

# Exploring the Perspectives of Graduate Students on Generative AI Chatbots in Educational Contexts

RUDAIBA ADNIN, Northeastern University, USA

HUAN-FENG YEH, Northeastern University, USA

ATHARVA, Northeastern University, USA

Generative AI chatbots (e.g., ChatGPT, Google Bard) are increasingly getting attention in educational contexts. However, there is limited research on how graduate students use generative AI chatbots and their experiences. Therefore, we investigate how graduate students use generative AI chatbots and their experiences with those chatbots. Consequently, we performed semi-structured interviews with 06 graduate students, designed user interface diagrams, and conducted user evaluation of those design diagrams with 05 graduate students. We discover that, although graduate students are using those chatbots frequently in educational tasks, they encounter several challenges, such as too-detailed responses, inaccurate responses, and privacy issues. Focusing on reducing those challenges, we create design diagrams. Our user evaluation of those design diagrams shows that students appreciated the design diagrams.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**.

Additional Key Words and Phrases: Education, Chatbots, Generative AI

## ACM Reference Format:

Rudaiba Adnin, Huan-Feng Yeh, and Atharva. 2018. Exploring the Perspectives of Graduate Students on Generative AI Chatbots in Educational Contexts. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 10 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 INTRODUCTION

The use of generative AI chatbots is increasingly being explored in educational contexts. In this light, ChatGPT, a generative AI Chatbot has gained significant attention [1]. Several studies [2, 3] demonstrated the potential benefits of generative AI chatbots, for example, ChatGPT can be used as a support tool that assists students in educational tasks [4], assists in completing assignments [2], and data analysis [5]. While ChatGPT has potential benefits in educational contexts, it also poses numerous challenges, such as providing an opportunity to cheat in assignments [2]. This raises concerns about losing creativity and critical skills in students [6, 7]. In addition, ChatGPT can assist students in answering assessment questions without having prior knowledge of the topic, however, cannot be fully relied on because it often generates invalid solutions [5, 8, 9].

However, none of the prior studies [5, 8, 9] focused specifically on graduate students and how those students use these generative AI chatbots. To address this, we explore the purpose of graduate students prior to adopting generative AI chatbots (e.g. ChatGPT, Google Bard) in educational contexts. We further uncover how graduate students understand the capabilities and limitations (e.g. hallucinations) of these chatbots. We also explore the future outlook and expectations of the students from those chatbots and uncover how we can design new solutions to improve their experiences.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

## 1.1 Research Questions

We explore the following research questions in our study.

- **RQ1:** How do graduate students use generative AI chatbots (e.g. ChatGPT, Google Bard) in educational contexts?
- **RQ2:** How do they understand the capabilities and limitations of these chatbots? What are their future outlook and expectations pertinent to these chatbots?
- **RQ3:** How might we design new solutions to enhance their experience with these chatbots in educational contexts?

## 1.2 Contribution

We make the following contributions through our study.

- Through semi-structured interviews, we investigated the existing practices and experiences of graduate students pertinent to generative AI chatbots. Thus, our study will inform the HCI community about how graduate students use generative AI chatbots and their perspectives.
- We provided novel design features for students' effective interaction with generative AI chatbots. We also performed a user evaluation of those design features. As a result, our study will further inform the HCI community about design implications pertinent to generative AI chatbots that focus on graduate students.

## 2 RELATED WORK

A feedback study [10] with children uncovered that ChatGPT has the potential to adapt its responses to suit students with a broad range to cover their specific requirements. To understand if ChatGPT can support children in discovering information for educational tasks, a study [11] analyzed the responses of ChatGPT and identified open challenges. The findings showed that although ChatGPT provides understandable responses to the prompts, it often does not generate reliable responses. Moreover, in [12], the authors used ChatGPT to correct open-ended questions and found that ChatGPT can assist with the correction process. A study [13] proposed a learning platform, RECIPE based on ChatGPT to facilitate EFL (English as a foreign language) education. The platform has two prompts: a prompt with the teacher role and a prompt for students to initiate dialogue. Moreover, a literature review [4] showed that ChatGPT has become a support tool that can assist in tasks, such as data compression and student queries. As a result, ChatGPT assists in performing tasks that can take a long time to complete for students and to discuss issues while completing an assignment [2]. ChatGPT has the potential to be integrated into data science courses as it assists in analyzing data without human effort [5]. Such generative AI chatbots also have the potential to be integrated into programming courses because they generate accurate and organized solutions [8]. However, it is concerning because the information provided by ChatGPT might not be valid [5, 8] and often ChatGPT hallucinates [14]. A study [9] further uncovered that ChatGPT is not fully reliable for learning about a topic for which students have no prior knowledge.

Although ChatGPT has the potential in educational contexts, however, it might cause various challenges and misuse, such as cheating by simply copying its outputs in assignments [2]. Such cheating is difficult to detect because ChatGPT generates human-like texts [6]. Therefore, students should point out how much assistance they had taken from these generative AI ChatGPT [4]. In this light, researchers [15, 16] suggested that although students might use generative AI tools, there should be guidelines for integrating generative AI tools into coursework to reduce the negative impact of these technologies in education. To understand the impact of generative AI chatbots, such as ChatGPT on assessment, a study [7] considered the perspectives of both students and educators. The results showed that educators believe that

these chatbots would promote critical thinking among students whereas students believe that these chatbots can be the cause of loss of creativity. Moreover, interpersonal interaction is important for interpersonal skill development and concept understanding. Here, integrating ChatGPT in educational contexts might reduce such interpersonal skills [3]. In [17], researchers explored how generative AI chatbots might automate the jobs of CS graduates which might cause their job displacement. However, none of these prior studies [5, 8, 9] focused on how graduate students use these generative AI chatbots and their unique challenges – which we will investigate in our study.

### 3 RESEARCH METHOD

We conducted our study in three phases: semi-structured interviews, designing user interface diagrams, and user evaluation.

#### 3.1 Semi-structured Interviews

We conducted semi-structured interviews with 06 graduate students. We recruited participants who provided their verbal consent to participate in the interviews. Table 1 shows the demographics of the interviewees. The interviews lasted approximately 40-60 minutes. We conducted the session over video-conferencing platforms (e.g., Zoom). Before the sessions, we stated that we were interested in learning about how they use generative AI chatbots (e.g. ChatGPT, Google Bard) in educational contexts. We also requested the participants to demonstrate how they use these chatbots. We developed an interview guide before conducting the interviews. We asked the following types of questions in the interviews.

- Practices: Can you tell me about your practices while using generative AI chatbots?
- Reflection: What is most important to you while using those chatbots? Do you understand the capabilities and limitations (e.g. hallucinations) of those chatbots?
- Challenges and Workarounds: What are the challenges you encounter while using those chatbots? What is your workaround to reduce those challenges?
- Expectations: What kinds of features do you wish you had while using those chatbots?

During interviews, we asked the students to prompt the generative AI chatbot (e.g., ChatGPT, Google Bard). For prompting, we focused on topics related to - generating creative responses, generating hallucinatory responses, and generating responses about a topic that the students have no prior knowledge of. Then, we asked the participants about their perspectives on the generated responses. After the interviews, we performed a reflexive thematic analysis [18] of the interview data and organized the findings into four high-level themes (Table 2).

#### 3.2 Designing user interface diagrams

We designed user interface diagrams (Figure 1, 2, and 3) that demonstrate possible workflows of new features for a generative AI chatbot. The new features were generated from the analysis of semi-structured interviews. The new features we designed include – (1) providing a concise response with options for details, (2) providing sources of a generated response, (3) providing prompt suggestions, (4) an option to run and modify generated code, and (5) multimodal forms of questioning opportunity. These design diagrams were used to provide participants with an understanding of how those features might be integrated into generative AI chatbots.

### 3.3 User Evaluation

We evaluated the user interface diagrams with 05 graduate students and gathered their feedback. We recruited participants who provided their verbal consent to participate in the evaluation. Table 3 shows the demographics of the participants. The interviews lasted approximately 30-40 minutes. We conducted the session over video-conferencing platforms (e.g., Zoom). During the evaluation, we asked the graduate students to prompt an existing generative AI chatbot (e.g., ChatGPT, Google Bard) with similar prompt scenarios that we used for designing the user interface diagrams. We demonstrated the user interface diagrams to the participants and asked their feedback. After the interview session, we provided participants with a questionnaire to provide their feedback about the usefulness of those design features.

Name	Education Level	Age	Pronouns
P1	PhD	25	He/Him
P2	PhD	26	He/Him
P3	PhD	27	He/Him
P4	PhD	25	She/Her
P5	MSc	23	He/Him
P6	PhD	26	She/Her

Table 1. Demographics of interviewed graduate students (n=06)

Themes	Description
Purpose and value of generative AI chatbots	Why graduate students use generative AI chatbots.
Assumptions about generating AI chatbots	The assumptions of graduate students related to the capabilities and limitations of generative AI chatbots.
Challenges while using generative AI chatbots	The issues graduate students encounter while using generative AI chatbots.
Expectations from generative AI chatbots	The expectations of graduate students while using generative AI chatbots.

Table 2. Themes and associated descriptions

**User:** I want to learn about Generative AI

**ChatBOT:** Generative AI is a fascinating field! It's a branch of artificial intelligence focused on creating or generating new content, often in the form of images, text, music, or other media. These AI models are designed to understand patterns from existing data and use that understanding to generate new, original content.

Click on the following options to learn more about generative AI.

Types of Generative Models
Applications of Generative AI
Learning Generative AI

**User:** What is AI?

**ChatBOT:** AI stands for artificial intelligence. It's a branch of computer science that focuses on creating systems or machines capable of performing tasks that typically require human intelligence. AI encompasses various techniques such as machine learning, where algorithms learn from data, natural language processing, which enables machines to understand and process human language, computer vision for interpreting and understanding visual information, and robotics for creating physical systems that can operate autonomously or semi-autonomously. AI aims to simulate intelligent behavior, enabling machines to perceive, reason, learn, and solve problems.

I received this information from the following sources

(1) <https://www.ibm.com/topics/artificial-intelligence>  
(2) [https://en.wikipedia.org/wiki/Artificial\\_intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence)

Fig. 1. From left to right, user interface diagrams for providing a concise response with options for details and providing sources of a generated response

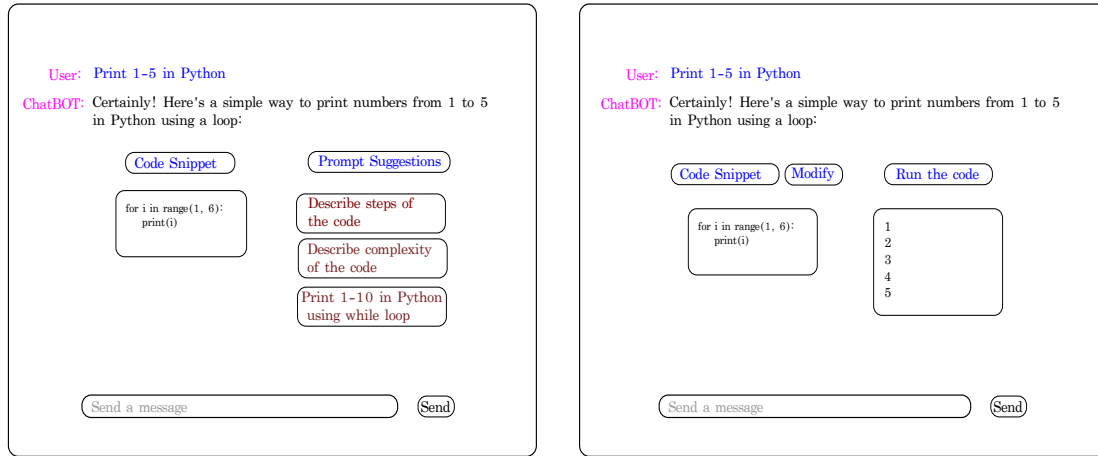


Fig. 2. From left to right, user interface diagrams for providing prompt suggestions and an option to run and modify generated code



Fig. 3. From left to right, user interface diagram for multimodal questioning opportunity.

Name	Education Level	Age	Pronouns
U1	PhD	26	He/Him
U2	PhD	27	He/Him
U3	MSc	23	He/Him
U4	PhD	25	She/Her
U5	PhD	28	He/Him

Table 3. Demographics of participants in user evaluation (n=05)

## 4 RESEARCH FINDINGS

### 4.1 Findings from the Interviews

Here, we elaborate findings from the interviews based on four high-level themes (Table 2)

**4.1.1 Purpose and value of generative AI chatbots.** Participants explained that generative AI chatbots assist them in performing educational tasks that would have taken a lot of time without the assistance of those chatbots. In this regard, P1 explained, *‘If I want to plot something really quick or if I don’t want to waste time in googling, then I can get help from them.’*

Participants explained that those generative AI chatbots have influenced a lot of aspects of their educational lives and brought a new shift in their lives. In this light, P4 said, *‘Previously, I used to read a lot of blogs or articles to implement a system. And now I spend a lot less time on the reading part rather I ask ChatGPT or BARD to write that code for me or explain that thing to me. So, I think my learning has gone through a different kind of shift due to those chatbots.’*

Participants also compared existing generative AI chatbots. Some felt that Google Bard is better because Bard provides several solutions for a question. Participants further mentioned their favorite features in those chatbots, for example, providing them with multiple solutions for a question. They also appreciated how they could have a conversation with those chatbots about accurate solutions. In this light, P6 mentioned, *'One feature that I like is that if we correct them, like if I say, your provided code is wrong, they give me a better solution. Another is they will provide me with multiple solutions and I can choose one which seems more appropriate.'*

**4.1.2 Assumptions about generating AI chatbots.** Participants assumed that generative AI chatbots often provide them with accurate responses, for example, P4 pointed out, *'In most cases, it gives me a kind of correct answer or close to accurate answer.'* However, some participants also believed that those generative AI chatbots will always skip some important information. In this light, P3 articulated, *'Those chatbots always miss some key points. Even if you mentioned them in the prompt, those chatbots are going to miss it.'*

Moreover, participants believed that those generative AI chatbots could not provide them with complete solutions, as P2 explained, *'The problem with generative AI like ChatGPT or BARD is that you cannot build complex or long systems. I mean it can generate something, but it messes up. The response is so generic that I need to modify a lot of things to make it work. It saves me a bit of time, but I don't think it's going to be the complete solution or the whole solution.'*

**4.1.3 Challenges while using generative AI chatbots.** Participants encountered several challenges while using generative AI chatbots, for example, not being able to implement complex or long systems, and not getting appropriate responses for information that is published after a certain date. A frequent challenge participants encounter is that those chatbots do not understand their questions and context. In this light, P6 explained, *'Sometimes they cannot exactly figure out the problem. I mean, if I give them the error warning, they will point out a lot of things, but they cannot choose which one is the major cause of that error. They will provide more like a generic thing. So, I have to ask the same question twice or thrice with more details to get an accurate answer.'*

Another significant challenge participants encounter is that those chatbots provide inaccurate information to them. In this light, P2 explained *'The issue is that often I have to make sure that ChatGPT is giving me the right information because it can give me wrong answers, or some made-up answers which can be harmful. So, I'm using it mostly for coding purposes, not for getting information about scientific articles or reading purposes.'* As a workaround, some participants do not use those chatbots for logical reasoning because those chatbots often provide inaccurate responses and participants cannot view the accuracy of the responses.

Participants also feel concerned about their privacy while using those chatbots. As a workaround, they do not provide their creative writing as input directly to those chatbots and do not use those chatbots when they have to share their personal information, such as their health-related information.

**4.1.4 Expectations from generative AI chatbots.** Participants desired that the generated response was not too detailed while they were looking for concise responses. In this regard, P6 mentioned, *'Often the responses are too detailed. I want the responses to be on point and concise.'*

Participants also expected that there would be multimodal forms of question-asking opportunities while using those chatbots. In this light, P5 articulated, *'I expect that there will be multimodal forms of asking question opportunities.'*

## 4.2 Design Implications for User Interface Diagrams

From the semi-structured interviews, we uncover the following design features.

- **Concise Response:** With this feature (Figure 1), users will always view concise answers. if they want more information, there will be options for them to select.
- **Showing the sources:** With this feature (Figure 1), users will view the sources of the generated responses of the generative AI chatbots, for example, references or links.
- **Prompt Suggestions:** With this feature (Figure 2), users will receive assistance on prompting the generative AI chatbots. They will view several prompt suggestions along with the generated response.
- **Run and modify the code:** With this feature (Figure 2), users will be able to run and modify the code snippets generated by the generative AI chatbots and view the output.
- **Multimodal forms of asking questions opportunities:** With this feature (Figure 3), users will be able to ask their questions to the generative AI chatbots in multimodal forms, such as Images, PDFs, etc.
- **Categorize Conversation:** With this implication, users will be able to have categorized conversations with generative AI chatbots based on specific needs, for example, Q/A, Fact-Checking, Creative tasks, etc.
- **Understanding Contexts:** With this implication, generative AI chatbots will remember the contexts of the whole chat session and provide a personalized answer. Here, the chatbot would clarify what users just said and form a real 'conversation'. There will also be an opportunity of not remembering contexts so that those chatbots would not remember the context of the whole chat session.

#### 4.3 Findings from the User Evaluation

	Concise Response	Mentioning Source	Prompt Suggestions	Run and modify the Code	Multimodal Question Asking Opportunity
The feature would enable me to accomplish tasks more quickly	4.2	4.0	4.6	4.0	4.4
The feature would improve my task performance	3.8	3.6	4.2	3.8	4.2
The feature in my task would increase my productivity	4.2	4.2	4.6	4.0	4.4
The feature would enhance my effectiveness of the task	4.0	4.0	4.6	4.2	4.8
The feature would make it easier to do my task	4.2	4.2	4.8	4.2	4.6
I would find the feature useful in my tasks	4.6	4.4	4.8	4.4	4.8
I want this feature to be integrated into the system	4.4	4.4	4.8	4.2	4.8
If the feature is available, I would use the system more	4.4	4.2	4.6	4.2	4.8

Table 4. Mean of Usefulness of features (n=05) (Response scale: 1=Strongly Disagree, ' 5=Strongly Agree)

During user evaluation, we asked participants to prompt ChatGPT with some specific questions. We decided to select specific questions to maintain consistency among all the participants and with the design diagrams.

All participants replied positively about the design features. Some provided suggestions regarding how we can improve the design feature, for example, regarding *'Run and Modify the Code'* feature, U2 suggested, *'Rather than showing side by side, the output should be shown after the code snippet section'*. Moreover, regarding *'Prompt Suggestions'* feature, U5 articulated, *'This feature will help me a lot because I often struggle to prompt properly to ChatGPT'*.

In addition, Table 4 shows the mean values of the eight questions pertaining to the design features. Here, we can observe that the *'Prommpt Suggesions'*, *'Multimodal Question Asking Opportunity'*, and *'Concise response'* features are the most desired among the participants. In addition, all the features were positively received by the participants.

## 5 DISCUSSION

In this section, we answer the research questions that we explored in our study.

### 5.1 Dependency on Generative AI Chatbots

In our study, we discovered that graduate students frequently use generative AI chatbots for their educational tasks. This supports prior research [10] on how ChatGPT has the potential to suit students with a broad range and their specific requirements. Students believed that the generative AI chatbots can assist them in performing educational tasks that would have taken a lot of time without the assistance of those chatbots. This extends prior work [2, 4, 12] on how ChatGPT has become a support tool that can assist in educational tasks [2, 4]. In addition, students mentioned their favorite features in those chatbots, for example, having a conversation with those chatbots and the opportunity to view multiple solutions for a question. As a result, students felt that those generative AI chatbots have brought a new shift in their lives, especially in terms of education. Students often assumed the capabilities of generative AI chatbots, for example, providing them with accurate responses, always skipping some important information, and not providing them with complete solutions. This extends prior work [7] on how students assume that those chatbots can be the cause of losing their creativity.

### 5.2 Challenges and Workarounds

Although students are frequently using generative AI chatbots, they encounter several challenges while using them. In this light, getting inaccurate responses is a significant challenge – which contributes to prior work [5, 8, 9, 11, 14] on how generative AI chatbots often do not generate reliable responses. As a workaround, students are using those chatbots for specific tasks, such as programming rather than information seeking, where they can validate the responses of those chatbots. This supports prior work [5, 8] on how ChatGPT has the potential to assist in specific tasks. Other challenges that students encounter include: not understanding the contexts of their questions, not being able to implement complex systems, and privacy concerns. As a workaround for privacy concerns, students do not provide their creative writing as input and do not share their personal information while questioning those chatbots. We further uncovered students expected some new features in these chatbots, for example, getting concise responses instead of detailed responses and multimodal forms of question-asking opportunities.

### 5.3 Appreciation for the New Features

During user evaluation, we discovered that students appreciated all five design features- concise responses, source mentioning, prompt suggestions, code modification, and multimodal question-asking. Here, the feature *'Prompt Suggestions'* was most appreciated among the participants and they wanted this feature to be integrated into the existing generative



AI chatbots. Students also provided suggestions regarding improving the design features. These findings contribute to prior work [13] that proposed a learning platform with new features to assist students.

## 6 LIMITATIONS OF THE STUDY

Our participants for semi-structured interviews and user evaluation were mostly CS majors which may have introduced selection bias in our study results. In addition, rather than all the features that were discovered from the semi-structured interviews, we drew diagrams for only five features. We also did not develop a fully working prototype where users can interact with those features.

## 7 FUTURE WORK OF THE STUDY

In the future, we want to increase and diversify participant demographics for a comprehensive understanding of how graduate students use generative AI chatbots. We also want to investigate impacts on their productivity, learning outcomes, and behavior of graduate students over time while using those chatbots. In addition, we want to focus on refining the design features based on user feedback from the user evaluation.

## 8 CONCLUSION

In our study, we uncovered that graduate students frequently use generative AI chatbots for their educational tasks and believed that these chatbots have brought about a new shift in their lives. However, they encountered several challenges during their interaction with those chatbots, for example, getting too detailed and inaccurate responses and privacy issues. As a workaround, students use generative AI chatbots for specific tasks where they can validate the responses and not provide their personal information while questioning. During user evaluation, we observed that students appreciated the design features that we generated from the findings of the interviews. As a result, our findings shed light on opportunities for the development of solutions to improve the experiences of graduate students while using generative AI chatbots.

## 9 EACH TEAM MEMBER'S CONTRIBUTION

This research work was a collaborative process. Everyone contributed to every phase of the work. Below we discuss each team member's contribution.

- **Rudaiba Adnin** formulated the research idea, research plan, and research team. She took 3 semi-structured interviews, designed and drew 3 user interface diagrams, and took 2 interviews for user evaluation. She also contributed to preparing the final project report and the presentation slide. She regularly attended weekly meetings with the research team.
- **Atharva** formulated the research idea, research plan, and research team. He took 1 semi-structured interview, designed and drew 1 user interface diagram, and took 2 interviews for user evaluation. He also contributed to preparing the final project report and the presentation slide. He regularly attended weekly meetings with the research team.
- **HuanFeng Yeh** formulated the research idea, research plan, and research team. He took 2 semi-structured interviews, designed and drew 1 user interface diagram, and took 1 interview for user evaluation. He also contributed to preparing the final project report and the presentation slide. He regularly attended weekly meetings with the research team.

## REFERENCES

- [1] Xiao Zhan, Yifan Xu, and Stefan Sarkadi. Deceptive ai ecosystems: The case of chatgpt. *arXiv preprint arXiv:2306.13671*, 2023.
- [2] Kamil Malinka, Martin Peresini, Anton Firc, Ondrej Hujnak, and Filip Janus. On the educational impact of chatgpt: Is artificial intelligence ready to obtain a university degree? In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, pages 47–53, 2023.
- [3] Michel Albonico and Paulo Júnior Varela. A report on the use of chatgpt in software engineering and systems analysis courses. In *Proceedings of the XXXVII Brazilian Symposium on Software Engineering*, pages 303–311, 2023.
- [4] Alfonso Renato Vargas-Murillo, Ilda Nadia Monica de la Asuncion Pari-Bedoya, and Francisco de Jesus Guevara-Soto. The ethics of ai assisted learning: A systematic literature review on the impacts of chatgpt usage in education. In *Proceedings of the 2023 8th International Conference on Distance Education and Learning*, pages 8–13, 2023.
- [5] Yong Zheng. Chatgpt for teaching and learning: An experience from data science education. *arXiv preprint arXiv:2307.16650*, 2023.
- [6] Fatemeh Mosaiyebzadeh, Seyedamin Pouriyeh, Reza Parizi, Nasrin Dehbozorgi, Mohsen Dorodchi, and Daniel Macêdo Batista. Exploring the role of chatgpt in education: Applications and challenges. In *The 24th Annual Conference on Information Technology Education*, pages 84–89, 2023.
- [7] Adele Smolansky, Andrew Cram, Corina Radulescu, Sandris Zeivots, Elaine Huber, and Rene F Kizilcec. Educator and student perspectives on the impact of generative ai on assessments in higher education. In *Proceedings of the Tenth ACM Conference on Learning@ Scale*, pages 378–382, 2023.
- [8] Eng Lieh Ouh, Benjamin Kok Siew Gan, Kyong Jin Shim, and Swavek Wlodkowski. Chatgpt, can you generate solutions for my coding exercises? an evaluation on its effectiveness in an undergraduate java programming course. *arXiv preprint arXiv:2305.13680*, 2023.
- [9] Abdulhadi Shoufan. Can students without prior knowledge use chatgpt to answer test questions? an empirical study. *ACM Trans. Comput. Educ.*, oct 2023. Just Accepted.
- [10] Emiliana Murgia, Maria Soledad Pera, Monica Landoni, and Theo Huibers. Children on chatgpt readability in an educational context: Myth or opportunity? In *Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, pages 311–316, 2023.
- [11] Emiliana Murgia, Zahra Abbasiantaeb, Mohammad Aliannejadi, Theo Huibers, Monica Landoni, and Maria Soledad Pera. Chatgpt in the classroom: A preliminary exploration on the feasibility of adapting chatgpt to support children's information discovery. In *Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, pages 22–27, 2023.
- [12] Gustavo Pinto, Isadora Cardoso-Pereira, Danilo Monteiro, Danilo Lucena, Alberto Souza, and Kiev Gama. Large language models for education: Grading open-ended questions using chatgpt. In *Proceedings of the XXXVII Brazilian Symposium on Software Engineering*, pages 293–302, 2023.
- [13] Jieun Han, Haneul Yoo, Yoonsu Kim, Junho Myung, Minsun Kim, Hyunseung Lim, Juho Kim, Tak Yeon Lee, Hwajung Hong, So-Yeon Ahn, et al. Recipe: How to integrate chatgpt into efl writing education. *arXiv preprint arXiv:2305.11583*, 2023.
- [14] Yuyan Chen, Qiang Fu, Yichen Yuan, Zhihao Wen, Ge Fan, Dayiheng Liu, Dongmei Zhang, Zhixu Li, and Yanghua Xiao. Hallucination detection: Robustly discerning reliable answers in large language models. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 245–255, 2023.
- [15] Marian Daun and Jennifer Brings. How chatgpt will change software engineering education. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, pages 110–116, 2023.
- [16] Parsa Rajabi, Parnian Taghipour, Diana Cukierman, and Tenzin Doleck. Exploring chatgpt's impact on post-secondary education: A qualitative study. In *Western Canadian Conference on Computing Education (WCCCE'23), May 04–05, 2023*. Simon Fraser University, 2023.
- [17] Nazmus Sakib, Fahim Islam Anik, and Lei Li. Chatgpt in it education ecosystem: Unraveling long-term impacts on job market, student learning, and ethical practices. In *The 24th Annual Conference on Information Technology Education*, pages 73–78, 2023.
- [18] Virginia Braun and Victoria Clarke. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4):589–597, 2019.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009