

RESEARCH

Free and Open Access

# Shared storytelling with a virtual learning companion: prospects in child-AI collaboration

Ethel Ong\*, Christine Rachel De Jesus, Luisa Katherine Gilig and Dionne Tiffany Ong

\*Correspondence:  
[ethel.ong@dlsu.edu.ph](mailto:ethel.ong@dlsu.edu.ph)  
College of Computer Studies,  
De La Salle University,  
2401 Taft Avenue, Manila, 1004,  
Philippines  
Full list of author information is  
available at the end of the article

## Abstract

Advances in natural language processing techniques enhance the capabilities of conversational agents to process input text and generate fluent responses in various genre. In the educational setting, these conversational agents can function as learning companions to promote peer learning, enhance social interaction, and motivate collaborative behavior. In this paper, we investigate how children collaborate with conversational agents in the context of story construction. We conducted user studies and administered survey forms to collect children's evaluation of their experience, captured the child-agent conversation logs, and used an observation checklist to monitor the interaction. We analyzed the logs to determine how children respond to varying roles portrayed by the agent and the dialogue moves that motivate collaborative behavior during storytelling. We correlated these with the children's reported perception and our own observation of the interaction based on performance, collaboration and humanity attributes. Drawing on these results, we identified the roles of conversational agents that align with children's traits and abilities as storytellers and propose dialogue strategies to facilitate collaborative storytelling. Findings from our study can inform future educational technology designers on the roles of conversational agents as collaborative learning companions who can adapt to the individual needs and preferences of the children.

**Keywords:** Conversational agent, Storytelling, Collaboration, Virtual learning companion

## Introduction

People naturally engage in storytelling when describing their daily life events and situations. Children in particular find storytelling both educational and entertaining as they learn about their environment and make sense of their experiences. Sharing stories also afford opportunities for the expression of thoughts and feelings while learning about language, values and social interaction.



© The Author(s). 2024 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The availability of digital, interactive and intelligent computing technologies has brought about changes in the way children create and share their stories (Coleman, 2012). Collaborative storytelling systems offer multiple avenues to support children as (co-)creators of their own stories. They employ conversational agents as part of the story construction process to provide cognitive and emotional support during the storytelling task (Robertson & Wiemer-Hastings, 2002). These agents can be valuable tools in the development of children's narratives and vocabulary skills through social interaction and practice. Collaboration with the agent can also help children develop critical listening skills as they take turns contributing to pieces of the story text (Ryokai et al., 2002).

Large language models (LLMs) expanded the role of computing technologies in education from merely serving as productivity tools to acting as intelligent learning companions with the ability to converse in natural language (Tlili et al., 2023). From answering simple queries to writing essays, these agents, in particular, ChatGPT (OpenAI, 2023), are being utilized to teach programming (Surameery & Shakor, 2023; Yilmaz & Yilmaz, 2023) and mathematics (Ellis & Slade, 2023; Li et al., 2023; Wardat et al., 2023), correct linguistic errors in a written text, and give assessment and feedback (Baskara & Mukarto, 2023; Klimova et al., 2023). Studies have reported mixed findings on whether such technologies should be admitted or banned in the teaching-learning process as their influence on students' learning remain unexplored while compromising academic integrity principles and values (Dehouche, 2021; Roe et al., 2023).

The 5th Industrial Revolution (5IR) may shed some insights on how we can reap the benefits and reduce the possible harm of conversational agents. 5IR puts forth the notion of AI as augmented intelligence, where humans and AI work together harmoniously and complement each other's strengths towards a shared goal (Davenport & Ronanki, 2016; Guszczka et al., 2017; Terveen, 1995). The idea of human-AI collaboration was first introduced in 1993 at the *AAAI Fall Symposium on Human-Computer Collaboration* to discuss the design of collaborative systems that can support planning, task delegation, coordination, communication, and adaptation of collaborating agents in the performance of shared tasks.

Shared storytelling is a mixed initiative task characterized by a fluid and natural interaction between the participants (Sun et al., 2017). Research work that leveraged AI with conversational interfaces, however, has failed to capture these human-like qualities of storytelling. Problems in natural language processing and dialogue management may render the virtual peer incapable of fully understanding the child's input and generating a relevant response. LLMs are showing great potential in addressing the challenges in text generation but the construction of stories are mostly delegated to the intelligent agent while the human participant focuses on formulating the appropriate prompts. We are primarily interested in collaborative storytelling environments where human users are considered as

the primary producers of their own stories and a conversational agent is utilized to play a collaborative yet supportive role towards the achievement of the task.

The use of conversational agents to engage learners in conversations as a means of achieving the target learning outcomes is not new. Called pedagogical agents in intelligent tutoring systems research (Kim et al., 2006; Schroeder, 2016), they have been employed as tutors (Graesser et al., 2005; Wik & Hjalmarsson, 2009), mentors (Zakharov et al., 2007), facilitators, learning companions (Cassell et al., 2007), and teachable agents (Zhao et al., 2012). Prior work utilized canned responses in supporting dialogue-based interaction during the delivery of the required learning content. With LLMs, renewed interest in pedagogical agent research has emerged (Hobert & Meyer von Wolff, 2019; Khosrawi-Rad et al., 2022; Wollny et al., 2021). LLMs open avenues to rethink the design of innovative approaches for these agents to perform their roles in a learning environment.

There is also little work that examined how children collaborate with conversational agents. This is particularly important in situations where social interaction and affinity to others are necessary for learning and play. Research in child-agent interaction has observed the development of relationships between children and the more social and intelligent conversational agents (Williams et al., 2019). In this paper, we seek to address the following research questions:

- RQ1. What strategies can conversational agents employ to motivate children to be more expressive in narrating their stories?
- RQ2. How do children perceive conversational agents that can collaborate with them in the construction of stories?
- RQ3. What roles do children expect these conversational agents to portray?

## **Related works**

Our research blends two fields of study - child-agent interaction and collaborative storytelling. We aim to better understand how children may respond to conversational learning companions in the latter's efforts at providing collaborative support during shared storytelling (RQ1, RQ2).

Child-agent interaction seeks to understand how children communicate with conversational agents during the performance of their learning tasks. This interaction entails an understanding of the roles that children expects the agent to portray, particularly as learning companions and AI assistants (RQ3). Utilizing free-form dialogue can set-up a more natural interaction that may encourage children to be more expressive in their storytelling. However, this can lead to communication breakdown that may disrupt the collaborative task and necessitates the utilization of repair strategies to recover from it. The whole child-agent interaction is also anchored on collaborative storytelling that employs

text understanding, text generation, and dialogue management techniques to enable the conversational agents to participate in the story construction process.

### **Conversational agents**

Conversational agents, through their collaborative behavior, are designed to assist the human users to think better while performing specific tasks (Davenport & Ronanki, 2016). This views AI as augmented intelligence, where machine intelligence and human intelligence are “*complementary, each bringing its own strength (to the collaboration circle) to reduce the gap of the other*” (Guszcza et al., 2017). With this view, research in conversational agents have portrayed them in differing roles depending on the types of collaborative interaction expected by the human users. These include co-author, critic, facilitator, tutor, and learning companion. Understanding the varying roles that a conversational agent may portray in the context of learning environments in general, and collaborative storytelling in particular, may inform the design of learning companions that can motivate children to exhibit collaborative behaviors in the performance of their learning tasks (RQ3).

Tracing their roots in virtual agents, conversational agents have been the subject of research in the field of education for decades. One example is Sam, an embodied virtual peer designed to engage children in collaborative storytelling tasks to develop their literacy skills (Ryokai et al., 2002; Wang & Cassell, 2003). As a co-author, Sam starts the story, then switches to facilitator role as it listens to children share their stories. The co-author role allows Sam to model literacy skills for children to emulate.

Conversational agents have also portrayed the role of story writing assistants with abilities beyond automated spelling and grammar correction. These agents can suggest story text (Chua et al., 2017; Swanson & Gordon, 2012) and perform narrative completion (Roemmele & Gordon, 2015). In Story Station (Halpin et al., 2004), the agent portrays the role of a critic that is capable of discerning between good and poor plots in children’s story writing. Feedback from teachers showed that they also consider the agent as a collaborative partner to help in identifying young writers who are struggling with the plot and need help.

Alice (Chua et al., 2017) is a story writing peer that can ask questions, give directives as well as recommend possible story text by serving either the facilitator or collaborator role. In the story writing space, the child remains in control and decides when it needs to ask Alice for help and to specify the type of help needed, i.e., ideas and suggestions. In Say Anything (Swanson & Gordon, 2012), the agent serves as a co-author and utilizes a simple turn-taking approach to alternately write story text with the child. The study of Sun et al. (2017) investigated the feasibility of a collaborative storytelling robot that uses conversation to insert new story content that is either related or not to the existing story.

## Communication and repair strategies

Utilizing free-form dialogue can set-up a more natural child-agent interaction that may encourage children to be more expressive in their storytelling and to collaborate with the conversational agent (RQ1, RQ2). Human-computer collaboration is a term used to describe how two entities – the human user (child) and the computer (conversational agent) – can work together on a shared task in order to achieve a common goal (Terveen, 1995). Communication is a key requirement for collaboration to take place, allowing the exchange of information needed to perform the shared task (Suchman, 1987; Terveen, 1995), in this case, the construction of stories.

Similar to human-to-human communication, breakdowns can manifest in human-to-agent communication. The design of conversational interfaces should include the identification of the causes of communication failure and the application of support structures to facilitate repair mechanisms when miscommunication occurs.

DeVito (2017) identified various types of noises that interfere with the interlocutor's abilities in sending and receiving messages thereby causing communication breakdown. These include physical noise from the environment; technological noise from computing devices used as the medium of communication; physiological noise caused by impairments in visual, auditory and speech functions; psychological noise due to biases, preconceived ideas and lack of interest in the conversation; and semantic noise due to gaps in language and the inherent ambiguity of words. Among these, we identified semantic noise and technological noise as the factors that are relevant to our study involving children.

Semantic noise emanates from an individual's proficiency in using the language and the extent of vocabulary and background knowledge about the subject matter. Children have different language proficiency skills that influence their ability to use the language to express their thoughts and emotions (Cohen, 2010; Habok & Magyar, 2018; Jurkic et al., 2023). Technology-based medium of communication, particularly voice and text interfaces, can also cause breakdowns when messages are not properly sent or received through the communication channel. The release of voice interfaces functioning as digital home assistants such as Siri and Amazon Echo has sparked research interest in investigating how families communicate with computers using these technologies (Beneteau et al., 2019; Cowan et al., 2017; Luger & Sellen, 2016; Sciuto et al., 2018; Ureta et al., 2020).

When there is breakdown in communication, interlocutors employ repair strategies to recover from the miscommunication (Beneteau et al., 2019). These include repetition, augmentation (Cheng et al., 2018), substitution or rephrasing an utterance (Most, 2002) possibly with the use of simpler or synonymous words, and linguistic code switching or adjusting one's speaking style to the given context or comprehension abilities of the listener (Wallach & Butler, 1994). Children 9 years old or older may employ context clues to define difficult terms (Brinton et al., 1986). Adults may engage in discourse scaffolding to help

children develop language skills, such as expanding the child's short single word or phrase input into a complete sentence (Wallach & Butler, 1994).

Proficiency in the English language, literacy skills and familiarity with the computing device can also pose communication challenges between the children and the conversational agent. Understanding these can shed insights in designing better conversational interfaces that include mechanisms for recovery when the child and the virtual agent fail to understand one another (Cheng et al., 2018). Hayes and Reddy (1983) further noted that employing conversational repair should be treated as a complementary task; both the child and the virtual agent should attempt to apply repair strategies in the event of miscommunication.

With LLMs, communication is typically in the form of a *prompt – response* pair. The learner issues a prompt representing a query or an instruction for the virtual agent to answer or perform. Communication breakdown occurs when the virtual agent fails to generate a correct response to the given query and is typically due to incorrect prompt formulation (Gregorio et al., 2023; OpenAI, n.d.). The only strategy to repair this miscommunication is through revising the prompt, i.e., giving a clearer instruction for the virtual agent to perform. Prompt engineering has thus become an important element when interacting with LLMs (GPTBot, n.d.; Grabb, 2023; Mayo, 2023).

### **Collaborative storytelling**

Collaborative or shared storytelling is characterized by two or more entities actively participating in the construction of a story. Three considerations must be addressed when designing collaborative storytelling systems: clarifying the purpose of storytelling; determining the forms of collaborative support that can be provided during a storytelling session (RQ2); and identifying the roles of the participants in the collaborative task (RQ3).

Collaborative storytelling systems provide interactive spaces for children to “*tell their own fantasy or personal stories as a way to explore issues of concern to them, and a way to practice language use and literacy skills*” (Cassell et al., 2000). Varied forms of interactive environments that address the different purposes of storytelling are found in literature. We broadly classified these into digital storytelling systems for creating digital stories, collaborative spaces for writing stories, and shared storytelling with a conversational agent. We are primarily interested in storytelling environments where human users are considered as the primary producers of their own stories while a conversational agent is utilized to play a collaborative yet supportive role towards the achievement of the task.

Digital storytelling environments lean towards the use of digital media technologies to support the creation of media-rich stories. Collaborative spaces for constructing stories, on the other hand, treat children as authors who need support during the creation of their

stories. They employ intelligent assistants meant to “*enhance the writing (storytelling) instruction, provide feedback, or encourage reflection*” (Warren et al., 2008), through collaborative strategies that include sharing of ideas, eliciting feedback and suggestion, demonstration, and co-authoring. Children’s perception of a conversational agent’s abilities to carry out these supporting tasks (RQ1) may affect their response to the agent’s attempts at collaboration (RQ2).

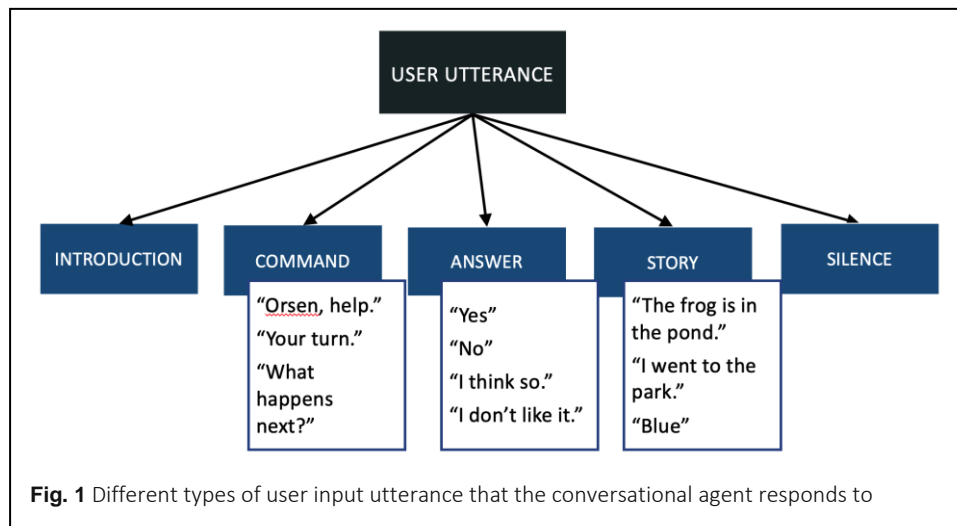
Language educators perform different mechanisms to help children when the latter encounters difficulty in starting and continuing their stories, such as asking questions, giving directives, and even recommending possible story text to move the story forward (Chua et al., 2017). Themes are sometimes used as story starters or prompts, to give children a certain topic that they may want to impart through their stories. Educators may also assist by suggesting different details of the story, and by prompting the child to expound on a story element. Prompts and suggestions in the form of story text can be about a story character’s attributes and interests (e.g., “*Please tell me more about Bob.*”, “*What does Bob like to play?*”), setting (e.g., “*Where did Jane go?*”, “*The children went to the park.*”), object descriptions (e.g., “*Tell me about the color of the car.*”), and explanations for the occurrence of events and their effect to the story world (e.g., “*What happens next?*”, “*Maybe the boy went to the market?*”). Examining how conversational agents employ some of these mechanisms to exhibit collaborative behaviors can shed insights and strategies to facilitate children’s construction of stories (RQ1).

## Agent design

Prior to conducting our experiments, we first designed our collaborative storytelling agent by taking into account the following: the roles that it will portray, the dialogue moves to be employed, and the commonsense knowledge base it utilizes during story construction.

## Agent roles

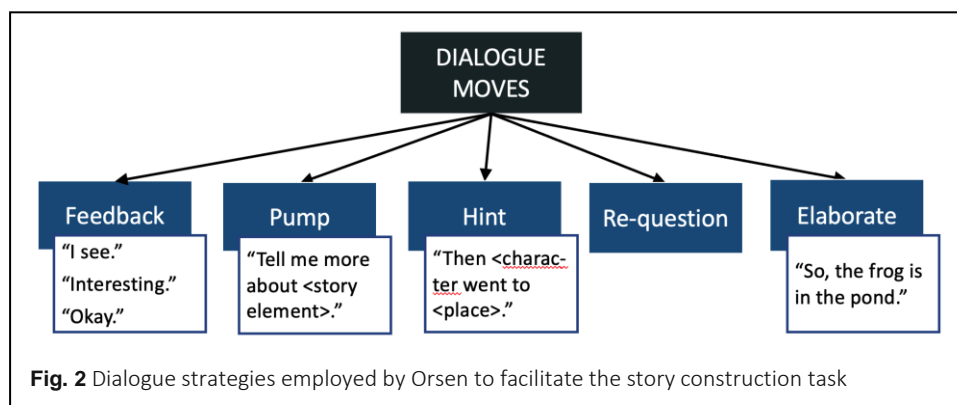
Our conversational agent, called **Orsen**, portrays two roles: a *facilitator* role to guide children in the story construction task, and a *co-author* role to collaborate in story writing by suggesting story text. The agent anticipates and responds to a variety of input utterances from children which we categorized into five (5): introduction, command, answer, story text, and silence. This is illustrated in Figure 1. Our preliminary data gathering revealed that some children introduced themselves to the agent. Such utterances are considered as *introduction*. The agent may also pose questions that children answer with “*yes*”, “*no*” or “*okay*”. These are classified as *answer*. Any utterances that contain the name of the agent is considered a *command* to perform a task. When the agent does not receive any input, this is flagged as *silence*. All other inputs are categorized as *story text*.



### Dialogue moves

Collaboration occurs when the virtual agent and the child engage in conversations to define the story characters and setting, rationalize characters' actions in relation to their motives and goals, imagine the varied events that may occur next, and even relate one's personal experiences to those of the characters. Following the approach of AutoTutor (Graesser et al., 2005), we utilize a number of dialogue strategies to facilitate the collaboration: feedback, pumps, hints, re-questioning and elaboration. This is illustrated in Figure 2.

*Feedback* is the response given to a child's utterance. Orsen currently does not assess the quality of the input, whether it leads to a good or bad story, nor does it check the correctness in using the language in terms of lexical choices and grammar. Instead, feedback is meant to reassure the child that the agent is listening and includes positive feedback ("Interesting", "Right") to portray the agent as a good critique, and neutral feedback ("Okay", "I see"). A feedback is usually accompanied by a *pump* that elicits further details regarding the





different story elements, i.e., character, setting and objects. Two types of pumps are used: general pumps to encourage sharing of story events (“*Tell me more.*”, “*What happens next?*”), and specific pumps (“*Tell me more about the prince.*”). Similar to writer’s block, a child may get stuck in the midst of telling their story. In this situation, *hints* are used to suggest story text that the child may opt to include, e.g., “*The prince went to the forest.*”

Children may sometimes stop responding to pumps. *Re-questioning* is used to pose the original pump to check if the child is still interested in continuing with the story. On the other hand, *elaboration* repeats the child’s utterance to give the perception that the agent understood the story text. For example, the child’s utterance could be “*The dog went to the park.*” and the agent could respond with “*Ah, so dog went to park.*”

### Commonsense knowledge

For collaborative storytelling to take place, Orsen must be equipped with a collection of commonsense knowledge containing concepts about our world and their semantic relations that children are familiar with or that may appear in children’s stories. Commonsense knowledge enables computers to “*reason in useful ways about ordinary human life*” (Singh et al., 2004) when it has information about our everyday activities and places where we perform them, objects found in our environment, the causal chain of events that can take place, and our social relationships with others.

Following the process of building a commonsense knowledge base described in Ong et al. (2018), we populated our knowledge base by extracting binary assertions from 11 Disney princess stories: *Snow White and the Seven Dwarves*, *Beauty and the Beast*, *The Little Mermaid*, *Sleeping Beauty*, *Cinderella*, *Brave*, *Mulan*, *Tangled*, *Pocahontas*, *Princess and the Frog*, and *Aladdin*. There were no specific inclusion criteria in selecting the candidate stories; rather it was based on finding a series of stories that had similar themes yet still had some variety. Other children’s stories that were considered include the *Winnie the Pooh* series and *Children’s Story Collections*. However, the *Winnie the Pooh* series has limited variance in locations, objects, and characters, while the *Children’s Story Collections* book does not have a particular theme that unifies the stories. The selected Disney princess stories were the ones featured on Disney’s princess website at the time of retrieval; the text used were also based on the website’s interpretation of each movie rather than a full movie script. Table 1 lists samples of seed concepts that were used to extract assertions from each of the 11 Disney princess stories.

### Crowdsourced commonsense knowledge

Children’s stories can cover broad topics, from everyday social situations to adventure stories, fairy tales, and even outer space. These stories contain information about children’s perception of the world they live in. For the conversational agent to remain a collaborative

**Table 1** Sample seed concepts that were used to extract assertions from the stories

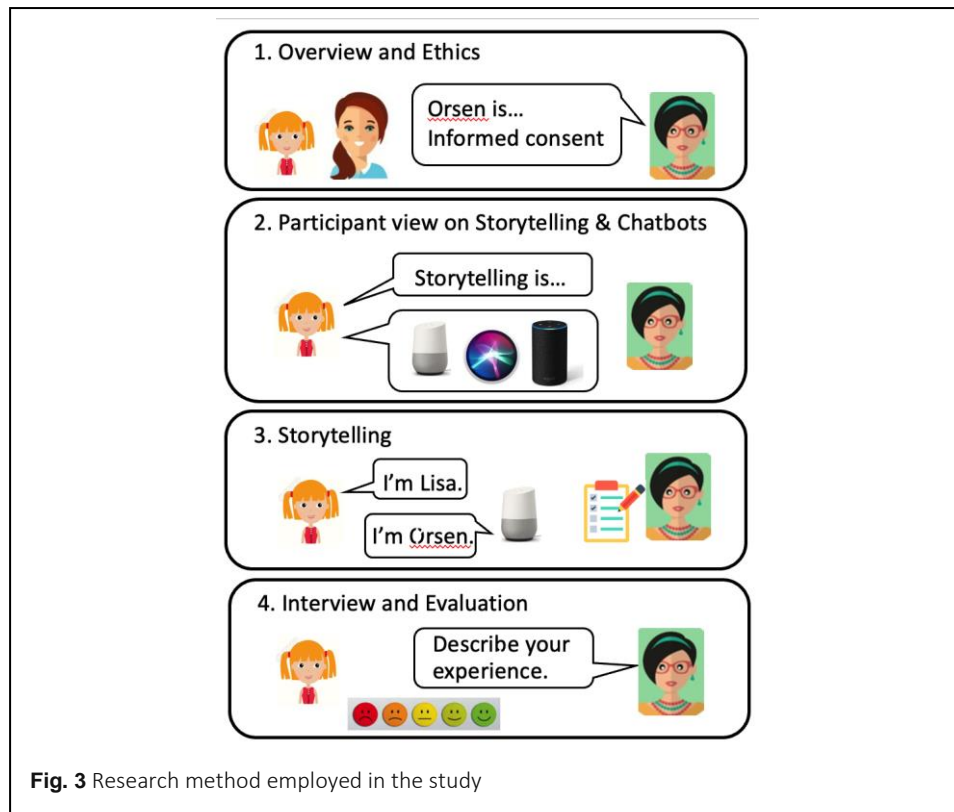
Story	Sample Seed	Assertions
Snow White and the Seven Dwarves	prince	[ prince IsA ruler ] [ prince AtLocation castle ]
	snow	[ snows UsedFor build snowman ] [ snow AtLocation mountain ]
Beauty and the Beast	enchantress	[ enchantress IsA woman ]
	beast	[ beast IsA attacker ] [ music UsedFor soothing savage beast ]
The Little Mermaid	mermaid	[ mermaid IsA imaginary being ] [ mermaid AtLocation sea ]
	sea	[ seagull AtLocation sea ] [ seal AtLocation sea ]
Cinderella	mice	[ mice CapableOf eat cheese ] [ mice HasProperty small ]
	bird	[ dove IsA bird ] [ bird AtLocation forest ]
Mulan	sword	[ sword HasProperty sharp ] [ sword IsA weapon ]
Pocahontas	canoe	[ canoe AtLocation water ] [ canoe UsedFor having fun ]

partner amidst the varying story themes, we expanded Orsen’s role to include that of a teachable agent. A *teachable agent* capitalizes on an instructional pedagogy where learning is achieved by teaching others (Zhao et al., 2012). Much like how children learn new concepts from stories, we leveraged on story-based conversation as a means for the agent to continuously learn new knowledge and validate this through collaborative storytelling with children.

Knowledge crowdsourced from users, in this case, children’s story text, must be validated (Dy et al., 2020). Two dialogue moves were added to enable Orsen to validate new assertions it sourced from children’s story text: suggesting and follow-up. *Suggesting* is similar to hinting except that the proposed story text is expressed as a question, i.e., “*What if ‘The prince went to the forest’?*”. Children are given control in affirming or refuting the correctness of the suggestion. If a child rejects this, the *follow-up* dialogue move is then used to determine if the reason for rejection is due to personal preference or if the given statement is incorrect, i.e., “*Why not? Don’t you like it or is it wrong?*”

## Method

Experiments with children proceeded in multiple stages, as shown in Figure 3, commencing with the orientation and administration of informed consent forms to both children and their parents or guardians. This is followed by a demonstration of how children can share stories with Orsen including special commands to seek assistance and to end the story. Each individual child-agent storytelling session lasts for about an hour, inclusive of the orientation, demonstration, and evaluation. The actual storytelling spans



10 to 30 minutes depending on the length of the shared story, the speed with which children can conceptualize their narratives, and the number of stories that children shared with Orsen (ranging from 1 to 3).

### Participants

Random sampling was employed in recruiting 44 children who are between 7 to 11 years old from three (3) basic education schools. These schools were selected based on geographical proximity and community engagement partnership with our university. While the selected schools follow the bilingual educational system, a key selection criterion is in the use of English as the primary medium of instruction. Children who were selected to participate in our study can speak and write English appropriate for their age level. This is a necessary requirement as the conversational agent is designed to understand and generate text only in the English language.

In accordance with university ethics policies, informed consent forms were administered to parents and guardians prior to the commencement of the experiments. They were given a briefing on the purpose of the study, the experiment process, the recording of the storytelling session, and research ethics protocol that includes statements on anonymity, use of collected data for report writing and dissemination of results, voluntary participation,

and the option to withdraw from the study without negative consequences. Informed assent forms were also given to children, thus allowing them free choice to participate in the study.

Children who agreed to participate were introduced to the conversational storytelling agent, its features and modes of communication. The set of keywords that can be used to instruct the agent to perform certain actions, such as starting and ending the session, and offering help, were also provided. No monetary compensation was given to the participants. Instead, the children received art materials as tokens for their participation.

## **Procedure**

Our study was conducted in three (3) iterations. In the first iteration, 11 participants communicated with Orsen using Google Home, a smart home device that utilizes Google Assistant as a speech-based virtual assistant. In the second iteration, 21 participants used Google Firebase as the text-based interface in communicating with Orsen. We originally intended to use speech as the primary mode for child-agent interaction. However, challenges encountered with the voice interface, particularly the disfluencies and grammatical errors typically found in children's speech (Ureta et al., 2020), led us to move to the Google Firebase platform for the remainder of our experiment. Since our study primarily aims to investigate children's perception of sharing stories with a conversational agent, dealing with problems in speech recognition and speech-to-text is currently not within our scope. In the third iteration, we invited 12 participants to share their stories with Orsen portraying the role of a teachable peer.

During the experiments, we used an observation checklist to monitor the child-agent interaction; one research proponent is assigned to accomplish the observation checklist for each participant. Afterwards, we assisted the participants in completing a structured survey form that asked them to use a 5-point Likert scale to rate their perception on the conversational agent's abilities in helping them create their stories. Qualitative feedback was also solicited to support the ratings.

## **Data collection instruments**

The observation checklist and the survey form used to collect children's perception of Orsen's collaborative behavior were structured using performance, collaboration and humanity attributes. *Performance* assesses the storytelling agent's ability to classify and interpret the input text and to apply an appropriate dialogue move to generate a relevant response. It is based on the performance attribute for evaluating conversational agents as defined by Radziwill and Benton (2017). This attribute also focuses on the robustness of the agent to unexpected input (Klüwer, 2011), and in generating inappropriate responses and performing repair strategies during communication breakdowns (Morrissey & Kirakowski, 2013).

The *collaboration* attribute looks at the structure of the conversation to assess how the child and the agent worked together to achieve their shared task. It requires “*a commitment to a common mission, where control and authority are determined by different collaborative structures*” (Fischer, 2015). For our study, the child has control and authority over the story flow, while the agent provides the support to help the child expound on elements of their stories. Conversation flow includes the ease with which the interlocutors are giving and taking turns, the correctness and alignment of the utterances which reflect the interlocutors’ understanding of each other’s dialogue intent, and the agent’s ability to suggest story text.

The *humanity* attribute assesses the quality of the conversation based on user satisfaction and engagement. This includes instances when the child showed evidence of treating Orsen as a social entity by using its suggested story text, responding to the agent’s (sometimes inappropriate) questions, accepting and correcting the agent’s mistakes, and utilizing repair strategies typically employed in human-to-human communication. Engagement is manifested through children’s expression of interest in pursuing and sustaining the conversation and collaboration with the storytelling agent.

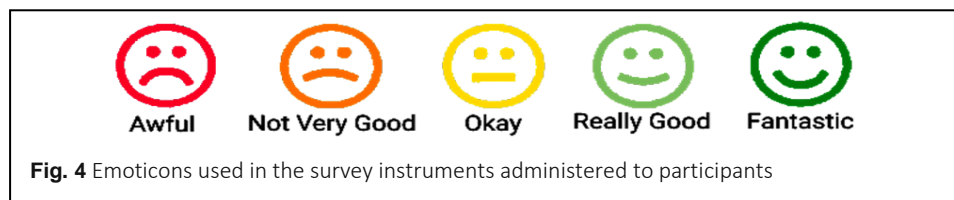
The observation checklist uses a simple Yes/No approach for the observers to indicate if a given scenario manifested during the child-agent interaction. Table 2 lists the items that were observed; one research proponent is assigned to observe one child-agent interaction. Items in the structured survey form are listed in Table 3. To help young participants better understand the items in the survey instrument, the 5-point Likert scale is represented using emoticons as illustrated in Figure 4.

**Table 2** Items in the observation checklist

<b>Performance</b>
1. The child takes Orsen’s suggested story text into consideration.
2. The child takes a long time in coming up with the story text.
3. Orsen follows the commands given by the child.
<b>Collaboration</b>
4. The child asks Orsen for feedback.
5. The child asks Orsen for suggestions.
6. The child ignores Orsen’s responses.
7. The child takes knowledge validation seriously.
8. The child helps Orsen understand new words when prompted.
<b>Humanity</b>
9. The child appears intimidated by Orsen.
10. The child does not lose interest in the conversation.
11. The child seems to enjoy sharing a story with Orsen.

**Table 3** Items in the structured survey form

<b>Performance</b>
1. I can understand Orsen's questions.
2. Orsen's questions are related to the story I am telling.
3. I can understand Orsen's suggestions.
4. Orsen's suggestions are related to the story I want to tell.
5. I can understand the words used by Orsen.
<b>Collaboration</b>
6. Orsen's questions help me continue the story.
7. Orsen's suggestions help me continue my story.
8. I see Orsen as someone I can teach.
9. I see Orsen as someone I can learn from.
<b>Humanity</b>
10. I can start chatting with Orsen on my own.
11. I find Orsen's voice to be friendly.
12. Orsen is fun to talk to.
13. I would like to talk to Orsen again.
14. I can talk to Orsen just like how I would talk to a classmate or friend.



## Data analysis

Qualitative analysis was performed on the collected data, specifically the conversation logs. The use of conversational analysis practices aims to “highlight opportunities for designers to improve interactions between humans and computers by detailing the structure of interactions from the user’s perspective” (Beneteau et al., 2019).

Conversation logs collected from child-agent interaction were reviewed to find specific instances of child-agent interaction that illustrate the three main attributes used in our observation checklist and survey form, namely task performance, collaboration, and humanity. Analysis includes finding evidences of (mis-)alignment of the child’s input and agent’s responses, (mis-)understanding that lead to communication breakdown, taking turns, and the child’s response to the agent’s suggested story text (i.e., accept, reject). A language educator evaluated all the logs using two guide questions.

1. Does the agent generate responses appropriate to the child’s utterances to encourage collaborative storytelling? Why or why not? Please highlight a portion of the log to support your answer.
2. Does the child’s response to Orsen express his/her view of Orsen as a collaborative peer? Why or why not? Please highlight a portion of the log to support your answer.

Two research proponents also reviewed the logs to find exemplars that the conversational agent is using its learned knowledge in suggesting relevant story text and instances of communication failure and repair strategies employed during collaborative storytelling.

## Results

Results from analyzing the data we collected through the observation checklist and conversation logs are organized following the three attributes: *task performance*, *collaborative behavior*, and *humanity*. We use the code *[In-Px]* to represent the *n*th iteration of the experiment involving the *x*th participant.

### Task performance

A conversational agent's *task performance* attribute is largely dictated by its ability to classify and interpret the input text and to apply an appropriate dialogue strategy to generate a relevant response. This is crucial to the success of the interaction in motivating children to be more expressive in their narrative (RQ1). Failure to employ an appropriate dialogue move may lead to a breakdown in communication that disrupts the collaborative storytelling session. Instances of these are noticeably evident in the first iteration of our experiments primarily due to problems with the Google Home voice interface.

**Difficulty in recognizing words.** Similar sounding words or phrases are difficult for Google Home to interpret. The problem is compounded by the lack of a Philippine English option in the device's settings. For example, when the child said "*She told Samson's secret to Philistines*," Home interpreted this as "*She told Samsung secret to the Philippines*." This issue is further exacerbated if the input is incomplete or lacks context. When a child said "*The witch was...*" and then paused to think, Google Home interpreted this as "*which was*". Uncommon character names are also misinterpreted, for example, the name "*Marson*" was interpreted as "*our son*" or "*Corazon*". Formulation of a relevant response is in part reliant on the correct extraction of story elements from the input. Errors such as those described here can lead the conversational agent to respond with "*What is secret?*" and "*Please describe Corazon*," leading to the child's confusion and further breakdown in communication.

**Ambiguity of words.** Words with multiple meanings may render the agent incapable of understanding the correct context of the input. Consider the example in Table 4. Orsen associated the word "*plate*" (line 1) with a "*mechanical plate*" (line 2a) instead of a "*dinner plate*", thus, generating story text about a repair shop. In Table 5, the agent made multiple attempts to contribute to the story by describing "*pepper*" in lines 2 and 4. But in its second attempt, the child reacted negatively when the agent mistook "*pepper*" as a spice instead of the name of a pet. To recover from its mistake, the agent stopped offering suggestions and opted for a general pump instead in line 6.

**Table 4** An excerpt of a conversation log showing an inappropriate and a corresponding appropriate response [I2-P3]

Line	Turn	Text
1	Child	We got a plate of food.
2a	Orsen	There was a mechanic at the repair shop. ( <i>inappropriate response</i> )
2b	Orsen	There was fish dinner on the plate. ( <i>appropriate response</i> )

**Table 5** An excerpt of a conversation log showing the child's reaction to the agent's misinterpretation of the context of "pepper" [I3-P12]

Line	Turn	Text
1	Child	the gang started bullying pepper and harry
2	Orsen	Hmm, so gang started bullying pepper and harry. Hmm, I think on a workday, pepper went to dinner
3	Child	No
4	Orsen	Hmm, I think the green pepper season sandwich
5	Child	No pepper is a dog
6	Orsen	I see, so what happens next?

**Processing language errors.** Performance can also be evaluated based on the conversational agent's ability to continue its operation despite the presence of incorrect inputs. Incorrect inputs are caused by the language proficiency of children and the medium of communication. Children's input text contains numerous misspellings and grammar errors, as illustrated in Table 6 where the starting input text should have been "once upon a". Regardless, the agent's text processing engine tried its best to parse and extract story elements in line 2.

**Foreign language.** Instances of mixing two or more languages were also encountered in the conversation logs. An example is shown in Table 7. Again, the agent resorted to using pumps to motivate the child to share more details regarding the unknown word "lola" (grandma) in line 4, which is evident in the exchange. Furthermore, the child continued with their story even when the agent made nonsense responses in lines 2 and 6.

**Table 6** An excerpt of a conversation log with misspellings [I2-P15]

Line	Turn	Text
1	Child	Ones <i>apana</i> time there was a girl she was cleaning her shirt
2	Orsen	Can you tell me more about <i>apana</i> ?

**Table 7** An excerpt of a conversation log illustrating the use of foreign language [I2-P7]

Line	Turn	Text
1	Child	It was since January in the house and my tita ( <i>aunt</i> ) was packing up because we were going to drive her to pampanga ( <i>a province</i> ).
2	Orsen	I see
3	Child	The ones who came with us are my mom, dad, cousins, tita, lola ( <i>grandma</i> ), lolo ( <i>grandpa</i> ), my sister and me
4	Orsen	Can you tell me more about lola?
5	Child	My lola is old but she was always happy but still a little sick
6	Orsen	Interesting
7	Child	So when we were in the road it took us about 2 hours to go there.



**Response to silence.** Long pauses are perceived by the agent as a child failing to hear its command or question. It then issues the “*I didn’t hear you. Could you repeat what you said?*” response. Surprised, children would tend to repeat their whole story, causing problems with the input text processing engine. Rephrasing the statement to “*I didn’t hear you. Could you repeat the last part?*” yielded more positive reactions from the children.

**Robustness to various input text.** Children often give input text that do not necessarily answer the agent’s questions. For instance, in line 1 of Table 8, Orsen asks for the child’s name. Instead of introducing themselves, the child went directly to their narrative in line 2. The agent coped with this unexpected input by responding with a feedback in line 3. This motivated the child to continue with the narrative in line 4. In line 5, the agent used the *elaborate* dialogue move to perform anaphora resolution (“*he*” => “*boy*”) and to restate the child’s input. Because the agent was not expecting an affirmation from the child in line 6, it repeated its statement in line 7. The child then continued by sharing the next event of their story in line 8.

**Task-Oriented Dialogue.** Commercial voice interfaces such as Google Home are designed to be predominantly used in mainstream task-oriented applications where the bulk of the conversation involves giving instructions (“*Play some music.*”) and inquiring for information (“*What is the weather tomorrow?*”). Children’s narratives, however, is a form of social talk that can range from single words to run-on sentences. Talkative children who already have a story in mind would give long sentences in one dialogue turn. Google Home cannot handle such input and would freeze. While the device is restarting, children became distracted, forgot their story, or lose interest in the conversation.

**Voice vs text interfaces.** Voice and text interfaces also have their share of challenges. Subdued voice, pronunciation issues, and run-on sentences lead to missing words in the resulting speech-to-text translation. Run-on sentences also pose difficulty for the agent to find and extract words representing characters, objects and events. We found that the quality of the interaction during the first iteration of our experiment is limited by the

**Table 8** An excerpt of a conversation log showing how Orsen responds to unexpected input text (line 2) and performs anaphora resolution (line 5) [I3-P4]

Line	Turn	Text
1	Orsen	Hi! What’s your name?
2	Child	The boy tipped ( <i>tripped</i> ) in school
3	Orsen	I see, that’s interesting
4	Child	He had a wound
5	Orsen	I see, so boy had wound.
6	Child	Yes
7	Orsen	Hmm, so boy had wound
8	Child	He put bandages on his wound
9	Orsen	Oh, so that’s what happens.
10	Child	He put cotton on it after

capabilities of Google Home. However, text interfaces, despite foregoing the translation process, is less intuitive for children. Their typing skills and their limited vocabulary words yielded utterances with misspellings and lack the proper punctuation marks needed for correctly parsing the input.

### Collaborative behavior

Collaborative storytelling requires both parties – the child and the agent – to commit to work together towards shared task of constructing a story. To be perceived as a good collaborator (RQ2), the conversational agent must exhibit some form of intelligence. This behavior can be manifested through evident understanding of the context of the input by properly detecting story elements. The agent’s ability to formulate and suggest relevant story text by sourcing concepts from its knowledge resource is also an indication of collaborative behavior, especially if the concepts are familiar to children, can continue the flow of the current story, and can trigger creativity to help children enhance their stories.

**Detecting Story Elements.** Orsen references previous utterances to determine the story element that will be the focus of its response. This is its mechanism for aligning its response to the child’s input. The process works by starting from the most recent utterance, then working backwards through the dialogue history to find a character, object or event that needs additional details.

Analysis of the conversation logs yielded differing reactions. When the focus of the agent’s response is based on the most recent input, it usually resulted in a positive input from the child, as illustrated in lines 1–4 of Table 9, where the child agreed with the Orsen’s viewpoint on the colors of the rainbow. This encouraged the child to think of possible events to add more details to their story.

**Table 9** An excerpt of a conversation log showing that the child ignored Orsen’s seemingly out-of-context question in line 19 [I2–P10]

Line	Turn	Text
1	Orsen	Describe rainbow
2	Child	Manu colors
3	Orsen	I see, that’s interesting
4	Child	I know
	:	:
	:	:
18	Child	when aya returned back to her home city Silla she was already 21 years old, while hansung was 23
19	Orsen	I see, that’s interesting. How old is I?
20	Child	both of them began bonding as soon as possible. both of them thought that they had now lost their feelings for each other but they were wrong. Over the weeks that they had spent together their feelings had come back except stronger.

**Table 10** An excerpt of a conversation log that illustrates how the child responded to the agent's reference to an old event. The log also contains misspelled words, *buyed* and *occasion* [I3-P11]

Line	Turn	Text
1	Child	after playing they went to the carnival and <i>ride the horse</i>
2	Orsen	Ah, I see, so tin to walk. What happens after tin to walk?
3	Child	then tin went to the ice crem stand and buyed cotton candy
4	Orsen	Oh, so that's what happens I think It was Citizenship Day, when tin went to museum.
5	Child	it was independence day and celebrate the occasion happily
6	Orsen	Then There was a horse in the gate.
7	Child	yes the horse was small and pretty and pink. <i>Tin ride the horse</i> with ymata and the dog

However, when the agent referenced older events, the child reacted in one of two ways: (i) The child perceived that Orsen needed help in understanding the story, thus, they patiently repeated the story again. Consider the excerpt in Table 10. In line 1, the child narrated that “(they) *ride the horse*”. The agent decided to use the “*horse*” to formulate its response only in line 6 but the child still continued to answer the agent in line 7. (ii) The child perceived that Orsen is not listening, therefore, they ignored the agent's response. In the excerpt in Table 9, the agent mistook “*I*” (line 19) as the name of a character, thus, the child ignored the question “*How old is I?*”

### Using knowledge resources

Our storytelling agent has at its disposal, a collection of commonsense knowledge containing words and their relations that are familiar to children or are typically found in children's stories. It uses this knowledge resource when processing input utterances to detect and differentiate characters from objects, and to identify events. The same knowledge resource is also used by the agent to contribute to the story text. The available words and the agent's proper utilization of the word relations found in the knowledge base to formulate its useful responses are indicators of its collaborative behavior.

***Lack of sufficient word relations for response generation.*** When the agent encounters a word that is not in its knowledge base, it becomes incapable of suggesting a story text and resorts to using general pumps (“*What happens next?*”, “*Tell me more.*”). Frequent usage of such pumps, however, is not indicative of a collaborative behavior and can also cause frustration if the child runs out of things to say. One way to address this is with the teachable peer role.

***Orsen as a teachable peer.*** In the third iteration of our experiments, the inclusion of a teachable peer role encouraged some children to assume the role of a tutor in providing descriptive answers to the agent's prompts. Table 11 gives an example of a child who is helping the agent learn about the concept of “*bandages*”. Table 12 illustrates various instances when the agent is validating crowdsourced knowledge. In (a), the child accepted

**Table 11** An excerpt of a conversation log illustrating the child answering the agent’s questions about the concept of “bandages” [I3-P4]

Line	Turn	Text
1	Orsen	I want to hear more about bandages.
2	Child	Bandages can heal your wound
3	Orsen	Ah I see. So bandages can heal. I wanna know more!
4	Child	Bandages can also be used for cuts
5	Orsen	Oh, so that’s what happens.
6	Child	There are different kinds of bandages.
7	Orsen	I see, that’s interesting. Tell me more then.
8	Child	Those are the cast, the gauge, and the first aid kit.

**Table 12** Sample dialogues depicting Orsen as a teachable peer

Child accepts suggestion	Orsen	What if “There was a dog in the park.”?
	Child	yes
	...	
	Child	harry and pepper went to the park to play with dogs
Child rejects suggestion	Orsen	What if “On a Saturday, the dog can walk.”?
	Child	The dog always walk.
	Orsen	Sorry, I don’t understand. Please answer with yes or no.
	Child	No.
	Orsen	Why not? Don’t you like it or do you think it’s wrong?
	Child	It is wrong.
	...	
	Orsen	What if “horse can be white”?
	Child	No, he is brown.
Child ignores suggestion	Orsen	What if “The ball was played.”?
	Child	Tin went to bed with the dog because shes happy if the dog is there. Tin played fetch with the dog and ate food.
	Orsen	Sorry, I don’t understand. Please answer with yes or no.

the suggested story text and even used this to add details to the story. In (b), the child rejected the suggested story text and may even give an alternative text. In (c), the child ignored the suggestion by sharing a different story text.

### Humanity attribute

Children exhibited certain behaviors as their way of recognizing the conversational agent as a social entity who portray certain roles throughout the collaborative task (Q3). This expression of *humanity* towards the agent includes interest in pursuing and sustaining the conversation, using turn-taking cues and interruptions, socializing through small talks, and treating the agent as a peer, a learning companion, a guide, or a storytelling coach (RQ3).

**Interest in the conversation.** Limitations in the text processing abilities of the conversational agent may lead to the generation of out-of-context questions and suggestions. Our results, however, showed that despite the misalignment of some of the agent’s responses to the child’s input, there were multiple instances when children

**Table 13** Examples of conversation logs where the children entertained the agent's questions and suggestions even if some are out of context

Line	Turn	Text
1	Orsen	Then, he went to a restaurant.
2	Child	Uh then he ate pasta
	:	:
11	Child	He's a merman he's a mermaid
12	Orsen	What does a man like to do?
13	Child	he likes too swim
	:	:
21	Agent	Tell me more about grocery store.
22	Child	He met his pet seahorse in a grocery store
	:	:
31	Orsen	What's gym like?
32	Child	The gym is where you exercise.

continued to entertain these pumps and hints. Consider the different examples in Table 13. In line 1, the agent suggested a story text that is not what the child had in mind, but still, the child went along and gave an input in line 2 that is aligned to the suggestion. In line 12, issues with text parsing due to lack of punctuation marks from voice inputs led the agent to only extract “*man*”. Regardless, the child still gave a description in line 13 when prompted. Similar behaviors can be observed in lines 21-22, lines 31-32, and throughout the conversation logs.

**Turn-taking.** Turn-taking cues and interruptions typically make up human conversations (Hirst, 1991). Interlocutors signify their intent to relinquish or to retain the dialogue turn through cues that include body signals, pauses and language. In voice interfaces, a pause in the child's input signals the agent that it has been granted a turn in the conversation. Some children explicitly used the trigger phrases “*Your turn*” and “*Give me an idea*” to relinquish their turn to the agent. The agent, on the other hand, used phrases such as “*So what happens next?*” and “*Tell me more.*” to give the turn back to the child.

**Interruptions.** Minimal interruptions were evident among the interlocutors using text-based interfaces. The child uses the ENTER key as a signal to give the turn back to the agent, while the agent displays its response when it is relinquishing the floor back to the child. On the other hand, the nature of voice interfaces allows interruptions to occur more naturally. In the first iteration of our experiment, four (4) children were observed to talk while the agent is talking. However, limitations in the device caused the agent to miss hearing and processing the utterances. Similarly, there were instances when the agent also spoke while the child is speaking. This is particularly evident when the children are shy and spoke in a voice which was inaudible to Google Home.

**Small talks.** Small talks are a natural part of human conversation to establish initial bonding and rapport (Olafsson & Bickmore, 2016). Small talks can start through simple introductions. During the early part of our experiments, we did not anticipate that children

**Table 14** An excerpt of a conversation log with the child introducing themselves as a response to the agent's self-introduction [I1-P6]

Line	Turn	Text
1	Orsen	Hi, I'm Orsen! Let's make a story. You start.
2	Child	Hi, I'm Pepper!
3	Orsen	That's interesting. Tell me more about Pepper.

**Table 15** An excerpt of the conversation log where the child treats Orsen as a teachable peer [I3-P2]

Line	Turn	Text
1	Orsen	I don't much about she. Please help me learn by telling more about she.
2	Child	she is a feminine
3	Orsen	I see, that's interesting.
4	Child	Do you want to know what is the meaning of feminine?
5	Orsen	Oh, so that's what happens.
6	Child	feminine is a long word for saying girl

will commence the conversation by introducing themselves to the agent. This led the agent to treat the child's introduction as part of the story, as illustrated in Table 14.

Another form of small talk is the embedding of sub-dialogues. This is particularly evident when the agent encounters a new word, such as "*bandage*", as shown in Table 11. While such deviations from the shared task are acceptable and typical in human-to-human conversations, one of the interlocutors must initiate a dialogue that will bring the focus back to the task at hand. Currently, the agent is unable to do so because of the way it identifies the topic of its response. It is left to the child to input the next story text that can resume the dialogue to their intended story.

**Orsen as a learning companion.** Analysis of our data showed multiple instances of the agent behaving as a teachable peer as a form of expanding its knowledge about the child's world. Examples of these are illustrated in Table 11 when Orsen tried to learn the concept of "*bandages*", and in Table 13 when it posed the pump "*What does a man like to do?*" (line 12). In Table 15, the child took the initiative to ask the conversational agent if it knows the meaning of the concept "*feminine*" (line 2). Prior research has reported increase in student's reflection, explanation, and knowledge organization abilities when teaching a peer (Roscoe et al., 2008).

## Discussion

Advances in the language abilities of conversational agents led to the generation of fluent text that can rival and even surpass those of human-generated text. However, educators are wary of utilizing the potential benefits of conversational agents as learners may become fully reliant on these tools when performing written assignments, raising concerns that the target learning competencies may not be achieved.

Augmented intelligence, on the other hand, posits the collaboration between human users and AI systems. Advances in text understanding and text generation techniques may address challenges in extracting relevant story elements from children's input text and formulating fluent responses that are congruent with the shared story. Beyond directly answering user inquiries and performing instructions, we discuss insights in designing and utilizing conversational agents as virtual learning companions. Drawing from our experiments where children engaged in shared storytelling with Orsen, we propose three (3) ideas for future designers of AI assistants: (i) the design of the dialogue structure that supports collaborative interaction (RQ1); (ii) how children perceived conversational agents as collaborators in the construction of stories (RQ2); and (iii) the roles of learning companions (RQ3).

### **Dialogue strategies for collaborative storytelling**

#### ***RQ1: What strategies can conversational agents employ to motivate children to be more expressive in narrating their stories?***

Our RQ1 seeks to identify strategies that conversational agents can employ to motivate children to be more expressive in narrating their stories. Research in designing dialogues for conversational agents have proposed the use of either a structured dialogue or a free-form dialogue. Task-oriented systems mostly follow a structured dialogue to allow the agent to collect the data it needs to complete a given task, e.g., online reservation systems. As alternatives to the more common form-filling interface, these conversational agents allow a natural form of interacting with virtual entities. However, everyday human-to-human conversations are not restricted to conform to a set structure. Instead, social communication abound with stories are free-form, allowing the interlocutors to control when to start, pause, resume, and end the interaction.

In order to expand the opportunities that collaborative agents provide for children, we need to set-up an environment where free-form storytelling is encouraged. Repair strategies are employed by the interlocutors to address communication breakdown when they occur.

***Communication breakdown repair strategies.*** Occurrences of misaligned child-input and agent-response are evident in all the recorded conversation logs. Primary causes of these are the imperfect nature of text processing and speech recognition technologies, and the inherent language and storytelling abilities of the participants. Our analysis of the logs showed that the agent formulates dialogues that mimic repair strategies described in prior works (Cheng et al., 2018; Most, 2022). Children also showed a high degree of persistence despite their difficulty in talking to Google Home or typing in Google Firebase; and tolerance despite the presence of errors in the agent's responses. Regardless, future designs should minimize such incidence to prevent the agent from demonstrating incorrect

storytelling practices and to increase opportunities for children to develop their language and literacy skills through the interaction.

***Linguistic code switching.*** The future design of conversational agents can also consider the ability of human interlocutors to adjust their speaking style to the given context or comprehension abilities of the listener (Wallach & Butler, 1994). To do this, the agent first has to associate a persona to the child it is collaborating with. It can then use this knowledge to assume a particular role (i.e., teachable agent, mentor, collaborator, facilitator), and to use particular language to address the specific needs and provide the necessary support of a particular persona. However, dynamically performing linguistic code switching would require the agent to have the ability to associate a persona from previous dialogue turns. The dialogue history, currently utilized for identifying the topic of an agent's response, may be used for such purpose.

### **Conversational agents in collaborative storytelling**

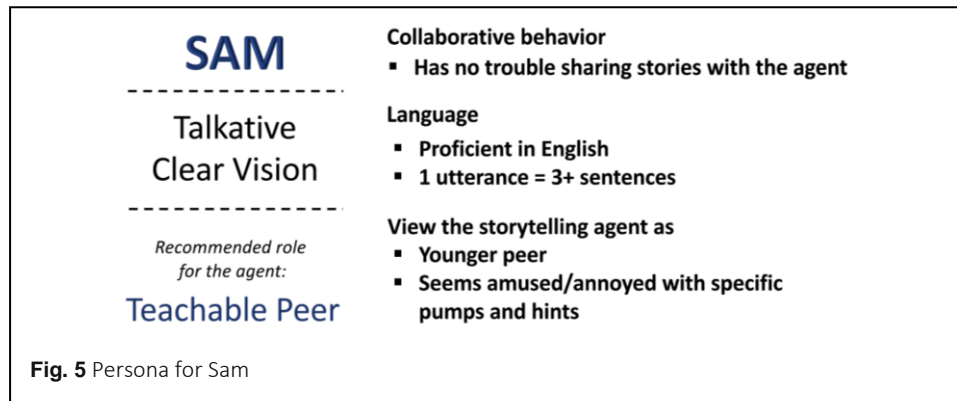
#### ***RQ2: How do children perceive conversational agents that can collaborate with them in the construction of stories?***

Our RQ2 focuses on children's perception of conversational agents that can collaborate with them during story construction. We approach this by characterizing children according to their traits and abilities in storytelling leading to four (4) personas of children as storytellers. These personas reflect children's level of comfort, engagement, clarity of vision for the story they want to share, and interest in communicating with the conversational agent while also setting the kinds of collaborative behavior they expect during the child-agent interaction.

Personas are fictional characters created by user interface / user experience (UI/UX) designers as a means of representing the different types of users of a particular software product or service (Dam & Siang, 2019). Among the different perspectives regarding personas described in Nielsen (2013), the *role-based* perspective focuses on user behavior, proficiency and skills. It can be combined with the *engaging* perspective to include user's background, psychological characteristics and emotional relationship. The resulting personas can help designers form a complete vision of their intended users even with only minimal and fragmented knowledge (Nielsen, 2013). We provide descriptions of the four (4) personas that we have identified from our data based on how participants perceived the conversational agent and which may have influenced their collaborative behavior.

**Sam** (Figure 5) is a talkative child who has a clear vision when it comes to telling her story. She is fluent in English and uses the language with confidence. She already likes storytelling even before interacting with Orsen. She knows the kinds of stories she wants to tell and does not really need the agent's help. She responds to the agent's questions like





she was simply entertaining the agent, a sign that she acknowledges Orsen's presence. For example, when the agent asked "How big is field?" Sam responded with "Well there wasn't a field but it was pretty big, I guess." then resumed with her story. She also has a tendency to combine multiple independent clauses as a single utterance, as seen in line 20 of Table 9. She treats Orsen more like a younger peer and becomes a bit frustrated or amused with some of the agent's more questionable responses, as illustrated in lines 2-5 of Table 16.

Sam prefers the storytelling agent to behave as a passive listener than as an active collaborator. She would respond well to the collaboration if she perceives the agent as a *teachable peer*. Sam would take on the responsibility of demonstrating the process of storytelling and clarifying misconceptions of the agent regarding story elements. She wants to remain in control of the storytelling task and expects the agent to provide frequent feedback such as "Interesting" and "I see", and general pumps such as "What happens next?" and "Tell me more." as illustrated in Table 17.

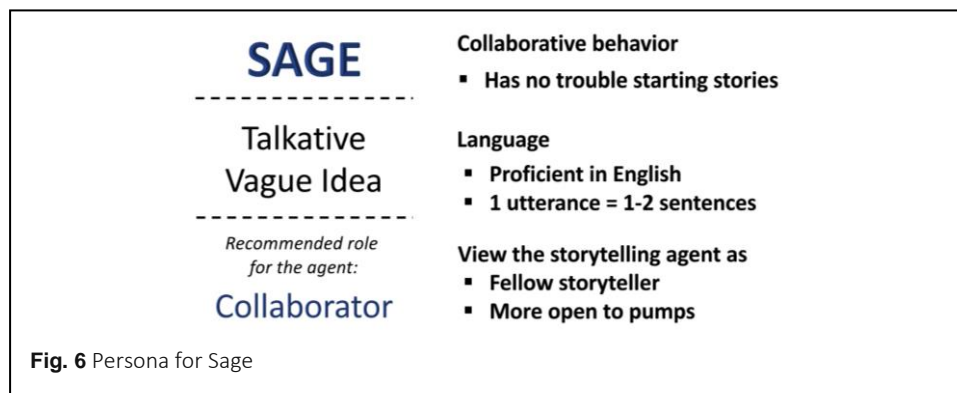
**Sage** (Figure 6) is also talkative and likes storytelling prior to interacting with the agent. However, he has a vague idea of the stories he would tell. He is more open to Orsen's questions and suggestions, often following along with the given hints. He prefers a slower pace when sharing his stories, using only one or two independent clauses before waiting

**Table 16** Sam responding to (amusing) questions for the sake of entertaining Orsen [I3-P12]

Line	Turn	Text
1	Child	harry and pepper went to the park to play with dogs
2	Orsen	Hmm, I think the pepper is a seasoning
3	Child	no
4	Orsen	What is dogs made of?
5	Child	Meat?
6	Orsen	I see, so harry and pepper went to park. What happens next?
7	Child	they met a gang of dogs

**Table 17** Another story excerpt depicting a Sam persona [I3-P2]

Line	Turn	Text
1	Child	she got tonsillitis
2	Orsen	Then what happens?
3	Child	she went to the doctor
4	Orsen	Oh so that's what happens.
5	Child	and then she got a antibiotic
6	Orsen	What happens next?
7	Child	and then she regret it
8	Orsen	Tell something about antibiotic
9	Child	it was color white and it taste bad

**Table 18** Sage's slow pace in storytelling, often waiting for the agent's response before continuing with his story [I3-P11]

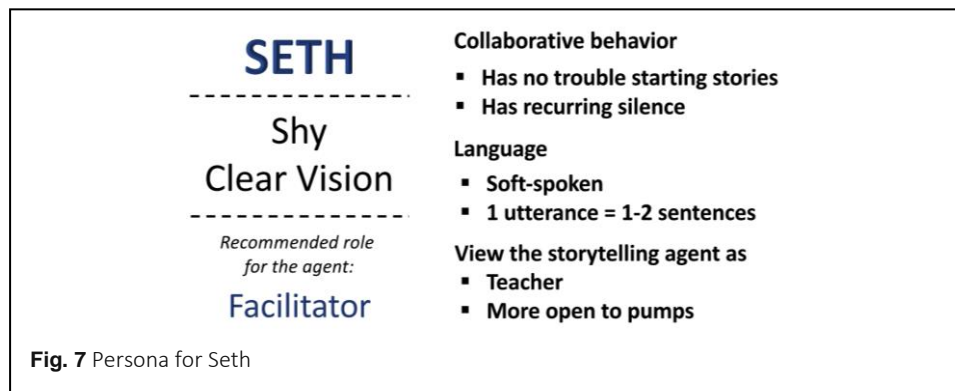
Line	Turn	Text
1	Child	tin went to the ice crem stand and buyed cotton candy
2	Orsen	Oh, so that's what happens I think It was Citizenship Day, when tin went to museum.
3	Child	it was independence day and celebrate the occassion happily
4	Orsen	Then There was a horse in the gate.
5	Child	yes the horse was small and pretty and pink. <i>Tin ride the horse</i> with ymata and the dog

for the agent's response, as illustrated in Table 18. Most of the time, Sage treats the agent as a peer who could help him with the story, though there are still situations when he would ignore the agent's irrelevant responses. For example, when the agent suggested the text "*Then, boy went to a restaurant.*" Sage responded with "*Uh then he ate pasta.*" (Table 13).

Sage perceives Orsen as a *collaborator* with the same developmental age and can understand the world in similar ways as he does. A collaborator sets up a co-authoring atmosphere where participants share the floor in either organized (role-play) or unorganized fashion (simultaneous turns) (Cassell et al., 2007). Dialogue turns of the agent is dominated by suggesting and hinting story text, with occasional specific pumps to fill-in missing information.

**Seth** (Figure 7), though quiet and shy, knows the story he wants to tell Orsen beforehand. He starts his story quickly but has trouble sustaining his narrative. He responds well to the agent's questions about the story but would also disregard hints at certain points in the conversation. Consider the log in Table 19. Seth started talking about a robot in line 1. The run-on sentence and lack of punctuation marks were a result of translation problems from speech to text by Google Assistant. When the agent suggested the story text, "*The robot is blue.*" in line 2, Seth disagreed but then encountered trouble in continuing with his story. It was not until the agent offered a pump "*How big is the robot?*" in line 6 that Seth was able to recover and continue with his story in line 7.

Seth perceives Orsen as an AI assistant that can guide him in knowing the part of the story he should talk about next and responds best when the agent acts as a *facilitator*. The agent can better support the needs of Seth to have an older peer who can give suggestions and hints to help move the story forward. The agent may utilize interaction resource tools, such as a checklist which is a computerized version of the everyday to-do-list (Terveen, 1995). The checklist, used in Alice (Chua et al., 2017), can benefit both the child and the agent in monitoring their progress by indicating the presence of different elements of a story as detected by the agent from the child's input.

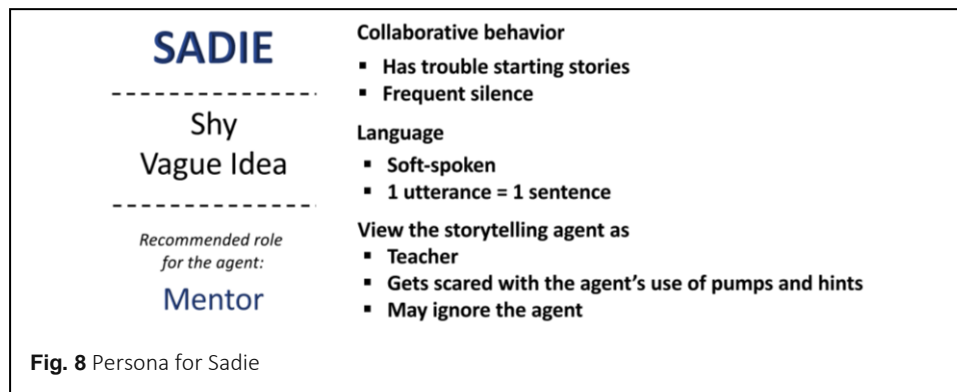


**Table 19** Seth's mixed responses to the agent's attempts at contributing to the story [I2-P18]

Line	Turn	Text
1	Child	once upon a time i look at the trash can cause i was about to throw away my bottle until i saw broken robot and then and i use it to make a super robot and then
2	Orsen	The robot is blue.
3	Child	No
4	Orsen	I want to know more.
5	Child	got famous cause it is hero
6	Orsen	How big is the robot?
7	Child	Robot is big since it is super

Quiet and soft-spoken, **Sadie** (Figure 8) has trouble starting her story because she does not have a clear idea of the story she would like to tell, thus, she would immediately pass the turn to Orsen (Table 20). There are instances when she would avoid starting the conversation altogether because she does not know what to say. Often, a third party (e.g., another human participant) needs to be present to encourage Sadie to start talking to Orsen. Sadie also appears to be more fearful of the agent's responses and feels pressured with the agent's questions. She prefers if the agent would adopt a friendlier tone and language and is visibly more comfortable with the interaction when the agent makes mistakes.

Sadie perceives Orsen as a mentor who can provide guidance to motivate her to share a story. Baylor and Kim (2005) posits that a mentor who can motivate learners will be “*more appropriate in constructivist learning environments, where learners' voluntary engagement is more critical than their knowledge acquisition*”. This aptly describes collaborative storytelling. As a mentor, the agent does not simply give out suggested story text during its dialogue turn nor does it assume an authoritarian figure (Baylor & Kim, 2005). Instead, the agent gives frequent feedback as a form of reassurance as seen in line 6 of Table 20. It balances the use of pumps to encourage independent thinking (lines 4 and 8) and hints (line 2) to demonstrate competence and knowledge in storytelling.



**Table 20** Sadie relies on the agent to guide her through storytelling. The missing words are caused by speech-to-text translation problems [I2-P13]

Line	Turn	Text
1	Child	help me start
2	Orsen	What if we had a story about a king?
3	Child	ok
4	Orsen	Can you give me more details?
5	Child	once there was a king who is so greedy
6	Orsen	Interesting, so, who was king?
7	Child	one day there was a rich coming to his house
8	Orsen	Is there anything else you can add?
9	Child	so the witch curse him

## Role of conversational agents

### ***RQ3: What roles do children expect these conversational agents to portray?***

RQ3 aims to identify the roles that children expect the conversational agents to portray. To answer this, we first draw from Terveen's (1995) classification of human-AI collaboration: human emulation and human complementary. In *human emulation*, the conversational agent mimics human capabilities and thought processes in performing story understanding and generation tasks. *Human complementary*, on the other hand, works on the premise that humans and AI have asymmetric abilities. Collaboration is needed to exploit the unique abilities of humans and AI while delegating responsibilities so that each participating entity is assigned appropriate and distinct roles. In our study, we are primarily interested in the human complementary approach where children retain control over the theme and flow of their narrative while the agent provides the necessary support to help in completing the storytelling or story construction task.

As a collaborator, conversational agents can portray varying roles depending on the context of their usage. In a learning environment, these pedagogical agents may serve as a facilitator, guide, tutor, learning companion, and teachable agent. A learning companion is a knowledgeable peer with the same, slightly above, or lower skill level than the student and who undergoes a similar learning experience (Theophilou et al., 2023). It is designed to accompany learners as they navigate through the learning environment and in the process of doing so, may establish a social relationship to foster interest in the learning task (Khosrawi-Rad et al., 2022).

As a learning companion, collaboration is a crucial trait of our storytelling agent. It works with children by co-constructing their stories and utilizing various dialogue strategies such as pumping for details and suggesting story text. Anchored on the idea that we learn by teaching others (Leikin, 2006; Muis et al., 2016; Tang et al., 2004), our storytelling agent can also function as a teachable peer, a type of learning companion that receives knowledge from the student (Biswas et al., 2005; Okita & Schwartz, 2006; Song, 2017).

With a view of conversational agents as long-term learning companions, it is worth investigating in future studies on how affinity and trust can be developed through frequent social interactions (Nißen et al., 2021; Skjuve et al., 2021; Wang et al., 2020). Conversational agents that can recognize and exhibit social cues and be attuned to the emotions of the learners may enable learning companions to offer emotional support that may in turn impact learners' motivation in performing the required learning tasks.

## Conclusion

In this study, we investigated the roles that conversational agents can portray and the forms of collaboration that may occur during storytelling. Through structured surveys with 44

children, observations of the child-agent interaction, and analysis of the conversation logs, we found that communication breakdown, limitations in the language abilities of the interlocutors, challenges in working with the input device (speech versus text interfaces), and children's attitude inherent storytelling abilities affected the child-agent collaboration. When communication breakdowns occur, the agent utilized repair strategies such as those suggested in previous work (Beneteau et al., 2019; Cheng et al., 2018; Golinkoff, 1986; Most, 2022). The conversational agent can have slightly above, similar, or slightly lower skills than the children they are collaborating with so they can function as tutors, learning companions, or teachable peers, respectively.

Future storytelling agents need to adapt to the dynamism of the ongoing conversation and to address the specific needs of each child-storyteller. A collaborative agent is suited for Sage who perceives an agent as a co-authoring peer with the same developmental age as he is. A mentor agent can address the need of Sadie for guidance and motivation to address her limited storytelling abilities. Seth needs a facilitator agent that can monitor his progress towards the completion of the required task. And finally, Sam can demonstrate her storytelling abilities to a teachable peer. These findings may be utilized in future studies that seek to expand the utility of collaborative learning companions. It would be interesting to explore how the use of agents embodying each of the specified roles may affect the openness of children to share their stories.

For further work, we intend to enhance the validity of our findings by including intercoder reliability scores to quantitatively measure the degree of agreement between independent coders. Employing a coding scheme may enhance the consistency of our analysis across different datasets which is necessary as we expand our user studies to a larger population base. We will also build upon the text understanding and generation abilities of large language models in designing conversational agents that embody different persona to investigate how their adaptive nature may influence children's narrative expression and their response to the agent's attempt at collaboration.

## Limitations

Our work has limitations. First, our prototyped storytelling agent uses a rule-based approach in extracting story elements from children's utterances and subsequently generating responses according to an identified dialogue move. This constrained the ability of the agent in input understanding and in generating variant responses. Second, there is disparity with our sample size for each iteration of our experiments which could affect the generalizability of our findings. Lastly, there were differences in the prototypes of Orsen that we utilized in each iteration of our experiments – voice interface, text interface, teachable peer. While we did our best to present common findings and nuanced differences through our analysis of the observation checklist, survey forms, and conversation logs, we

may have missed less evident but still important considerations in the design of future learning companions that anchor its conversation on storytelling.

#### Abbreviations

AAAI: Association for the Advancement of Artificial Intelligence; AI: Artificial Intelligence; SIR: 5<sup>th</sup> Industrial Revolution; LLM: Large Language Models; RQ: Research question; UI/UX: User interface / user experience.

#### Acknowledgments

We would like to acknowledge the funding we received from the Department of Science and Technology – Philippine Council for Industry, Energy, and Emerging Technology Research and Development and from De La Salle University during our software development and data collection. We would also like to extend our gratitude to our participants (children from different local schools) for spending time to share their stories with our conversational agent. We would also like to thank their parents, guardians, and teachers for giving their informed consent allowing their children, wards, and students to participate in our study.

#### Authors' contributions

All authors equally shared in the research conceptualization, design, evaluation, and documentation of the results and findings.

#### Authors' information

Ethel Ong is a Professor of the College of Computer Studies at De La Salle University – Manila. A senior researcher for the University's Center for Language Technologies, she develops automated story generation systems and narrative-based conversational interfaces with applications in learning and healthcare. Christine Rachel De Jesus, Luisa Katherine Gilig, and Dionne Tiffany Ong are former undergraduate students of De La Salle University. They are now working in different computing industries in the country.

#### Funding

No funding was received in the preparation of this manuscript. During the course of the software development and data collection, funding was provided through (1) a research grant from the Department of Science and Technology – Philippine Council for Industry, Energy, and Emerging Technology Research and Development, and (2) counterpart funding from De La Salle University.

#### Availability of data and materials

We are bound by university ethics policies and guidelines to safeguard the raw data we collected from our participants. The ethics clearance provided by the university's Research Ethics Review Committee extends only to the use of anonymized and aggregated data in the generation of technical reports for the funding agency and for publication of research findings. The results we presented in our paper provide a synthesis of the data from which we drew our analysis and recommendations. Excerpts of conversation logs collected during the experiments can also be seen in the different parts of our paper. The full conversation logs cannot be shared with other researchers who are not part of the original team.

#### Declarations

##### Competing interests

The authors declare that they have no competing interests.

##### Author details

All authors are from the College of Computer Studies of De La Salle University, Manila, Philippines.

Received: 31 December 2023 Accepted: 20 August 2024

Published online: 1 January 2025 (Online First: 14 October 2024)

#### References

- Baskara, R., & Mukarto, M. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Journal of English Language Teaching and Applied Linguistics*, 7(2). <http://dx.doi.org/10.21093/ijeltal.v7i2.1387>
- Baylor, A., & 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>

- Beneteau, E., Richards, O. K., Zhang, M., Kientz, J. A., Yip, J., & Hiniker, A. (2019). Communication breakdowns between families and Alexa. In S. Brewster, G. Fitzpatrick, A. Cox & V. Kostakos (Eds.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). ACM. <https://doi.org/10.1145/3290605.3300473>
- Biswas, G., Leelawong, K., Belyne, K., & Adebisi, B. (2005). Case studies in learning by teaching behavioral differences in directed versus guided learning. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 828–833). Cognitive Science Society.
- Brinton, B., Fujiki, M., Loeb, D. F., & Winkler, E. (1986). Development of conversational repair strategies in response to requests for clarification. *Journal of Speech and Hearing Research*, 29(1), 75–81. <https://doi.org/10.1044/jshr.2901.75>
- Cassell, J., Ryokai, K., Klaff, J., & Tauzero, B. (2000). Story Spaces: Interfaces for children's voices. In M. Tremaine (Ed.), *Proceedings of CHI '00 Extended Abstracts on Human Factors in Computing Systems* (pp. 243–244). ACM. <https://doi.org/10.1145/633292.633434>
- Cassell, J., Tartaro, A., Rankin, Y., Oza, V., & Tse, C. (2007). Virtual peers for literacy learning. *Educational Technology*, 47(1), 39–43.
- Cheng, Y., Yen, K., Chen, Y., Chen, S., & Hiniker, A. (2018). Why doesn't it work? Voice-driven interfaces and young children's communication repair strategies. In M. N. Giannakos, L. Jaccheri & M. Divitini (Eds.), *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 337–348). ACM. <https://doi.org/10.1145/3202185.3202749>
- Chua, H. G., Cu, G. E., Ibarrientos, C. P., Paguiligan, M. D., & Ong, E. (2017). Towards a virtual peer that writes stories with children. In W. Chen et al. (Eds.), *Proceedings of the 25th International Conference on Computers in Education* (pp. 95–97). Asia-Pacific Society for Computers in Education.
- Cohen, N. J. (2010). The impact of language development on the psychosocial and emotional development of young children. In R. E. Tremblay, M. Boivin & R. D. Peters (Eds.), *Encyclopedia on early childhood development*. <https://www.child-encyclopedia.com/language-development-and-literacy/according-experts/impact-language-development-psychosocial-and>
- Coleman, J. (2012, March 30). Why collaborative storytelling is the future of marketing. *Fast Company*. <https://www.fastcompany.com/1826645/why-collaborative-storytelling-future-marketing>
- Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., Earley, D., & Bandeira, N. (2017). "What can I help you with?" Infrequent users' experiences of intelligent personal assistants. In M. Jones, M. Tscheligi, Y. Rogers & R. Murray-Smith (Eds.), *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 1–12). ACM. <https://doi.org/10.1145/3098279.3098539>
- Dam, R., & Siang, T. (2019). Personas – A simple introduction. *Interaction Design Foundation*. <http://www.interaction-design.org/literature/article/personas-why-and-how-you-should-use-them>
- Davenport, T., & Ronanki, R. (2016). The rise of cognitive agents: Will humans prefer to talk with machines? *Deloitte Insights*. <https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/rise-of-cognitive-agents-artificial-intelligence-applications.html>
- Dehouche, N. (2021). Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). *Ethics in Science and Environmental Politics*, 21, 17–23. <https://doi.org/10.3354/esep00195>
- DeVito, J. (2017). *Essentials of human communication* (9th ed.). Pearson.
- Dy, J., Brito, C., Tan, V., Lola, J., & Ong, E. (2020). Acquiring commonsense knowledge during collaborative storytelling. *IOP Conference Series: Materials Science and Engineering*, 1077, 012023. <https://doi.org/10.1088/1757-899X/1077/1/012023>
- Ellis, A. R., & Slade, E. (2023). A new era of learning: Considerations for ChatGPT as a tool to enhance statistics and data science education. *Journal of Statistics and Data Science Education*, 31(2), 128–133. <https://doi.org/10.1080/26939169.2023.2223609>
- Fischer, G. (2015). Information, participation, and collaboration overload - A design trade-off analysis. In B. R. Barricelli, G. Fischer, A. Mørch, A. Piccinno & S. Valtolina (Eds.), *Proceedings of the 3rd Edition of the International Workshop on Cultures of Participation in the Digital Age: Coping with Information, Participation, and Collaboration Overload Co-Located with the 5th International Symposium on End-User Development (IS-EUD 2015) (CEUR Workshop Proceedings)*, 1641 (pp. 1–9). CEUR-WS.org
- Golinkoff, R. (1986). 'I beg your pardon?' The preverbal negotiation of failed messages. *Journal of Child Language*, 13(3), 455–476. <https://doi.org/10.1017/S0305000900006826>
- GPTBot (n.d.). *Mastering ChatGPT: How to craft effective prompts (full guide)*. <https://gptbot.io/master-chatgpt-prompting-techniques-guide/>
- Grabb, D. (2023). The impact of prompt engineering in large language model performance: A psychiatric example. *Journal of Medical Artificial Intelligence*, 6. <https://doi.org/10.21037/jimai-23-71>
- Graesser, A., Chipman, P., Haynes, B., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612–618. <https://doi.org/10.1109/TE.2005.856149>
- Gregorio, A. M., Manual, S. J., Palmaares, A. J., Tang, S. M., & Ong, E. (2023). Generating interactive stories with ChatGPT to teach Filipino values. In J. L. Shih et al. (Eds.), *Proceedings of the 31st International Conference on Computers in Education*, vol. 2 (pp. 52–61). Asia-Pacific Society for Computers in Education.
- Guszcza, J., Lewis, H., & Evans-Greenwood, P. (2017, January 23). Cognitive collaboration: Why humans and computers think better together. *Deloitte Insights*. <https://www2.deloitte.com/us/en/insights/deloitte-review/issue-20/augmented-intelligence-human-computer-collaboration.html>



- Habok, A., & Magyar, A. (2018). The effect of language learning strategies on proficiency, attitudes, and school achievement. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.02358>
- Halpin, H., Moore, J. D., & Robertson, J. (2004). Towards automated story analysis using participatory design. In B. Barry & K. Brooks (Eds.), *Proceedings of the 1st ACM Workshop on Story Representation, Mechanism and Context* (pp. 75–83). ACM. <https://doi.org/10.1145/1026633.1026652>
- Hayes, P. J., & Reddy, R. (1983). Steps toward graceful interaction in spoken and written man-machine communication. *International Journal of Man-Machine Studies*, 19(3), 231–284. [https://doi.org/10.1016/S0020-7373\(83\)80049-2](https://doi.org/10.1016/S0020-7373(83)80049-2)
- Hirst, G. (1991). Does conversation analysis have a role in computational linguistics? Review article for computers and conversations. *Computational Linguistics*, 17(2), 211–227.
- Hobert, S., & Meyer von Wolff, R. (2019). Say hello to your new automated tutor – A structured literature review on pedagogical conversational agents. *Wirtschaftsinformatik*. Siegen, Germany.
- Jurkic, A., Halliday, S. E., & Hascher, T. (2023). The relationship of language and social competence of preschool- and kindergarten-age single and dual language learners in Switzerland and Germany. *Early Childhood Research Quarterly*, 63, 72–83. <https://doi.org/10.1016/j.ecresq.2023.02.003>
- Khosrawi-Rad, B., Schlimbach, R., Strohmman, T., & Robra-Bissantz, S. (2022). Design knowledge for virtual learning companions. In *Proceedings of the International Conference on Information Systems Education and Research*.
- Kim, Y., Baylor, A. L., & Group, P. (2006). Pedagogical agents as learning companions: The role of agent competency and type of interaction. *Educational Technology Research and Development*, 54(3), 223–243. <https://doi.org/10.1007/s11423-006-8805-z>
- Klimova, B., Pikhart, M., Polakova, P., Cerna, M., Yayilgan, S. Y., & Shaikh, S. (2023). A systematic review on the use of emerging technologies in teaching English as an applied language at the university level. *Systems*, 11(1), 42. <https://doi.org/10.3390/systems11010042>
- Klüwer, T. (2011). From chatbots to dialogue systems. In D. Perez-Marin & I. Pascual-Nieto (Eds.), *Conversational agents and natural language interaction: Techniques and effective practices* (pp. 1–22). IGI Global. <https://doi.org/10.4018/978-1-60960-617-6.ch001>
- Leikin, R. (2006). Learning by teaching: The case of Sieve of Eratosthenes and one elementary school teacher. In R. Zazkis & S. Campbell (Eds.), *Number theory in mathematics education: Perspectives and prospects* (pp. 115–140). Erlbaum.
- Li, P.-H., Lee, H.-Y., Cheng, Y.-P., Starčič, A. I., & Huang, Y.-M. (2023). Solving the self-regulated learning problem: Exploring the performance of ChatGPT in mathematics. In Y. M. Huang & T. Rocha (Eds.), *Innovative Technologies and Learning. ICITL 2023. Lecture Notes in Computer Science*, vol 14099 (pp. 77–86). Springer, Cham. [https://doi.org/10.1007/978-3-031-40113-8\\_8](https://doi.org/10.1007/978-3-031-40113-8_8)
- Luger, E., & Sellen, A. (2016). “Like having a really bad PA”: The gulf between user expectation and experience of conversational agents. In J. Kaye, A. Druin, C. Lampe, D. Morris & J. P. Hourcade (Eds.), *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5286–5297). ACM. <https://doi.org/10.1145/2858036.2858288>
- Mayo, M. (2023, July 20). Prompt engineering for effective interaction with ChatGPT. *Machine Learning Mastery*. <https://machinelearningmastery.com/prompt-engineering-for-effective-interaction-with-chatgpt/>
- Morrissey, K., & Kirakowski, J. (2013). ‘Realness’ in chatbots: Establishing quantifiable criteria. In M. Kurosu (Ed.), *Human-Computer Interaction. Interaction Modalities and Techniques. HCI 2013. Lecture Notes in Computer Science*, vol 8007 (pp. 87–96). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-39330-3\\_10](https://doi.org/10.1007/978-3-642-39330-3_10)
- Most, T. (2002). The use of repair strategies by children with and without hearing impairment. *Language, Speech, and Hearing Services in Schools*, 33(2), 112–123. [https://doi.org/10.1044/0161-1461\(2002\)009](https://doi.org/10.1044/0161-1461(2002)009)
- Muis, K. R., Psaradellis, C., Chevrier, M., Di Leo, I., & Lajoie, S. P. (2016). Learning by preparing to teach: Fostering self-regulatory processes and achievement during complex mathematics problem solving. *Journal of Educational Psychology*, 108(4), 474–492. <https://doi.org/10.1037/edu0000071>
- Nißen, M., Selimi, D., Janssen, A., Cardona, D., Breitner, M., Kowatsch, T., & Wangenheim, F. (2021). See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons. *Computers in Human Behavior*, 127(2), 107043. <https://doi.org/10.1016/j.chb.2021.107043>
- Nielsen, L. (2013). Personas. In *The Encyclopedia of Human-Computer Interaction*, 2nd Ed. Interaction Design Foundation. <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/personas>
- Okita, S. Y., & Schwartz, D. L. (2006). When observation beats doing: Learning by teaching. In S. Barab, K. Hay & D. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences* (pp. 509–515). International Society of the Learning Sciences.
- Olafsson, S., & Bickmore, T. (2016). “That reminds me...” Towards a computational model of topic development within and across conversations. In *Proceedings of the 10th International Conference on language Resources and Evaluation* (pp. 5–8). European Language Resources Association.
- Ong, D. T., De Jesus, C. R., Gilig, L. K., Alburo, J. B., & Ong, E. (2018). Building a commonsense knowledge base for a collaborative storytelling agent. In K. Yoshida & M. Lee (Eds.), *Knowledge Management and Acquisition for Intelligent Systems. PKAW 2018. Lecture Notes in Computer Science*, vol 11016 (pp. 1–15). Springer, Cham. [https://doi.org/10.1007/978-3-319-97289-3\\_1](https://doi.org/10.1007/978-3-319-97289-3_1)
- OpenAI. (2023). *Introducing ChatGPT*. OpenAI. <https://openai.com/blog/chatgpt>

- OpenAI. (n.d.). ChatGPT prompt engineering for developers. OpenAI. <https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/>
- Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *Software Quality Professional*, 19(3), 25–36.
- Robertson, J., & Wiemer-Hastings, P. (2002). Feedback on children's stories via multiple interface agents. In S. A. Cerri, G. Gouardères & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems. ITS 2002. Lecture Notes in Computer Science*, vol 2363 (pp. 923–932). Springer Berlin Heidelberg. [https://doi.org/10.1007/3-540-47987-2\\_92](https://doi.org/10.1007/3-540-47987-2_92)
- Roe, J., Renandya, W. A., & Jacobs, G. M. (2023). AI-powered writing tools and their applications for academic integrity in the language classroom. *Journal of English and Applied Linguistics*, 2(1), 3. <https://doi.org/10.59588/2961-3094.1035>
- Roemmele, M., & Gordon, A. S. (2015). Creative Help: A story writing assistant. In H. Schoenau-Fog, L. Bruni, S. Louchart & S. Baceviciute (Eds.), *Interactive Storytelling. ICIDS 2015. Lecture Notes in Computer Science*, vol 9445 (pp. 81–92). Springer International Publishing. [https://doi.org/10.1007/978-3-319-27036-4\\_8](https://doi.org/10.1007/978-3-319-27036-4_8)
- Roscoe, R. D., Wagster, J., & Biswas, G. (2008). Using teachable agent feedback to support effective learning-by-teaching. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 2381–2386). Cognitive Science Society.
- Ryokai, K., Vaucelle, C., & Cassell, J. (2002). Literacy learning by storytelling with a virtual peer. In G. Stahl (Ed.), *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community* (pp. 352–360). International Society of the Learning Sciences. <https://doi.org/10.3115/1658616.1658666>
- Schroeder, N. L. (2016). Pedagogical agents for learning. In D. H. Choi, A. Dailey-Hebert & J. S. Estes (Eds.), *Emerging tools and applications of virtual reality in education* (pp. 216–238). IGI Global.
- Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018). “Hey Alexa, What’s up?” A mixed-methods studies of in-home conversational agent usage. In I. Koskinen, Y.-k. Lim, T. Cerratto-Pargman, K. Chow & W. Odom (Eds.), *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 857–868). ACM. <https://doi.org/10.1145/3196709.3196772>
- Singh, P., Barry, B., & Liu, H. (2004). Teaching machines about everyday life. *BT Technology Journal*, 22(4), 227–240. <https://doi.org/10.1023/B:BTJ.0000047601.53388.74>
- Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My chatbot companion—A study of human-chatbot relationships. *International Journal of Human-Computer Studies*, 149, 102601. <https://doi.org/10.1016/j.ijhcs.2021.102601>
- Song, D. (2017). Designing a teachable agent system for mathematics learning. *Contemporary Educational Technology*, 8(2), 176–190.
- Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.
- Sun, M., Leite, I., Lehman, J. F., & Li, B. (2017). Collaborative storytelling between robot and child: A feasibility study. In P. Blikstein & D. Abrahamson (Eds.), *Proceedings of the 2017 Conference on Interaction Design and Children* (pp. 205–214). ACM. <https://doi.org/10.1145/3078072.3079714>
- Surameery, N. M. S., & Shakor, M. Y. (2023). Use ChatGPT to solve programming bugs. *International Journal of Information Technology and Computer Engineering*, 3(1), 17–22. <https://doi.org/10.55529/ijitc.31.17.22>
- Swanson, R., & Gordon, A. S. (2012). Say Anything: Using textual case-based reasoning to enable open-domain interactive storytelling. *ACM Transactions on Interactive Intelligent Systems*, 2(3), 1–35. <https://doi.org/10.1145/2362394.2362398>
- Tang, T. S., Hernandez, E. J., & Adams, B. S. (2004). “Learning by Teaching”: A peer-teaching model for diversity training in medical school. *Teaching and Learning in Medicine*, 16(1), 60–63. [https://doi.org/10.1207/s15328015tlm1601\\_12](https://doi.org/10.1207/s15328015tlm1601_12)
- Terveen, L. G. (1995). Overview of human-computer collaboration. *Knowledge-Based Systems*, 8(2-3), 67–81. [https://doi.org/10.1016/0950-7051\(95\)98369-H](https://doi.org/10.1016/0950-7051(95)98369-H)
- Theophilou, E., Schwarze, V., Börsting, J., Sánchez-Reina, R., Scifo, L., Lomonaco, F., Aprin, F., Ognibene, D., Taibi, D., Hernández-Leo, D., & Eimler, S. (2023). Empirically investigating virtual learning companions to enhance social media literacy. In G. Fulantelli, D. Burgos, G. Casalino, M. Cimitile, G. Lo Bosco & D. Taibi (Eds.), *Higher Education Learning Methodologies and Technologies Online. HELMeTO 2022. Communications in Computer and Information Science*, vol 1779 (pp. 345–360). Springer, Cham. [https://doi.org/10.1007/978-3-031-29800-4\\_27](https://doi.org/10.1007/978-3-031-29800-4_27)
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Ureta, J., Brito, C. I., Dy, J. B., Santos, K. A. F., Villaluna, W. L., & Ong, E. (2020). At home with Alexa: A tale of two conversational agents. In P. Sojka, I. Kopeček, K. Pala & A. Horák (Eds.), *Text, Speech, and Dialogue. TSD 2020. Lecture Notes in Computer Science*, vol 12284 (pp. 495–503). Springer. [https://doi.org/10.1007/978-3-030-58323-1\\_53](https://doi.org/10.1007/978-3-030-58323-1_53)
- Wallach, G., & Butler, K. (1994). *Language learning disabilities in school-age children and adolescents: Some principles and applications*. Pearson.

- Wang, A. H., & Cassell, J. (2003). Co-authoring, corroborating, criticizing: Collaborative storytelling for literacy learning. In *Proceedings of the Vienna Workshop '03: Educational Agents - More than Virtual Tutors*. Vienna, Austria.
- Wang, Q., Jing, S., Camacho, I., Joyner, D., & Goel, A. (2020). Jill Watson SA: Design and evaluation of a virtual agent to build communities among online learners. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguy, P. Bjørn, S. Zhao, B. P. Samson & R. Kocielnik (Eds.), *Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–8). ACM. <https://doi.org/10.1145/3334480.3382878>
- Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2286. <https://doi.org/10.29333/ejmste/13272>
- Warren, S. J., Dondlinger, M. J., & Barab, S. A. (2008). A MUVE towards PBL writing: Effects of a digital learning environment designed to improve elementary student writing. *Journal of Research on Technology in Education*, 41(1), 113–140. <https://doi.org/10.1080/15391523.2008.10782525>
- Wik, P., & Hjalmarsson, A. (2009). Embodied conversational agents in computer assisted language learning. *Speech Communication*, 51(10), 1024–1037. <https://doi.org/10.1016/j.specom.2009.05.006>
- Williams, R., Park, H. W., & Breazeal, C. (2019). A is for artificial intelligence: The impact of artificial intelligence activities on young children's perceptions of robots. In S. Brewster, G. Fitzpatrick, A. Cox & V. Kostakos (Eds.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–11). ACM. <https://doi.org/10.1145/3290605.3300677>
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet? A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, 4, 18. <https://doi.org/10.3389/frai.2021.654924>
- Yilmaz, R., & Yilmaz, F. G. K. (2023). Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. *Computers in Human Behavior: Artificial Humans*, 1(2), 100005. <https://doi.org/10.1016/j.chbah.2023.100005>
- Zakharov, K., Mitrovic, A., & Johnston, L. (2007). Pedagogical agents trying on a caring mentor role. In R. Luckin, K. R. Koedinger & J. Greer (Eds.), *Proceedings of the 2007 Conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work* (pp. 59–66). IOS Press.
- Zhao, G., Ailiya, & Shen, Z. (2012). Learning-by-Teaching: Designing teachable agents with intrinsic motivation. *Journal of Educational Technology and Society*, 15(4), 62–74.

## Publisher's Note

The Asia-Pacific Society for Computers in Education (APSCE) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Research and Practice in Technology Enhanced Learning (RPTEL)**  
is an open-access journal and free of publication fee.