



Conversational Agents to Support Remote Personalized Instruction for Diverse Learners

Richard Allen Carter Jr¹ · Ling Zhang² · Tiffany L. Hunt¹ · Christopher R. Emerling³ · Sohyun Yang⁴ · James Rujimora⁵

Accepted: 29 June 2023 / Published online: 18 July 2023
© Association for Educational Communications & Technology 2023

Abstract

Recent developments in Conversational Agents (CAs) have offered researchers new ways to develop personalized learning support for students that is accessible in various environments. Traditionally, CA learner experiences are accessible through home devices, including Alexa Echo and Google Nest, for tasks such as setting reminders, ordering products, and controlling smart home functions. However, there is an opportunity to expand CA's range to investigate the role of CAs in guiding instruction as a novel digital learning modality for learners with diverse needs. This article introduces considerations for designing CAs that leverage aspects of personalized student learning paired with common pedagogical instruction accessible on demand. Guided by classical theories of multimodal learning and self-regulated learning, we describe how to design CAs to capture personalized student variables throughout the voice instruction. Finally, we discuss the implications and potential for CAs moving forward to meet diverse learning needs of all learners.

Keywords Conversational agent · Multimedia learning · Personalized learning · Self-regulated learning

With a shift in the delivery of instructional practices in evolving learning environments (e.g., online modalities), researchers and practitioners seek innovative ways to create personalized learning (PL) experiences that are accessible to all learners with diverse learning needs (Zhang et al., 2020a, b). Conversational Agents (CAs) have emerged as a type of technology that demonstrates the potential to facilitate PL experiences. The emergence of CAs can be found in embedded technologies in most of the personal devices we use throughout our day, such as cellphones, personal computers, and tablets. Interactions with these devices are achieved

through Apple's Siri, Microsoft's Cortana, and Amazon's Alexa, depending upon which device you are using. Additionally, CAs enable individuals to interact with devices such as home speakers (e.g., Amazon Alexa, Google Assistant, Samsung Bixby), computers, televisions, and mobile phones through voice, which creates hands-free human-computer interaction (Myers et al., 2019). The application of CAs has permeated such societal sectors as smart homes, transportation, marketing, and healthcare, which has exhibited higher perceived efficiency, lower cognitive effort, higher enjoyment, and higher service satisfaction as opposed to

✉ Richard Allen Carter Jr
rcarter8@uwyo.edu

Ling Zhang
ling.zhang@unc.edu

Tiffany L. Hunt
thunt@uwyo.edu

Christopher R. Emerling
christopheremerling@gmail.com

Sohyun Yang
s_yang12@fhsu.edu

James Rujimora
jrujimora@Knights.ucf.edu

- ¹ School of Counseling, Leadership, Advocacy, and Design, University of Wyoming, 1000 E. University Ave., Laramie, WY 82071, USA
- ² School of Education, University of North Carolina at Chapel Hill, 1000 East Cameron Ave., Chapel Hill, NC 27514, USA
- ³ College of Education, University of Illinois at Chicago, 1040 W Harrison St., Chicago, IL 60607, USA
- ⁴ College of Education, Fort Hays State University, 600 Park Street Rarick Hall 257, Hays, KS 67601, USA
- ⁵ College of Community Innovation and Education, University of Central Florida, 12494 University Blvd., Orlando, FL 32816, USA

text-based interaction (Rzepka et al., 2022). The opportunity to harness CAs as an instructional technology expands the educational modalities through which personalized instruction, guidance, and support may occur. This opportunity is contingent upon effectively integrating CAs' unique technological affordances for engaging with individual learners and robust pedagogical designs driven by established learning theories.

This article aims to provide theory-driven guidance on designing CAs that facilitate PL experiences and describe specific design considerations for achieving such desired outcomes. We begin by providing an overview of CAs and research investigating the use of CAs in education. We then draw upon theories of multimodal learning and self-regulated learning (SRL) to discuss PL features that CAs can leverage, such as providing multiple instructional resources and paths, enhancing on-demand access to preferred learning resources, and offering customized guidance. For example, when learners first access the CA, they are prompted to answer questions related to their preferences (e.g., goals, preferred work times, interests, strengths). Next, we detail the process of designing a CA experience by embedding pedagogical practices and personalization features from two aspects: (a) predesign of instructional materials and resources and (b) design for learner interactions with CAs. We conclude by offering implications for advancing research on design and discussion of the possibilities of this technology moving forward.

CAs and Related Research

Mortensen (n.d.) described CAs' fundamental function as the ability to use voice as the access point to invoke smart devices through CAs to make commands, seek information, and create interactive experiences. CAs are dynamic systems that simulate human conversation using language (Kocaballi et al., 2020). CA devices use cloud services to process voice commands that are then analyzed, prompting the system to respond appropriately to individual and unique interactions (Chung, 2019). The integration of CAs into almost every professional discipline and everyday function has increased awareness of artificial intelligence (AI) and speech recognition software to meet the needs of users for both professional and personal activities (Nguyen & Vo, 2018). Additionally, CA devices have become more integral in the home setting (Chkroun & Azaria, 2019), seamlessly interweaving into aspects of daily activities. The initial prototypes of CAs were restrained to responding to only basic one-word commands that limited user capabilities and experiences (Li et al., 2004). However, CAs have evolved to provide experiences that can offer nuanced and increasingly complex interactions.

CAs in Education

CAs can be applied across various learning contexts to simulate human instruction and play in multiple instructional roles, such as providing immediate and personalized feedback to learners (Sharma et al., 2019; Winkler & Roos, 2019). Utilizing AI techniques, CAs can consistently capture learner progress monitoring data more efficiently and effectively that teachers are unable to, with commands prompting specific responses focused on the individual learner's capabilities (Akyuz, 2020). These data and prompts allow CAs to provide structured feedback, assist learners with tasks, adapt learning tasks and materials to learner preferences, and support conversation abilities (Kulik & Fletcher, 2016; Ma et al., 2014; Rus et al., 2013). Winkler and Roos (2019) described the pedagogical benefits to adopt CAs into practice as (a) adapting the devices to support different contexts to meet the individual needs of the learner, b) allowing teachers the autonomy to create relevant interactions without relying on developers, and (c) learners' gaining familiarity with the devices from continued exposure.

In an early study, Baylor and Kim (2005) investigated and validated the effectiveness of three instructional roles (i.e., expert, motivator, and mentor) that pedagogical agents played in improving college students' learning. Over the past decades, research on pedagogical agents have made significant progress, providing insights into the design and applications of CAs by leveraging more recent technological advances. In particular, advances in natural language processing and speech recognition technology allow for natural verbal interactions between individual learners and agents, thus better simulating human-to-human conversations (Kim & Baylor, 2016). In a more recent study conducted in a high school and a vocational business school, Winkler et al. (2021) found a greater improvement in problem-solving skills for students who used Alexa-based CAs to complete assignments than their peers using traditional paper-based methods. In other studies, researchers created and evaluated CAs deployed on Google Home that engage young children in joint reading by asking questions, providing feedback, and adapting scaffolding (e.g., Xu et al., 2021; Xu & Warschauer, 2020).

Several studies investigated teachers' perceptions of using CAs to support student learning in such ways as completing homework and acquiring knowledge (e.g., Dousay & Hall, 2018; Incerti et al., 2017; Jean-Charles, 2018). For example, a study surveying in-service teachers' experiences of integrating Alexa Echo in the classroom showed that a successful CA integration could help teachers improve classroom practices and yield positive student learning outcomes (Dousay & Hall, 2018). Additionally,

Incerti et al. (2017) found that a high percentage of surveyed pre-service teachers (83%) would consider utilizing CAs on Alexa Echo in their classroom. However, these teachers indicated multiple challenges related to CA use, such as technical issues, students' misuse of CAs devices, and lack of capacity to design CA-based instruction. To extend the challenges that educators may face, data privacy compliance must be addressed in any use of classroom CAs. The Family Educational Rights and Privacy Act (FERPA) contains pointed language on the rights of children and families regarding the data that is collected and how that data is stored and analyzed. The protections brought forth by FERPA are critical to ensure that student privacy is protected and, therefore, should be central to any design related to CAs.

The Potential of CAs for Learners with Diverse Learning Needs

Previous research investigated the usage of CAs for diverse learners, including young children (e.g., Lovato & Piper, 2019), English language learners (e.g., Xu et al., 2021), and individuals with visual impairments (e.g., Jariwala et al., 2021), learning disabilities, or any difficulty using a keyboard/mouse (e.g., Marvin, 2020; Lister et al., 2020). For example, research showed that voice invocation has made it easy for young children, particularly those who are not fluent readers or writers, to utilize devices to seek information (Lovato & Piper, 2019). In reviewing related literature, Lovato and Piper (2019) recognized the value of CAs in alleviating some technical obstacles young children faced when seeking information through search engines that required typing and spelling skills. Moreover, Jariwala et al. (2021) developed an intelligent CA system that could provide personalized responses for students with visual impairments when learning new mathematical concepts and engaging in self-directed learning. Importantly, Lovato and Piper (2019) note that while students could benefit from CA-based instruction, more research is needed to investigate skills and knowledge necessary for meaningful interactions with CAs. For example, the skill to ask CAs questions and problem-solve if CAs do not recognize the response has potential impact on interaction experiences for students.

The existing research revealed preliminary evidence on the effects and potentials of utilizing CAs to support learning experiences for learners with diverse needs. However, challenges emerging from the literature suggested that more research is needed to investigate the design of CA-based learning from both technical and pedagogical perspectives. On the technical side, the rapid development of CA devices and voice-building apps (e.g., Alexa Skill Blueprints, Voiceflow) has made it easy for teachers or other stakeholders

with limited training to create CAs without coding required (Emerling et al., 2020; Winkler & Roos, 2019). While these blueprints offer easy access to creating interactive skills, they only produce one to one spoken skills with limited functionality. On the pedagogical side, there is a lack of clear guidance on how to design CAs given that research on CA-based instruction is still in its infancy. To fully leverage the potential of CAs to provide PL for diverse learners, the design and implementation should be grounded in human learning theories, instructional design principles, and pedagogies (Xu & Warschauer, 2020). To fill this gap, we provide guidance on combining pedagogical practices with PL features when designing CAs in the following sections.

PL Definition and Research

Understanding the concept of PL serves as an initial step of designing CAs to ensure and increase personalization for diverse learners. Researchers from across disciplines have defined or applied PL as emerging technologies (e.g., Chen, 2009), instructional approaches (e.g., Walkington & Bernacki, 2015), or systematic learning designs (e.g., Zhang et al., 2020a, b, 2022). In this article, we refer to Zhang et al. (2020a, b, 2022) definition of PL as a systematic learning design that focuses on tailoring instruction to individual students' strengths, needs, preferences, interest, prior knowledge, and goals that leads to well-rounded educational experiences including increased access to disciplines and twenty-first-century work skills. This PL definition indicates the complexity of PL implementation for individual learners. A large body of research tied PL closely to adaptive learning systems, intelligent tutoring systems, ubiquitous learning systems, and robotics (Xie et al., 2019). These technologies with varying features afford the ability for learners to engage in learning activities customized to their needs in relation to cognition, metacognition, motivation, and affect (e.g., Arroyo et al., 2014). Focusing on different functionalities, researchers applied varied learning and instructional design theories to the development and implementation of PL.

Most current PL technologies have functionalities to guide learners through each step of a learning process by providing immediate hints and feedback, the design of which were grounded in AI concepts and cognitive theories (Kulik & Fletcher, 2016). For example, Chen (2009) developed a personalized intelligent tutoring system with SRL-assisted mechanisms based on Zimmerman's (2000) cyclical model of SRL theory. Mayer's (2005) cognitive theory of multimedia learning (CTML) was used to guide the design of adaptive learning systems that support learners in processing information through multiple modalities (e.g., Arroyo et al., 2014). Several ubiquitous learning systems were developed

based on Vygotsky's (1978) social constructivism theory that enabled learners to interact with authentic learning environments to enhance meaningful learning (e.g., Hwang et al., 2012). Development and applications of these PL technologies exemplify the translation of classical educational theories to evolving learning contexts, providing implications for CA designs.

Instructional Design and Learning Theories Guiding CA Designs

As an emergent technology, CAs have functionalities that can be leveraged to facilitate varying aspects of PL. In a recent study, Winkler et al. (2021) found that CA provided PL experiences by allowing students to learn at their own pace and receive individual guidance. Drawing upon constructivist learning theories, Winkler and colleagues explained that students benefited from interactions with CAs as individual coaches who provide dynamic scaffoldings and step-by-step problem-solving guidance. To advance the conversation around the potential of CAs to facilitate PL, we draw upon CTML (Mayer, 2005; Moreno, 2005) to discuss the design of CAs. Additionally, we adopt Zimmerman's (2000) cyclical model of SRL theory to demonstrate the affordance of CAs to provide personalized guidance on individual learning processes.

Cognitive Theories of Multimodal Learning

According to Mayer's CTML (2005), humans possess a dual-channel, limited-capacity, and active information processing system. Effective instructional design, therefore, needs to consider learners' dual channels for visual/pictorial and auditory/verbal information processing. Each channel is assumed to have limited processing capacity, making it critical to chunk instructional information, build connections among pieces of incoming information, and connect to existing knowledge. Learners actively engage in cognitive processing to understand and organize incoming information into a coherent representation of learning experiences. Thus, multimedia instructional materials should be designed to guide appropriate information processing without overloading learners' cognitive systems.

Currently, many CA devices do not have screens to support visual engagement. However, new iterations of CA devices have large screens and developer features that allow for multimodal representations of information. As such, guided by CTML, CAs can combine verbal and non-verbal representations of knowledge to prime learners' auditory and visual information processing channels. Expanded on CTML, Moreno's (2005) cognitive-affective theory of learning with media (CATLM) provides multiple instructional

design principles that specifically apply to the design of agent-based multimedia learning. CAs present information in a conversational style, which aligns with the personalization aspect of instructional design principles supported by CATLM. Based on CATLM, Moreno and Mayer (2007) highlighted the importance of embedding interactivity into multimedia learning environments that facilitate multidirectional communication and guide learners' active cognitive processing. They suggested creating opportunities for learners to ask questions and receive answers (dialoguing), determine the pace and/or order of the learning content presented in segments (controlling), control aspects of presented information (manipulating), seeking information through multiple options (searching), and select from various available sources to determine the learning content (navigating).

The above-mentioned instructional principles and considerations provide guidance on how to design CAs. Some concrete examples will be described in the following sections. However, it is important to acknowledge the debatable effects of agent-based technologies according to cognitive load theory (Sweller et al., 1998). In guiding the design of pedagogical agents, Louwerse et al. (2009) argued that multiple sources of information might add to an extraneous cognitive load to learners when these sources convey similar information. On the other hand, the researchers suggested that well-designed agents could reduce cognitive load by directing students toward specific tasks and resources as well as by providing multiple modalities that reinforce each other to generate a modality effect. Additionally, previous research on CAs indicated that if users were allowed to end the conversation at any time when interacting with CAs, they would be less likely to experience high cognitive loads (Rzepka et al., 2022). Thus, these design elements need to be considered when designing CAs to facilitate individual learning experiences.

Self-Regulated Learning

As a classic learning theory, SRL provides an avenue to understand individual learners' cognitive, motivational, and emotional aspects of learning (Panadero, 2017). Effective learning occurs in a structured learning environment that minimizes the impact of cognitive load by supporting learners in developing SRL skills (Kirschner, 2002). With structured support, students can utilize SRL skills to better allocate cognitive resources to the learning tasks (Park et al., 2015). This is especially important for technology-enhanced PL environments embedded with flexible activities and multimodal materials (Basham et al., 2016).

According to Zimmerman's (2000) cyclical model, SRL includes three phases of the metacognitive process: forethought, performance, and self-reflection. In the forethought phase, learners engage in task analysis (e.g., goal setting,

strategic planning) and activation of motivational beliefs (e.g., self-efficacy, outcome expectancies, intrinsic interest, and goal orientations) that influence the use of learning strategies. In the performance phase, learners perform the task, use self-control strategies (e.g., imagery, time management, help-seeking), and monitor progress to keep cognitively engaged and motivated. In the self-reflection phase, learners self-evaluate how they have performed the task and make attributions of performance (e.g., success, failure) to perceived causes. Additionally, learners generate self-reactions, such as self-satisfaction and adaptive or defensive responses that can positively or negatively influence future task performances.

Researchers have utilized SRL to theoretically underpin the development of PL technologies (e.g., Chen, 2009; Desai & Chin, 2020). For CAs, the AI-based conversational interface can be programmed with prompts and structure for learners to practice SRL skills when interacting with the CA. Desai and Chin (2020) discussed the feasibility of implementing multiple SRL strategies in CAs, such as providing hints, motivational prompts, feedback, teach-back, and gauging deep questions. These strategies are designed by leveraging CAs' functionality of consistently capturing and analyzing learner progress monitoring data. Based on these data, CAs can provide scaffolds and feedback tailored to individual learners. We recommend that CA designers harness this functionality and embed prompts that guide learners in setting goals, performing tasks to achieve the goals, and self-evaluating performance to practice their SRL skills. Detailed design examples aligned to SRL will be discussed in the following sections.

The design considerations grounded in CTML/CATLM and SRL can serve as a starting point for connecting personalization design principles with the current functionalities afforded by CAs. CTML/CATLM as instructional design theories offer a lens into how to design CAs that supports individual learners' information processing; SRL as a learning theory provides guidance on embedding supports and structures for improving individual learners' self-regulatory actions when interacting with the CA. While learning theories deepen the knowledge about how to design CAs, it is critical to highlight how they transfer to or undergird pedagogical practices delivered by CAs to facilitate PL across evolving learning environments.

Identifying Pedagogical Practices for Learners with Diverse Learning Needs

There is an array of pedagogical practices, such as explicit instruction and inquiry-based learning, that were developed based on learning theories and proven to be effective in improving student learning across content areas and contexts (e.g., Pedaste et al., 2015; Rupley et al., 2009). Previous

research suggests that learners vary in pre-existing knowledge, cognitive capacities, and metacognitive skills; thus, it is important to provide personalized support and guidance for learners across learning contexts (Cantor et al., 2018). In a brick-and-mortar classroom setting, teachers select and implement specific pedagogical practices that address the needs of learners. The educational decisions that teachers make are then delivered by a structure for providing direct, explicit, and systematic guidance which is critical for empowering learners to succeed in varying contexts, especially in online learning environments (Alfieri et al., 2011; Kirschner et al., 2006).

In a brick-and-mortar environment, teachers can break down complex skills (e.g., math concepts, literacy strategies), model learning processes, provide graduated prompts or feedback, allow for self-monitoring, and apply other explicit instruction practices to facilitate student learning (Hughes et al., 2017). For example, research has substantiated that explicit instruction could help improve mathematical learning outcomes (Rupley et al., 2009; Woodward et al., 2018). These studies emphasize the importance of explicitly teaching math words, visually illustrating connections of math concepts, modeling the use of strategies and tools, and guiding learners in monitoring learning progress. However, there are multiple challenges pertaining to modeling, step-by-step guidance, and monitoring in traditional face to face (f2f) settings. For example, classroom group instruction is usually delivered at a brisk pace, which may fall short of meeting individual learner needs, such as differing needs for time, resources, guidance, or other instructional support (Archer & Hughes, 2010).

Another challenge may arise when step-by-step guidance or modeling diminishes in an environment without physical teacher presence, such as online learning settings (Carter et al., 2020). To address these challenges, CA provides the possibility of translating proven pedagogical practices into a technology-based learning setting where immediate teacher support is absent and more personalization can be embedded in individual learning processes. Specifically, the CA can simulate teacher-led instruction by modeling the learning process and providing personalized feedback for learners with diverse needs when learning independently in the classroom or at home.

An Illustration of how to Design CAs for PL

Before creating CAs, it is critical to understand local and state laws that address student privacy especially in online settings. As of 2019, child advocate groups have brought attention to the concern about violations against the Child's Online Privacy Protection Act (COPPA) when data on children's interactions with smart speakers was recorded

(Kaufman, 2022). Further, some states have specific laws that place needed protections for processing data from children under 13. For instance, the Washington State Data Privacy Act forbids processing children's data without parental permission. Therefore, any potential benefit of CA-based instruction is bound to protections afforded to learners. On the pedagogical side, it is critical to carefully create content and arrange resources prior to CA use (i.e., predesign) and design prompts supporting learner interactions with CAs to maximize the potential of CAs. In this section, we demonstrate how CTML/CATLM and SRL guide the predesign and design for interactions of CAs, respectively.

Predesign of CAs

It is important to note that as of the writing of this manuscript, designers have the capability to create a fully functional CA that is contained within a developer platform (e.g., Alexa, Google Assistant). When designers want to incorporate other resources into a CA (e.g., YouTube), they must strategically plan for how resources can be integrated into the CA. Currently, interoperability is not seamless. This means that in order for learners to access external resources, they may need to exit the skill. This presents a challenge because learners will have to re-enter the skill once they have finished with the external resource.

Building a CA starts with creating a dialogue flow, which is a script illustrating the conversation between the learner and the CA. Given learner variability, designers of CAs are suggested to consider all supports needed for all learners to succeed in PL experiences. In a f2f classroom, teachers decide which skills, strategies, and concepts to be taught and match learning content to learner needs. After determining the content and considering learner characteristics, teachers break down complex learning skills and strategies into smaller segments, sequence skills logically, and provide learners with distributed and cumulative practices over time (Archer & Hughes, 2010). Similarly, carefully considering and arranging these instructional strategies and resources ensures content programmed into CAs is organized and supportive of learners with diverse learning needs.

Guided by CTML and CATLM, there are two major considerations for incorporating personalization features in the predesign of CAs. Like many PL technologies (e.g., Abawi, 2015; Looi et al., 2009), CAs can host various multimedia resources that provide flexibility for student learning. For example, a CA device with a screen (e.g., Alexa Echo Show) supports the delivery of instruction in multiple modalities that facilitate voice-, image-, and text-supported learning which helps learners seek information

through diverse options from various sources. This provides a way of personalizing content based on learner preferences for perceiving information and cognitive abilities to process the presented information (Mayer, 2005; Moreno, 2005).

Additionally, CAs can voice the connection of new content to previous content. If the learner is accessing the CA through a device with a screen, a visualization of the connections can be made available throughout the learning experience. A visualization, such as a graphic organizer, can show the relatedness of the content. This, paired with the CA talking through the connections highlighted on the screen, provides multiple ways in which the learner can make note of the relatedness of the content, be reminded of what content preceded the current learning, and what instruction will occur next.

The second personalization consideration for pre-designing CAs involves the development of multiple entry points and prompts guiding appropriate choices based on learners' prior knowledge. As discussed above, CATLM-guided instructional design principles indicate that learners process information better by engaging in dialogue as well as determining the pace and/or order of the learning content. Learners are guided through a self-paced learning process through dialoguing with the CA. To enhance personalization in the dialoguing process, a menu of options to learn components of a skill can be programmed into the CA. This design provides learners with autonomy over the order of learning content.

Interactions with CAs

Provided that the CA has been pre-designed with considerations for offering multiple options and resources, CA can guide individual learners through instruction step by step. Here it is critical to ensure that students have been prepared to interact with CAs. Students should be familiar with how to invoke a CA, how to ask questions, how to navigate through the choices offered in the CA, and how to stop a CA when they are finished or need a break. Preparing students to engage with CAs has the potential to ensure that learners follow the structure of CAs. Compared to traditional f2f settings, the individual guidance process provided by technology can be embedded with more opportunities for learners to practice self-regulation skills during forethought, performance, and self-reflection phases (Romero et al., 2019; Zimmerman, 2000). Based on SRL, we provide the following scenario for designing a CA. The authors note that although the current technological structures and capabilities allow for the creation of theoretical based CAs to exist, future advancements will provide opportunities to extend beyond the scope of this article.

Forethought Phase

After a learner invokes the CA, they hear a welcome message and introduction. To provide the highest level of PL, the learner will be asked several questions that will guide their interaction. These questions may include the learner's name, grade level, amount of time they plan to engage with the learning content, interests that could be used to provide rewards as learners interact with CAs and goals related to learning. This information will be captured as variables that support the learner by providing reminders of goals, alerts to time spent with the skill, and next steps. Many of these aspects align with Zimmerman's (2000) forethought phase of SRL in that learners will set goals and plan to act upon these goals.

After the variables have been captured, the CA will introduce the topic with a statement of the goal of instruction and learner expectations. In traditional f2f environments, teachers will acknowledge the aim of the lesson by stating the goals, projecting written statements, or talking individually with learners (Archer & Hughes, 2010). In addition, teachers may employ environmental reminders that learners can refer to during instruction. In CA-based learning, once the learner hears the expectations she or he can determine if they are prepared to begin the lesson. If they are not, the learners can ask the CA to take them back to review content that can promote learner success. CAs can provide verbal prompts that guide the learner through planning how to be successful with the learning activity. These prompts have the potential to provide meaningful structures that afford learners multiple opportunities to strengthen their forethought phase of SRL (Chen, 2009).

Once the learner has heard the goals and expectations, CAs can lead the learner through a review of prior skills that are critical for success with the new content. In a f2f setting, teachers can prompt learners to reflect on prior learning and use this time to connect previous learning to the content that will be covered during the lesson. This allows learners to choose how they want to review the content. For instance, the learner will hear a guided reflection question to engage their previous understandings of crucial content. When interacting with the CA, learners review skills with prompts to self-assess prior knowledge. This allows learners to determine if they are confident in their ability to move forward with instruction, or if they need to revisit pre-requisite skills. If they request more information to support their understanding, they can select to be taken to a video, hear a podcast, or send documents to their email for review. This feature assures that learners have multiple ways to review content and that instruction is supported beyond CA reading a script to the learner (Mayer, 2005).

Performance Phase

Providing clear step-by-step directions is critical for learning to occur. In a traditional f2f setting, teachers can model problem-solving by performing the skill while talking aloud the importance of each step and how they are making decisions within each step. To facilitate learning in online settings where immediate teacher support is absent, CAs can model individual problem-solving processes (Winkler et al., 2021). With previously collected learner characteristics variables, this modeling could fold in aspects of learner preferences including the type of media that is offered to the learner to solve the problem.

In f2f settings, the process of modeling is often extended to include Gradual Release of Responsibility (GRR; Fisher & Frey, 2013). With GRR, the teacher will first model the skill independently, then work with the learners to perform the skill together, then ultimately have the learner perform the skill on their own. The GRR strategy can be employed in CAs by the learner selecting how they want the skill modeled through a preset list of options (i.e., YouTube, visuals on screen accompanied by voice, hard copy to email). In the case of choosing YouTube videos, learners will need to exit the skill and re-invoke the skill to come back. As CA technology advances, challenges like lack of interoperability between platforms will need to be addressed. After the skill has been modeled, CAs can guide the learner through the skill and offer suggestions based on learner responses. Finally, CAs can offer a prompt that asks the learner to complete the skill independently.

Additionally, to assist the learner in contextualizing new information, CAs can provide a range of examples and non-examples of using strategies to learn the skill. This range of (non-)examples supports the learner in developing understandings of when, and when not, to apply the content being learned. These modeling features align to the performance phase of SRL by promoting access to task strategies through verbal reminders to use strategies when the learner experiences challenges with the academic task.

Self-Reflection Phase

In order for learners to show their mastery of content, it is critical that they are provided multiple opportunities to practice that are woven throughout instruction (Archer & Hughes, 2010). While teachers and CA designers determine the essential skills that will drive instruction, they should also determine how to build in checks for understanding that verify the learner is ready to proceed. In a f2f classroom, teachers do this by prompting learners to answer questions related to the content. This may occur through choral responding, peer responses, or individual responses (Archer & Hughes, 2010).

When engaging with a CA, learners will have multiple occasions to show their learning as well as reflect on their performance. This feature of CAs can be personalized by offering learners tasks and paths based on their responses to checks for understanding. These prompts are crucial for understanding where learners are in the learning process as well as for generating the opportunities for CAs to make decisions on how to best support engagement and progress.

Using CAs, learners can be asked specific questions related to content. CAs may offer behavior-specific praise based on correct answers or support the learner with further instruction to reach the correct answer. In the f2f classroom, teachers are mediating learning for multiple learners, which makes it difficult to always maintain awareness of the instructional needs of each learner and provide immediate feedback based on learner performance. Therefore, with well-designed CA instruction, learners can receive increased levels of support that are based on their preferences, can be restated at any point, and are available on demand.

Feedback is essential to providing learners the support needed to maximize the benefit of instruction (Hattie & Timperley, 2007). To be most effective, feedback should be delivered immediately in order for the learner to make corrections to their thinking and apply new understandings to finish academic tasks (McLeskey et al., 2017). With CAs, the learner responds to prompts and receives immediate feedback. If the learner answers the prompt incorrectly, the CA can offer the learner support in correcting the error, including restating the prompt, modeling how to solve the problem, directing the learner to media to revisit the content, or taking the learner back to the previous section of the skill. Feedback from the CA can be used to guide the learners' reflection on their performance. This could occur through verbal prompts that structure learner reflection to include satisfaction with their performance as well as how the learner may adapt their approach to the next task.

Discussion

As an emerging technology, CAs are presenting new possibilities of personalizing learning experiences for all learners. In this article, we provided guidance on the design of CAs that integrate personalization features guided by CTML/CATLM and SRL. Well-designed CAs may harness PL features that provide tailored learning support and content for learners in a wide variety of settings, including online learning and f2f environments. Research on using CAs in education is emerging (e.g., Jariwala et al., 2021; Lister et al., 2020; Xu et al., 2021). However, more efforts are needed to investigate how CA informs PL experiences and enhances learning outcomes. Additionally, due to the variance of capabilities found in CA hardware (e.g., Alexa Echo Show with

screen, Alexa Echo Dot without screen), further research is needed to investigate how learners interact with different CA devices.

One major consideration that has emerged for further investigation is data privacy. Data privacy has been critically discussed with the increased introduction of CA devices in K-12 classrooms and home learning environments (Riddell, 2019). CAs can collect users' auditory data to provide relevant feedback, accordingly; therefore, the collected data are transferred to databases based on shared agreements made between company and users (Kelly & Statt, 2019). However, users often skip reading data agreements thoroughly, trusting that the devices are being used for educational purposes so their data are safe (Riddell, 2019). Therefore, teachers, parents, and learners will have to consider the use of CA devices and decide to what extent they will use them. Taking this into consideration, our recommendation is that CAs be accessed in online learning settings with the learner's family/caretaker present until further research has occurred on data privacy in f2f settings.

With the new emphasis on providing PL experiences to learners in online learning environments, new technologies are emerging. It is clear that no single technology will meet the diverse needs of all learners. The potential of CA lies in integrating personalization features and common pedagogical practices in a way that enhances optimal learning experiences for all learners. Although CA is in its infancy, advancements in terms of design, usability, and implementation are occurring rapidly. In order to maximize the potential of CA, the field should continually monitor growth in key areas such as developments in AI and data privacy as well as the impact of PL on learning. Further, research should continue on how CAs can support learners with diverse needs, including elements such as delivering voice-based instruction, guiding goal setting and monitoring, and transmission of meaningful feedback.

Authors' Contribution Richard Carter and Tiffany Hunt contributed to the study conception. All authors contributed to the literature review, the first draft of the manuscript, and revisions of the manuscript. All authors read and approved the final manuscript.

Data Availability and Code Availability No consent nor ethical approval was required.

Declarations

Ethics Approval This is a conceptual manuscript. No data was collected for this study; therefore, this study did not involve human participants, and no ethical approval is required.

Competing Interests All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

- Abawi, L. A. (2015). Inclusion “from the gate in”: Wrapping students with personalized learning support. *International Journal of Pedagogies and Learning*, 10, 47–61. <https://doi.org/10.1080/22040552.2015.1084676>
- Akyuz, Y. (2020). Effects of intelligent tutoring systems (ITS) on personalized learning (PL). *Creative Education*, 11(06), 953–978. <https://doi.org/10.4236/ce.2020.116069>
- Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103(1), 1–18. <https://doi.org/10.1037/a0021017>
- Archer, A. L., & Hughes, C. A. (2010). *Explicit instruction: Effective and efficient teaching*. Guilford Press.
- Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, 24, 387–426. <https://doi.org/10.1007/s40593-014-0023-y>
- Basham, J. D., Hall, T. E., Carter, R. A., Jr., & Stahl, W. M. (2016). An operationalized definition of personalized learning. *Journal of Special Education Technology*, 31(3), 126–136. <https://doi.org/10.1177/0162643416660835>
- Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, 15(2), 95–115. <https://doi.org/10.1007/s40593-015-0055-y>
- Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2018). Malleability, plasticity, and individuality: How children learn and develop in context. *Applied Developmental Science*, 1–31. <https://doi.org/10.1080/10888691.2017.1398649>
- Carter, R. A., Jr., Rice, M., Yang, S., & Jackson, H. A. (2020). Self-regulated learning in online learning environments: strategies for remote learning. *Information and Learning Sciences*, 121(5/6), 321–329. <https://doi.org/10.1108/ILS-04-2020-0114>
- Chen, C. M. (2009). Personalized E-learning system with self-regulated learning assisted mechanisms for promoting learning performance. *Expert Systems with Applications*, 36, 8816–8829. <https://doi.org/10.1016/j.eswa.2008.11.026>
- Chkroun, M., & Azaria, A. (2019). LIA: A virtual assistant that can be taught new commands by speech. *International Journal of Human-Computer Interaction*, 35(17), 1596–1607. <https://doi.org/10.1080/10447318.2018.1557972>
- Chung, K. S. (2019). Sensor based affective learning information process system for intelligent tutoring system. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/1284/1/012033>
- Desai, S., & Chin, J. (2020). An explorative analysis of the feasibility of implementing metacognitive strategies in self-regulated learning with the conversational agents. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 64, No. 1, pp. 495–499). <https://doi.org/10.1177/1071181320641112>
- Dousay, T. A., & Hall, C. (2018). Alexa, tell me about using a virtual assistant in the classroom. In T. Bastiaens, J. Van Braak, M. Brown, L. Cantoni, M. Castro, R. Christensen, G. Davidson-Shivers, K. DePryck, M. Ebner, M. Fominykh, C. Fulford, S. Hatzipanagos, G. Knezek, K. Kreijns, G. Marks, E. Sointu, E. Korsgaard Sorensen, J. Viteli, J. Voogt, ... O. Zawacki-Richter (Eds.), *Proceedings of EdMedia: World Conference on Educational Media and Technology* (pp. 1413–1419). Amsterdam, Netherlands: Association for the Advancement of Computing in Education (AACE). Retrieved July 6, 2023 from <https://www.learnedtechlib.org/primary/p/184359/>.
- Emerling, C. R., Yang, S., Carter, R. A., Jr., Zhang, L., & Hunt, T. (2020). Using amazon alexa as an instructional tool during remote teaching. *TEACHING Exceptional Children*, 53(2), 164–167. <https://doi.org/10.1177/0040059920964719>
- Fisher, D., & Frey, N. (2013). *Better learning through structured teaching: A framework for the gradual release of responsibility*. ASCD.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hughes, C. A., Morris, J. R., Therrien, W. J., & Benson, S. K. (2017). Explicit instruction: Historical and contemporary contexts. *Learning Disabilities Research and Practice*, 32(3), 140–148. <https://doi.org/10.1111/ldrp.12142>
- Hwang, G. J., Tsai, C. C., Chu, H. C., Kinshuk, K., & Chen, C. Y. (2012). A context-aware ubiquitous learning approach to conducting scientific inquiry activities in a science park. *Australasian Journal of Educational Technology*, 28(5), 931–947. <https://doi.org/10.14742/ajet.825>
- Incerti, F., Franklin, T. J., & Kessler, G. K. (2017). Amazon Echo: Perceptions of an Emerging Technology for Formal and Informal Learning. In Y. Baek (Ed.), *Game-based learning: Theory, strategies and performance outcomes*. NovaScience publishers.
- Jariwala, A., Marghitu, D., & Chapman, R. (2021). A multimodal platform to teach mathematics to students with vision-impairment. In *International Conference on Human-Computer Interaction* (pp. 109–117). https://doi.org/10.1007/978-3-030-78095-1_10
- Jean-Charles, A. (2018, March). Internet of things in education: Artificial intelligence voice assistant in the classroom. In *Society for Information Technology & Teacher Education International Conference* (pp. 883–885). Association for the Advancement of Computing in Education (AACE).
- Kaufman, S. (2022). The invisible yet omnipresent ear: The insufficiencies of the Children’s online privacy protection act. *NYU Ann. Surv. Am. L.*, 78, 101.
- Kelly, M., & Statt, N. (2019, July 3). *Amazon confirms it holds on to Alexa data even if you delete audio files*. The Verge. Retrieved March 15, 2020, from <https://www.theverge.com/2019/7/3/20681423/amazon-alexa-echo-chris-coons-data-transcripts-recording-privacy>
- Kim, Y., & Baylor, A. L. (2016). Research-based design of pedagogical agent roles: A review, progress, and recommendations. *International Journal of Artificial Intelligence in Education*, 26, 160–169. <https://doi.org/10.1007/s40593-015-0055-y>
- Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, 12, 1–10. [https://doi.org/10.1016/S0959-4752\(01\)00014-7](https://doi.org/10.1016/S0959-4752(01)00014-7)
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1
- Kocaballi, A. B., Quiroz, J. C., Rezazadegan, D., Berkovsky, S., Magrabi, F., Coiera, E., & Laranjo, L. (2020). Responses of conversational agents to health and lifestyle prompts: Investigation of appropriateness and presentation structures. *Journal of Medical Internet Research*, 22(2), e15823. <https://doi.org/10.2196/15823>
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42–78.
- Li, B., Li, L., Li, B., Sivalingam, K. M., & Cao, X. R. (2004). Call admission control for voice/data integrated cellular networks: Performance analysis and comparative study. *IEEE Journal on Selected Areas in Communications*, 22(4), 706–718. <https://doi.org/10.1109/jsac.2004.825987>
- Lister, K., Coughlan, T., Iniesto, F., Freear, N., & Devine, P. (2020). Accessible conversational user interfaces: considerations for

- design. In *Proceedings of the 17th International Web for All Conference* (pp. 1–11).
- Looi, C. K., Wong, L. H., So, H. J., Seow, P., Toh, Y., Chen, W., Zhang, B., Norris, C., & Soloway, E. (2009). Anatomy of a mobilized lesson: Learning my way. *Computers & Education*, *53*, 1120–1132. <https://doi.org/10.1016/j.compedu.2009.05.021>
- Louwerse, M. M., Graesser, A. C., McNamara, D. S., & Lu, S. (2009). Embodied conversational agents as conversational partners. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(9), 1244–1255. <https://doi.org/10.1002/acp.1527>
- Lovato, S. B., & Piper, A. M. (2019). Young children and voice search: What we know from human-computer interaction research. *Frontiers in Psychology*, *10*, 1–5. <https://doi.org/10.3389/fpsyg.2019.00008>
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, *106*(4), 901. <https://doi.org/10.1037/a0037123>
- Marvin, E. (2020). Digital assistant for the visually impaired. *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pp. 723–728. <https://doi.org/10.1109/ICAIIIC48513.2020.9065191>
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press.
- McLeskey, J., Barringer, M-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., Lewis, T., Maheady, L., Rodriguez, J., Scheeler, M. C., Winn, J., & Ziegler, D. (2017). High-leverage practices in special education. Council for Exceptional Children & CEDAR Center.
- Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 507–524). Cambridge University Press.
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, *19*(3), 309–326. <https://doi.org/10.1007/s10648-007-9047-2>
- Mortensen, D. (n.d.). *How to design voice user interfaces*. Interaction Design Foundation. Retrieved March 16, 2020. <https://www.interaction-design.org/literature/article/how-to-design-voice-user-interfaces>
- Myers, C. M., Furqan, A., & Zhu, J. (2019). The impact of user characteristics and preferences on performance with an unfamiliar voice user interface. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. <https://doi.org/10.1145/3290605.3300277>
- Nguyen, M. S., & Vo, T. L. (2018). Resident identification in smart home by voice biometrics. In *International conference on future data and security engineering* (pp. 433–448). Springer.
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422.
- Park, B., Flowerday, T., & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, *44*, 267–278. <https://doi.org/10.1016/j.chb.2014.10.061>
- Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T., ... & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61.
- Riddell, R. (2019, January 28). *FETC '19: 'Alexa, tell me what you know about FERPA.'* EducationDive. Retrieved on March 15, 2020, from <https://www.educationdive.com/news/fetc-19-alexa-tell-me-what-you-know-about-ferpa/546957/>
- Romero, L., Saucedo, C., Caliusco, M. L., & Gutiérrez, M. (2019). Supporting self-regulated learning and personalization using ePortfolios: a semantic approach based on learning paths. *International Journal of Educational Technology in Higher Education*, *16*, 1–16.
- Rupley, W. H., Blair, T. R., & Nichols, W. D. (2009). Effective reading instruction for struggling readers: The role of direct/explicit teaching. *Reading & Writing Quarterly*, *25*(2–3), 125–138. <https://doi.org/10.1080/10573560802683523>
- Rus, V., D’Mello, S., Hu, X., & Graesser, A. (2013). Recent advances in conversational intelligent tutoring systems. *AI Magazine*, *34*(3), 42–54. <https://doi.org/10.1609/aimag.v34i3.2485>
- Rzepka, C., Berger, B., & Hess, T. (2022). Voice assistant vs. Chatbot—examining the fit between conversational agents’ interaction modalities and information search tasks. *Information Systems Frontiers*, *24*(3), 839–856. <https://doi.org/10.1007/s10796-021-10226-5>
- Sharma, K., Papamitsiou, Z., & Giannakos, M. (2019). Building pipelines for educational data using AI and multimodal analytics: A “grey-box” approach. *British Journal of Educational Technology*, *50*(6), 3004–3031. <https://doi.org/10.1111/bjjet.12854>
- Sweller, J., Van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296. <https://www.jstor.org/stable/23359412>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Walkington, C., & Bernacki, M. (2015). Students authoring personalized “algebra stories”: Problem-posing in the context of out-of-school interests. *The Journal of Mathematical Behavior*, *40*, 171–191. <https://doi.org/10.1016/j.jmathb.2015.08.001>
- Winkler, R., & Roos, J. (2019). Bringing AI into the classroom: Designing smart personal assistants as learning tutors. *ICIS 2019 Proceedings*. https://aisel.aisnet.org/icis2019/learning_environ/learning_environ/10/
- Winkler, R., Söllner, M., & Leimeister, J. M. (2021). Enhancing problem-solving skills with smart personal assistant technology. *Computers & Education*, *165*, 104148. <https://doi.org/10.1016/j.compedu.2021.104148>
- Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendr, A., Koedinger, K. R., & Ogbuehi, P. (2018). *Improving mathematical problem solving in Grades 4 through 8: A practice guide (NCEE 2012-4055)*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wvc/Docs/PracticeGuide/MPS_PG_043012.pdf
- Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, *140*, 103599. <https://doi.org/10.1016/j.compedu.2019.103599>
- Xu, Y., & Warschauer, M. (2020). Exploring young children’s engagement in joint reading with a conversational agent. In *Proceedings of the Interaction Design and Children Conference* (pp. 216–228). <https://doi.org/10.1145/3392063.3394417>
- Xu, Y., Wang, D., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing Children’s reading with a conversational agent vs. a human partner. *Computers & Education*, *161*. <https://doi.org/10.1016/j.compedu.2020.104059>
- Zhang, L., Basham, J. D., & Yang, S. (2020a). Understanding the implementation of personalized learning: A research synthesis. *Educational Research Review*, *33*. <https://doi.org/10.1016/j.edurev.2020.100339>
- Zhang, L., Yang, S., & Carter, R. A., Jr. (2020b). Personalized learning and ESSA: What we know and where we go? *Journal of*

Research on Technology in Education, 52(3), 253–274. <https://doi.org/10.1080/15391523.2020.1728448>

- Zhang, L., Basham, J. D., & Carter, R. A., Jr. (2022). Measuring personalized learning through the Lens of UDL: Development and content validation of a student self-report instrument. *Studies in Educational Evaluation*, 72, 101121.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.