WIP: Generative and Custom Chatbots in Computer Programming Education and their Effectiveness A Systematic Literature Review

Syed Hassan Tanvir
Engineering Education Department
University of Florida
Gainesville, Florida, USA
tanvir.shassan@ufl.edu

Gloria J Kim
Engineering Education Department
University of Florida
Gainesville, Florida, USA
gloriakim@ufl.edu

Abstract – This work-in-progress research-to-practice paper describes a systematic literature review on the use of chatbots in programming courses. We scoped existing peer-reviewed and published articles focusing on the strength of custom-built chatbots and generative AI chatbots in computer engineering education – specifically, courses that employ C++, Java, and Python programming languages. The purpose of this work is to obtain in-depth information about chatbots, with a focus on their functionality, application, and effectiveness. We were also interested in the extent to which learning taxonomies and frameworks are integrated with chatbots. This study adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) framework to review n=20 articles. Based on the reviews, we provide a set of guidelines on how learning taxonomies and instructional design frameworks can be mapped while developing or implementing chatbots in programming courses.

Keywords—Chatbot, Computer Engineering, Generative-AI chatbots, Pedagogical agents, Learning taxonomies.

I. Introduction and Background

Integrating AI-based chatbots has emerged as a transformative force in the contemporary education landscape. Chatbots, commonly known as conversational agents, simulate human conversation with smart devices using natural language processing techniques [1].

Chatbots use two types of communication channels. The first involves text-based platforms like WhatsApp, Slack, Snapchat, Facebook Messenger, and Instagram. The second encompasses voice-based virtual assistants like Siri, Alexa, Cortona, and Google Assistant. Most chatbots use both channels to fulfill diverse user needs [2], [3]. This technology is no longer limited to human conversations only. It has moved ahead in responding to answer questions, either in academic or professional work settings.

The use of artificial intelligence in education is rapidly increasing [4]. Chatbots offer personalized assistance to students and educators at any time, which aids in reducing wait times for responses and fostering a dynamic learning environment. Especially in the aftermath of the pandemic, chatbots are invaluable in easing educators' workload by offering personalized support to students. Recent studies claimed that ChatGPT has the capability to clear the United States Medical Licensing Exam (USMLE) without any specialized training [5] and Uniform Bar Exam [6]. Additionally, researchers are investigating ChatGPT potential in other standard tests and state exams, such as Quantitative reasoning in the GRE [7]. Currently, AI chatbots are among the most widely utilized technologies to assist in teaching and learning [8].

Several studies have been conducted on the use of chatbots for answering students' queries related to the subject [9]–[11], providing student course content [12], home-works and assignments deadlines, advice [13], subject-related practice questions [11] and supportive study materials. An intelligent system provides a strategy for involving students to deliver individualized instructions [12], [14]. Students can interact with chatbots individually [15], or they can also support collaborative learning activities, as stated by [16]–[19].

Cunningham et al. [12] described how chatbots could be used in educational settings. On the other hand, Thomas [20] discussed previous literature to report how chatbots offer more effective education. Previous research has tried to provide a holistic view of existing knowledge in the application of chatbot technology in education [8], [12], [21]. However, there is a need for a systematic literature review (SLR) that focuses on chatbots specifically designed for teaching, learning, or mentoring on C++, Java, and Python programming languages.

The overall objective of this systematic literature review is to understand the prevalence of these chatbots in educational settings and their effectiveness in leveraging educational theories to improve learning outcomes in programming courses.

II. Method

This work-in-progress systematic literature review involves an intensive process, as described by Snyder [24], to identify relevant findings on any phenomenon. We have adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) framework as outlined in [22], which includes 27 items and a four-phase flow diagram. This framework helps researchers enhance systematic review and meta-analysis reporting by clarity and transparency, although it is not a quality assessment tool. Perez et al. and El Azhari et al. [23], [24] have also used the PRISMA framework to conduct a systematic literature review on the application of chatbots in education. For this study, we scoped existing peer-reviewed and published articles in Scopus, Ebsco Host, Wiley Online Library, and Web of Science. The limited scholarly articles of interest provided by these databases underscore a significant gap in this area, highlighting the need for more extensive research into educational chatbots.

A. Research Questions

This work is guided by the following three research questions (RQs):

- RQ-1 What strengths do different types of chatbots offer in fundamental programming, specifically

C++, Java, and Python programming language courses, in enhancing student learning outcomes?

- RQ-2 In what ways do various types of chatbots incorporate instructional design models to enhance instructional strategies in software development courses?
- RQ-3 To what extent do the chatbots employed in those courses incorporate learning taxonomies?

RQ-1 explores the various strengths of chatbots in teaching or learning C++, Java, and Python Programming languages. In this context, "strengths" are defined as specific attributes of the chatbots that positively enhance student learning outcomes. These attributes included the chatbot's effectiveness in facilitating understanding of the programming topic, the ability to engage and motivate students, the adaptability of various learning styles, and the quality of feedback provided. RQ-2 probes how identified chatbots in RQ-1 incorporate instructional design models to enhance instructional strategies by providing personalized learning paths, interactive and engaging content delivery, and feedback on programming tasks. The RQ-3 aims to assess how effectively chatbots used in programming courses support educational frameworks that guide students from basic understanding to more complex skills.

B. *Search Syntax*

We employed several keywords encompassing various facets of education and chatbot technology. These keywords included "Education", "Educational Chatbot", "learning", "teaching", "Supportive learning", "Undergraduate", "High School", "Chatbot", "Smart Bots", "Smart Agents", "Intelligent Agents", "Pedagogical agents", "Programming", "Fundamentals of programming", "Object Oriented Programming", "Generative -AI", "ChatGPT". Then, we further refined the search query using Boolean operators Like "AND" and "OR" to ensure the retrieval of relevant literature. We have used a combination of several search strings to find the relevant articles in different databases. An additional two filters were applied to refine the search further.

(TITLE-ABS-KEY (chatbot) OR TITLE-ABS-KEY (pedagogical AND agents) OR TITLE-ABS-KEY (smart AND bots) OR TITLE-ABS-KEY (intelligent AND bot) AND TITLE-ABS-KEY (learning) OR TITLE-ABS-KEY (teaching) OR TITLE-ABS-KEY (assisting) AND TITLE-ABS-KEY (fundamental AND programming) OR TITLE-ABS-KEY (programming AND languages) AND TITLE-ABS-KEY (education)

C. *Inclusion & Exclusion Criteria*

The articles for systematic literature review must have inclusion and exclusion criteria as stated by Snyder [25]. We established this study's inclusion and exclusion criteria to ensure systematic review relevance and rigor.

1) *Inclusion Criteria (IC)*

The articles had to be published in English from 2015 to 2024 and must focus on educational context, used explicitly for learning and teaching C++, Java, and Python programming languages. The search was limited to full-text journal and conference articles only. This systematic literature review focuses on those chatbots used to assist, improve skills, and increase engagement in C++, Python, or fundamental programming courses.

2) *Exclusion Criteria (EC) Filter 1 & Filter 2*

The exclusion criteria for this work-in-progress study were rigorously defined due to the increasing development and use of chatbots in education. We excluded studies that were duplicated in databases. Research articles that focused on non-educational subjects or were not set within an educational context were excluded. Additionally, we removed articles that lacked distinct methods, results, and conclusion sections. Research studies that used C++, Java, and Python programming languages solely for developing chatbots rather than for educational purposes in teaching or learning these programming languages were excluded as well. Finally, we removed other studies that did not address our research questions.

D. *Quality Evaluation (QE)*

The systematic literature review process is rigorous, so it is critical to examine and evaluate the quality of the articles chosen for the final sample [22].In relation to achieving our research objective, the quality of all the selected articles was evaluated using the criteria stated in Table I.

TABLE I. QUALITY EVALUATION (QE) CRITERIA

QE #	Criteria
QE 1	Does the selected study comply with the established inclusion & exclusion criteria?
QE 2	Does the study use chatbots to learn, teach, and assist in fundamental programming courses, specifically C++, Java, or Python programming languages?
QE 3	Does the study include experimental or any other quantitative or qualitative evaluation?
QE 4	Are the study results consistent with the stated aims?
QE 5	Was the article published in an authentic and reputable source?

Due to the small final number of articles, the authors applied the inter-rater reliability (IRR) method for the quality evaluation process, using a scale from 1-3 (1- Not good, 2- good, and 3- Very good). To include the paper in the dataset, a minimum score of 7.5 is needed, which is 50 % of the total. Two researchers performed this process; one extracted the articles, while the other double-checked it. They discussed and resolved any discrepancies. They agreed on the final data set before proceeding with further analysis.

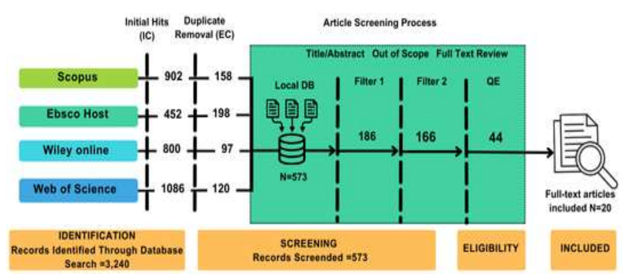

Fig. 1. Prisma Flow Diagram for SLR

E. *Data Extraction*

Based on the inclusion and exclusion criteria designed by the authors and the quality criterion mentioned by [22]. Fig 1 shows the data extraction process, which includes the manual and automatic search processes, the selection process, and the final data set of selected articles. After applying inclusion, exclusion criteria, and quality evaluation criteria, the total number of articles was N=20, Scopus=10, Ebsco Host=4, Web of Science=4, and Wiley Online=2.

III. RESULTS

Our results are tied to each research question as follows:*RQ-1: What strengths do different types of chatbots in C++ and Python programming language courses offer in enhancing student learning outcomes?*

The shortlisted articles for this systematic literature review have underscored the varied and effective use of chatbots in programming education, highlighting their capacity to enhance both learning outcomes and student engagement. Ismail and Ade-Ibijola [13] highlighted that chatbots assist novice programmers by helping them understand fundamental programming concepts rather than simply memorizing code to pass the exam. Similarly, Ibrahim Teo [26], [27] took the initiative to develop chatbots that assist students by providing real-time support and feedback on programming structures, thereby enhancing their learning experience. Chatbots, such as those mentioned Kerly and Bull [28], use conversational agent-based chatbots to interact with learners, allowing them to reflect on their understanding and negotiate the content of their learner model.

In terms of technical enhancement and fostering personalized learning, Hai and Hung [29] designed a unique "CodeBot', which assists Vietnamese students with C++ and Python-related questions with remarkable accuracy F-1 scores of (0.96 and 0.99), ensuring that the Codebot correctly interpreted the intent behind the learner inquiry. On the other hand, Kerly and Bull [32] implemented a rule-based chatbot to inform students about adherence to Python's PEP-8 coding standards, although their study reported no significant difference in learning outcomes between the control and treatment groups ($p=0.958$). Benotti et al. [14] introduced an interactive chatbot from a gamification perspective that increases learner participation and retention. Their study reported that female students were more motivated to learn (Chi-square test, $p=0.01$). Moreover, [23], [30] chatbots focused on syntax error corrections of C++ and Python programming languages by providing immediate feedback. Carreira [31] deployed a Portuguese chatbot, Pyo, which assists the students in understanding the fundamental programming concepts through explaining and providing code examples, though their study findings reported that students preferred real human interaction over chatbots. Nishane and Rajendran [32] introduced "PyGuru 2.0" conversational agent to investigate learner and bot interaction for a wide spectrum of interaction categories such as error correction, code assistance, and conceptual explanation. Cabada et al. [33] Introduced "Java Sensei" smart chatbot to support Java programming language learning. It adopts learners' emotional and cognitive states to provide personalized learning by utilizing fuzzy logic and neural network concepts.

Expanding the utility of the chatbot Auccahuasi et al. [34] has applied NLP and AI-based chatbot technology to teach mathematics concepts through Python programming, fostering the development of logical, analytical, and computational skills.

Most of the selected articles for this systematic literature review have used chatbots for formative assessment. Studies by S. Hobert [19] revealed that chatbots could effectively perform formative and summative assessments, significantly improving learning outcomes. Only two articles [15], [35] described their use for formative and summative assessment.

RQ-2: In what ways do various types of chatbots incorporate instructional design models to enhance instructional strategies in software development courses?

Instructional design models play a critical role in developing educational chatbots, ensuring that these chatbots are not just technically sound but also educationally effective with additional alignment of educational objectives. There are several instructional design models that focus on creating effective and engaging learning experiences, such as ADDIE, Dick and Carey Systems, etcetera. One of the renowned models is the four-component instructional design system (4C/ID model) proposed by Van Merrienboer et al. [36], which provides blueprints for training complex skills. In our search, we found only two articles [27], [37] that used the 4C/ID model to design their chatbot scripts. Both articles reported that the 4C/ID model increases overall engagement.

RQ-3: To what extent do the chatbots employed in those courses incorporate learning taxonomies or frameworks?

Each shortlisted chatbot is primarily designed to enhance learners' learning experience; however, most of the shortlisted articles did not describe the integration of learning taxonomies in their chatbots. PyGuru 2.0 [32] and Code Tutor [15] are the two chatbots that stand out as more explicit integrations of learning taxonomy. In contrast, other chatbots might incorporate pedagogical roles that align with educational goals without specific reference to learning taxonomies. Similarly, a notable contribution by Subramaniam [34] involved developing seven Java-centric chatbots based on rule-based and machine learning to tailor student-bot interaction, mapping the seven learning objectives of Java to seven distinct chatbots, which resulted in significantly improved post-test scores compared to the pre-test. In our shortlisted articles, only two studies describe and categorize different types of cognitive engagement in their chatbot [15], [32] have used the Interactive, Constructive Active, and Passive (ICAP) framework.

IV. DISCUSSION

This systematic literature review analyzed 20 articles focusing on using chatbots to facilitate learning fundamental programming, specifically C++, Java, and Python. Our results revealed a significant gap in the application of chatbots as a comprehensive tool. To support novice learners, it is important for developers and instructors to consider scientific literature when designing, developing, and applying chatbots in educational settings. The chatbot should be built on education conceptual frameworks such as (ICAP), Universal Design for Learning (UDL), and learning taxonomies e.g. Blooms Taxonomy [38] to explore its true potential in education. Additionally, another important and usually overlooked aspect is to train instructors on effectively integrating chatbots in their classrooms. This can be covered by the Technology Pedagogical Content Knowledge (TPACK) Framework, which equips instructors with the knowledge and skills needed to integrate technology into their teaching in a balanced and effective manner. Aligning custom-built chatbots with these frameworks will ensure they are pedagogically sound, achieving the intended learning outcomes.

Based on our findings, we propose a set of guidelines in Table II to help future researchers design a chatbot that aligns with learning taxonomies. Chatbot scripts incorporating the guidelines using Python coding as an example are available upon contacting the authors.

TABLE II. STEPS TO ALIGN CHATBOTS WITH LEARNING TAXONOMIES

Serial #	***Guidelines***	***Description***
1.	Learning Objectives	Articulate learning objectives for each level of Bloom's taxonomy separately, from basic knowledge to creativity.
2.	Alignment with Curriculum/course	Chatbot content and interaction patterns are consistent with the course's learning objective.
3.	Content for each Cognitive level	Create customized scripts that facilitate learning all cognitive levels of Bloom's taxonomy (remembering, understanding, applying, analyzing, evaluating, and creating.
4.	Interactive and Adaptive Dialogue	Natural language processing enables the chatbot to respond dynamically to learners' queries. Scripts can be designed [27], [37]
5.	Integrate Assessment	Integrate tools within the chatbot to assess learners' progress and adjust the difficulty of programming tasks according to their performance by using (adaptive algorithm)
6.	Personalized experience	Align learners' personalized experiences by implementing adaptive algorithms that provide customized interactions and educational pathways tailored to each learner's response and progress.
7.	Chatbot optimization	Regularly update the chatbot based on the learner feedback and performance data to refine instructional strategies.

V. LIMITATIONS AND FUTURE WORK

Many custom-built chatbots are developed using Python language. Therefore, when creating and refining our search query, we came across a large number of papers that did not address our research questions. Future work will involve developing guidelines for aligning instructional strategies with the types of chatbots and their roles. As groundwork, we have identified eight (8) different types of chatbots based on their design, purpose, and the educational needs they address. We have further classified the chatbots into three distinct categories according to their roles: (1) Supporter for teaching programming concepts; (2) Assistant for summarizing or simplifying students' learning tasks; and (3) Mentor for developing students' skills. We are also currently examining the effectiveness of generative AI-based chatbots in facilitating programming education.

REFERENCES

[1] L. Bradeško and D. Mladenić, 'A Survey of Chabot Systems through a Loebner Prize Competition', 2012.

[2] M. Aleedy et al., 'Generating and analyzing chatbot responses using natural language processing', *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 9, 2019[Online]. Availablehttps://dx.doi.org/10.14569/IJACSA.2019.0100910[Accessed: 27November2023].

[3] G. Zhao et al., 'MOLI: Smart Conversation Agent for Mobile Customer Service', *Information*, vol. 10, no. 2, p. 63, Feb. 2019[Online]. Availablehttp://www.mdpi.com/2078-2489/10/2/63[Accessed: 27November2023].

[4] S. Roos and R. Lochan, 'A passing trend or a valuable pedagogical tool?', 2018.

[5] T. H. Kung et al., 'Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models', *PLOS Digital Health*, vol. 2, no. 2, p. e0000198, Feb. 2023[Online]. Availablehttps://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000198[Accessed: 6August2024].

[6] D. M. Katz et al., 'GPT-4 passes the bar exam', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 382, no. 2270, p. 20230254, Feb. 2024[Online]. Availablehttps://royalsocietypublishing.org/doi/full/10.1098/rsta.2023.0254[Accessed: 6August2024].

[7] U. Farooq and S. Anwar, 'ChatGPT Performance on Standardized Testing Exam –A Proposed Strategy for Learners'.

[8] C. W. Okonkwo and A. Ade-Ibijola, 'Chatbots applications in education: A systematic review', *Computers and Education: Artificial Intelligence*, vol. 2, p. 100033, Jan. 2021[Online]. Availablehttps://www.sciencedirect.com/science/article/pii/S2666920X21000278[Accessed: 25November2023].

[9] F. Clarizia et al., 'Chatbot: An Education Support System for Student', in *Cyberspace Safety and Security*, Cham, 2018, pp. 291–302 [Online]. Available: 10.1007/978-3-030-01689-0_23.

[10] B. R. Ranoliya et al., 'Chatbot for university related FAQs', in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2017, pp. 1525–1530[Online]. Availablehttps://ieeexplore.ieee.org/document/8126057[Accessed: 25November2023].

[11] S. Sinha et al., *Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018*, vol. 937. Singapore: Springer Singapore, 2020[Online]. Availablehttp://link.springer.com/10.1007/978-981-13-7403-6[Accessed: 25November2023].

[12] S. Cunningham-Nelson et al., 'A Review of Chatbots in Education: Practical Steps Forward', 2019.

[13] M. Ismail and A. Ade-Ibijola, 'Lecturer's Apprentice: A Chatbot for Assisting Novice Programmers', in *2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, Vanderbijlpark, South Africa, 2019, pp. 1–8[Online]. Availablehttps://ieeexplore.ieee.org/document/9015857/[Accessed: 25November2023].

[14] L. Benotti et al., 'Engaging high school students using chatbots', in *Proceedings of the 2014 conference on Innovation & technology in computer science education - ITiCSE '14*, Uppsala, Sweden, 2014, pp. 63–68[Online]. Availablehttp://dl.acm.org/citation.cfm?doid=2591708.2591728[Accessed: 10May2024].

[15] S. Hobert, 'Say Hello to "Coding Tutor"! Design and Evaluation of a Chatbot-based Learning System Supporting Students to Learn to Program', 2019.

[16] S. Chaudhuri et al., 'Engaging Collaborative Learners with Helping Agents', 2009.

[17] R. Kumar and C. P. Rosé, ‘Architecture for Building Conversational Agents that Support Collaborative Learning’, *IEEE Trans. Learning Technol.*, vol. 4, no. 1, pp. 21–34, Jan. 2011[Online]. Availablehttp://ieeexplore.ieee.org/document/5669250/[Accessed: 25November2023].
[18] G. Stahl, *Group cognition: computer support for collaborative knowledge*. Cambridge, Mass.: MIT Press, 2006.
[19] E. Walker et al., ‘Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity’, *Computer Supported Learning*, vol. 6, no. 2, pp. 279–306, Jun. 2011[Online]. Availablehttps://doi.org/10.1007/s11412-011-9111-2[Accessed: 25November2023].
[20] H. Thomas, ‘Critical Literature Review on Chatbots in Education’, vol. 4, no. 6, 2020.
[21] J. Q. Pérez et al., ‘Rediscovering the use of chatbots in education: A systematic literature review’, *Computer Applications in Engineering Education*, vol. 28, no. 6, pp. 1549–1565, 2020[Online]. Availablehttps://onlinelibrary.wiley.com/doi/abs/10.1002/cae.22326[Accessed: 8May2024].
[22] D. Moher et al., ‘Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement’, *Ann Intern Med*, vol. 151, no. 4, pp. 264–269, Aug. 2009[Online]. Availablehttps://www.acpjournals.org/doi/full/10.7326/0003-4819-151-4-200908180-00135[Accessed: 1May2024].
[23] J. E. Perez et al., ‘JEPPY: An Interactive Pedagogical Agent to Aid Novice Programmers in Correcting Syntax Errors’, *IJACSA*, vol. 11, no. 2, 2020[Online]. Availablehttp://thesai.org/Publications/ViewPaper?Volume=11&Issue=2&Code=IJACSA&SerialNo=7[Accessed: 10May2024].
[24] K. EL Azhari et al., ‘SMART Chatbots in the E-learning Domain: A Systematic Literature Review’, *International Journal of Interactive Mobile Technologies*, vol. 17, no. 15, pp. 4–37, Aug. 2023[Online]. Availablehttp://lp.hscl.ufl.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=aci&AN=169896228&site=eds-live[Accessed: 11May2024].
[25] H. Snyder, ‘Literature review as a research methodology: An overview and guidelines’, *Journal of Business Research*, vol. 104, pp. 333–339, Nov. 2019[Online]. Availablehttps://www.sciencedirect.com/science/article/pii/S0148296319304564[Accessed: 6May2024].
[26] N. H. Ibrahim Teo, ‘e JAVA Chatbot for Learning Programming’, vol. 8, pp. 3290–3298, Sep. 2020.
[27] Y.-H. Lin and T. Tsai, ‘A Conversational Assistant on Mobile Devices for Primitive Learners of Computer Programming’, in *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, Yogyakarta, Indonesia, 2019, pp. 1–4[Online]. Availablehttps://ieeexplore.ieee.org/document/9226015/[Accessed: 10May2024].
[28] A. Kerly and S. Bull, ‘The Potential for Chatbots in Negotiated Learner Modelling: A Wizard-of-Oz Study’, in *Intelligent Tutoring Systems*, vol. 4053, M. Ikeda, K. D. Ashley, and T.-W. Chan, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 443–452[Online]. Availablehttp://link.springer.com/10.1007/11774303_44[Accessed: 10May2024].
[29] N. X. Hai and N. V. Hung, ‘CODEBOT – A VIETNAMESE CHATBOT SYSTEM FOR ANSWERING C++ AND PYTHON-RELATED QUESTIONS’, *Tạp chí Khoa học*, vol. 15, no. 3, p. 48, Sep. 2019[Online]. Availablehttp://journal.hcmue.edu.vn/index.php/hcmuejos/article/view/145[Accessed: 10May2024].
[30] S. I. Malik et al., ‘Fostering the Learning Process in a Programming Course With a Chatbot’:, *International Journal of Online Pedagogy and Course Design*, vol. 12, no. 1, pp. 1–17, Aug. 2022[Online]. Availablehttps://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/IJOPCD.306686[Accessed: 8May2024].
[31] G. M. Carreira, ‘Development of a chatbot to assist introductory programming students’, in *Development of a chatbot to assist introductory programming students*, 2022[Online]. Availablehttps://estudogeral.uc.pt/handle/10316/102157[Accessed: 10May2024].
[32] D. Singh et al., ‘Catalyzing Python Learning: Assessing an LLM-based Conversational Agent’, 2023.
[33] R. Z. Cabada et al., ‘An Affective Learning Environment for Java’, in *2015 IEEE 15th International Conference on Advanced Learning Technologies*, Hualien, Taiwan, 2015, pp. 350–354[Online]. Availablehttp://ieeexplore.ieee.org/document/7265347/[Accessed: 10May2024].
[34] W. Auccahuasi et al., ‘Interactive online tool as an instrument for learning mathematics through programming techniques, aimed at high school students’, in *Proceedings of the 6th International Conference on Information Technology: IoT and Smart City*, Hong Kong Hong Kong, 2018, pp. 70–76[Online]. Availablehttps://dl.acm.org/doi/10.1145/3301551.3301580[Accessed: 10May2024].
[35] S. Abbasi and H. Kazi, ‘Measuring Effectiveness of Learning Chatbot Systems on Student’s Learning Outcome and Memory Retention’, *Asian Journal of Applied Science and Engineering*, vol. 3, no. 7, p. 57, Aug. 2014[Online]. Availablehttp://ajase.weebly.com/uploads/1/3/4/5/13455174/ajase_7.7.pdf[Accessed: 10May2024].
[36] J. J. G. van Merriënboer et al., ‘Blueprints for complex learning: The 4C/ID-model’, *ETR&D*, vol. 50, no. 2, pp. 39–61, Jun. 2002[Online]. Availablehttps://doi.org/10.1007/BF02504993[Accessed: 11May2024].
[37] Y.-H. Lin, ‘Chatbot Script Design for Programming Language Learning’, in *2022 IEEE 5th Eurasian Conference on Educational Innovation (ECEI)*, Taipei, Taiwan, 2022, pp. 123–125[Online]. Availablehttps://ieeexplore.ieee.org/document/9829460/[Accessed: 9May2024].
[38] B. S. Blooms, ‘Handbook on Formative and Summative Evaluation of Student Learning’, McGraw-Hill Book Company, 330 West 42nd Street, New York, New York 10036 ($11, Feb. 1971.