogiorgaki et al., 2023; Parvin et al., 2022; Soleiman et al., 2014; Zhang et al., 2021), while seven studies examined off-the-shelf CAs (Balasuriya et al., 2018; Lewis & Vellino, 2021; Masina et al., 2020; Smith et al., 2021, 2023; van Wingerden et al., 2023; Zubatiy et al., 2021). Google Home was the most used CA, appearing in six studies. In addition, eight studies developed customized virtual CAs featuring on-screen animated characters to enhance user interaction (Ali et al., 2020; Aljameel et al., 2019; Garcia-Pi et al., 2023; Gianotti et al., 2023a; Mower, Black, et al., 2011; Mower, Lee, et al., 2011; Tanaka et al., 2017; Tarpin-Bernard et al., 2021), while others focused on voice based chatbots (El Rhatassi et al., 2023). Moreover, some studies focused on identifying user needs and gathering requirements for future voice based systems to enhance user interactions with CAs (Koushik & Kane, 2023; Spitale et al., 2020). By examining these various types of CAs, we aimed to provide a comprehensive understanding of their capabilities and practical implications for improving interactions for individuals with IDD.

3.3. Spectrum and severity levels of IDD

We identified various terms used by researchers to describe IDD in the analyzed papers, including intellectual disability, autism spectrum disorder, neurodevelopmental disorders, and cognitive disorders (Table 3). Additionally, these conditions sometimes co-occurred with other disabilities, such as motor, mental, physical, linguistic, emotional, and sensory disabilities (Bérubé et al., 2021; El Rhatassi et al., 2023; Masina et al., 2020; Tavares et al., 2022; Volochtchuk et al., 2023). We found that autism spectrum disorder emerged as the primary focus in 14 studies, while intellectual disabilities were analyzed in nine studies. Both cognitive disabilities and neurodevelopmental disabilities were addressed in seven studies each, showing equal representation from the article set. Overall, the diversity of disabilities within the IDD spectrum highlights the need for tailored interventions and support systems to address the unique challenges faced by individuals with these conditions.

Although IDD is typically classified into four severity levels: mild, moderate, severe, and profound (Carulla et al., 2011) the studies we reviewed generally categorized disabilities into only three levels: mild, moderate, and severe (Table 4). Those with mild IDD require minimal support and can often complete tasks independently. Individuals with moderate IDD need significant assistance from caregivers for daily activities. In contrast, those with severe IDD require continuous, intensive support in all aspects of life, typically provided by caregivers. A limited number of studies have addressed interventions across all three severity levels (Catania et al., 2020; Catania & Garzotto, 2023; Papadogiorgaki et al., 2023; Parvin et al., 2022). Most studies focused primarily on mild and moderate levels (Ali et al., 2020; Catania & Garzotto, 2023; Garcia-Pi et al., 2023; Greuter et al., 2022; Smith et al., 2021; Smith et al., 2023; Soleiman et al., 2014; van Wingerden et al., 2023), while a significant number did not specify the severity levels of users (Allen et al., 2018; Arya et al., 2023; Balasuriya et al., 2018; Cha et al., 2021; Masina et al., 2020; Mower, Black, et al., 2011; Mower, Lee, et al., 2011; Spitale et al., 2020;

Table 3. Range of disabilities in IDD.

Disability type	Count of papers	References
Autism Spectrum Disorder	14	(Ali et al., 2020; Aljameel et al., 2019; Allen et al., 2018; Bakhai et al., 2020; Cha et al., 2021; Gianotti et al., 2023a, 2023b; Mower, Black, et al., 2011; Mower, Lee, et al., 2011; Parvin et al., 2022; Soleiman et al., 2014; Spitale et al., 2020; Tanaka et al., 2017; Zhang et al., 2021)
Intellectual Disability	9	(Arya et al., 2023; Balasuriya et al., 2018; Bérubé et al., 2021; Even et al., 2022; Greuter et al., 2022; Papadogiorgaki et al., 2023; Smith et al., 2021, 2023; van Wingerden et al., 2023)
Neurodevelopmental Disorder	7	(Catania et al., 2019, 2020, 2023; Catania & Garzotto, 2023; El Rhatassi et al., 2023; Garcia-Pi et al., 2023; Volochtchuk et al., 2023)
Cognitive Disability	7	(Koushik & Kane, 2023; Masina et al., 2020; Syed Mahmudul et al., 2022; Tarpin-Bernard et al., 2021; Tavares et al., 2022; Volochtchuk et al., 2023; Zubatiy et al., 2021)

Table 4. Severity levels of the disabilities.

Disability type	Mild	Moderate	Severe	Not specified
Autism Spectrum Disorder	(Ali et al., 2020; Gianotti et al., 2023a, 2023b; Parvin et al., 2022; Soleiman et al., 2014)	(Gianotti et al., 2023a, 2023b; Parvin et al., 2022)	(Parvin et al., 2022)	(Allen et al., 2018; Cha et al., 2021; Mower, Black, et al., 2011; Mower, Lee, et al., 2011; Spitale et al., 2020; Tanaka et al., 2017; Zhang et al., 2021)
Intellectual Disability	(Papadogiorgaki et al., 2023; Smith et al., 2021, 2023; van Wingerden et al., 2023)	(Greuter et al., 2022; Papadogiorgaki et al., 2023; Smith et al., 2021, 2023; van Wingerden et al., 2023)	(Papadogiorgaki et al., 2023)	(Arya et al., 2023; Balasuriya et al., 2018)
Neurodevelopmental Disorder	(Catania et al., 2019, 2020; Catania & Garzotto, 2023)	(Catania et al., 2019, 2020; Catania & Garzotto, 2023)	(Catania et al., 2020; Catania & Garzotto, 2023)	–
Cognitive Disability	(Zubatiy et al., 2021)	–	–	(Koushik & Kane, 2023; Lewis & Vellino, 2021; Masina et al., 2020)

Tanaka et al., 2017; Tavares et al., 2022; Zhang et al., 2021). In addition, some studies were excluded from the table because, although designed for the target population, they neither involved users in requirements gathering nor evaluated the system with participants.

3.4. Age group distribution

We categorized the reviewed studies into three broader age groups for clarity (Table 5). The majority of the studies (17 total) concentrated on adults (18–64 years), demonstrating a prominent research focus on this age group. Research involving children (0–17 years) was limited, with only ten studies. Notably, there were only three studies focusing on older adults (65+ years), revealing a significant research gap and highlighting the need for further exploration in this area. In contrast, research on CAs for older adults without IDD has grown substantially (Even et al., 2022).

3.5. Data analysis

We conducted a five stage analysis to synthesize the corpus and organize the findings, comprising: (i) study type and application domain (subsection 4.1); (ii) contribution types, design methods, and evaluation methods (subsection 4.2); (iii) synthesis of cross-study design considerations (subsection 4.3); (iv) limitations reported in the studies (subsection 4.4); and (v) quality appraisal (Section A). Across stages, categories and labels were developed iteratively from recurring patterns in the extraction spreadsheet (subsection 2.4), rather than applying a pre-defined codebook. The first author conducted the initial clustering and coding, and the second author independently reviewed the codes and categories across all analysis sections described below. Any disagreements or ambiguous cases were resolved through

Table 5. Age group distribution of studies.

Age group	Count of papers	References
Adults (18–64 years)	17	(Bakhai et al., 2020; Balasuriya et al., 2018; Catania et al., 2020; Catania & Garzotto, 2023; Garcia-Pi et al., 2023; Gianotti et al., 2023a, 2023b; Greuter et al., 2022; Koushik & Kane, 2023; Lewis & Vellino, 2021; Masina et al., 2020; Papadogiorgaki et al., 2023; Smith et al., 2021, 2023; Tanaka et al., 2017; van Wingerden et al., 2023), (Cha et al., 2021)*
Children (0–17 years)	10	(Ali et al., 2020; Aljameel et al., 2019; Allen et al., 2018; Mower, Black, et al., 2011; Mower, Lee, et al., 2011; Parvin et al., 2022; Soleiman et al., 2014; Spitale et al., 2020; Zhang et al., 2021), (Cha et al., 2021)*
Older Adults (65+ years)	3	(Smith et al., 2023; van Wingerden et al., 2023; Zubatiy et al., 2021)

*Represents studies considering multiple age groups.

discussion to reach consensus, and the full research team then reviewed the final categorization for consistency. Although the corpus included six literature reviews (Bérubé et al., 2021; Catania et al., 2023; Even et al., 2022; Syed Mahmudul et al., 2022; Tavares et al., 2022; Volochtchuk et al., 2023), our analysis focuses on the 31 primary studies because the review papers synthesized prior work and did not report study level CA designs, deployments, or user-study procedures.

3.5.1. Study type and application domain

We analyzed each study along two dimensions: study type (i.e., what the study primarily focused on in CA for IDD) and application domain (i.e., what the CA was primarily intended to support) (Table 6). For study type, we reviewed the extracted spreadsheet data, identified recurring patterns in study focus, and grouped studies into three categories: understanding user needs (examining experiences, needs, barriers, or preferences of people with IDD), prototype development and evaluation (describing a CA design/implementation and reporting user-facing testing), and performance evaluation (focusing primarily on quantitative system or interaction measures such as recognition accuracy or task completion). For the application domain dimension, we categorized each study by identifying the CAs primary intended purpose as described by the authors. This yielded five high level domains: social interactions, daily activities, health, education, and employment. We further specified subdomains (e.g., emotional communication training, general activity support, digital therapies) to differentiate distinct forms of support within each domain.

3.5.2. Contribution types and method mapping

To characterize the types of contributions in prior work and how they were developed and evaluated, we coded (i) each study's primary contribution type and (ii) the design and evaluation methods it reported. We coded each study's primary contribution type using the HCI contribution classification framework proposed by Wobbrock and Kientz (2016). Using this framework, we categorized studies as artifact driven (focus on a CA system, tool, or prototype), empirical (reporting findings from user studies or field deployments), methodological (proposing a new method or analytic approach), or survey (synthesising prior work). No included studies fell into the theoretical, dataset, or opinion categories. We then categorized whether each study explicitly reported specific design methods (co-design, interviews, focus groups, questionnaires, other) and evaluation methods (interviews, observations, questionnaires, pilot studies, surveys, quantitative metrics). A method was categorized as present only if it was explicitly reported as an activity undertaken by the authors. When reporting detail was insufficient to identify a structured design process, we categorized the methodology as other (e.g., off-the-shelf deployments, adaptations of existing systems, or methodologically unclear approaches). These binary codes were compiled into a study by method matrix (Table 7), which we used to derive the descriptive distributions reported in findings and to inspect common method combinations across studies.

3.5.3. Cross-study design considerations

To derive the cross study design considerations summarized in Figure 3, we used a conceptual mapping approach consistent with HCI literature (Long & Magerko, 2020), in which recurring design relevant insights are distilled and organized into a framework. For each study, we extracted discrete design insight statements into the spreadsheet, including (i) reported interaction challenges and breakdowns, (ii) design strategies used to address them, and (iii) author stated recommendations or implications for

Table 6. Overview of study type and application domain.

Application area	Sub domain	Articles per category		
		Understanding users	Prototype development and evaluation	Performance evaluation
Social Interactions	Verbal Communication Trainer	(Spitale et al., 2020)	(Ali et al., 2020; Mower, Black, et al., 2011; Soleiman et al., 2014)	(Mower, Lee, et al., 2011; Tanaka et al., 2017)
	Emotional Communication Trainer	(Catania et al., 2020)	(Catania et al., 2019; Catania & Garzotto, 2023)	–
	Games	–	(Greuter et al., 2022; Zhang et al., 2021)	–
Daily Activities	General Activities	(Balasuriya et al., 2018; Cha et al., 2021; Koushik & Kane, 2023; Lewis & Vellino, 2021; Smith et al., 2021; Zubatiy et al., 2021)	(Allen et al., 2018; Arya et al., 2023; Papadogiorgaki et al., 2023)	(Masina et al., 2020)
	Transportation	–	(Gianotti et al., 2023a, 2023b)	–
	Agency and Well-being.	(Smith et al., 2023; van Wingerden et al., 2023)	–	–
Health	Oral Care	–	(Parvin et al., 2022)	–
	Digital Therapies	–	(Tarpin–Bernard et al., 2021)	–
	Medical Assistance	–	(El Rhatassi et al., 2023)	–
Employment	Job Interviews	–	(Garcia–Pi et al., 2023)	–
Education	Higher Education	–	(Bakhai et al., 2020)	–
	STEAM	–	(Aljameel et al., 2019)	–

Table 7. Matrix of contribution type, design methods, and evaluation methods across included studies. A filled square (■) indicates method usage.

	Contribution type			Design methods					Evaluation methods						
Paper	Emp.	Art.	Meth.	Co.	Int.	Focus.	Quest.	Other.	Int.	Obs.	Questi.	Pilot.	Survey	Quant.	Other.
Ali et al., 2020		■						■	■				■		
Papadogiorgaki et al., 2023		■			■		■		■		■	■			
Gianotti et al., 2023a		■		■						■	■				
Balasuriya et al., 2018	■							■	■	■					
Cha et al., 2021	■			■	■				■						
Greuter et al., 2022		■		■					■						
Lewis & Vellino, 2021	■							■	■						
Zhang et al., 2021		■						■					■	■	
Parvin et al., 2022		■			■				■		■	■			
Bakhai et al., 2020		■				■					■				
Garcia-Pi et al., 2023		■		■								■			
Aljameel et al., 2019		■			■						■			■	
Zubatiy et al., 2021	■							■	■				■		
Mower, Black, et al., 2011		■						■				■			
Mower, Lee, et al., 2011		■						■						■	
Catania et al., 2019		■				■				■				■	
Catania & Garzotto, 2023		■				■				■				■	
Catania et al., 2020	■					■				■					
Koushik & Kane, 2023	■				■				■	■					
Tarpin-Bernard et al., 2021		■						■							■
El Rhatassi et al., 2023		■						■							■
Masina et al., 2020	■							■	■					■	
Smith et al., 2021	■							■						■	
Smith et al., 2023	■							■	■	■			■	■	
van Wingerden et al., 2023	■							■		■	■			■	
Tanaka et al., 2017		■						■						■	
Soleiman et al., 2014		■						■		■					
Allen et al., 2018		■						■						■	
Spitale et al., 2020			■					■							■
Gianotti et al., 2023b		■		■						■		■			
Arya et al., 2023		■						■				■		■	

Column abbreviations: Emp.=Empirical, Art.=Artifact, Meth.=Methodological, Co.=Co-design, Int.=Interviews, Focus.=Focus groups, Quest.=Questionnaires, Other.=Other activities, Obs.=Observations, Pilot.=Pilot studies, Survey.=Surveys, Quant.=Quantitative evaluations.

Figure 3. Conceptual framework showing the five synthesized design considerations and the associated design strategies identified across the reviewed studies.

future CA design. We then conducted an iterative clustering process: insights were grouped by conceptual similarity, clusters were merged or split where needed, and cluster labels were refined through repeated comparison across studies to ensure that categories were internally coherent and distinct. The resulting structure was reviewed through author team discussion until consensus was reached on a stable set of categories. This synthesis yielded five recurring, design considerations that organize our findings: personalization and user adaptation; multimodal and multisensory support; reducing cognitive load through interaction simplicity; proactive guidance and communication repair; and supporting engagement and motivation.

3.5.4. Reported study limitations

After extracting author reported limitations from each study into the structured spreadsheet, we inductively grouped similar limitation statements into higher level categories. We then iteratively clustered limitations by conceptual similarity, compared clusters across studies to merge overlapping categories and refine boundaries, and agreed on category labels that captured recurring constraints. The resulting structure was reviewed through research team discussion until consensus was reached. This synthesis produced four recurring limitation areas that organize our findings: sample and representativeness constraints; methodological constraints (e.g., short duration, single session evaluations, controlled settings); communication and interaction challenges (e.g., recognition errors, pacing, wake word recall); and ethical constraints (e.g., restrictions on log retention, shared device risks).

3.5.5. Quality appraisal

We conducted a quality appraisal using concepts from the Critical Appraisal Skills Programme (CASP) checklists (Critical Appraisal Skills Programme, 2018) and the Mixed Methods Appraisal Tool (MMAT)

(Nha Hong et al., 2018), given the heterogeneous mix of qualitative, quantitative, and mixed-methods designs in the corpus. Rather than using these criteria to exclude studies, we used them to guide interpretation when methodological reporting was limited (e.g., clarity of procedures, appropriateness of participants, transparency of data collection and analysis, and ethical reporting). This appraisal helped us calibrate the strength of claims reported in the findings when reporting quality varied across studies. Full appraisal criteria and per study summaries are provided in Appendix A.

4. Findings

In this section, we analyze the key research categories, beginning with the study type and application domains of CAs, followed by an examination of contribution types and design and evaluation methods, cross-study design considerations and reported study limitations. Each application area opens with an overview and is followed by an analysis of relevant studies organized into three categories: understanding user needs, prototype development, and performance evaluation. The section concludes with a summary of the main findings and insights emerging from each subdomain.

4.1. Study type and application domain classification

4.1.1. Social interactions

4.1.1.1. Verbal communication trainer. Individuals with IDD often face challenges in social interactions due to limited exposure to interpersonal communication, which can lead to social exclusion, loneliness, and potential health issues (Abbott & McConkey, 2006; Walsh et al., 2019). The social interactions category is divided into three subdomains: verbal communication trainer, emotional communication trainer, and games. The verbal communication trainer aims to improve speech skills, whereas the emotional communication trainer focuses on supporting individuals to recognize and express emotions in communication. Meanwhile, games provide an engaging, low-pressure environment for practising communication and promoting social interaction. Together, these subdomains highlight the role of CAs in promoting social engagement and supporting various aspects of interpersonal communication.

A multi-criterion decision analysis (Spitale et al., 2020) was conducted to explore the needs and preferences of children with autism spectrum disorder for different types of CAs, including social robots, human-moderated interactions, virtual agents, and off-the-shelf CAs such as Google Home and Amazon Echo. Their findings revealed that virtual agents were the most preferred, as they helped reduce feelings of alienation, enhance relational skills, and improve therapy engagement.

In the category of prototype development and evaluation, three studies focused on verbal communication training. Two of these studies assessed the effectiveness of virtual agents: LISSA (Ali et al., 2020) and Rachel (Mower, Black, et al., 2011), while another study evaluated a custom parrot themed physical CA (Robo parrot) (Soleiman et al., 2014). LISSA facilitated informal dialogues with children with high functioning autism spectrum disorder, providing feedback on nonverbal cues such as eye contact, smile, volume, and body movements. In contrast, Rachel guided the children through structured activities, including emotional problem-solving tasks and storytelling exercises, while analyzing multimodal data such as audio, video, and physiological signals like skin temperature and heart rate. The Robo Parrot, designed for children with mild to very mild autism (Soleiman et al., 2014), features sensors, motors, and microphones that enable it to engage in unstructured conversations and respond dynamically to user interactions through a Wizard-of-Oz approach. Findings from Robo Parrot suggest that although some children were initially hesitant to engage, their interaction levels increased significantly over time with parental presence.

Two studies (Mower, Lee, et al., 2011; Tanaka et al., 2017) quantitatively analyzed speech characteristics in children with autism spectrum disorder. One study used the previously developed virtual agent Rachel (Mower, Black, et al., 2011) to examine differences in speech patterns by comparing children's interactions with the Rachel system to their interactions with their parents during verbal activities. The results indicated that the interactions with Rachel closely mirrored the natural communication patterns between parents and children, suggesting that parental participation did not significantly influence the interaction. In another study (Tanaka et al., 2017), a virtual agent was developed to assist children with autism spectrum disorder in participating in role-play scenarios on specific topics. The children received feedback on their verbal performance as well as non-verbal behaviors, such as head movements and smiling frequency. These features were quantitatively analyzed and presented to users to enhance their social interactions.

The studies highlighted that virtual agents were more effective than voice-only CAs for verbal communication training in children with autism spectrum disorder, as their multimodal nature enhanced engagement and enthusiasm (Ali et al., 2020; Mower, Black, et al., 2011; Mower, Lee, et al., 2011; Spitale et al., 2020). Caregivers and participants also preferred virtual agents over human-moderated interactions, as they provided a safe and non-intimidating environment (Ali et al., 2020; Spitale et al., 2020). Two studies (Mower, Black, et al., 2011; Soleiman et al., 2014) explored the impact of parental involvement on children's engagement but reported contrasting findings. This suggested that virtual agents were generally more user-friendly than physical CAs, as they required less parental support. However, the impact of this difference should be explored in future studies. Moreover, studies (Ali et al., 2020; Mower, Lee, et al., 2011) emphasized the importance of real-time positive reinforcement in encouraging behavioral adjustments and highlighted the need for a wider range of conversational topics beyond predefined scenarios to better accommodate individual preferences (Ali et al., 2020; Mower, Black, et al., 2011). Studies in this category typically involved small, relatively homogeneous samples of individuals with autism spectrum disorder and reported mostly single-session evaluations.

4.1.1.2. Emotional communication trainer. Emotional communication training is essential for fostering empathy, building relationships, and enhancing social understanding (Hostyn & Maes, 2009). However, individuals with IDD often face challenges in expressing emotions, which can affect their social interactions. To address this, three studies (Catania et al., 2019, 2020; Catania & Garzotto, 2023) by Catania et al. examined Emoty, a web-based customized CA designed to help adults with neurodevelopmental disorders (mild to severe) to express emotions through voice.

The study Catania et al. (2020) on understanding user needs explored participants' preferences regarding Emoty's anthropomorphic traits, including its use of an Italian feminine voice by Google, during extended interactions. The findings revealed Emoty's dual nature: while certain aspects, such as its precise performance and the comfort users experienced while interacting with it, led to a perception of it as more machine-like, other features, such as its human-like feminine voice and the comfort users felt sharing personal thoughts, contributed to a perception of it as more human-like.

In prototype development and evaluation studies (Catania et al., 2019; Catania & Garzotto, 2023), participants were prompted to read a sentence and express a corresponding emotion (e.g., happy, sad, angry, fearful, surprised) using their voice, as if acting out the emotion. Emotion probabilities were then determined using the IBM Watson Tone Analyzer, and real-time feedback was provided based on these probabilities. Furthermore, Emoty provided visual feedback on its status such as idle, listening, speaking, or pending, helping users manage the interaction and better understand the system's responses. Catania et al. extended their study (Catania & Garzotto, 2023) by introducing a graphical interface for caregivers in Emoty, enabling them to classify emotions in users' utterances. This modification aimed to address the inaccuracies of the model, which failed to accurately recognize the emotions of the participants in some cases. These two studies (Catania et al., 2019; Catania & Garzotto, 2023) highlighted that although participants initially struggled, they showed significant improvement in emotional communication with Emoty over time, requiring progressively less assistance from caregivers, suggesting a positive role for Emoty in enhancing emotional communication.

The studies mentioned above suggest (Catania et al., 2019, 2020; Catania & Garzotto, 2023) that web-based CAs with visual elements, such as emojis, were more effective than voice-only modalities in emotional communication training, even for adult users, as they helped maintain interaction and engagement. Further, studies (Catania et al., 2019; Catania & Garzotto, 2023) highlighted the importance of implementing real-time feedback and conducting longer-term evaluations in future research. Long-term evaluations are particularly crucial, as they allow individuals with neurodevelopmental disorders the time needed to adapt to technology effectively. Additionally, studies (Catania et al., 2019;

Catania & Garzotto, 2023) pointed out the inconvenience of requiring users to click a button each time they wanted to speak; however, with advances in voice assisted technology, this issue has been addressed by modern CAs using wake words to activate the system. As these results were generated within structured, task based settings, they offer limited insight into how emotional communication would unfold in less constrained, everyday environments.

4.1.1.3. Games. Games play a crucial role in supporting individuals with IDD by promoting social interactions, enhancing cognitive skills, and providing an engaging learning environment (Tsikinas & Xinogalos, 2019). To examine the impact of CAs in collaborative gameplay, two studies (Greuter et al., 2022; Zhang et al., 2021) developed and evaluated prototypes to explore user engagement through games. Social Adventure (Greuter et al., 2022), a narrative-based game, allowed adults with mild intellectual disabilities to give verbal instructions through customized CA by pronouncing letters or forming sentences. The other study involving the virtual agent ICON2 (Zhang et al., 2021), focused on collaborative problem-solving, where children with autism spectrum disorder worked with the CA to move puzzle pieces and form shapes. The findings from the Social Adventure study showed that 88% of users successfully completed all stories, indicating high engagement levels. Similarly, children interacting with ICON2 rated their enjoyment on a 1–5 Likert scale, with an average score of 4.4, reflecting overall satisfaction.

These findings suggest the potential of collaborative gameplay in fostering communication and social interactions. Both studies utilized periodic voice commands to encourage interactions, while one study incorporated hand movements with a virtual agent to enhance engagement and create a more immersive experience for children. However, some participants struggled with the audio-only narrative (Greuter et al., 2022), highlighting the limitations of relying solely on auditory input. These findings indicate that virtual agents, which incorporate multiple interaction methods, may be more user-friendly than voice-only CAs. However, further research is needed to determine whether this approach sustains long-term involvement, particularly among adult users in the gaming domain.

4.1.2. Daily activities

Daily activities are essential for individuals with IDD, as they promote independence and enhance life skills (Pearce, 2013). In our analysis, we found that daily activities span multiple domains, including general activities, transportation, agency and well-being, and verbal communication. This area has been widely studied for people with IDD, as engaging in daily activities enhances their quality of life, improves communication, and supports overall well-being, making it a crucial aspect of their development and independence.

4.1.2.1. General activities. Six studies examined general activities to understand the needs of people with IDD, exploring how individuals engage with tasks such as searching for information, checking weather updates, setting reminders, and following routines. Among them, five studies investigated the utilization of off-the-shelf CAs such as Google Home and Amazon Echo (Balasuriya et al., 2018; Cha et al., 2021; Lewis & Vellino, 2021; Smith et al., 2021; Zubatiy et al., 2021), while another study gathered future requirements for CAs through interviews (Koushik & Kane, 2023). These investigations were conducted across varying environmental settings, encompassing both disability care environments (Balasuriya et al., 2018; Cha et al., 2021) and participants' home environments (Koushik & Kane, 2023; Lewis & Vellino, 2021; Smith et al., 2021; Zubatiy et al., 2021).

Balasuriya et al. (2018) conducted a study in a disability care environment, exploring how adults with intellectual disabilities interacted with off-the-shelf CAs, specifically Siri or Google Assistant. Similarly, Cha et al. investigated the efficacy of a Korean CA (NUGU Candle), in supporting adolescents with autism spectrum disorder in their daily activities (Cha et al., 2021). Balasuriya et al. noted that 72% of the participants preferred using voice interactions over typing, and half of them successfully completed the four tasks assigned to them. Despite the benefits, both studies identified several challenges with CAs, including difficulties in pronouncing the wake words, the need for extra time to formulate questions, stammering and difficulty with specific words. The majority of participants preferred breaking complex queries into smaller segments to prevent issues with communicating longer sentences. Furthermore, participants who spoke softly, even without speech impediments, had difficulty being understood, leading to incorrect search queries. Cha et al. reported that some participants even experienced frustrations and yelled at the device due to pronunciation issues such as, stammering and difficulty with specific words.

Four studies were conducted in home settings (Koushik & Kane, 2023; Lewis & Vellino, 2021; Smith et al., 2021; Zubatiy et al., 2021); among them Lewis and Vellino (2021) investigated the interaction between CA and adults with cognitive disabilities for one month. In their study, participants were provided with Amazon Echo Dot devices, and feedback was collected from both participants and caregivers. While the feedback was generally positive, some participants expressed concerns about the lack of Alexa skills in French. Zubatiy et al. (2021) examined the interaction between ten older adults with mild cognitive impairment and their caregivers using a Google Home device over ten weeks, reporting an average of six daily interactions per dyad. A semi-randomized controlled trial was conducted with an intervention group of adults with ID using Amazon Alexa and Google Home (Smith et al., 2021). The participants' speech was recorded both prior to and following a 12-week period, during which they were instructed to say phrases related to the smart speaker (e.g., "turn on the light") as well as unrelated words (e.g., "hat"). The findings indicated a significant improvement in speech intelligibility ratings, with approximately 80% of users and caregivers reporting increased enjoyment and a greater sense of independence. However, the three studies noted challenges with participants remembering wake words and connecting the CA to other smart devices. Additionally, Zubatiy et al. found that some participants had difficulty recalling the smart speaker's presence and expressed doubts about the device's reliability. Another study Koushik and Kane (2023), gathered information on future design suggestions for context-aware CAs tailored to support individuals with cognitive impairments in their daily activities. Remote interviews with participants and their caregivers revealed desired features, including task structuring, motivation enhancement, and community support integration, to be incorporated into future CAs.

In the area of prototype development and evaluation, Papadogiorgaki et al. (2023) developed an integrated customized CA, consisting of both a web-based and a mobile application, designed to support adults with intellectual disabilities in performing general activities. These activities included accessing information such as encyclopedia topics, weather forecasts, and entertainment through a CA. The study showed that the system improved independence in participants with mild intellectual disabilities, enabling them to complete tasks without external help. However, users with severe intellectual disabilities encountered difficulties with voice recognition, primarily due to unclear speech, which necessitated caregiver intervention to formulate questions effectively. Similarly, Arya et al. (2023) developed a customized CA that featured separate applications for children and parents. The children's app included an adaptive microphone that adjusted to speech patterns, enabling activities such as daily habit tasks and follow schedules set by their parents. The parental app, on the other hand, provided real-time alerts for safety concerns, such as falls or unusual behavior. Notably, the system achieved an 80% task completion rate, demonstrating its effectiveness. The study Allen et al. (2018), examined whether the Amazon Echo could function as a speaker-independent device to retrieve visual scene cues (static images and videos) to help children with autism follow spoken directives (e.g., "Put the boy behind the lamp"). The Echo processed voice commands and retrieved cues from the Symboltalk app, displaying them on an iPad. The initial accuracy was low (8.6%), improved to 45.7% (p. 73) after customizing the vocabulary, but remained inconsistent (0%−44.4%) in clinical settings due to background noise and unintentional vocalizations.

Masina et al. (2020) quantitatively examined the impact of cognitive and linguistic functions on the interaction of adults with cognitive disabilities when using Google Home. Participants initially completed tasks (e.g., changing colors of lamps, interacting with the TV (e.g., YouTube, Spotify, and Netflix, setting alarms) with a CA in a simulated environment, after which their vocal intensities were analyzed using regression models. The study revealed that while participants encountered difficulties with phrasing, speech rate, comprehension, and pronunciation, those with a certain level of linguistic ability were still able to communicate effectively with the CA.

From the above studies, it was concluded that CAs played a supportive role in supporting general activities and enhancing the independence of individuals with IDD. Contrary to previous suggestions by prior studies, the multimodal interactive nature of CAs was found to be less critical in daily activities, as most studies focused on adults and involved relatively simple tasks (Balasuriya et al., 2018; Lewis & Vellino, 2021; Masina et al., 2020; Zubatiy et al., 2021). Furthermore, studies highlighted the benefits of using CAs in home environments, where participants engaged in more frequent interactions and performed a variety of tasks that extended beyond basic activities, such as connecting to smart devices (Lewis & Vellino, 2021; Zubatiy et al., 2021). Unlike other subdomains, caregivers played an influential role in helping with some tasks, such as supporting certain activities and responding to emergencies (Arya et al., 2023; Papadogiorgaki et al., 2023). The study that involved individuals with cognitive disabilities and their caregivers in all phases concluded that CAs fostered greater independence for both groups (Zubatiy et al., 2021). The findings suggested that CAs should support multiple speakers by integrating caregivers more actively and assessing the broader impact of their role in interacting with the CA.

4.1.2.2. Transportation. Transportation is crucial in facilitating daily activities, yet there is a significant gap in research related to this domain. Gianotti et al. (2023a) developed a customized CA to assist young adults with high-functioning autism spectrum disorder in navigating train transportation in real world settings. This application combined virtual reality, augmented reality, and a CA into a mobile application, augmented by wearable devices. The CA offered real time navigational guidance, assisted users in navigating paths, and sent proactive alerts if users failed to interact. In emergencies, the CA could contact caregivers, ensuring timely support and enhancing travel safety.

In a follow-up study (Gianotti et al., 2023b), the authors enhanced their original system by introducing two methods for triggering CA responses. These included a user initiated mode, where the agent responds to direct input or queries, and a proactive mode, where the agent triggers responses based on detecting specific behavioral cues. The study concluded that the proactive mode was particularly effective for individuals with autism spectrum disorder, improving navigation and overall interaction with the system. However, some participants felt that the CA lacked a personal touch and was uninviting, highlighting the need to improve its ability to engage users. This highlights a wider limitation in the system, indicating that its responses lacked the level of personalization and social engagement needed by users.

The findings indicated that combining CAs with virtual reality and augmented reality enhances user interactions in transportation by providing essential visual and auditory cues for navigation. It was suggested that CAs should be more interactive, with a focus on their appearance. Additionally, participants preferred text messaging over voice responses in noisy, crowded environments, likely due to social acceptability concerns. This emphasizes the limitations of voice-based CAs and the need for adaptable interaction strategies based on the context.

4.1.2.3. Agency and well-being. Agency and well-being are essential for supporting individuals with IDD, as they promote autonomy, confidence, and engagement (Mpofu et al., 2020). Two mixed-method studies (Smith et al., 2023; van Wingerden et al., 2023) utilized the Warwick-Edinburgh Mental Well-being Scale to assess the well-being needs of participants while performing activities (e.g., entertainment, asking help). Smith et al. (2023) conducted a 12 week study involving 22 adults with intellectual disabilities. They also noted that ethical and privacy constraints restricted access to detailed device logs, which limited the behavioral data for analysis. Wingerden et al. (2023) studied CA usage patterns over 20 weeks during the COVID − 19 pandemic among individuals with visual impairments and/or mild to moderate intellectual disabilities. Both studies identified challenges, including issues with speech intelligibility, difficulty remembering phrases, problems accessing device features and unintentionally external activations.

The findings suggested that off-the-shelf CAs are a cost-effective and easily deployable solution, making them suitable for large-scale trials. However, their lack of customization for individuals with IDD can lead to challenges, particularly in terms of speech intelligibility. As a result, longer-term evaluations are needed to better understand how users adapt and to identify strategies for improving interactions over time.

4.1.3. Health

Despite experiencing higher rates of physical and mental health challenges, individuals with IDD often have limited access to healthcare needs (Reppermund et al., 2019). This makes it crucial to leverage computer-assisted interventions to raise awareness and address these healthcare disparities. This review focuses on research across various health domains, including oral care, therapy sessions, and medical assistance.

4.1.3.1. Oral care. While several interventions have been developed to improve oral care in clinical settings, there is a reported gap in support for oral care at home. The Alexa skill prototype, called MyDentist (Parvin et al., 2022) was designed to assist children with autism spectrum disorder to perform oral care routines at home over a three week period. The skill was developed to assist with the transition to routine oral care activities by serving as both a motivator and a timer. To enhance engagement during toothbrushing, the application played songs during the timer, making the experience more pleasant. The study suggested using short sentences, preferably single words, to effectively interact with Alexa Skills and emphasized the importance of diversifying the dialogue based on the child's autistic profile.

4.1.3.2. Digital therapies and medical assistance. In prototype development and evaluation, Theradia focused on digital therapies (Tarpin-Bernard et al., 2021), while another study explored a voice-based chatbot for medical assistance (El Rhatassi et al., 2023). Theradia is a virtual agent created to assist adults with neurodevelopmental disorders during digital therapies. It provides social presence, coaching, and support using a Wizard-of-Oz approach, and incorporates speech features and facial analysis to enhance interaction and engagement. The voice-based chatbot aims to offer personalized assistance to adults with autism spectrum disorder and mental health concerns, including detecting developmental and mental health disorders and helping users manage their conditions. However, neither of these applications has been evaluated with end users in real-world settings.

The findings highlighted a preference for customized CAs in the healthcare domain, as they are better equipped to meet the unique needs of individuals, providing more personalized and effective support. Although two studies developed prototypes with this goal in mind, neither evaluated them with individuals with intellectual disabilities, revealing a significant gap in understanding their real-world applicability. Additionally, interactive features, such as playing songs, have been shown to enhance engagement, particularly for children, by creating a more personalized and dynamic experience. Despite the potential benefits of customized CAs for individuals with IDD, their deployment faces significant challenges, mainly due to the need for extensive customization to effectively meet the diverse needs of users.

4.1.4. Employment

Individuals with IDD face significant challenges in accessing employment opportunities and are often under-represented in open employment compared to individuals with other disabilities (Wilson & Campain, 2020). Despite these challenges, only one study has explored the use of CAs to support employment.

4.1.4.1. Job interviews. AllyChat (Garcia-Pi et al., 2023) is a virtual agent designed to assist adults with intellectual disabilities in practising job interviews within a virtual reality environment. A post-pilot study with neurotypical students provided positive feedback on the visual design of the environment. However, concerns were raised regarding the agent's speech speed and language complexity. Therefore, further research is needed to assess its effectiveness for individuals with intellectual disabilities.

The integration of virtual reality with CAs has the potential to significantly enhance interview training by providing a more engaging, contextually relevant, and realistic environment. Compared to voice only systems, this multimodal approach may improve training effectiveness, offering users a dynamic and interactive experience. Future research should investigate how this multimodal interaction impacts individuals with intellectual disabilities and its potential to improve learning outcomes.

4.1.5. Education

Education is crucial for supporting individuals with IDD by providing opportunities for skill development, independence, and social integration. However, many face challenges in educational institutes due to the lack of trained teachers and adequate resources (Hornby & Kauffman, 2024).

4.1.5.1. Higher education. Motivate Me (Bakhai et al., 2020) is a customized CA designed to support higher education students with autism spectrum disorder to improve their motivation while studying. The system generates personalized study routines and includes virtual characters that express emotions when tasks are not completed correctly. The prototype was evaluated with neurotypical students, achieving a system usability scale score of 71.66. However, its effectiveness with students with autism spectrum disorder has not yet been assessed.

4.1.5.2. Steam. LANA-I (Aljameel et al., 2019) is a web based Arabic virtual agent designed to teach science tutorials to children with autism spectrum disorder. The evaluation was conducted in two phases with neurotypical students: the first phase focused on the design and implementation of the Arabic CA, while the second phase involved the development of an educational science tutorial. This tutorial incorporated audio and video features and was integrated into a web application. The results demonstrated that LANA-I was an effective computer-assisted intervention, as most users provided correct responses after learning.

These studies revealed some uncertainty about the effectiveness of the above CAs, as they have not yet been evaluated with people with IDD, similar to their use in employment settings, leaving their impact on end users unclear. However, the studies demonstrated that interactive elements like audio, video, and visuals, successfully engage participants across various age groups. Since multimodal interactions may increase cognitive load, future research should focus on how these features specifically affect individuals with IDD.

4.2. Contribution types, design methods, and evaluation methods

Using the contribution type and method classification outlined in our data analysis, we first classified all the studies according to their primary contribution (Table 7). Across the 31 primary studies (excluding surveys (Bérubé et al., 2021; Catania et al., 2023; Even et al., 2022; Syed Mahmudul et al., 2022; Tavares et al., 2022; Volochtchuk et al., 2023)), the majority were artifact-driven, with 20 papers (64.5%) presenting a CA system, tool, or prototype as the main contribution. The remainder reported empirical investigations (10 papers; 32.3%) or proposed a methodological contribution (1 paper; 3.2%). In terms of design methods, 13 of the 31 studies (42%) reported a clearly identifiable, structured design process. The remaining 18 studies (58%) were coded as other because they did not report sufficient detail to determine a specific design method. In these papers, authors used off-the-shelf CAs (Balasuriya et al., 2018; Lewis & Vellino, 2021; Masina et al., 2020; Smith et al., 2021, 2023; van Wingerden et al., 2023; Zubatiy et al., 2021), adapted designs from existing systems or expert-led designs (Ali et al., 2020; Allen et al., 2018; Mower, Black, et al., 2011; Mower, Lee, et al., 2011; Soleiman et al., 2014; Zhang et al., 2021), design thinking (Arya et al., 2023), referenced a multi criteria decision approach (Spitale et al., 2020), or methodologically unclear approaches (El Rhatassi et al., 2023; Tanaka et al., 2017; Tarpin-Bernard et al., 2021).

Within the studies that reported design methods, interviews and co-design were the most frequently reported, each appearing in five studies (39%). Focus groups were reported in four studies (31%), while questionnaires appeared in a single study (8%). Across the co-design studies, participation typically involved three groups: people with IDD, caregivers (e.g., parents, support staff), and expert stakeholders (e.g., psychologists, therapists). A subset of studies centered people with IDD in co-design, collaborating with them to reflect on prior CA use and to envision future roles, personalities, and functions for CAs

(Cha et al., 2021; Greuter et al., 2022). Other studies reported expert led co-design (Gianotti et al., 2023a, 2023b), in which experts primarily shaped tasks, visual content, and dialogue flows for augmented reality and virtual reality tasks. A third group adopted mixed co-design, bringing together people with IDD and experts to review early prototypes, suggest scenarios, and refine agent behaviors (Garcia-Pi et al., 2023). Interviews were predominantly semi-structured and conducted either with people with IDD or with caregivers. Most studies interviewed people with IDD directly, while two relied on caregiver interviews (Aljameel et al., 2019; Parvin et al., 2022) to elicit user needs, preferences, and contextual constraints shaping CA use. One study complemented interviews with short questionnaires (Papadogiorgaki et al., 2023) to capture communication preferences for the proposed CA directly from people with IDD. Another study engaged both people with IDD and their caregivers (Koushik & Kane, 2023) to inform shared design decisions. In contrast to co-design and interviews, focus groups were conducted with expert stakeholders. Focus groups were reported in four studies, three of which were by Catania et al. (2019, 2020; Catania & Garzotto, 2023); in these studies, focus groups were used across the exploration, prototyping, and validation stages. Taken together, these design methods indicate that design activities were more often mediated through experts and caregivers than based on sustained, direct engagement with people with IDD. Of the 14 studies that reported design methods, six (42.9%) involved people with IDD. Across these six studies, IDD participation ranged from 2 to 17 participants: 33.3% involved 1–5 participants with IDD, 50% involved 6–10 participants, and 16.7% involved more than 10 participants, with the largest design cohort including 17 participants.

Compared with design methods, evaluation methods were more consistently reported, with 16 studies (52%) using multiple evaluation methods (two or more) and involving people with IDD directly. Across the studies, quantitative performance metrics were reported in 12 studies (39%), interviews in 11 studies (36%), and observations in 10 studies (32%). Questionnaires were used in six studies (19%), pilot study evaluations in six studies (19%), and surveys in four studies (13%). Most interviews were conducted directly with people with IDD; in some cases, interviews included both people with IDD and caregivers (Cha et al., 2021; Lewis & Vellino, 2021; Zubatiy et al., 2021), and only one study conducted interviews with caregivers (Parvin et al., 2022). Observations were particularly important, as several studies noted that they enabled researchers to capture non-verbal cues and interaction breakdowns when participants found it difficult to verbally articulate their experiences (Balasuriya et al., 2018; Catania et al., 2019, 2020; Catania & Garzotto, 2023; Smith et al., 2023; Soleiman et al., 2014; van Wingerden et al., 2023). Compared with the design stage, questionnaires appeared more often in evaluation, likely because they are well suited to capturing standardized outcome measures (e.g., satisfaction, perceived usefulness), whereas design work more often relied on co-design and interviews. Pilot evaluations were reported in six studies and were primarily used to gather initial feedback on usability, acceptability, and feasibility (Arya et al., 2023; Garcia-Pi et al., 2023; Gianotti et al., 2023b; Papadogiorgaki et al., 2023; Parvin et al., 2022). Many studies complemented these qualitative data with quantitative analysis, for example by reporting model or system performance metrics (e.g., recognition accuracy, task completion rates) or summary statistics derived from standardized scales and task outcomes (Aljameel et al., 2019; Allen et al., 2018; Arya et al., 2023; Catania et al., 2019; Catania & Garzotto, 2023; Masina et al., 2020; Mower, Lee, et al., 2011; Smith et al., 2021, 2023; Tanaka et al., 2017; van Wingerden et al., 2023; Zhang et al., 2021). Across the 28 evaluation studies, 22 involved people with IDD (78.6%). Among studies reporting participant counts, 20% included 1–5 participants, 35% included 6–10 participants, 30% included 11–20 participants, and 15% included more than 20 participants.

Several patterns emerged when cross-analyzing design and evaluation methods. Studies that reported co-design did not report quantitative metrics; instead, co-design was typically paired with qualitative evaluation, most commonly interviews and/or observations. Interviews also appeared to play an end-to-end role: 80% (four studies) of studies that used interviews during design also used interviews during evaluation, and three studies (60%) additionally reported evaluation questionnaires. Focus groups showed a different pairing, most often co-occurring with observation based evaluation in three studies (75%). Pilot evaluations were reported exclusively in artifact-focused papers in six studies (100%), likely reflecting the use of pilots to assess the feasibility and usability of newly developed prototypes. Given the small corpus and heterogeneous reporting, we interpret these results as descriptive co-occurrence patterns that highlight common pairings and gaps rather than implying causal relationships.

Some studies also highlighted several challenges in using some evaluation methodologies such as questionnaires and interviews, with people with IDD. For example, in Social Adventure (Greuter et al., 2022), researchers noted that eliciting detailed feedback was difficult due to limited expressive communication, often requiring additional probing, repetition, and cross checking to obtain interpretable responses. They also reported that administering structured questionnaires was challenging, as participants found the System Usability Scale difficult to complete and it was not used for the remainder of the study. Similarly, in Alexism (Parvin et al., 2022), the researchers reported that direct interviews with children were difficult to conduct and therefore interviewed parents as proxy informants to capture children's experiences and perspectives. Some studies (Soleiman et al., 2014; Zubatiy et al., 2021) noted that having parents present during evaluation sessions could improve engagement by helping to scaffold interactions and support children who required additional prompting or reassurance. Taken together, these examples indicate that, although studies applied a range of evaluation methods, their practical utility depended on the communication demands placed on participants.

We also classified each paper by its dominant study design method (qualitative, quantitative, or mixed-methods). Across the 31 primary papers (excluding literature reviews), 10 studies (32.3%) used primarily qualitative designs (e.g., interview based case studies, observational fieldwork, or usability/prototype pilots) (Balasuriya et al., 2018; Catania et al., 2020; Cha et al., 2021; Garcia-Pi et al., 2023; Greuter et al., 2022; Koushik & Kane, 2023; Lewis & Vellino, 2021; Mower, Black, et al., 2011; Papadogiorgaki et al., 2023; Soleiman et al., 2014). Six studies (19.4%) used quantitative designs (e.g., controlled comparisons, pre/post testing, or semi-randomised trials) (Aljameel et al., 2019; Allen et al., 2018; Mower, Lee, et al., 2011; Smith et al., 2021; Tanaka et al., 2017; Zhang et al., 2021). Twelve studies (38.7%) used mixed-methods approaches, combining qualitative feedback with quantitative performance metrics, logs, or scale-based measures (Ali et al., 2020; Arya et al., 2023; Bakhai et al., 2020; Catania et al., 2019; Catania & Garzotto, 2023; Gianotti et al., 2023a, 2023b; Masina et al., 2020; Parvin et al., 2022; Smith et al., 2023; van Wingerden et al., 2023; Zubatiy et al., 2021). In addition, three papers (9.7%) proposed concepts or decision frameworks without an end-user evaluation and are therefore reported separately (El Rhatassi et al., 2023; Spitale et al., 2020; Tarpin-Bernard et al., 2021).

4.3. Synthesis of cross-study design considerations

Across the included studies, we identified five cross-study design considerations that inform CA design for individuals with IDD (Figure 3): personalization and user adaptation; multimodal and multisensory support; reducing cognitive load through interaction simplicity; proactive guidance and communication repair; and supporting engagement and motivation. Together, these considerations form a conceptual framework that synthesizes the review's application areas, usability issues, and methodological patterns into actionable design priorities.

4.3.1. Personalization and user adaptation

Personalization was a recurring design strategy for aligning CA interactions with the diverse communication abilities and preferences of people with IDD (Greuter et al., 2022; Koushik & Kane, 2023; Parvin et al., 2022). Reported personalization strategies primarily targeted pacing and communication style (e.g., slowing or pausing CA speech (Greuter et al., 2022; Masina et al., 2020), extending listening windows (Masina et al., 2020; Zubatiy et al., 2021), and using short, concrete phrasing (Greuter et al., 2022; Masina et al., 2020; Parvin et al., 2022)), alongside dialogue adjustments that varied response length and vocabulary and modulated the level of scaffolding provided. Researchers also highlighted personalizing ways of initiating interaction and surface characteristics including alternative wake words (Catania & Garzotto, 2023; Parvin et al., 2022), preferred voices (Gianotti et al., 2023b; Greuter et al., 2022), embodied characters (Gianotti et al., 2023a), and customizable narrative elements to better reflect users' preferences and support continued engagement (Ali et al., 2020; Greuter et al., 2022). Across the corpus, personalization was primarily implemented through configurable routines and scripted interactions (Koushik & Kane, 2023; Parvin et al., 2022), with some work exploring behavior responsive or predictive adaptation to reduce recurring breakdowns (Ali et al., 2020; Masina et al., 2020).

4.3.2. Multimodal and multisensory support

Multimodal and multisensory interaction was frequently used to accommodate variation in how people with IDD perceive, express themselves, and sustain attention during CA use (Gianotti et al., 2023a; Masina et al., 2020; Soleiman et al., 2014; Tanaka et al., 2017). Multimodality shaped both how users expressed intent and how systems delivered feedback and guidance. To support user expression, several systems offered speech alternatives such as written text entry or simple on-screen buttons when speech was difficult or unreliable, reducing dependence on voice only interaction (Catania et al., 2019; Gianotti et al., 2023a, 2023b). To support feedback and guidance, studies described pairing spoken responses with visual prompts or on-screen text and using expressive visual cues to help users interpret system responses and maintain focus (Catania & Garzotto, 2023; Masina et al., 2020; Tanaka et al., 2017). A small subset further leveraged augmented reality and virtual reality scenarios to embed prompts and tutorials in concrete environments, supporting in-context guidance and transfer across settings (Gianotti et al., 2023a, 2023b). Some studies (Aljameel et al., 2019) also suggested short video tutorials for onboarding as a more accessible alternative to written manuals, enabling users to learn one interaction mechanic at a time and practise it immediately. These design suggestions reduced reliance on speech alone and improved the legibility of system state and guidance, highlighting the value of flexible, switchable channels for both input and output.

4.3.3. Reducing cognitive load through interaction simplicity

Simplifying interaction patterns was a recurring approach for reducing linguistic and cognitive demands during CA use (Gianotti et al., 2023b; Masina et al., 2020; Parvin et al., 2022). A common strategy was to minimize verbosity by using short, clear utterances (including single-word triggers) and avoiding lengthy or compound command formulations (Gianotti et al., 2023b; Masina et al., 2020; Parvin et al., 2022). Several systems reduced interaction steps by constraining choices at decision points and breaking tasks into smaller, sequential units, including through shallow option structures and routine automation that reduced the need to manage long instruction chains (Greuter et al., 2022; Parvin et al., 2022). Where learning was a focus, step-by-step walk-throughs that introduced one feature at a time further supported users to follow and remember task sequences (Gianotti et al., 2023a). In addition, a small number of studies highlighted the role of low complexity interfaces (e.g., simple touch targets and minimal screens) in reducing interaction overhead alongside simplified language (Balasuriya et al., 2018; Catania et al., 2019).

4.3.4. Proactive guidance and conversational repair

Individuals with IDD often experienced hesitation, incomplete requests, or difficulty initiating help seeking, making it difficult to get back on track after the interaction broke down (Gianotti et al., 2023b; Koushik & Kane, 2023; Masina et al., 2020). In response, several systems implemented proactive guidance and repair strategies, including prompts that periodically checked whether users needed help, clarification questions when commands were ambiguous or incomplete, and scaffolded interaction flow that guided users through next steps (Cha et al., 2021; Gianotti et al., 2023b; Koushik & Kane, 2023). Some studies also emphasized the importance of clear system status cues (e.g., indicating when the CA is listening or responding) to reduce uncertainty during breakdowns (Masina et al., 2020). Researchers also described caregiver mediated adjustments and mitigation of improper or unachievable requests, where caregivers could revise prompts or the CA redirected requests that were inappropriate or not supported by the system (Cha et al., 2021; Koushik & Kane, 2023). These forms of proactive scaffolding and error handling were described as helping maintain conversational flow and reduce user frustration, highlighting the value of CAs that can monitor for breakdowns and offer timely, transparent pathways to repair.

4.3.5. Supporting engagement and motivation

Across the reviewed systems, motivational support emerged as a recurring design strategy for sustaining participation and encouraging repeated interaction. Studies commonly incorporated immediate feedback, positive reinforcement, and celebratory responses when users completed tasks or made progress, often embedded within short narratives featuring recurring characters or game like activities (Ali et al., 2020; Gianotti et al., 2023b; Greuter et al., 2022; Koushik & Kane, 2023). In addition, several systems employed expressive or embodied agents and other affective cues to signal successful input recognition and support users' confidence during interaction (Tanaka et al., 2017). Taken together, these approaches point to a shared emphasis on making progress visible and socially meaningful. Authors consistently associated such motivational features with increased willingness to engage, greater persistence with challenging tasks, and more positive attitudes toward CAs (Ali et al., 2020; Bakhai et al., 2020).

4.4. Limitations reported in the studies

Across the studies, we identified recurring patterns in the reported limitations, most commonly relating to four areas: sample sizes, methodological constraints, communication challenges with CAs, and ethical considerations. Limitations related to sample size were the most frequently acknowledged and were consistently described as constraining the robustness and generalizability of the findings. Nine studies reported issues related to sample size (Catania & Garzotto, 2023; Garcia-Pi et al., 2023; Greuter et al., 2022; Masina et al., 2020; Papadogiorgaki et al., 2023; Parvin et al., 2022; Soleiman et al., 2014; van Wingerden et al., 2023; Zhang et al., 2021), limiting both statistical confidence and the diversity of behavioral observations that could be drawn. Participant pools were also described as homogeneous in four studies (Catania & Garzotto, 2023; Cha et al., 2021; Masina et al., 2020; Papadogiorgaki et al., 2023), for example, being drawn from a single organization, educational setting, or geographic region, which restricted demographic and contextual variation and limited ecological validity. Due to these constraints, two studies (Aljameel et al., 2019; Bakhai et al., 2020) relied on proxy participants, such as caregivers or neurotypical students, to approximate the perspectives of individuals with IDD. Although framed as a pragmatic response to recruitment barriers, the authors acknowledged that this approach may reduce the authenticity and nuance of perspectives directly captured from people with IDD.

Methodological limitations primarily concerned short study durations (Catania & Garzotto, 2023; Lewis & Vellino, 2021; Zhang et al., 2021) and single session evaluations (Ali et al., 2020; Bakhai et al., 2020; Soleiman et al., 2014), which restricted opportunities to examine sustained engagement, learning processes, or behavioral change over time. Researchers also noted that evaluations were often conducted in controlled or institutional settings, such as laboratories, clinics, or training centers (Greuter et al., 2022; Papadogiorgaki et al., 2023), rather than in naturalistic environments, such as participants' homes. These settings were described as limiting ecological validity and offering only a partial representation of real-world interaction demands.

Communication related limitations formed another recurrent category, with researchers consistently highlighting barriers inherent to speech centered interaction. Pronunciation variability was reported as a key factor contributing to recognition errors and inconsistent system responses in nine studies (Allen et al., 2018; Arya et al., 2023; Gianotti et al., 2023b; Masina et al., 2020; Papadogiorgaki et al., 2023; Smith et al., 2023; Soleiman et al., 2014; van Wingerden et al., 2023; Zhang et al., 2021). Two studies described difficulties arising from unclear or stammered utterances (Balasuriya et al., 2018; Smith et al., 2021) that reduced intelligibility, and two further studies highlighted soft spoken speech as a barrier (Arya et al., 2023; Smith et al., 2021), often resulting in incomplete detection. Timing related communication challenges were also noted: some studies reported that participants needed extended time to formulate requests before speaking, or slow articulation could trigger premature system timeouts or cutoffs (Balasuriya et al., 2018; Masina et al., 2020). Difficulties recalling wake words were mentioned in three studies (Greuter et al., 2022; Parvin et al., 2022; Smith et al., 2021), further interrupting conversational flow. A subset of three studies highlighted issues related to interface complexity (Ali et al., 2020; Balasuriya et al., 2018; Catania et al., 2020), such as navigating multi-step menus or configuration options. In response to these communication and pacing challenges, several studies recommended personalizing CA behavior based on the user's profile (Bakhai et al., 2020; Catania et al., 2019; Greuter et al., 2022; Koushik & Kane, 2023; Mower, Black, et al., 2011; Parvin et al., 2022), positioning personalization as a mitigation strategy.

Ethical concerns were reported in most studies, typically framed in terms of institutional approval, recruitment, and consent processes, with an emphasis on guardian or caregiver consent. A small number of papers (two studies), noted that, even when participants and their guardians had consented, it was not ethically permissible to retain detailed interaction logs because continuous audio capture could also record co-residents, support staff, or visitors who had not provided consent (Smith et al., 2021, 2023). Consequently, these studies did not retain logging data and instead relied on self-report and staff observations rather than fine-grained usage traces (e.g., frequency, duration, utterance types), limiting the level of detail they could provide about everyday device use. Beyond these constraints, one study (Cha et al., 2021) reported that, in shared or group settings, some participants were concerned that multiple users could access a common device history (e.g., media requests), creating risks of inadvertent disclosure of personal information.

5. Discussion

We reviewed 37 studies examining how individuals with IDD engage with voice based CAs. Although research in this space is growing, individuals with IDD remain under-represented as direct users of these technologies. In this section, we discuss more effective ways to report on IDD, key usability challenges in current CA designs, the shift toward competency based systems, inclusive methodologies, and the ethical considerations and safeguards highlighted in this review, and we propose future directions for advancing the accessibility and relevance of CAs for individuals with IDD.

5.1. Reporting on IDD

As highlighted in our review, IDD encompasses a broad spectrum of disabilities, with varying degrees of cognitive functioning. IDD is classified into mild, moderate, severe, and profound severity levels based on standardized IQ tests in clinical settings (Boat & Wu, 2015). However, only a few studies (Catania et al., 2020; Catania & Garzotto, 2023; Papadogiorgaki et al., 2023; Parvin et al., 2022) specifically referenced severity levels, while the majority described IDD in more general terms. This lack of specificity often complicates the interpretation and comparison of findings, as the effectiveness of interventions can vary significantly across different levels of cognitive functioning. A key issue with categorizing studies based on severity, such as clinical IQ scores, is that it often overlooks the significance of adaptive functioning and the ability to perform daily life skills. This is further complicated by the fact that the term IDD varies across countries. For example in UK, the term intellectual disability is also known as general learning disabilities (Department of Health & Social Care, 2025). In contrast, in the United States and Australia, IDD is generally distinguished from specific learning disabilities (e.g., dyslexia, dysgraphia, dyscalculia), although terms are sometimes used inconsistently in educational and policy contexts (Grigorenko et al., 2020; Todd et al., 2022).

Instead of broadly categorizing individuals by severity levels, another approach is to use an ability based classification system for reporting. This approach corresponds with the frameworks utilized by DSM – 5 (American Psychiatric Association, 2013) and AAIDD (American Association on Intellectual and Developmental Disabilities) (Shorgen & Turnbull, n.d), which are based on daily living skills and support levels. By emphasizing specific capabilities, such as whether participants can independently communicate, follow instructions, or perform daily activities with varying levels of support, this system allows for a more nuanced understanding of individual needs. We believe that shifting the focus from severity levels to individual abilities will provide researchers with a clearer picture of the diverse needs within the IDD community. This approach could lead to more effective solutions that better address the unique challenges faced by individuals with IDD. It also corresponds with Wobbrock et al.'s (2011) argument that assistive tools should be designed based on users' specific abilities, tailoring solutions to individual traits rather than relying on generalized disability levels. Adopting this perspective could lead to more personalized and impactful interventions, ultimately improving outcomes for individuals with

IDD. Reporting participants' abilities and support needs more clearly provides essential context for interpreting the interaction breakdowns observed during CA use, which we discuss next.

5.2. Improving usability of CAs

Individuals with IDD encounter recurring usability challenges when interacting with voice based CAs, which can limit effective use in everyday contexts. Addressing these breakdowns is important for reducing frustration, supporting confidence, and enabling more reliable interactions. In our review, we summarize the main usability issues reported in the IDD focused CA literature (Table 8) and outline potential mitigation strategies. These strategies are informed by adjacent CA research with other underserved user groups, including older adults and blind users, where similar interaction challenges have been documented.

Pronunciation issues are a significant barrier for individuals with IDD. These difficulties can include challenges with specific word articulation, such as lisping or stammering (Balasuriya et al., 2018; Masina et al., 2020; Smith et al., 2023; van Wingerden et al., 2023; Zubatiy et al., 2021), inconsistent speech rates (Balasuriya et al., 2018; Masina et al., 2020), difficulties in formulating questions (Balasuriya et al., 2018), and difficulties remembering wake words or commands (Lewis & Vellino, 2021; Zubatiy et al., 2021). These issues often result in unclear or misinterpreted conversations, which can significantly impact the user experience. More recently, researchers developed a random forest classifier to identify issues in speech features, which accurately detects lisped words and provides real-time corrections, thereby improving speech accuracy for individuals with speech impairments (Itagi et al., 2019). This technique could be extended to address pronunciation and lisping challenges faced by individuals with IDD when interacting with CAs. Research shows that stammering leads to a reduction in voice amplitude, which can help identify and remove repetitions, elongations, and pauses, thereby improving speech recognition. Mishra et al. developed a deep learning algorithm that enhances speech recognition for individuals who stammer by utilizing these amplitude changes (Mishra et al., 2021). This method could also improve speech detection and recognition for individuals with IDD when interacting with CAs.

Variations in speech rate can complicate interactions with CAs, as these systems may struggle to process commands when the speech rate is too slow or too fast. This issue is also prevalent among older adults, who may experience a slower speech rate due to age related factors (Lee, 2015). Some studies have suggested that allowing for customizable speech rates in CAs, ranging from slow to fast, can improve speech detection and response accuracy (Dowding et al., 2024; Wang et al., 2023). Several studies have shown that some individuals with IDD require additional time to formulate their questions, which often leads to the device exiting its listening mode during pauses. As a result, this can lead to incomplete queries and inaccurate responses. To address issues caused by long pauses, recent research suggests extending the listening window or allowing users to set a specific duration (e.g., three seconds or more) based on their typical voice interaction patterns. This approach is also used in turn-taking models for conversational systems with recurrent neural networks, ensuring that pauses are not mistakenly interpreted as the end of a sentence (Skantze, 2017). Additionally, the time taken to formulate questions can be indirectly mitigated by offering prompts or hints to help users frame their questions more quickly, such as showing possible matches or questions on screen based devices, thus reducing waiting time for users (Budzinski et al., 2019).

Table 8. Summary of usability issues and potential mitigations.

Usability issues	Mitigation strategies
Pronouncing specific / lisp words	Detect speech features related to lisping using a random forest algorithm (Itagi et al., 2019)
Stammering	Detect stammering using amplitude threshold level and eliminate repetitions using deep learning methods (Mishra et al., 2021)
Soft speaking	Train the models using soft speech (Lin et al., 2023; Zhu et al., 2023)
Varied speech rates	Customizing the speech rate by adjusting it from slow to fast (Dowding et al., 2024; Wang et al., 2023)
Extended question formulation	Set a recording window and provide hints and suggestions (Budzinski et al., 2019; Skantze, 2017)
Remembering wake actions	Use familiar words, employ single word phrases, and prompt actions automatically (Parvin et al., 2022; Zubatiy et al., 2021)
Complex user interfaces	Adapt principles from the WCAG (Caldwell et al., 2008)

Memory and recall can present significant challenges for individuals with IDD, impacting their ability to remember specific commands or wake words needed to activate the CA (Zubatiy et al., 2021). Simplifying activation processes, such as using a single wake phrase or familiar words (Parvin et al., 2022), or adapting interfaces to prompt relevant actions automatically, can enhance usability and ensure the CA remains effective even when users struggle to remember commands (Zubatiy et al., 2021). For example, if a CA learns that a user typically wakes up around 10 a.m. and frequently requests a morning routine, it could proactively offer this routine without requiring a specific command. Another major issue with CAs is that even users without speech issues may speak very softly (Cha et al., 2021), making it difficult for the CA to detect their commands. To address this issue, some researchers suggest training models on soft speech to improve detection accuracy for typical users (Lin et al., 2023; Zhu et al., 2023).

Research studies also highlight the importance of using clear and straightforward language in CAs and avoiding exaggerations, as complex or figurative language may confuse individuals with IDD (Cha et al., 2021; Karwai, 2016). This is particularly crucial, as studies involving children with IDD have raised concerns about the tone of CA voices. Research in this context has focused on enhancing the naturalness of CA voice by analyzing tone and speech pattern rates to improve communication effectiveness (Ali et al., 2020). In addition to addressing pronunciation issues, future research should also focus on improving user interfaces for screen based CAs to reduce cognitive load for individuals with IDD. The touchscreens on many off-the-shelf voice CAs could incorporate principles from the Web Content Accessibility Guidelines (WCAG) (Caldwell et al., 2008), particularly the principles of operability and understandability. This would involve simplifying long, complex sentences and providing large, easily accessible buttons on screen based devices to enhance usability and make information more comprehensible.

5.3. Toward competency based CAs

Personalization has been consistently emphasized as a key design goal in CAs for individuals with IDD, aiming to accommodate diverse cognitive, communicative, and contextual needs across domains such as education, health, social participation, and employment (Koushik & Kane, 2023; Parvin et al., 2022; Smith et al., 2023). However, the heterogeneous and context-dependent nature of IDD presents persistent challenges for conventional personalization strategies, which typically rely on static user profiles or generalized ability classifications. These approaches often fail to capture how users' communication patterns, attention, and comprehension vary across contexts, tasks, and emotional states, which can limit long-term usability and inclusivity. The competency based design framework (Bayor et al., 2021) proposed by Bayor et al. developed with individuals with intellectual disabilities, provides a valuable conceptual foundation for rethinking personalization in this space. It shifts the focus from static representations of ability what a user is assumed to be capable of in general to observable competencies, or the behaviors users actually demonstrate in specific interactions. By emphasizing contextual and situational performance rather than fixed ability levels, this framework encourages systems to adapt to demonstrated behaviors, such as conversational initiation, response pacing, or attention to prompts, rather than relying solely on predefined profiles. This orientation aligns with broader ability based and adaptive system paradigms (Shneiderman, 2000; Wobbrock et al. 2011) that advocate designing around users' actual strengths and strategies rather than their difficulties, while extending them toward continuous, context aware adaptation.

In a CA context, such an approach would enable systems to act as adaptive partners that evolve alongside the user. For instance, a competency aware CA could adapt its dialogue flow based on observed interaction patterns simplifying turn-taking for users who respond briefly, slowing its speech rate for those who pause frequently, or supplementing voice interaction with visual cues when verbal engagement decreases. This dynamic adaptation resonates with Loitsch's (Loitsch, 2018) Adaptive Inclusive Interactive Systems (AIIS) framework. Developed in a broader accessibility context, AIIS shows how passive observation and adaptive inference can reduce the need for manual configuration. It emphasizes automatic adaptation of interface parameters such as modality switching, interaction pacing, or visual contrast based on observed user behavior, thereby reducing the cognitive and manual effort required to configure accessibility settings. Collectively, these approaches highlight a shift from one time personalization toward continuous, interaction driven adaptation. By grounding system behavior in observed competencies, rather than assumed abilities, CAs can become more responsive to the lived experiences of individuals with IDD.

5.4. Encouraging inclusive methodologies

Across both design and evaluation, the corpus relied heavily on researcher-led methods, with limited use of participatory approaches that position people with IDD as partners in the work. Although co-design is well recognized in inclusive and disability research as a valuable method for centering the perspectives of people with IDD (Scior & Werner, n.d; Wang, 2022), it appeared in only 5 out of 31 studies, with only two directly involving individuals with IDD as co-designers (Cha et al., 2021; Greuter et al., 2022), while one involved people with IDD alongside experts (Garcia-Pi et al., 2023). The remaining co-design papers primarily worked with experts (Gianotti et al., 2023a, 2023b). This limited use both in terms of how few studies employed co-design at all and how infrequently people with IDD were involved directly stands in contrast to the centrality of co-design in accessibility and disability research, and indicates that participatory approaches remain the exception rather than the norm in CA work for this population.

Authors who attempted co-design pointed to several practical constraints that shaped these methodological choices, including the time and staffing required to run participatory sessions, limited funding for repeated workshops, and the challenges of engaging non verbal or minimally verbal participants in methods that rely heavily on spoken interaction (Banire et al., 2024; Kirk et al., 2021; Robb et al., 2021). To navigate these constraints, researchers adopted alternative techniques to support more accessible involvement, such as observations to capture non verbal cues (Ravn et al., 2022), proxy interviews with caregivers or support staff (Banire et al., 2024), prototyping activities that made abstract concepts tangible (Robb et al., 2021), and scenario based prompts to help participants explore ideas (Colin Gibson et al., 2020; Spencer González et al., 2020). In several cases, combining multiple methods provided the flexibility needed to accommodate diverse communication styles and cognitive profiles (Banire et al., 2024). These patterns highlight an ongoing challenge for CA research with individuals with IDD: while participatory and design led approaches are recognized as valuable, they have yet to be consistently embedded in practice. Developing more accessible, personalized, and communication sensitive adaptations of participatory methods may support deeper involvement in both the design and evaluation of future CAs (Kenny et al., 2023; McDonald et al., 2022). Such efforts may help align methodological practice with the broader goals of inclusive research and contribute to technologies that more accurately reflect the needs, contexts, and preferences of people with IDD.

5.5. Strengthening ethical considerations and safeguards

Across our corpus, many IDD studies acknowledged ethical considerations, but discussion was typically framed as procedural compliance (e.g., institutional approval, recruitment, and consent), often foregrounding guardian or caregiver consent (Balasuriya et al., 2018; Lewis & Vellino, 2021; Masina et al., 2020; Zubatiy et al., 2021). Only a small subset considered ethics as an ongoing concern in everyday deployment, including what interaction data are captured and reused, how system capabilities and limitations are communicated to users, and how sustained use might shape decision making and help seeking practices (Cha et al., 2021; Smith et al., 2021, 2023). By contrast, CA research in mental health (Bérubé et al., 2021; Car et al., 2020; Li et al., 2023) and aging contexts (Huang et al., 2025; Pradhan et al., 2018; Spangler et al., 2022) engages more directly with risks such as surveillance, misrepresentation, and dependency, and proposes concrete mitigation strategies. We draw on these safeguards and discuss how they may need to be adapted for IDD focused CA design and evaluation.

Surveillance emerged as a concern, with some IDD focused studies noting risks associated with storing voice data or interaction logs (Smith et al., 2021, 2023). Adjacent work with older adults and mental health individuals further details how these issues can become more salient in routine, in-the-home use, where CAs may be experienced as always on or where logging and sharing practices are unclear (Car et al., 2020; Huang et al., 2025; Spangler et al., 2022). For people with IDD who may have limited control over device settings or stay in shared arrangements, unintentional surveillance may pose a heightened risk. Safeguards discussed in adjacent domains emphasize making data practices transparent and easy to control, for example through accessible onboarding about what is captured and why, prominent controls to pause or disable logging, and retention/sharing options that can be configured with supporters where appropriate (Huang et al., 2025; Li et al., 2023; Spangler et al., 2022).

Some papers in our corpus also highlighted misinterpretation risks, noting that confident CA responses may lead users to overestimate what the system understands or is able to perform (Cha et al., 2021; Smith et al., 2021, 2023). For individuals with IDD, this may be compounded by communication difficulties and the authority often attributed to digital systems. Misheard input and unpredictable responses are frequently reported across CA research (Balasuriya et al., 2018; Qin et al., 2025) and can distort meaning or shift control away from the user. Prior work recommends safeguards that calibrate trust by making uncertainty and system boundaries explicit, for example by signaling when the system is unsure, clarifying when it is providing general rather than personalized information, and supporting repair through clarification prompts, confirmation steps for consequential actions, and simple correction/undo mechanisms (Car et al., 2020; Huang et al., 2025; Qin et al., 2025).

A further ethical consideration concerns how sustained CA use may reshape support seeking in ways that increase reliance on the system. In IDD contexts, this may be reinforced when a CA is treated not only as a functional tool but also as a social companion (Cha et al., 2021; Lewis & Vellino, 2021; van Wingerden et al., 2023). Work in mental health and aging similarly notes that frequent CA use can shift decision making, emotional regulation, or social contact onto the system, sometimes at the expense of human support (Car et al., 2020; Huang et al., 2025; Qin et al., 2025). Although IDD focused papers discussed dependency less often, many envisioned supports (e.g., daily routines, communication, self-management) that could become embedded in everyday practice. Safeguards discussed in adjacent domains emphasize preserving autonomy and maintaining routes to human assistance, including framing the CA as a helper rather than a decision-maker, prompting involvement of trusted supporters where appropriate, and, in higher risk situations, encouraging escalation to caregivers or professionals (Huang et al., 2025; Li et al., 2023; Qin et al., 2025). Taken together, these points suggest treating ethics in IDD focused CA research as an ongoing design and evaluation lens explicitly addressing surveillance, misrepresentation, and dependency rather than as a one-off procedural step at recruitment.

5.6. Future directions for CA design

Our findings highlight several research directions for advancing CA design with people with IDD. Across the corpus, reported usability issues repeatedly centered on interaction breakdowns misrecognition, hesitation, and incomplete requests and the effort required to recover from them. This points to several priorities for future CA design with people with IDD. In particular, breakdown recovery should be treated as a primary interaction goal. Rather than optimizing only for recognition success, CAs should make system state legible and provide low effort pathways to recover from misrecognition, hesitation, and incomplete requests. This direction also requires evaluations that report breakdown types, repair attempts, time-to-recovery, and the role of supporters during recovery, in addition to task outcomes. Recent work in socially assistive robotics points to technical advances that could inform this direction (Guhr et al., 2024). For example, one study developed an on-device voice interface for a socially assistive robot and improved recognition accuracy by fine-tuning a wav2vec model on task specific speech data. Adapting similar approaches for IDD focused CAs may improve recognition robustness, reduce breakdowns, and better support privacy preserving use in everyday environments.

Personalization often needs to move beyond one time configuration toward support that can evolve with the user. Rather than relying on extensive upfront setup, future CAs could enable gradual, competency aware adjustments as communication patterns, attention, and context change while keeping data collection minimal, privacy preserving, and under user control (e.g., lightweight on-device signals or simple summaries that can be reviewed and edited). Multimodal support may be particularly important where voice only interaction is a barrier; pairing speech with simple visual cues, short text, or structured choices can improve comprehension by making responses more concrete, whereas adding parallel channels can also increase interaction complexity. These possibilities should be developed and validated in everyday settings (e.g., homes and community environments), where noise, distraction, and shared device use routinely shape interaction. Finally, methodological practice may need to better align with inclusive design goals. Participatory involvement remains limited in the corpus, despite the prevalence of mediated and shared use contexts; future work could therefore benefit from engaging people with IDD and their support networks throughout design and evaluation so that adaptation and repair strategies reflect lived communication needs rather than expert assumptions. In practice, this includes co-designing repair strategies and pacing controls with users and caregivers, and reporting how support roles (prompting, mediation, shared-device management) shaped both design decisions and evaluation outcomes.

Beyond these needs, several application specific opportunities emerge in education, employment, and healthcare. Work on inclusive learning (Loitsch & Striegl, 2024) demonstrates how artificial intelligence can scaffold participation through captioning, text simplification, image description generation, and other accessibility oriented mechanisms. Although developed for higher education, these principles translate directly to educational contexts where voice based CAs could support low barrier learning, structured task guidance, and adaptive content delivery. Early CA studies with individuals with IDD show encouraging effects on engagement and comprehension (Aljameel et al., 2019; Bakhai et al., 2020), but further work is needed to understand how these accessibility mechanisms can be tuned to different learner profiles.

In employment settings, early prototypes such as AllyChat (Garcia-Pi et al., 2023) highlight the potential of CAs to scaffold job readiness tasks by modeling appropriate responses, rehearsing interview skills, and breaking down complex communication demands into manageable steps. Extending accessibility driven techniques such as simplified language, multimodal prompts, or adaptive pacing could further support users with varied cognitive and communication profiles, but empirical evidence on how to balance assistance with autonomy is still limited. Similar opportunities exist in healthcare, where voice interfaces may help moderate information load, support symptom communication, or scaffold everyday self management tasks (Syed Mahmudul et al., 2022). Despite promising results in adjacent health domains, voice based CAs have not been systematically evaluated with IDD users, leaving open questions about accessibility, comprehension, and interaction safety. Collectively, these trajectories point to a broader challenge, translating inclusive artificial intelligence principles into application areas CA designs that enhance accessibility, autonomy, and comprehension for individuals with IDD. Establishing this foundation will be critical for ensuring that future voice based CAs provide safe, effective, and contextually relevant support across learning, employment, and healthcare settings.

6. Conclusion

Our literature review provides valuable insights into the use of CAs by individuals with IDD. We identified various types of CAs and their applications across different application areas, contribution types and prevailing design and evaluation methods, limitations of the reviewed studies and synthesis of cross-study design considerations. The findings emphasize the promising potential of CAs to support individuals with IDD, while also highlighting the challenges in addressing their diverse needs and the limited involvement of this user group in the design process. These insights suggest the need for more inclusive, user-centered design approaches to ensure CAs are both effective and accessible for individuals with IDD. Future research should prioritize enhancing the usability of CAs, focusing on the development of competency based, personalized systems and incorporating co-design approaches to better engage individuals with IDD. Researchers must also adopt more robust design methodologies to improve the usability of CAs for individuals across the entire spectrum of IDD, accommodating varying competencies and support requirements. Moreover, there is a significant need for research in underexplored areas such as education, health, and employment to provide better support for individuals with IDD in these critical domains.

Notes

1. https://www.amazon.com/echo.
2. https://home.google.com/intl/en_au/about-google-home/.
3. https://www.apple.com/siri/.
4. https://assistant.google.com/
5. https://www.covidence.org/.

Author contributions

CRediT: **Madhuka Nadeeshani**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **Jacqueline Johnstone**: Data curation, Software, Validation, Writing – review & editing; **Kirsten Ellis**: Conceptualization, Data curation, Supervision, Writing – original draft, Writing – review & editing; **Swamy Ananthanarayan**: Conceptualization, Data curation, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Funding

This research is funded by the Monash University Faculty of Information Technology International Postgraduate Research Scholarship.

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ORCID

Madhuka Nadeeshani http://orcid.org/0000-0002-7312-1984
Jacqueline Johnstone http://orcid.org/0009-0006-1015-7169
Kirsten Ellis http://orcid.org/0000-0002-7570-0939
Swamy Ananthanarayan http://orcid.org/0000-0002-9808-5844

References

Abbott, S., & McConkey, R. (2006). The barriers to social inclusion as perceived by people with intellectual disabilities. *Journal of Intellectual Disabilities*, *10*(3), 275–287. https://doi.org/10.1177/1744629506067618

Ali, M. R., Razavi, S. Z., Langevin, R., Al Mamun, A., Kane, B., Rawassizadeh, R., Schubert, L. K., & Hoque, E. (2020). *A Virtual Conversational Agent for Teens with Autism Spectrum Disorder: Experimental Results and Design Lessons* [Paper presentation]. Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, IVA 2020, In. Association for Computing Machinery, Inc. https://doi.org/10.1145/3383652.3423900

Aljameel, S., O'Shea, J., Crockett, K., Latham, A., & Kaleem, M. (2019). LANA-I: An Arabic conversational intelligent tutoring system for children with ASD. In *Advances in Intelligent Systems and Computing*., Vol. 997. Springer Verlag. 498–516. https://doi.org/10.1007/978-3-030-22871-2_34

Allen, A. A., Shane, H. C., & Schlosser, R. W. (2018). The Echo™ as a speaker-independent speech recognition device to support children with autism: An exploratory study. *Advances in Neurodevelopmental Disorders*, *2*(1)3 2018), 69–74. https://doi.org/10.1007/s41252-017-0041-5

American Psychiatric Association. (2013). *Diagnostic and statistical manual of DSM-5*. Technical Report https://doi.org/10.1176/appi.books.9780890425787

Arya, L., Ranade, P., & Goswami, T. (2023). *Intellectually Disabled Child Holistic Growth Management Platform: Voice Interaction and Scheduling* [Paper presentation].2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development (OTCON)., In IEEE1–6. https://doi.org/10.1109/OTCON56053.2023.10113913

Bakhai, A., Constantin, A., & Alexandru, C. A. (2020). MotivateMe!: An Alexa Skill to Support Higher Education Students with Autism. In *Proceedings of the 14th International Conference on Interfaces and Human–Computer Interaction (IADIS I-HCI 2020)*. International Association for Development of the Information Society (IADIS), 187–191. https://www.iadisportal.org/digital-library/motivateme-an-alexa-skill-to-support-higher-education-students-with-autism

Balasuriya, S. S., Sitbon, L., Bayor, A. A., Hoogstrate, M., & Brereton, M. (2018). *Use of voice activated interfaces by people with intellectual disability* [Paper presentation]. *ACM International* [Paper presentation]. Conference Proceeding Series. Association for Computing Machinery, 102–112. https://doi.org/10.1145/3292147.3292161

Banire, B., Neves, J., Qaraqe, M., Othman, A., & Al-Thani, D. (2024). Co-design of technology involving autistic children: A systematic literature review. *International Journal of Human–Computer Interaction*, *40*(22), 7498–7516. https://doi.org/10.1080/10447318.2023.2266248

Bayor, A. A., Brereton, M., Sitbon, L., Ploderer, B., Bircanin, F., Favre, B., & Koplick, S. (2021). Toward a competency-based approach to co-designing technologies with people with intellectual disability. *ACM Transactions on Accessible Computing*, *14*(2), 1–33. https://doi.org/10.1145/3450355

Bérubé, C., Schachner, T., Keller, R., Fleisch, E., Wangenheim, F. V., Barata, F., & Kowatsch, T. (2021). Voice-based conversational agents for the prevention and management of chronic and mental health conditions. *: Systematic Literature Review*, *23*(3), 933. https://doi.org/10.2196/25933

Boat, T., & Wu, J. (2015). *Mental disorders and disabilities among low-income children* (p. 451). National Academies Press. https://doi.org/books/NBK332877

Brill, T. M., Munoz, L., & Miller, R. J. (2019). Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, *35*(15-16)10 2019), 1401–1436. https://doi.org/10.1080/0267257X.2019.1687571

Bruce, J. M., McQuiggan, M., Williams, V., Westervelt, H., & Tremont, G. (2008). Burden among spousal and child caregivers of patients with mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, *25*(4), 385–390. https://doi.org/10.1159/000122587

Budzinski, O., Noskova, V., & Zhang, X. (2019). The brave new world of digital personal assistants: Benefits and challenges from an economic perspective. *NETNOMICS: Economic Research and Electronic Networking*, *20*(2-3)), 177–194. https://doi.org/10.1007/s11066-019-09133-4

Caldwell, B., Micheal, C., Reid, L., & Vanderheiden, G. (2008). *Web Content Accessibility Guidelines (WCAG) 2.0*. Technical Report. https://www.w3.org/TR/WCAG20/

Car, L. T., Dhinagaran, D. A., Kyaw, B. M., Kowatsch, T., Joty, S. R., Theng, Y. L., & Atun, R. A. (2020). Conversational agents in health care: Scoping review and conceptual analysis. *Journal of Medical Internet Research* 22 (2020). https://doi.org/10.2196/17158

Carulla, L. S., Reed, G. M., Vaez-azizi, L. M., Cooper, S., Leal, R. M., Bertelli, M., Adnams, C., Cooray, S., Deb, S., Dirani, L. A., Girimaji, S. C., Katz, G., Kwok, H., Luckasson, R. U. T. H., Simeonsson, R. U. N. E., Walsh, C., Munir, K., & Saxena, S.,. (2011). Intellectual developmental disorders: Towards a new name, definition and framework for "mental retardation/intellectual disability" in ICD-11. *World Psychiatry*, *10*(3), 175–180. https://doi.org/10.1002/j.2051-5545.2011.tb00045.x

Catania, F., & Garzotto, F. (2023). A conversational agent for emotion expression stimulation in persons with neurodevelopmental disorders. *Multimedia Tools and Applications*, *82*(9), 12797–12828. 4 2023),. https://doi.org/10.1007/s11042-022-14135-w

Catania, F., Beccaluva, E., & Garzotto, F. (2020). The conversational agent "Emoty" Perceived by people with neurodevelopmental disorders: Is it a human or a machine? In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 11970, LNCS, pp. 65–78). Springer. https://doi.org/10.1007/978-3-030-39540-7_5

Catania, F., Di Nardo, N., Garzotto, F., Occhiuto, D. (2019). *Emoty: An emotionally sensitive conversational agent for people with neurodevelopmental disorders*. https://hdl.handle.net/10125/59641 https://doi.org/10.24251/HICSS.2019.244

Catania, F., Spitale, M., & Garzotto, F. (2023). Conversational agents in therapeutic interventions for neurodevelopmental disorders: A survey. *ACM Computing Surveys*, *55*(10), 1–34. https://doi.org/10.1145/3564269

CDC Child Development. (2024). Retrieved February 20, 2025, from https://www.cdc.gov/child-development/about/developmental-disability-basics.html

Cha, I., Kim, S. I., Hong, H., Yoo, H., & Lim, Y. K. (2021). *Exploring the use of a voice-based conversational agent to empower adolescents with autism spectrum disorder* [Paper presentation]. Conference on Human Factors in Computing Systems – Proceedings. Association for Computing Machinery. https://doi.org/10.1145/3411764.3445116

Cho, E. (2019). *Hey Google, Can I Ask You Something in Private?* [Paper presentation]. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19)., Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3290605.3300488

Choi, D., Kwak, D., Cho, M., & Lee, S. (2020). *"Nobody Speaks that Fast!" An Empirical Study of Speech Rate in Conversational Agents for People with Vision Impairments* [Paper presentation]. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20),. Association for Computing Machinery, New York, NYUSA, pp. 1–13. https://doi.org/10.1145/3313831.3376569

Colin Gibson, R., D. Dunlop, M., & Bouamrane, M.-M. (2020). *Lessons from Expert Focus Groups on how to Better Support Adults with Mild Intellectual Disabilities to Engage in Co-Design* [Paper presentation]. Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (Virtual Event, Greece) (ASSETS). Association for Computing Machinery, New York, NY, USA, p. 12. https://doi.org/10.1145/3373625.3417008

Critical Appraisal Skills Programme. (2018). Retrieved December 01, 2025, from CASP Cohort Study Checklist. https://casp-uk.net/casp-tools-checklists/

Department of Health and Social Care. (2025). Learning disability: Applying All Our Health. Retrieved April 21, 2025, from https://www.gov.uk/government/publications/learning-disability-applying-all-our-health Accessed: 2025-04-21.

Dowding, S., Gutwin, C., & Cockburn, A. (2024). User speech rates and preferences for system speech rates. *International Journal of Human-Computer Studies*, *184*(C), 103222. https://doi.org/10.1016/j.ijhcs.2024.103222

El Rhatassi, F. E., El Ghali, B., & Daoudi, N. (2023). A Chatbot's architecture for customized services for developmental and mental health disorders: Autism. In *Lecture Notes in Networks and Systems* (Vol. 669 LNNS, pp. 134–141). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-031-29860-8_14

Eunice Keneddy Shriver. (n.d.). Eunice Keneddy Shriver National Institute of Child Health and Human Development. Retrieved February 20, 2025, from https://www.nichd.nih.gov/health/topics/idds/conditioninfo

Even, C., Hammann, T., Heyl, V., Rietz, C., Wahl, H. W., Zentel, P., & Schlomann, A. (2022). Benefits and challenges of conversational agents in older adults: A scoping review. *Zeitschrift Fur Gerontologie Und Geriatrie*, *55*(5)8 2022), 381–387. https://doi.org/10.1007/s00391-022-02085-9

Fogg, B. J. (1998). *Persuasive computers: Perspectives and research directions* [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Los Angeles, California, USA) (CHI '98), In. ACM Press/Addison-Wesley Publishing Co., USA, 225–232. https://doi.org/10.1145/274644.274677

Garcia-Pi, B., Chaudhury, R., Versaw, M., Back, J., Kwon, D., Kicklighter, C., Taele, P., & Seo, J. H. (2023). *AllyChat: Developing a VR Conversational AI Agent Using Few-Shot Learning to Support Individuals with Intellectual Disabilities* [Paper presentation]. York, In Human-Computer Interaction – INTERACT 2023: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part IV, (United Kingdom). Springer-Verlag, Berlin, Heidelberg, 402–407. https://doi.org/10.1007/978-3-031-42293-5_43

Gianotti, M., Patti, A., Vona, F., Pentimalli, F., Barbieri, J., & Garzotto, F. (2023a). *Multimodal interaction for persons with autism: The 5A case study* [Paper presentation]. International Conference on Human-Computer Interaction, In. Springer, 581–600.

Gianotti, M., Patti, A., Vona, F., Pentimalli, F., Barbieri, J., & Garzotto, F. (2023b). *Multimodal Interaction for Persons with Autism: The 5A Case Study* [Paper presentation]. In Universal Access in Human-Computer Interaction: 17th International Conference, UAHCI 2023, Held as Part of the 25th HCI International Conference, HCII 2023, Copenhagen, Denmark, July 23–28, 2023, Proceedings, Part I, (Copenhagen, Denmark). Springer-Verlag, Berlin, Heidelberg, 581–600. https://doi.org/10.1007/978-3-031-35681-0_38

Greuter, S., Watson, J., & Balandin, S. (2022). Social Adventure. *International Journal of Gaming and Computer-Mediated Simulations*, *14*(1)7 2022), 1–21. https://doi.org/10.4018/IJGCMS.303107

Grigorenko, E. L., Fuchs, L. S., Willcutt, E. G., Compton, D. L., Wagner, R. K., & Fletcher, J. M. (2020). Understanding, educating, and supporting children with specific learning disabilities: 50 Years of science and practice. *The American Psychologist*, *75*(1), 37–51. https://doi.org/10.1037/amp0000452

Guhr, O., Loitsch, C., Weber, G., & Böhme, H.-J. (2024). *Enhancing usability of voice interfaces for; socially assistive robots through deep learning: A German Case Study*. Springer-Verlag. 231–249. https://doi.org/10.1007/978-3-031-60615-1_15

Hobert, S., Von Wolff, R. M. (2019). *Say Hello to Your New Automated Tutor-A Structured Literature Review on Pedagogical Conversational Agents*. Technical Report. https://api.semanticscholar.org/CorpusID:201114924

Hornby, G., & Kauffman, J. M. (2024). Inclusive Education, Intellectual Disabilities and the Demise of Full Inclusion. *Journal of Intelligence*, *12*(2), 20. 2 2024). https://doi.org/10.3390/jintelligence12020020

Hostyn, I., & Maes, B. (2009). Interaction between persons with profound intellectual and multiple disabilities and their partners: A literature review. *Journal of Intellectual & Developmental Disability*, *34*(4), 296–312. https://doi.org/10.3109/13668250903285648

Huang, Y., Zhou, Q., & Piper, A. M. (2025). *Designing Conversational AI for Aging: A Systematic Review of Older Adults' Perceptions and Needs* [Paper presentation]. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 181, p. 20. https://doi.org/10.1145/3706598.3713578

Itagi, A., Baby, C. J., Rout, S., Bharath, K. P., Karthik, R., & Rajesh Kumar, M. (2019). Lisp detection and correction based on feature extraction and random forest classifier. In *Lecture Notes in Electrical Engineering* (Vol. 521, pp. 55–64). Springer Verlag. https://doi.org/10.1007/978-981-13-1906-8_6

Karwai, P. (2016). Dos and don'ts on designing for accessibility: Accessibility in government. https://accessibility.blog.gov.uk/2016/09/02/dos-and-donts-on-designing-for-accessibility/.

Kenny, N., Doyle, A., & Horgan, F. (2023). Transformative inclusion: Differentiating qualitative research methods to support participation for individuals with complex communication or cognitive profiles. *International Journal of Qualitative Methods*, *22*. https://doi.org/10.1177/16094069221146992

Kirk, J., Bandholm, T., Andersen, O., Husted, R. S., Tjørnhøj-Thomsen, T., Nilsen, P., & Pedersen, M. M. (2021). Challenges in co-designing an intervention to increase mobility in older patients: A qualitative study. *Journal of Health Organization and Management*, *35*(9)2021), 140–162. https://doi.org/10.1108/JHOM-02-2020-0049

Koushik, V., & Kane, S. (2023). *Ability + Motivation: Understanding Factors that Influence People with Cognitive Disabilities in Regularly Practicing Daily Activities* [Paper presentation]. ACM International Conference Proceeding Series., In Association for Computing Machinery, 122–133. https://doi.org/10.1145/3587281.3587295

Lee, J. Y. (2015). Aging and speech understanding. *Journal of Audiology & Otology*, *19*(1), 7–13. https://doi.org/10.7874/jao.2015.19.1.7

Lewis, L., & Vellino, A. (2021). *Helping Persons with Cognitive Disabilities using Voice-Activated Personal Assistants* [Paper presentation]. 2021 8th International Conference on ICT and Accessibility, ICTA 2021. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICTA54582.2021.9809777

Li, H., Zhang, R., Lee, Y.-C., Kraut, R. E., & Mohr, D. C. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *NPJ Digital Medicine*, *6*(1), 236. https://doi.org/10.1038/s41746-023-00979-5

Lin, Z., Patel, T., & Scharenborg, O. (2023). *Improving Whispered Speech Recognition Performance Using Pseudo-Whispered Based Data Augmentation* [Paper presentation]. 2023 IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2023. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ASRU57964.2023.10389801

Loitsch, C. (2018). *Designing Accessible User Interfaces for All by Means of Adaptive Systems* [Ph. D. Dissertation]. https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-319846

Loitsch, C., & Striegl, J. (2024). *AI for inclusive learning in higher education: Diversity, accessibility, and mental health*. 595. –612 https://doi.org/10.17877/DE290R-24364

Long, D., & Magerko, B. (2020). *What is AI Literacy? Competencies and Design Considerations* [Paper presentation]. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20). In Association for Computing Machinery, New York, NYUSA1–16. https://doi.org/10.1145/3313831.3376727

Masina, F., Orso, V., Pluchino, P., Dainese, G., Volpato, S., Nelini, C., Mapelli, D., Spagnolli, A., & Gamberini, L. (2020). Investigating the accessibility of voice assistants with impaired users: Mixed methods study. *Journal of Medical Internet Research*, *22*(9), e18431. https://doi.org/10.2196/18431

McDonald, K. E., Gibbons, C., Conroy, N., & Olick, R. S. (2022). Facilitating the inclusion of adults with intellectual disability as direct respondents in research: Strategies for fostering trust, respect, accessibility and engagement. *Journal of Applied Research in Intellectual Disabilities*, *35*(1)2022), 170–178. https://doi.org/10.1111/jar.12936

Mishra, N., Gupta, A., & Vathana, D. (2021). Optimization of stammering in speech recognition applications. *International Journal of Speech Technology*, *24*(3)9 2021), 679–685. https://doi.org/10.1007/s10772-021-09828-w

Mower, E., Black, M., Floresa, E., Willams, M., & Narayanan, S. (2011). *Design of emotionally targeted interactive agent for children with autism* [Paper presentation]. 2011 IEEE International Conference on Multimedia and Expo, In. https://doi.org/10.1109/ICME.2011.6011990

Mower, E., Lee, C.-C., Gibson, J., Chaspari, T., Williams, M., & Narayanan, S. (2011). Analyzing the Nature of ECA Interactions in Children with Autism. In INTERSPEECH 2989–2993.

Mpofu, E., Houck, E., Linden, A., & Fernandez, C. (2020). Intellectual and developmental disabilities wellbeing. *Sustainable Community Health: Systems and Practices in Diverse Settings.*(pp. 461–497). https://doi.org/10.1007/978-3-030-59687-3_14

Nha Hong, Q., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). Mixed methods appraisal tool (MMAT), version 2018. *Registration of Copyright*, *1148552*(10), 1–7.

Papadogiorgaki, M., Grammalidis, N., Grammatikopoulou, A., Apostolidis, K., Bei, E. S., Grigoriadis, K., Zafeiris, S., Livanos, G., Mezaris, V., & Zervakis, M. E. (2023). An Integrated Support System for People with Intellectual Disability. *Electronics*, *12*(18), 3803. https://doi.org/10.3390/electronics12183803

Parvin, P., Manca, M., Senette, C., Buzzi, M. C., Buzzi, M., & Pelagatti, S. (2022). *Alexism: ALEXa supporting children with autISM in their oral care at home* [Paper presentation]. ACM International Conference Proceeding Series. Association for Computing Machinery. https://doi.org/10.1145/3531073.3531157

Pearce, D. (2013). *Intellectual Disability: Promoting daily living skills in adults Adults living at home with parents*. Technical Report. https://www.ideas.org.au/uploads/resources/247/intellectual-disability-promoting-daily-skills-adults.pdf [Database].

Pradhan, A., Mehta, K., & Findlater, L. (2018). "Accessibility came by accident" use of voice-controlled intelligent personal assistants by people with disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

Qin, J., Nan, Y., Li, Z., & Meng, J. (2025). Effectiveness of communication competence in AI conversational agents for health: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *27*, e76296. https://doi.org/10.2196/76296

Ravn, J. H., Møberg Jacobsen, R., & Brodersen Hansen, N. (2022). *Co-designing with adult people with ASD: A review of applied tools and techniques* [Paper presentation]. NordiCHI '22 Adjunct) Association for Computing MachineryNew YorkNYUSA, Article 46, p. 5. https://doi.org/10.1145/3547522.3547690

Reinders, S., Ananthanarayan, S., Butler, M., & Marriott, K. (2023). *Designing Conversational Multimodal 3D Printed Models with People who are Blind* [Paper presentation]. In Proceedings of the 2023 ACM Designing Interactive Systems Conference (Pittsburgh, PA, USA) (DIS '23). Association for Computing Machinery, New York, NY, USA, pp. 2172–2188. https://doi.org/10.1145/3563657.3595989

Reppermund, S., Heintze, T., Srasuebkul, P., Reeve, R., Dean, K., Smith, M., Emerson, E., Snoyman, P., Baldry, E., Dowse, L., Szanto, T., Sara, G., Florio, T., Johnson, A., Clements, M., McKenzie, K., & Trollor, J. (2019). Health and wellbeing of people with intellectual disability in New South Wales, Australia: A data linkage cohort. *BMJ Open*, *9*(9), e031624. https://doi.org/10.1136/bmjopen-2019-031624

Robb, N., Boyle, B., Politis, Y., Newbutt, N., Kuo, H. J., & Sung, C. (2021). Participatory technology design for autism and cognitive disabilities: A narrative overview of issues and techniques. *Recent advances in technologies for inclusive well-being: Virtual patients, gamification and simulation*., 469–485. https://doi.org/10.1007/978-3-030-59608-8

Sarkis-Onofre, R., Catalá-López, F., Aromataris, E., & Lockwood, C. (2021). How to properly use the PRISMA Statement. *Systematic Reviews*, *10*(1)2021), 117. https://doi.org/10.1186/s13643-021-01671-z

Scior, K., Werner, S. (n.d.). *Intellectual disability and stigma Stepping Out from the Margins*. Technical Report. https://link.springer.com/book/10.1057/978-1-137-52499-7

Shneiderman, B. (2000). Universal usability. *Communications of the ACM*, *43*(5), 84–91. https://doi.org/10.1145/332833.332843

Shorgen, K. A., & Turnbull, H. R. (n.d). Public policy and outcomes for persons with intellectual disability: Extending and expanding the public policy framework of AAIDD's 11th Edition of Intellectual Disability: Definition, Classification, and Systems of Support. ([n. d.]). https://doi.org/10.1352/1934-9556-48.5.375

Simpson, J., Gaiser, F., MacÍk, M., & Breßgott, T. (2020). *Daisy: A Friendly Conversational Agent for Older Adults* [Paper presentation]. ACM International Conference Proceeding Series. Association for Computing Machinery. https://doi.org/10.1145/3405755.3406166

Skantze, G. (2017). *Towards a General, Continuous Model of Turn-taking in Spoken Dialogue using LSTM Recurrent Neural Networks* [Paper presentation]. Technical Report. pp. 15–17. https://doi.org/10.18653/v1/W17-5527

Smith, E., Sumner, P., Hedge, C., & Powell, G. (2021). Smart speaker devices can improve speech intelligibility in adults with intellectual disability. *International Journal of Language & Communication Disorders*, *56*(3), 583–593. https://doi.org/10.1111/1460-6984.12615

Smith, E., Sumner, P., Hedge, C., & Powell, G. (2023). Smart-speaker technology and intellectual disabilities: Agency and wellbeing. *Disability and Rehabilitation. Assistive Technology*, *18*(4), 432–442. https://doi.org/10.1080/17483107.2020.1864670

Soleiman, P., Salehi, S., Mahmoudi, M., Ghavami, M., Moradi, H., & Pouretemad, H. (2014). *RoboParrot: A robotic platform for human robot interaction, case of autistic children* [Paper presentation]. 2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM), pp. 711–716. https://doi.org/10.1109/ICRoM.2014.6990987

Spangler, H. B., Driesse, T. M., Lynch, D. H., Liang, X., Roth, R. M., Kotz, D., Fortuna, K., & Batsis, J. A. (2022). Privacy concerns of older adults using voice assistant systems. *Journal of the American Geriatrics Society*, *70*(12)), 3643–3647. https://doi.org/10.1111/jgs.18009

Spencer González, H., Vega Córdova, V., Exss Cid, K., Jarpa Azagra, M., & Álvarez Aguado, I. (2020). *Including intellectual disability in participatory design processes: Methodological adaptations and supports* [Paper presentation]. Proceedings of the 16th Participatory Design Conference 2–20 – Participation(s) Otherwise – Volume 1 (Manizales, Colombia) (PDC '20)., In Association for Computing Machinery, New York, NY, USA, 55–63. https://doi.org/10.1145/3385010.3385023

Spitale, M., Catania, F., Crovari, P., & Garzotto, F. (2020). *Multicriteria decision analysis and conversational agents for children with autism* [Paper presentation]. Proceedings of the Annual Hawaii International Conference on System Sciences, Vol. 2020-January. IEEE Computer Society, pp. 1005–1014. https://doi.org/10.24251/HICSS.2020.125

Syed Mahmudul, H., Rytis, M., & Robertas, D. (2022). Dialogue agents for artificial intelligence-based conversational systems for cognitively disabled: A systematic review. https://doi.org/10.1080/17483107.2022.2146768

Tanaka, H., Negoro, H., Iwasaka, H., & Nakamura, S. (2017). Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PLoS One*, *12*(8), e0182151. https://doi.org/10.1371/journal.pone.0182151

Tarpin-Bernard, F., Fruitet, J., Vigne, J. P., Constant, P., Chainay, H., Koenig, O., Ringeval, F., Bouchot, B., Bailly, G., Portet, F., Alisamir, S., Zhou, Y., Serre, J., Delerue, V., Fournier, H., Berenger, K., Zsoldos, I., Perrotin, O., Elisei, F., … Ghenassia, D. (2021). THERADIA: Digital Therapies Augmented by Artificial Intelligence. In *Lecture Notes in Networks and Systems* (Vol. 259, pp. 478–485). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-80285-1_55

Tassé, M. J., Schalock, R. L., Luckasson, R., Tassé, J. D. J. M., Schalock, R. L., Luckasson, R., & Tassé, M. J. (n.d). []. *American Journal on Intellectual and Developmental Disabilities An Overview of Intellectual Disability: Definition, Diagnosis, Classification, and Systems of Supports*. (12 th ed.). Technical Report.

Tavares, R., Sousa, H., & Ribeiro, J. (2022). Smart Speakers and Functional Diversity: A Scoping Review. In A. P. Costa António Moreira, M. C. SánchezGómez, & S. Wa-Mbaleka (Eds.), *Computer Supported Qualitative Research* (pp. 48–64). Springer International Publishing. https://doi.org/10.1007/978-3-031-04680-3_4

Todd, N., Gaunt, L., & Porta, T. (2022). Terminology and provision for students with learning difficulties: An examination of australian state government education department websites. *Australian Journal of Teacher Education*, *47*(7), 21–37. https://doi.org/10.14221/ajte.2022v47n7.2

Tsikinas, S., & Xinogalos, S. (2019). Studying the effects of computer serious games on people with intellectual disabilities or autism spectrum disorder: A systematic literature review. *Journal of Computer Assisted Learning*, *35*(1)2019), 61–73. https://doi.org/10.1111/jcal.12311 van Wingerden, E., Vacaru, S. V., Holstege, L., & Sterkenburg, P. S. (2023). Hey Google! Intelligent personal assistants and well-being in the context of disability during COVID-19. *Journal of Intellectual Disability Research*, *67*(10), 973–985. https://doi.org/10.1111/jir.13064

Volochtchuk, A. V. L., Leite, H., & Vieira, A. D. (2023). Voice assistant technology applied to populations with developmental and physical disabilities. *Behaviour and Information Technology*, *43*(11), 2300–2322. https://doi.org/10.1080/0144929X.2023.2243343

Walsh, E., Holloway, J., Lydon, H., McGrath, A., & Cunningham, T. (2019). An exploration of the performance and generalization outcomes of a social skills intervention for adults with autism and intellectual disabilities. *Advances in Neurodevelopmental Disorders*, *3*(4), 372–385. https://doi.org/10.1007/s41252-019-00125-x

Wang, J., Yang, S., & Xu, Z. (2023). Talk like me: Exploring the feedback speech rate regulation strategy of the voice user interface for elderly people. *Frontiers in Psychology*, *14*, 1119355. https://doi.org/10.3389/fpsyg.2023.1119355

Wang, Z. (2022). How might we evaluate co-design? A literature review on existing practices. https://doi.org/10.21606/drs.2022.774

Wilson, E., Campain, R. (2020). *Fostering employment for people with intellectual disability: The evidence to date*. Technical Report. https://www.everyonecanwork.org.au/wp-content/uploads/2020/11/Fostering-employment-for-people-with-intellectual-disability-Accessible.pdf

Wobbrock, J. O., & Kientz, J. A. (2016). Research contributions in human-computer interaction. *Interactions*, *23*(3), 38–44. https://doi.org/10.1145/2907069

Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing*, *3*(3), 1–27. https://doi.org/10.1145/1952383.1952384

Wohlin, C. (2014). *Guidelines for snowballing in systematic literature studies and a replication in software engineering* [Paper presentation]. Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (London, England, United Kingdom) (EASE '14). Association for Computing Machinery, New York, NY, USA, Article 38, p. 10. https://doi.org/10.1145/2601248.2601268

Zhang, L., Amat, A. Z., Zhao, H., Swanson, A., Weitlauf, A., Warren, Z., & Sarkar, N. (2021). Design of an intelligent agent to measure collaboration and verbal-communication skills of children with autism spectrum disorder in collaborative Puzzle Games. *IEEE Transactions on Learning Technologies*, *14*(3), 338–352. https://doi.org/10.1109/TLT.2020.3029223

Zhu, Z., Zhang, L., Pei, K., & Chen, S. (2023). A robust and lightweight voice activity detection algorithm for speech enhancement at low signal-to-noise ratio. *Digital Signal Processing*, *141*(C), 104151. https://doi.org/10.1016/j.dsp.2023.104151

Zubatiy, T., Vickers, K. L., Mathur, N., & Mynatt, E. D. (2021). Empowering dyads of older adults with mild cognitive impairment and their care partners using conversational agents. https://doi.org/10.1145/3411764

About the authors

Madhuka Nadeeshani is a PhD candidate in the Faculty of IT at Monash University. Her research focuses on human–computer interaction and accessibility, using participatory methods to design and evaluate systems that support people with disabilities, integrating natural language processing and machine learning.

Jacqueline Johnstone is a PhD candidate with the Faculty of IT, Monash University. Her research sits at the intersection of interactive technologies and collaborative design, focusing on human-centered co-design practices and the development of accessible tools and systems that leverage emerging technologies to support people with diverse needs.

Kirsten Ellis has a creative and technical background, with extensive experience designing, developing, and evaluating systems with and for people with disabilities. Her research interests include accessible and adaptive technology, human-computer interaction, and how technology can support a more inclusive society.

Swamy Ananthanarayan is a Senior Lecturer at the Faculty of Information Technology, Monash University. His research spans human-computer interaction and accessibility, using inclusive design, participatory methods, and emerging technologies to create equitable interactive experiences for people with disabilities.

Appendix A. Quality appraisal using CASP and MMAT informed criteria

The quality appraisal in Table 9 is based on criteria systematically adapted from the CASP and MMAT checklists. Rather than applying the full CASP and MMAT tools separately to each study type, we combined overlapping concepts into five dimensions that could be used consistently across qualitative, quantitative and mixed-methods studies. "Methodological clarity" covers the clarity of research aims, the appropriateness of the study design, and the alignment between methods, data, and research questions. "Participant appropriateness" reflects recruitment pathways, sampling strategies, and the suitability or representativeness of participants for the stated aims. "Data collection quality" captures the description and adequacy of data collection procedures and measurements, including completeness of outcome data. "Ethics and analysis" combines reporting of ethical approval and procedures with transparency and rigor in qualitative, quantitative, or mixed-methods analysis. "Reporting transparency" assesses the clarity with which findings are presented, the use of supporting evidence, and the acknowledgement of limitations. The resulting CASP and MMAT informed summary is used only to contextualize the strength and limitations of the evidence base and does not determine which studies are included in the synthesis.

Table 9. Quality appraisal across included studies using CASP and MMAT informed criteria (Y = yes, N = No, CT = can't tell).

Study	Methodological clarity	Participant appropriateness	Data collection quality	Ethics and analysis	Reporting transparency
Ali et al., 2020	CT	Y	Y	CT	Y
Papadogiorgaki et al., 2023	Y	Y	Y	Y	Y
Gianotti et al., 2023a	Y	Y	Y	Y	Y
Balasuriya et al., 2018	Y	Y	Y	Y	Y
Cha et al., 2021	Y	Y	Y	Y	Y
Greuter et al., 2022	Y	Y	Y	Y	Y
Lewis & Vellino, 2021	Y	Y	Y	CT	Y
Zhang et al., 2021	Y	Y	Y	Y	Y
Parvin et al., 2022	Y	Y	Y	Y	Y
Bakhai et al., 2020	Y	Y	Y	CT	Y
Garcia-Pi et al., 2023	Y	Y	Y	Y	Y
Aljameel et al., 2019	Y	N	Y	Y	Y
Zubatiy et al., 2021	Y	Y	Y	Y	Y
Mower, Black, et al., 2011	Y	Y	N	CT	Y
Mower, Lee, et al. 2011	Y	Y	N	CT	Y
Catania et al., 2019	Y	Y	Y	Y	Y
Catania & Garzotto, 2023	Y	Y	Y	Y	Y
Catania et al. 2020	Y	Y	Y	Y	Y
Koushik & Kane, 2023	Y	Y	CT	CT	Y
Tarpin-Bernard et al., 2021	CT	N	Y	CT	Y
El Rhatassi et al., 2023	CT	N	Y	CT	Y
Masina et al., 2020	Y	Y	Y	Y	Y
Smith et al., 2021	Y	Y	Y	Y	Y
Smith et al., 2023	Y	Y	Y	Y	Y
van Wingerden et al., 2023	Y	Y	Y	Y	Y
Tanaka et al., 2017	CT	Y	Y	Y	CT
Soleiman et al., 2014	Y	Y	N	CT	Y
Allen et al., 2018	Y	Y	Y	Y	Y
Spitale et al., 2020	Y	CT	N	Y	Y
Gianotti et al., 2023b	Y	Y	CT	Y	Y
Arya et al., 2023	Y	Y	CT	CT	Y