2024

Conversational Agents in a Game Context

Fiona Lermen
Universität des Saarlandes, German, fiona.lermen@gmail.com

Andreas Kilian
Universität des Saarlandes, German, andreas.kilian@uni-saarland.de

Conversational Agents in a Game Context

Research Paper

Fiona Lermen[1] and Andreas Kilian[1]

[1] Universität des Saarlandes, Department, Saarbrücken, Germany
{file00001, andreas.kilian}@uni-saarland.de

Abstract. Conversational Agents (CAs) are systems that use natural language to interact with users. These interactions do not always proceed smoothly, which negatively impacts continued usage. Games provide both entertainment and function as a bonding activity, creating intrinsic motivation. Accordingly, CAs and users connecting in a game context could solve awkward interactions. In this paper, we conduct a systematic literature review, differentiating game-based CA applications in different areas. Our research found that gameful CA applications are most commonly found in education, use gamified features or appear as in-game characters, and are designed for children or general audiences. The paper supplies promising implications for CA design and presents a research agenda, including directions for further studies on a common theoretical framework, CA behavior, interaction design, and other target audiences.

Keywords: Conversational Agents, Games, Gamification, Systematic Literature Review

1 Introduction

Conversational agents (CAs) are software-based systems that use natural language to interact with users (McTear et al., 2016). Their history stretches far back, with ELIZA (Weizenbaum, 1966), a chat program designed to mimic a psychotherapist, being the first of its kind. Since 2016, which has been termed "the year of the chatbot" (Dale, 2016), CAs have been increasingly popular. Today, virtual assistants like Amazon's Alexa or Apple's Siri are present in many people's homes and CAs have various application areas, e.g., in education, healthcare, or customer service (Feine et al., 2019). While early pattern-matching agents were limited in their responses, current large language models, which are capable of intelligent human-like answers, allow for open-topic conversations (Strohmann et al., 2023). One of the most recent developments is the rise of generative artificial intelligence, i.e., OpenAI's ChatGPT, which is capable of producing multiple languages as well as programming code and can simulate various moods and relationships as a conversation partner (Rajabi et al., 2023). Both the long history of CAs, as well as recent renewed interest within the research community, showcase the relevance of the field. However, interactions between CAs and human users are not always successful. Users struggle when CAs are limited in their responses,

19th International Conference on Wirtschaftsinformatik,
September 2024, Würzburg, Germany

making conversations challenging, frustrating, and distracting (Samrose et al., 2020). Furthermore, CAs who are unable to engage their users risk being forgotten about (Cho et al., 2019). Therefore, a more interesting design presents an opportunity to increase CA appeal and retain users' attention.

Games provide entertainment and are a popular pastime as well as bonding opportunity. Building on the intrinsic motivation created by games, the concept of gamification aims to create engaging experiences by incorporating game elements in non-game contexts, i.e., in education (Barata et al., 2013). One solution to the awkwardness of interactions between CAs and users could therefore be to facilitate meetings through a game context, creating more engaging CAs. Game-based CA applications already exist, i.e., the "CiboPoliBot" tutor, which teaches school children how to lead a healthy life through an interactive messenger-based game (Fadhil & Villafiorita, 2017) or "Anne", a CA able to entertain older adults with card or puzzle games (Stara et al., 2021). However, research on game-based CAs has focused on specific applications thus far and there is a lack of overview and theory across different areas. Broader insights are needed to design more interesting CAs, leveraging the intrinsic motivation sparked by games. Accordingly, this paper aims to extend the knowledge of the field by providing an overview of existing research as well as design recommendations and future research directions. This paper raises the following research question:
RQ: *Which conversational agents in a game context exist, how can they be categorized, and what is their impact on the user?*

2 Background Literature

2.1 Playing Games

Playing games is an integral part of human life and interactions. Games can be traced back several millennia and enjoy never-changing popularity, showcased by the multi-billion-dollar video game industry (Predescu & Mocanu, 2020). They promote engaging experiences and can create intrinsic motivation, which refers to people engaging in an activity for itself, rather due to external influences. Intrinsic motivation is sparked through the factors challenge, fantasy, and curiosity, all or some of which may be present in games (Malone, 1981). Additionally, playfulness (Serenko, 2008) and perceived fun (Igbaria, 1994) are significant factors in system adoption and increase users' enjoyment, which is why investigating interactions in a game context may supply promising insights for system design.

There are no universal definitions of "game" or "play", but rather multiple theories in play. Deterding, Dixon, et al. (2011) distinguish the domains of "playing", which has a more improvisational and explorative nature, and "gaming", which is more structured and features specific rules and goals. A related concept is "gamification", referring to the use of game elements in non-gaming contexts, to improve user experience and engagement (Deterding, Sicart, et al., 2011). One popular application area for gamified systems is education, i.e., students being able to earn badges or compete against each other to enhance learning and improve participation (Barata et al., 2013).

2.2 Conversational Agents

McTear et al. (2016) define CAs as software-based systems that interact with users through natural language. Starting from ELIZA (Weizenbaum, 1966), CAs have garnered much interest over the years and today, virtual assistants have become part of people's daily lifes. Current large language models are capable generating human-like answers, allow for open-topic conversations (Strohmann et al., 2023) and can produce multiple languages as well as programming code (Rajabi et al., 2023). There are many different application areas for CAs. In education, they enhance the learning experience and support teachers in a variety of subjects, i.e., science, safety, and the English language (Ruan et al., 2019). CAs also function as buddies or support systems, i.e., to children with autism (Cha et al., 2021). Businesses employ CAs in customer service to answer common requests and focus their human capital on solving unusual issues, increasing customer satisfaction while reducing operational costs (Kongthon et al., 2009).

Multiple determinants of CA usage have been examined before: Samrose et al. (2020) show that users prefer interacting with an empathetic CA over a non-empathetic CA during a boring car drive. Furthermore, humans respond socially to CAs, and an appropriate CA design can guide users' reactions (Feine et al., 2019). Users also have individual preferences towards CA behavior, i.e., some might be more comfortable with a humorous agent while others prefer serious behavior (Li et al., 2017). Preferences may vary depending on the context of the interaction. Promoting successful long-term adoption of CAs requires a deeper understanding of users' preferences and should take user-specific characteristics, such as age, into account (Riefle & Benz, 2021).

2.3 CAs in a Game Context

Common applications of CAs in a game context include the use of pedagogical agents, that use gamification to improve teaching and learning. Katchapakirin and Anutariya (2018) suggest "ScratchThAI", a CA that can be used in schools to teach students the programming language "Scratch", supplementing a lack of teachers. Students are motivated through gamification methods, i.e., by receiving points or badges for successful task completion. Benner, Schöbel, Süess, et al. (2022) review text-based pedagogical CAs while Khosrawi-Rad et al. (2023) offer an extensive review of gamified pedagogical CAs targeted at adults. This demonstrates that the idea of using gamification to enhance pedagogical CAs has gained traction in research. Within healthcare and assistance, CAs can act as caregivers or buddies for the elderly. Stara et al. (2021) propose a CA called "Anne", which can provide medication reminders, play music and engage users with card or puzzle games, which unfolded to become one of the most used features. Within entertainment, CAs enhance games by enabling open dialog: for example, the user could play as a barkeeper and interact with non-playable character (NPC) CAs, increasing immersion (Aljammaz et al., 2020). Thus, multiple application areas for CAs in a game context exist already. However, a scoping review revealed that they have not been overarchingly considered beyond the area of education yet. Examining CAs in a

game context provides the opportunity to derive design recommendations for enhancing engagement. As a first step to this goal, this paper presents an overview of the field, showcases different applications of CAs and derives a research agenda.

3 Method

To answer the research question and provide a comprehensive overview, a systematic literature review was conducted, based on the approaches proposed by Webster and Watson (2002), Kitchenham et al. (2009), as well as vom Brocke et al. (2009, 2015). The review provides solid foundations for further research (Webster & Watson, 2002).

3.1 Search Strategy

To gather relevant literature, a suitable search string was developed and applied to scientific databases. The string was derived from the research question and consists of the two main aspects of the topic: CAs and playing games. The full search string is presented in Table 1.

Table 1. Search String

Conversational Agent		Playing Games
"conversational agent*" OR "digital agent*" OR "virtual agent*" OR "conversational interface" OR "chatbot*" OR "chat bot*" OR "chatterbot*" OR "smart assistant*" OR "smart personal assistant*" OR "virtual assistant*" OR "virtual personal assistant*" OR "digital assistant*"	AND	"game*" OR "playful" OR "gameful" OR "gamification" OR "gamified"

The string was applied within five major databases, which are ACM Digital Library, AIS eLibrary, IEEE Explore, Wiley Online Library, and EBSCOhost. The research focused on two different types of outlets: Peer-reviewed journal articles were included due to their high quality while the current rejuvenation of the field made the inclusion of conference proceedings reasonable and necessary (vom Brocke et al., 2015).

3.2 Paper Selection Strategy

To achieve an objective selection of papers and ensure the quality and validity of the findings, selection criteria were applied. The criteria are presented in Table 2.

Table 2. Filter Criteria

Inclusion	Exclusion
Peer-reviewed journal article or conference proceeding written in English.	The CA does not match the selected definition.
The publication features a specific CA interacting with the user in a game context.	The full text is not available.
	No information on the type of CA is provided.
There is a discussion of the interaction and its impact.	The game is not examined in detail.

Only publications that featured CAs in a game or game-adjacent context, i.e., storytelling, were included. Gamified CAs were only included if they featured multiple elements of gamification, such as feedback, badges, or progress bars (Schöbel et al., 2020). Furthermore, works that merely featured games as a backdrop to other research objectives were excluded. After applying the search string to the scientific databases and using basic filter criteria, such as publication type and language, the remaining results were screened for eligibility. First, the title and abstract were examined, and, barring suitability, a full-text review was conducted under the application of the selection criteria, in line with the approach by Brereton et al. (2007). An additional backward search, as recommended by Webster and Watson (2002) was carried out as well. The search and selection process is presented in Figure 1. The search was conducted in the fall of 2023 and yielded 5,045 initial results in total, which were narrowed down in a three-step process. The reference management program Mendeley was used to keep track of the results. In the end, a final set of 33 papers was included in the review.

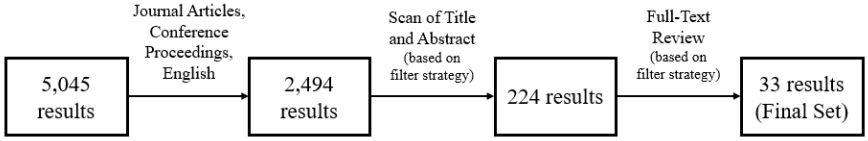

Figure 1. Search Process

3.3 Data Extraction Strategy

All papers of the final set were analyzed and important data was extracted, including bibliographical information such as title, author, year, and publication type. We analyzed the mode of communication such as text-based, voice-based, or mixed (Gnewuch et al., 2017), and whether a Wizard-of-Oz setting was applied. Furthermore, the application was categorized into the areas education, edutainment, entertainment, and health, which were derived from the results in an inductive approach. We also sorted publications by their domain, and differentiated between exploratory and open settings referred to as "playing" and rule-based settings referred to as "gaming" (Deterding et al., 2011). Furthermore, we analyzed the target audience of the CAs, as well as their impact, referring to the effects observed through the CAs' use.

4 Results

4.1 Bibliographical Information and Agent Type

The publications included in the review were released between 2008 and 2023, with a sharp uptick in activity during the past five years. Conference proceedings were the dominant publication type, with 28 out of 33 results compared to just five journal articles. There is an even distribution of different CA types, with 12 out of 33 results being purely text-based, 11 being voice-based, and a further 10 mixed-type results featuring a combination of voice, text, or interface-based interactions. Furthermore, four applications were tested in a Wizard-of-Oz setting.

4.2 Application Area

Figure 2: Application AreasFigure 2 provides a graphical overview of the categories: "education" features the most results, with a total of 18 publications, while "edutainment" counts six. The "entertainment" category encompasses seven results and the final two works belong to the "health" category.

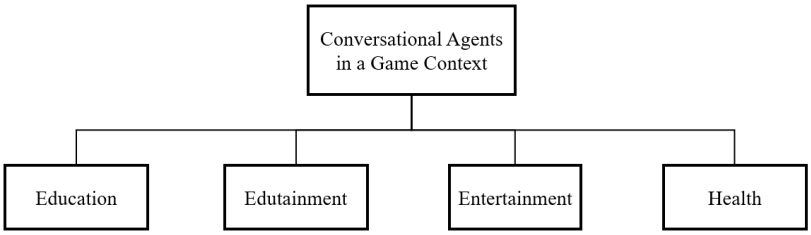

Figure 2: Application Areas

Education CAs in education support teaching or learning, both classroom-based and self-directed. They also enhance early learning and allow young children to acquire early literacy skills or language skills through activities such as storytelling. For example, Tewari and Canny (2014) tested a question-answering CA, which engages children in a picture and question game in a Wizard-of-Oz setting, supporting early literacy development and communication skills. The CA successfully engaged the children and even outperformed a human conversation partner. CAs can also be used to support storytelling activities between parents and children: "Floppy" (Lin et al., 2022) and "StoryBuddy" (Zhang et al., 2022) are examples of CAs that enhance the reading experience by either interjecting with questions or providing suitable questions to parents. CAs can also be used to induce creativity, i.e., by guiding children through crafting an object and then helping them develop a fitting story starring their character, to reflect on the activity (Hubbard et al., 2021). Another application is CA "Rachel", designed to help autistic children learn emotions via picture games and stories, guiding the children to describe and choose suitable emotions for a given scenario (Mower et al., 2011).

In general, educational CAs feature a variety of subjects, such as computational thinking and programming: "TurtleTalk" allow children to grasp basic coding concepts like loops by guiding a CA via a voice user interface (Jung et al., 2019). Dietz et al. (2023) combine programming concepts with literacy education and leverage both stories and drawings to create a personally meaningful learning experience. Katchapakirin and Anutariya (2018) propose "ScratchThAI", a CA designed to teach the block-based language "Scratch", making up for a shortage of teachers versed in the subject. The CA would use gamification, assigning missions and giving out points and badges as rewards to engage students in learning. Another example of a gamified CA is "CiboPoliBot" (Fadhil & Villafiorita, 2017), which teaches children about leading healthy lifestyles through a mini-game in which they must select healthy food options while limiting food waste. Students earn points depending on their performance and compete with their classmates, creating additional motivation to engage with the material. CAs can also be used to teach basic mathematical concepts, such as shapes and patterns (Xu et al., 2023). Finally, there are general approaches such as quiz games, which can be applied to a variety of subjects, i.e., to support an online Java programming course (Aguirre et al., 2018). Benner, Schöbel, and Süess (2022) create a gamified pedagogical CA and found that it offered an enjoyable self-study experience. Graesser et al. (2014) showcase learning in a trialogue setting with the CA "AutoTutor": users are guided by a CA tutor and compete against another CA student.

Edutainment The "edutainment" sub-category was introduced for publications combining elements of entertainment and education equally and features serious games for different subjects, i.e., cultural heritage (Mori et al., 2013), language learning (Sagae et al., 2010), coastal safety (Bellotti et al., 2011), and health (Gonzalez-Rodríguez et al., 2020). Serious games are an attractive educational tool, due to their broad appeal. Gonzalez-Rodríguez et al. (2020) argue that they are a more effective form of learning for children than static online courses, due to being inherently entertaining. According to Mori et al. (2013), the learning effect of interacting with their CA, acting as an Italian 16th-century painter, is comparable to reading a text on the subject. Interface design ranges from VR (Bellotti et al., 2011) to a messenger-style application (Gonzalez-Rodríguez et al., 2020). CAs usually feature as NPCs within the game setting and interact with the player, providing information and guiding them through the game world. Neto et al. (2011) present "Solis'Curse", developed for the Pantheon Museum in Portugal and meant to enrich visitors' experiences by recapping their visit with a quiz game. Visitors were prone to revisit the exhibition to get the answers correct and paid more attention, making the most of their visit.

Entertainment The "entertainment" category encompasses recreational applications and primarily features CAs integrated into game worlds as NPCs, making games more interesting by going beyond pre-defined dialogue options (Fraser et al., 2018). For example, the player may act as a barkeeper and converse with a CA patron to gather knowledge about the game world and delve deeper into characters' backstories (Aljammaz et al., 2020). Augello et al. (2009) propose an in-game CA to provide decision support in games like SimCity, expanding on a classic hint system and adding to

the enjoyment. Within virtual reality (VR), players were found to prefer voice interaction over a graphical interface when conversing with an assistant character (Buchta et al., 2022). Next to in-game CAs, there are also real-world approaches: an early example is a tabletop CA offering to play a card game within a community center for the elderly (Gabrielli et al., 2008). CAs can also serve as playmates and companions for children, reading them books and playing games (Zhao & McEwen, 2022).

Health Represents the smallest category, grouping two CAs targeted at seniors suffering from dementia. Embodied CA "Anne", designed to facilitate patients' daily lives, includes a schedule and reminders, serves as a communication partner, and supplies entertainment through games, which were deemed an important feature by users (Stara et al., 2021). The second application, "Magic Brush", features two CAs guiding a user through a drawing-based dementia prevention game (Yang et al., 2022).

4.3 Domain

The "gaming" domain features prominently, encompassing 25 works, ranging from educational quizzes, over interactive tabletop card games, to virtual game worlds. The applications have a set scenario, making interactions more structured and rule-based. This domain also includes gamification-based applications, such as the use of badges or leaderboards (i.e., Katchapakirin & Anutariya, 2018). The eight works assigned to the "playing" domain encompass game-adjacent activities, such as storytelling (i.e., Lin et al., 2022), drawing (i.e., Yang et al., 2022), or crafting (Hubbard et al., 2021). Users, often children, were in control of the activity and had more influence on the progression of the interaction. For example, they could interject with questions during storytelling, making the activity more exploratory.

4.4 Target Audience

13 applications are targeted at children, featuring various sub-groups and age ranges, though mostly preschoolers between the ages two and six, or elementary school children, up until the age of 10, were addressed explicitly. All these works belong to the "education" category and among them, two publications featured dyads of young children and a parent interacting with the CA together, i.e., during storytelling activities (Lin et al., 2022). The remaining applications featured a single user interacting with one or multiple CAs. There are 12 general applications, which do not specify a target audience or age range, mostly found in serious games and entertainment, and four applications aimed at students. Four applications are targeted at older adults.

4.5 Impact

The impact of the various CA applications was overwhelmingly positive: across all areas, an increase in engagement could be observed. Within education, the learning experience was enhanced by integrating a CA and became more fun. Children's understanding of different subjects also increased, i.e., of programming concepts (Jung et al.,

2019) or mathematical language (Xu et al., 2023). Generally, users were able to achieve target learning outcomes. Children were successfully engaged by CAs, allowing them to bond (Lin et al., 2022), and CAs demonstrated their ability to support early literacy, language learning, and communication. Regarding entertainment, users also considered the CAs to be engaging and enjoyable, adding depth to the game experience (Aljammaz et al., 2020). According to Fraser et al. (2018), users spent more time talking to NPC-CAs, and rated the game experience as more immersive and enjoyable, if the CA was capable of processing emotions. Furthermore, users were generally comfortable in the presence of CAs. 10 of the CAs remained in the conceptual phase, meaning their impact could not yet be observed in interaction with a user.

5 Discussion

5.1 Implications

Education is the most prominent application area, which is somewhat surprising, as the initial expected context of games would be entertainment. "Engagement is known to be a key factor in student performance and achievement" (Ruan et al., 2020, p. 1), resulting in continuous efforts to design engaging learning experiences. Gamified CAs have emerged as promising tools towards this goal, which explains the popularity of educational applications. They are also comparatively well-researched, though other reviews have set a closer focus, i.e., on text-based CAs (Benner, Schöbel, Süess, et al., 2022) or gamified CAs for adults (Khosrawi-Rad et al., 2023). CAs offer a way to enhance education by designing a fun learning experience and can compensate for a lack of human teachers. Nevertheless, entertainment still features prominently in the results, mostly in the form of complex in-game NPCs. The category "edutainment" combines both aspects and features serious games starring CAs, simultaneously providing entertainment and education. There is a distinction to be made between gamified CAs making non-game contexts more fun, and in-game CAs used to provide support and create more immersive game experiences. Furthermore, CAs may offer game-adjacent activities like storytelling or crafting. However, there is a lack of applications that focus on CAs playing games with users, potentially due to the existing rich literature on human-CA interactions.

Voice-based CAs are especially popular in applications targeted to children, which can be attributed to younger children's reliance on speech. In general, voice is an accessible medium of interaction and does not require advanced technological skills, making it attractive to inexperienced users (Aguirre et al., 2018). Children between the ages of 2 and ten were the most frequent target audience next to general applications. However, there is a lack of research on teenagers, who might consider gamified approaches to learning overly juvenile, if they were designed for children. Furthermore, only a couple of applications focus on older adults, who are also interested in games and fun (Matulík et al., 2020).

The overwhelmingly positive impact of applications that went through a trial phase shows the potential of CAs in a game context. It is recommended that practitioners

leverage the intrinsic motivation provided by games in their CA design. They may either incorporate game elements or add games to the CAs' services to facilitate breaking the ice between users and CAs. The success of in-game CAs also provides implications for game designers, presenting them with an opportunity to create more immersive experiences and retain players' attention. While the gaming domain featured more prominently in the results, the CAs grouped in the playing domain were also well-received, showing that CAs can successfully support activities like storytelling. This provides implications for the design of companion CAs for children or those with special needs and shows that the inclusion of games or game elements may increase usage and decrease awkwardness. Considering children's inclination to bond with the CAs (Lin et al., 2022), use should be monitored, e.g., by a parent or guardian, and an ethical design needs to be pursued, to protect this vulnearable user group from manipulation (Chubb et al., 2022). Nevertheless, games will not singularly solve the problem of users ceasing to interact with CAs. If users do not see their benefits, they might relegate them to a fancy toy (Trajkova & Martin-Hammond, 2020). Thus, designers must also ensure that CAs offer inherent value.

5.2 Future Research

Multiple promising avenues for future research on CAs in a game context were uncovered during the review. First, they require a robust theoretical foundation, to capitalize on best practices and optimize both effectiveness and user satisfaction. Benner, Schöbel, Süess, et al. (2022) highlight a lack of in-depth knowledge or framework for gamified CAs, leading researchers to design without a proper theoretical foundation. Therefore, both gamified and in-game applications should be categorized according to the frameworks or design guidelines employed in their development (D1.1). In the absence of a common denominator, the feasibility of a unified framework, applicable across different areas, should be evaluated (D1.2), to foster a consistent approach.

For gamified CAs, further research is needed to contrast the impact of different game elements (D2.1), allowing for a derivation of more specific design recommendations. Applications could be categorized according to their elements of gamification (Khosrawi-Rad et al., 2023). Seering et al. (2020) investigated a CA's evolution in an online community: throughout the experiment, the CA adjusted its behavior from a quizzical toddler to an adult community member. This raises the idea of companion CAs for children developing alongside them, adjusting their behavior over time. Lin et al. (2022) also suggest the development of a CA as a same-aged peer to children. Similarly, it would be interesting to examine whether older adults prefer an evenly matched CA or one emulating a young grandchild (D2.2).

Our results reveal that further research is needed on CAs and users playing a game together (D3.1). While CAs offering games are well received, the games are often not subjected to a comprehensive analysis. Research in which a game merely serves as a backdrop to investigate other variables may be replicated with a focus on the game itself (i.e., Samrose et al., 2020). Users' behavior towards a CA partner and a human partner should be compared (i.e., Tewari & Canny, 2014) (D3.2), and interactions should be

examined over a longer period (Elgarf et al., 2021) (D3.3), to ensure that benefits do not stem from the novelty of the experience (Ruan et al., 2020).

The preferences of various demographic groups, such as teenagers (D4.1) or seniors (D4.2), should be explored in depth, as understanding their varying needs is crucial for inclusive CA design. Some applications already feature a parent-child dyad interacting with CA (i.e., Zhang et al., 2022). Further settings in which one CA interacts with multiple users concurrently could be examined, i.e., siblings simultaneously demanding the CAs' attention, or classroom settings (D4.3). While this review focuses on games and game-adjacent activities, further research is needed on other playful behaviors, such as humor (Ceha et al., 2021), or flirting, and users' perceptions of various playful actions should be contrasted across different areas (D5.1). Table 3 summarizes directions for future research and presents the reviewed publications they were developed from

Table 3. Research Agenda

Avenue	Directions for Future Studies
RA1: Theoretical Foundations	D1.1: Investigate and categorize the theoretical foundations on which CAs in a game context are built (based on Benner, Schöbel, Süess, et al. (2022)).
	D1.2: Propose a common theoretical framework (based on Benner, Schöbel, Süess, et al. (2022)).
RA2: CA Behavior	D2.1: Determine which game elements are especially impactful. (based on Khosrawi-Rad et al. (2023)).
	D2.2: Analyze users' preferences toward the CA's maturity level (based on Seering et al. (2020) and Lin et al. (2022)).
RA3: User-CA Interaction	D3.1: Investigate users and CAs playing games together (based on Stara et al. (2021) and Samrose et al. (2020)).
	D3.2: Compare users' interactions with a CA partner and a human partner (based on Tewari & Canny (2014)).
	D3.3: Examine the long-term impact of game interactions between a CA and a user (Elgarf et al., 2021).
RA4: Target Audience	D4.1: Conduct empirical studies on teenagers' preferences for CA interactions in a game context (based on Jung et al. (2019)).
	D4.2: Conduct empirical studies on seniors' preferences for CA interactions in a game context (based on Matulík et al. (2020)).
	D4.3: Examine interactions between multiple users and one CA. (based on Zhang et al. (2022)).
RA5: Playfulness	D5.1: Explore other playful CAs and users' reactions in different settings (based on Ceha et al. (2021) and Shani et al. (2022)).

Overall, this paper provides an overview of the current state of research on CAs in a game context. It identifies popular applications and denotes their overwhelmingly positive impact, highlighting the potential of the field. Multiple research gaps are presented, allowing it to serve as a starting point for further research.

5.3 Limitations

Although our analysis provides valuable insights, it is not without limitations. The review was conducted by one research team, which poses the risk of subjective judgment. Findings were discussed within the team of authors and selection criteria were strictly applied, to limit potential biases. The category "edutainment" was created due to the increased cohesiveness of the results contained within but the publications could also be grouped within the categories "education" and "entertainment". Overall, the categorization serves as an initial distinction, which should be refined through further research. In line with the defined methodology, works that only featured games as a setting for other research objectives were not considered for the review. Nevertheless, these works may provide valuable insights. Furthermore, this review only considers applications that intentionally include multiple gamified features. Similarly, Khosrawi-Rad et al. (2023) only featured CAs making use of two or more elements of gamification in their review. Finally, there is a limited number of databases included in this review, meaning some relevant publications might have been missed.

6 Conclusion

Interactions between CAs and users do not always proceed smoothly, which negatively impacts adoption and usage. One potential solution to this issue is to incorporate game elements in design, sparking intrinsic motivation and creating more engaging experiences. While there are a multitude of applications featuring CAs in a game context, the field had not been overarchingly examined before. This paper proposes an initial categorization of such applications into the areas of education, edutainment, entertainment, and health. The most prominent category is education, featuring gamified CAs supporting teaching and learning via quizzes and tutoring. Regarding entertainment, there are efforts to create more immersive game experiences by including in-game CAs as NPCs, which allows for more open dialogue and in-depth characters. There is also a distinction to be made between the more open and exploratory "playing" domain and the stricter and rule-based "gaming" domain. Furthermore, there are gamified CAs, in-game CAs and CAs taking part in games or game-adjacent activities, i.e., crafting. These CAs are overwhelmingly perceived positively and can increase user engagement, highlighting their potential. This paper contributes to existing research through its categorization of game-based applications of CAs and deriving a multi-faceted research agenda: A robust theoretical framework and a closer analysis of different game elements are needed. Additionally, the paper highlights previously neglected target audiences. Overall, CAs in a game context encompass a wide spectrum of applications and offer many opportunities in an evolving field of research.

References

Aguirre, C. C., Kloos, C. D., Alario-Hoyos, C., & Munoz-Merino, P. J. (2018). Supporting a MOOC through a Conversational Agent. Design of a First Prototype. *2018 International Symposium on Computers in Education (SIIE)*, 1–6. https://doi.org/10.1109/SIIE.2018.8586694

Aljammaz, R., Oliver, E., Whitehead, J., & Mateas, M. (2020). Scheherazade's Tavern: A Prototype For Deeper NPC Interactions. *International Conference on the Foundations of Digital Games*, 1–9. https://doi.org/10.1145/3402942.3402984

Augello, A., Pilato, G., & Gaglio, S. (2009). A Conversational Agent to Support Decisions in SimCity like Games. *2009 IEEE International Conference on Semantic Computing*, 367–372. https://doi.org/10.1109/ICSC.2009.74

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving participation and learning with gamification. *Proceedings of the First International Conference on Gameful Design, Research, and Applications*, 10–17. https://doi.org/10.1145/2583008.2583010

Bellotti, F., Berta, R., De Gloria, A., & Lavagnino, E. (2011). Towards a conversational agent architecture to favor knowledge discovery in serious games. *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, 1–7. https://doi.org/10.1145/2071423.2071444

Benner, D., Schöbel, S., & Süess, C. (2022, December 12). Towards Gamified Conversational Agents for Self-Regulated Learning in Digital Education. *ICIS 2022 Proceedings*. https://aisel.aisnet.org/icis2022/learning_iscurricula/learning_iscurricula/11

Benner, D., Schöbel, S., Süess, C., Baechle, V., & Janson, A. (2022, January 17). Level-Up your Learning – Introducing a Framework for Gamified Educational Conversational Agents. *Wirtschaftsinformatik 2022 Proceedings*. https://aisel.aisnet.org/wi2022/hci/hci/5

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. https://doi.org/10.1016/j.jss.2006.07.009

Buchta, K., Wojcik, P., Nakonieczny, K., Janicka, J., Galuszka, D., Sterna, R., & Igras-Cybulska, M. (2022). Modeling and optimizing the voice assistant behavior in Virtual Reality. *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 397–402. https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00086

Ceha, J., Lee, K. J., Nilsen, E., Goh, J., & Law, E. (2021). Can a Humorous Conversational Agent Enhance Learning Experience and Outcomes? *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3411764.3445068

Cha, I., Kim, S.-I., Hong, H., Yoo, H., & Lim, Y. (2021). Exploring the Use of a Voice-based Conversational Agent to Empower Adolescents with Autism Spectrum

Disorder. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3411764.3445116

Cho, M., Lee, S., & Lee, K.-P. (2019). Once a Kind Friend is Now a Thing. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 1557–1569. https://doi.org/10.1145/3322276.3322332

Chubb, J., Missaoui, S., Concannon, S., Maloney, L., & Walker, J. A. (2022). Interactive storytelling for children: A case-study of design and development considerations for ethical conversational AI. *International Journal of Child-Computer Interaction*, *32*, 100403. https://doi.org/10.1016/J.IJCCI.2021.100403

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, *22*(5), 811–817. https://doi.org/10.1017/S1351324916000243

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, 2425–2428. https://doi.org/10.1145/1979742.1979575

Dietz, G., Tamer, N., Ly, C., Le, J. K., & Landay, J. A. (2023). Visual StoryCoder: A Multimodal Programming Environment for Children's Creation of Stories. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3544548.3580981

Elgarf, M., Skantze, G., & Peters, C. (2021). Once Upon a Story: Can a Creative Storyteller Robot Stimulate Creativity in Children? *Proceedings of the 21th ACM International Conference on Intelligent Virtual Agents*, *21*, 60–67. https://doi.org/10.1145/3472306.3478359

Fadhil, A., & Villafiorita, A. (2017). An Adaptive Learning with Gamification & Conversational UIs: The Rise of CiboPoliBot. *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*, 408–412. https://doi.org/10.1145/3099023.3099112

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational Agents. *International Journal of Human-Computer Studies*, *132*, 138–161. https://doi.org/10.1016/j.ijhcs.2019.07.009

Fraser, J., Papaioannou, I., & Lemon, O. (2018). Spoken Conversational AI in Video Games. *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, *18*, 179–184. https://doi.org/10.1145/3267851.3267896

Gabrielli, S., Bellutti, S., Jameson, A., Leonardi, C., & Zancanaro, M. (2008). A single-user tabletop card game system for older persons: General lessons learned from an in-situ study. *2008 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems*, 85–88. https://doi.org/10.1109/TABLETOP.2008.4660188

Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards Designing Cooperative and Social Conversational Agents for Customer Service. *International Conference on Interaction Sciences*.

Gonzalez-Rodríguez, C., Herzog-Cruz, D., & Gonzalez-González, C. (2020). Conversational game to improve digital health literacy. *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality*, 677–681. https://doi.org/10.1145/3434780.3436606

Graesser, A. C., Li, H., & Forsyth, C. (2014). Learning by Communicating in Natural Language With Conversational Agents. *Current Directions in Psychological Science*, *23*(5), 374–380. https://doi.org/10.1177/0963721414540680

Hubbard, L. J., Chen, Y., Colunga, E., Kim, P., & Yeh, T. (2021). Child-Robot Interaction to Integrate Reflective Storytelling Into Creative Play. *Creativity and Cognition*, 1–8. https://doi.org/10.1145/3450741.3465254

Igbaria, M. (1994). An examination of the factors contributing to microcomputer technology acceptance. *Accounting, Management and Information Technologies*, *4*(4), 205–224. https://doi.org/10.1016/0959-8022(94)90023-X

Jung, H., Kim, H. J., So, S., Kim, J., & Oh, C. (2019). TurtleTalk: An Educational Programming Game for Children with Voice User Interface. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3290607.3312773

Katchapakirin, K., & Anutariya, C. (2018). An Architectural Design of ScratchThAI. *Proceedings of the 10th International Conference on Advances in Information Technology*, 1–7. https://doi.org/10.1145/3291280.3291787

Khosrawi-Rad, B., Grogorick, L., & Robra-Bissantz, S. (2023). Game-inspired Pedagogical Conversational Agents: A Systematic Literature Review. *AIS Transactions on Human-Computer Interaction*, *15*(2), 146–192. https://doi.org/10.17705/1thci.00187

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, *51*(1), 7–15. https://doi.org/10.1016/J.INFSOF.2008.09.009

Kongthon, A., Sangkeettrakarn, C., Kongyoung, S., & Haruechaiyasak, C. (2009). Implementing an online help desk system based on conversational agent. *Proceedings of the International Conference on Management of Emergent Digital EcoSystems*, 450–451. https://doi.org/10.1145/1643823.1643908

Li, J., Zhou, M. X., Yang, H., & Mark, G. (2017). Confiding in and Listening to Virtual Agents. *Proceedings of the 22nd International Conference on Intelligent User Interfaces*, 275–286. https://doi.org/10.1145/3025171.3025206

Lin, G. C., Schoenfeld, I., Thompson, M., Xia, Y., Uz-Bilgin, C., & Leech, K. (2022). "What color are the fish's scales?" Exploring parents' and children's natural interactions with a child-friendly virtual agent during storybook reading. *Interaction Design and Children*, 185–195. https://doi.org/10.1145/3501712.3529734

Malone, T. W. (1981). Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science*, *5*(4), 333–369. https://doi.org/10.1207/s15516709cog0504_2

Matulík, M., Vavrečka, M., & Vidovićová, L. (2020). Edutainment Software for the Pepper Robot. *Proceedings of the 2020 4th International Symposium on Computer Science and Intelligent Control*, 1–5. https://doi.org/10.1145/3440084.3441194

McTear, M., Callejas, Z., & Griol, D. (2016). *The Conversational Interface*. Springer International Publishing. https://doi.org/10.1007/978-3-319-32967-3

Mori, D., Berta, R., De Gloria, A., Fiore, V., & Magnani, L. (2013). An easy to author dialogue management system for serious games. *Journal on Computing and Cultural Heritage*, *6*(2). https://doi.org/10.1145/2460376.2460381

Neto, J. N., Silva, R., Neto, J. P., Pereira, J. M., & Fernandes, J. (2011). Solis'Curse - A Cultural Heritage Game Using Voice Interaction with a Virtual Agent. *2011 Third International Conference on Games and Virtual Worlds for Serious Applications*, 164–167. https://doi.org/10.1109/VS-GAMES.2011.31

Predescu, A., & Mocanu, M. (2020). A data driven survey of video games. *2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 1–6. https://doi.org/10.1109/ECAI50035.2020.9223203

Rajabi, P., Taghipour, P., Cukierman, D., & Doleck, T. (2023). Exploring ChatGPT's impact on post-secondary education: A qualitative study. *Proceedings of the 25th Western Canadian Conference on Computing Education*, 1–6. https://doi.org/10.1145/3593342.3593360

Riefle, L., & Benz, C. (2021). *User-Specific Determinants of Conversational Agent Usage: A Review and Potential for Future Research* (pp. 115–129). https://doi.org/10.1007/978-3-030-86797-3_8

Ruan, S., He, J., Ying, R., Burkle, J., Hakim, D., Wang, A., Yin, Y., Zhou, L., Xu, Q., Abuhashem, A., Dietz, G., Murnane, E. L., Brunskill, E., & Landay, J. A. (2020). Supporting children's math learning with feedback-augmented narrative technology. *Proceedings of the Interaction Design and Children Conference, IDC 2020*, 567–580. https://doi.org/10.1145/3392063.3394400

Ruan, S., Jiang, L., Xu, J., Tham, B. J. K., Qiu, Z., Zhu, Y., Murnane, E. L., Brunskill, E., & Landay, J. A. (2019). QuizBot: A Dialogue-based Adaptive Learning System for Factual Knowledge. *Conference on Human Factors in Computing Systems - Proceedings*, *13*. https://doi.org/10.1145/3290605.3300587

Sagae, A., Johnson, W. L., & Row, R. (2010). Serious Game Environments for Language and Culture Education. *Proceedings of the NAACL HLT 2010: Demonstration Session*, 29–32. https://doi.org/10.5555/1855450.1855458

Samrose, S., Anbarasu, K., Joshi Affectiva, A., & Mishra Affectiva, T. (2020). *Mitigating Boredom Using An Empathetic Conversational Agent*. https://doi.org/10.1145/3383652.3423905

Schöbel, S. M., Janson, A., Söllner, M., Lowry, P. B., Petter, S., & Leimeister, J. M. (2020). Capturing the complexity of gamification elements: a holistic approach for analysing existing and deriving novel gamification designs. *European Journal of Information Systems*, *29*(6), 641–668. https://doi.org/10.1080/0960085X.2020.1796531

Seering, J., Luria, M., Ye, C., Kaufman, G., & Hammer, J. (2020). It Takes a Village: Integrating an Adaptive Chatbot into an Online Gaming Community. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376708

Serenko, A. (2008). A model of user adoption of interface agents for email notification. *Interacting with Computers*, *20*(4–5), 461–472. https://doi.org/10.1016/j.intcom.2008.04.004

Shani, C., Libov, A., Tolmach, S., Lewin-Eytan, L., Maarek, Y., & Shahaf, D. (2022). "Alexa, Do You Want to Build a Snowman?" Characterizing Playful Requests to Conversational Agents. *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 1–7. https://doi.org/10.1145/3491101.3519870

Stara, V., Vera, B., Bolliger, D., Paolini, S., de Jong, M., Felici, E., Koenderink, S., Rossi, L., Von Doellen, V., & Di Rosa, M. (2021). Toward the Integration of Technology-Based Interventions in the Care Pathway for People with Dementia: A Cross-National Study. *International Journal of Environmental Research and Public Health 2021, Vol. 18, Page 10405*, *18*(19), 10405. https://doi.org/10.3390/IJERPH181910405

Strohmann, T., Siemon, D., Khosrawi-Rad, B., & Robra-Bissantz, S. (2023). Toward a design theory for virtual companionship. *Human–Computer Interaction*, *38*(3–4), 194–234. https://doi.org/10.1080/07370024.2022.2084620

Tewari, A., & Canny, J. (2014). What did spot hide? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1807–1816. https://doi.org/10.1145/2556288.2557205

Trajkova, M., & Martin-Hammond, A. (2020). "Alexa is a Toy": Exploring Older Adults' Reasons for Using, Limiting, and Abandoning Echo. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376760

vom Brocke, J., Niehaves, B., Riemer, K., Chair, M. H., & Simons, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. *ECIS*, 2206–2217.

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems*, *37*. https://doi.org/10.17705/1CAIS.03709

Webster, J., & Watson, R. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly 26(2)*, xiii–xxiii. https://www.jstor.org/stable/4132319

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45. https://doi.org/10.1145/365153.365168

Xu, W., Ma, J., Yao, J., Lin, W., Zhang, C., Xia, X., Zhuang, N., Weng, S., Xie, X., Feng, S., Ying, F., Hansen, P., & Yao, C. (2023). MathKingdom: Teaching Children Mathematical Language Through Speaking at Home via a Voice-Guided Game. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, *23*, 1–14. https://doi.org/10.1145/3544548.3581043

Yang, M., Lee, K., Kim, E., Song, Y., Lee, S., Kang, J., Han, J., Song, H., & Kim, T. (2022). Magic Brush: An AI-based Service for Dementia Prevention focused on Intrinsic Motivation. *Proc. ACM Hum.-Comput. Interact*, *6*(CSCW2), 21. https://doi.org/10.1145/3555549

Zhang, Z., Xu, Y., Wang, Y., Yao, B., Ritchie, D., Wu, T., Yu, M., Wang, D., & Li, T. J.-J. (2022). StoryBuddy: A Human-AI Collaborative Chatbot for Parent-Child Interactive Storytelling with Flexible Parental Involvement. *CHI Conference on Human Factors in Computing Systems*, 1–21. https://doi.org/10.1145/3491102.3517479

Zhao, Z., & McEwen, R. (2022). "Let's read a book together": A Long-term Study on the Usage of Pre-school Children with Their Home Companion Robot. *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 24–32. https://doi.org/10.1109/HRI53351.2022.9889672