The Analysis of Student Errors in ARIN-561 – An Educational Game for Learning Artificial Intelligence for High School Students

Ning Wang[1], Eric Greenwald[2], Ryan Montgomery[2], and Maxyn Leitner[1]

[1] University of Southern California, Los Angeles, CA USA
[2] University of California, Berkeley, CA USA

Abstract. Around the world, the prevalence of artificial intelligence (AI) in global citizens' work and life has been recognized by government and intergovernmental agencies. Efforts by researchers, practitioners, and policy-makers are underway to develop guidelines, curriculum and best practices to help the future workforce who are youth in today's classrooms develop basic AI competencies. While there is growing attention to broadening AI educational opportunities and, especially, to providing learning experiences for younger students, relatively little is currently known about how to most effectively provide AI education to K-12 (kindergarten through 12th grade) students. In this paper, we discuss the design and evaluation of an educational game for high-school AI education called ARIN-561. Drawing on feedback from over 1,200 students, we conducted analysis on student performances in the game, particularly student errors. Results show the relationship and dependencies of different activities within the game and shed light on the design instructional support to help students build AI knowledge to succeed in the game.

Keywords: educational game · K-12 AI education.

The critical role of artificial Intelligence (AI) in the future of work and life has been recognized around the globe. While some of today's youth will become the future AI workforce and a majority of them will join a workforce that utilizes AI, all will become end-users, such as consumers of AI [7]. It is critical, therefore, to prepare future generations with basic knowledge of AI, not just through higher education, but beginning with childhood learning. The United Nations and governments around the world are racing to develop strategies in response to the growing need of youth AI education (e.g., [20]), including establishing guidelines [1, 7] and curriculum [4] for K-12 (Kindergarten through 12th grade) AI education. While AI's impact on society is deepening and expanding in myriad ways, there has been little research into how students, especially pre-college students, construct an understanding of and gain practice with core ideas in the field. As a result, there is yet little possibility of grounding the design of learning experiences in evidence-based accounts of how youth learn AI concepts, how understanding progresses across concepts, or what concepts are most appropriate for what age-levels.

In this paper, we discuss our research into a game-based approach towards AI education for high school students. Decades of research evidence point to the efficacy of game-based learning in promoting student learning [21]. Most recently, there is research into utilizing educational games to help elementary school students learn about AI [12]. In our work, we developed an educational game called ARIN-561. ARIN-561 is a 3D role-playing game designed to teach high-school students AI concepts, prompt them to apply their math knowledge, and develop their AI problem-solving skills. In the game, students play as a space-faring scientist who has crash landed on an alien planet, named ARIN-561 (Figure 1). In order to safely return home, the scientist begins exploring the planet to gather resources needed to repair the broken ship while uncovering the mystery of the planet. The activities for survival and for exploration form the basis for students to learn and apply AI algorithms to solve problems. Previously, we have conducted a small-scale pilot study to evaluate the efficacy of ARIN-561 in learning AI for high school student population [13]. While results from the pilot study suggested the efficacy of ARIN-561 for learning AI, the sample size was relatively small. Subsequently, we conducted a larger scale evaluation study with over 1,200 students. At the conclusion of the study, students provided feedback on the improvements they would like to see in the future releases of the ARIN-561. Among them are suggestions to provide additional instructional support when students make errors and get stuck within the game. In this paper, we discuss the design and evaluation of ARIN561. In particular, we zoom in the analysis on student errors made in the game using data gathered from the large-scale evaluation study. Results show the relationship and dependencies of different activities within the game and shed light on the design instructional support to help students build knowledge to succeed in the game.

1 Related Work

AI education has long been absent from K-12 classrooms. Recent efforts are beginning to investigate the integration of AI into K-12 schools, including defining AI literacy [16] and developing curricula and guidelines [7, 17]. Researchers in youth AI education have been experimenting with teaching AI, including machine learning [23, 30] and ethics [5], within the context of computational thinking [22] through conversational agents [14], dance [18], and game-based learning [12]. Discussions on youth AI education are heating up in Europe [10, 2], China [19], Israel [26], and around the world [28, 29]. For example, researchers in Thailand have designed an agricultural-based AI challenge to foster middle-school students' learning of the machine learning process in the form of a game [25], where students build machine learning models to classify ripe or unripe mangoes. In Australia, researchers have designed and implemented classroom activities for teaching fundamental concepts of AI to Year 6 students to demystify AI through activities such as an unplugged activity on facial recognition and a simple robotic exercise that introduces the concept of machine learning [9].

The work presented here aims to uncover how to design an educational game to meet the challenges of teaching AI to K-12 students. This work builds upon explorations into how K-12 students approach AI concepts, what obstacles they face, and how to guide them through obstacles [8]. This work also draws upon previous investigations into linking AI to the K-12 math curriculum to identify AI concepts suitable for high school students [27], as well as work investigating the learning of computational thinking [11] and seminal research into comprehension of mathematical representations (e.g., [3, 6]).

2 ARIN-561 Game-Based Learning Environment

ARIN-561 is a 3D role-playing game designed to teach high-school students AI concepts, prompt them to apply their math knowledge, and develop their AI problem-solving skills. The game currently covers three classical search algorithms: breadth-first search (BFS), depth-first search (DFS), and greedy search. Each topic consists of two modules: a tutorial module (e.g., Figure 1 bottom left) and a transfer module (e.g., Figure 1 bottom right). Embedded in all the tutorial and transfer modules are quizzes that help students pause and self-assess (Figure 1 top right). In-game challenges, such as searching for missing spaceship parts or cracking passwords, serve as natural opportunities for the introduction of search algorithms as a topic. The essential concepts such as space and time complexity also lend opportunities to connect math knowledge familiar to high school students and these AI concepts that are usually taught in higher education. The integrated educational content in ARIN-561 leverages this opportunity by supporting the students' application of math knowledge to the evaluation of each algorithm as they progress through the game. In addition to the learning modules, students can also explore the game environment for "off-task" activities [24], such as gathering minerals around the planet.

Activities in the game aim to achieve three learning goals: understanding how AI algorithms are used to solve problems in the real world, learning the strengths and weaknesses of different AI algorithms, and developing a working understanding of how AI algorithms work. In ARIN-561, game modules are organized by learning topics, such as BFS and DFS. After scaffolding students through the first AI algorithm (such as BFS), each new AI algorithm (e.g., DFS) is introduced through an example problem that the previous algorithms fail to solve (e.g., computer runs out of memory when using BFS for route planning). The students are guided through the analysis to uncover why the previous algorithm failed (e.g., storing too many nodes in computer memory) and how to modify it to address its weakness (e.g., prioritizing expanding child nodes instead of sibling nodes in the search tree).

3 Methods

Recruitment Twenty-three math, science, and computer science teachers from a school district in a major metropolitan area in the United States participated

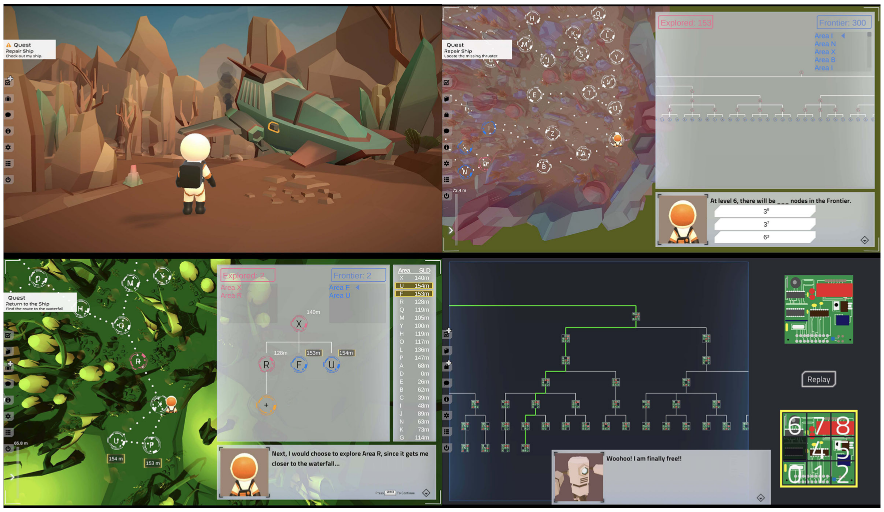

Fig. 1: Screen capture from ARIN-561. Top-left: The player (i.e. the student) crash landed on a foreign planet. Top-right: student is presented with a quiz question about estimating the complexity of search algorithms. Bottom left: student think-alouds through the greedy search algorithm. Bottom-right: the student solves an 8-puzzle using one of the search algorithm to fix their companion robot's circuit board.

in the study. 1274 high school-aged youth from classes taught by participating teachers were recruited for the study.

Procedure Participating teachers were provided an overview of the game, learning goals, and study procedure a few months before the study began. A few weeks prior to the study, students were given an online parental consent form and a youth assent form. Only students who consented participated in the study. The study was carried out over 4 class sessions, each lasting 45-55 minutes long, with at least 2 class sessions dedicated to individual gameplay for students. During the first session, students were first assigned IDs to protect their identity throughout the study, and then completed the pre-survey online. At the end of the first session, students logged into the ARIN-561 game online via a web browser. Any technical difficulties encountered were addressed during the first session, via support from the research team. During the second and third sessions, students continued to interact with ARIN-561 at their own pace. Game progression, play time, and answers to in-game questions were recorded for each participant. During the fourth session, students completed the post-survey online.

With restricted access to school campuses due to COVID-19, the study was carried out entirely by the participating teachers. The research team did not participate in the data collection. The study was carried out within a month span.

Additionally, because students were not required to answer all the questions on the pre- and post-surveys, there are is data at the item level for some students.

Measures The pre-survey consisted of items about students' demographic background, AI Use Type, Interest in AI, AI Knowledge (15 questions), Math Self-efficacy [15], and Math Knowledge. All scales except the Math Self-efficacy were developed by the research team. The AI Use Type included items such as "When I think about how I'd like to interact with AI in the future, I expect that: I will use AI systems in my everyday life as a consumer, and I expect to USE AI systems as a part of my job." The Interest in AI scale included questions such as "Outside of school I try to learn a lot about AI." The assessment of AI knowledge and math knowledge specifically focused on the content covered in ARIN-561, in the format of multiple-choice questions. The AI questions were set in the context of solving AI problems similar to those encountered in the game. The questions assessed students' understanding of, for example, pros and cons of the search algorithms, search algorithms most applicable to specific types of problems, etc. In the post-survey, the same items on interest in AI and AI knowledge from the pre-survey were included. In addition to the surveys, game logs from ARIN-561 were collected. The logs included the in-game click-stream data and responses to in-game quizzes.

In-game errors occur during three different activities within the modules for learning BFS, DFS, and Greedy algorithms.

- **Errors in In-Game Quizzes**: The quizzes are embedded within the tutorial modules for BFS, DFS, and Greedy algorithms. If the students answer the quiz question incorrectly, they are immediately offered the opportunity to try again. The tutorial only moves on until the students answer the question correctly. Thus errors in in-game quizzes are the incorrect attempts to answer the question.
- **Errors in Search Tree Expansion**: the last activity in the tutorial module (usually after the presentations of all the quiz questions) is asking the students to expand the search tree by, for example, clicking on a location on the map (e.g., Top-right and lower-left screens in Figure 1). This occurs after the students have completed the portion of the tutorial that walks them through how to expand the search tree step by step. In this activity, students are asked to demonstrate their understanding of how different search algorithms expand the search space. Errors in this activity are mouse clicks when students click on the wrong locations on the map, thus attempting to expand the wrong nodes in the search tree.
- **Errors in Transfer Puzzles**: After completing a tutorial module for learning, for example BFS, students are presented with a transfer problem to demonstrate their problem-solving skills using the search algorithm. The transfer problems are various puzzles (e.g., an 8-sliding puzzle, or cracking a password), depends on the search algorithm under discussion. To solve the puzzles, students expand a search tree by clicking on the node to be expanded next (e.g., which piece to move in an 8-sliding puzzle), until a solution

is found (e.g., the series of moves to solve the 8-sliding puzzle; lower-right screen in Figure 1). Errors in this activity are similar to those in search tree expansion in the tutorial module – mouse clicks when students click on the wrong node to expand next.

The activities for learning each of the three algorithms (BFS, DFS, and Greedy) include these three activities discussed here.

4 Results

Of the 1274 participating students, 1014 completed the post-survey. The research team was able to match pre-, post- surveys, and game logs for 764 students. Other than normal attrition (e.g., students absent at either pre, post administration, or game play class), additional data loss was primarily due to errors in student ID entries on the survey platform, which resulted in mismatches of student IDs between both surveys and game logs. We conducted ANOVA analyses to ensure the final sample of 764 students was not significantly different from the full participant sample in terms of background, such as gender, race/ethnicity, and prior mathematical knowledge.

The participants' average age was 16, with 18% 12th graders, 30% 11th graders, 23% 10th graders, and 29% 9th graders. A total of 46% of the students identified as male, 48% identified as female and 6% identified as other categories or preferred not to disclose. 27% of the students speak English at home, 67% speak both English and a second language at home, and 6% speak only a language other than English at home. Spanish is reported as the non-English language for those students. Interestingly, even though ARIN-561 and the surveys are offered in both English and Spanish, and the teachers were briefed about the language choice prior to the study, all the students chose to use the English version of the surveys and the game.

4.1 Descriptive statistics

Overall, the number of student errors in the game varied greatly between students and between game modules. From the descriptive analysis presented in Figure 2, we can see that there are several extreme outliers where, for example, a student made over 1700 errors in the BFS tree expansion activities. Additional examination of the game logs revealed that students often revisit a specific module in the game (e.g., taking the tutorial of BFS multiple times). This means that students may answer the same quiz questions correctly multiple times or commit the same error in search tree expansion multiple times. Thus, for the analysis presented here, we define in-game errors as *error rates*, which is the number of errors divided by the sum of errors and correct actions (e.g., mouse clicks correctly expanding the search tree).

For student errors occurred within modules to learn each of the search algorithm (i.e., BFS, DFS, greedy algorithm), we analyzed how often errors were

	Quiz (BFS)	Quiz (DFS)	Quiz Greedy)	Tree Expansion (BFS)	Tree Expansion (DFS)	Tree Expansion (Greedy)	Transfer Puzzle (BFS)	Transfer Puzzle (DFS)	Transfer Puzzle (Greedy)
N	762	762	762	762	762	762	762	762	762
Mean	2	6	3	104	51	50	14	61	6
Median	2	6	2	44	29	17	8	39	6
Std. Deviation	3	7	3	172	69	82	23	77	6
Minimum	0	0	0	0	0	0	0	0	0
Maximum	37	123	30	1782	597	941	328	700	73

Fig. 2: Descriptive statistics of the number of errors made in different activities in game modules.

made during the in-game quizzes, the search tree expansion, and in transfer puzzles. ANOVA with repeated measures show that there are significant differences between student error rates in quizzes, tree expansions and transfer puzzles within the BFS learning modules ($p < .001$, paired contrasts are statistically significant at $p < .001$ as well). Figure 3 shows the means of student errors within each module. Within the game modules for learning DFS and greedy algorithms, we found the same sets of statistically significant results. We also conducted another set of ANOVA with repeated measures comparing the student errors within the same type of activity (e.g., in-game quizzes) across modules learning different algorithms. Again, we found the same set of statistically significant results ($p < .001$, for paired contrasts as well). For example, from Figure 3, we can see that students made errors more often when learning DFS, compared to BFS and greedy.

4.2 Student background and In-Game Errors

In the pre-survey, we gathered data on students' demographic background, such as gender, grade level, language spoken at home, and video game experiences. Since a vast majority (94%) of the participating students speak English as one of the languages at home, we did not perform analysis on the relationship of this variable and dependent variables. We then performed analysis on how the rest of the demographic background variables impact overall student error rates in BFS, DFS, and Greedy modules, and in in-game quizzes, search tree expansions, and transfer puzzles. Overall, in-game performances in terms of student errors made did not differ significantly between students from different grade levels, of different gender, or with different video gaming experiences.

The pre-survey also includes items that measure Math Self-Efficacy, (relevant) Math Knowledge, and Interest in AI. We then conducted correlational analysis on how these factors impacted student in-game errors. Measured in both the pre- and post-survey, the AI knowledge scale includes 3 sub-scales for each of the search algorithms covered in the game (BFS, DFS, and Greedy search). We performed pairwise correlation tests of these variables and overall errors made in BFS, DFS, and greedy modules, and in in-game quizzes, search tree expansions, and transfer puzzles. Overall, existing AI and math knowledge are significantly

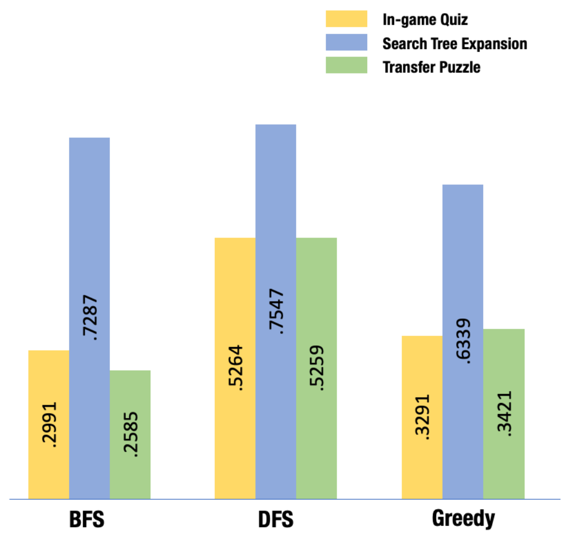

Fig. 3: Error rates (between 0 and 1) in different activities in ARIN-561 modules for learning BFS, DFS and greedy search algorithm.

and negatively correlated with the dependent variables (Table 1). Interestingly, there was no statistically significant correlation between knowledge of the three AI knowledge sub-scales and student errors made in the corresponding game modules (e.g., pre-AI knowledge in BFS is not significantly correlated with errors in BFS modules with in-game quiz, search tree expansion, and transfer puzzle combined).

In-Game Progress and Errors The pedagogical design of ARIN-561 is based on the hypothesis that AI algorithms build on each other. Algorithms, such as DFS introduced later in the game are discussed in comparison to previously introduced algorithms, such as BFS. While students can jump between different modules by going through the menu selection screen in the game (only enabled after the module is completed), overall, students took a relatively linear path through the game, by going through BFS, DFS, then Greedy game modules. Thus, as students progress through the game, mastering previously discussed algorithms should help students' learning of the new ones, while learning the new algorithms should help reinforce the learning of the previously discussed ones. While we can analyze how reaching milestones in the game, such as completing the BFS module (both tutorial and transfer problem modules), impacts errors made in subsequent modules (e.g., DFS, Greedy), in our sample, we have very few students did *not* complete a specific game module. Thus, we don't have a sufficiently large contrasting sample size to perform the analysis.

	PreAI_knowledge	PreMath_knowledge
In-game Quizzes	-.116 (P=.0002)	-.132 (p<.001)
Search Tree Expansions	-.199 (p<.001)	-.094 (p=.012)
Transfer Puzzles	-.162 (p<.001)	-.084 (p=.028)
BFS	-.093 (p=.011)	-0.043 (p=.24)
DFS	-.207 (p<.001)	-.122 (p=.002)
Greedy	-.199 (p<.001)	-.156 (p<.001)

Table 1: Correlations between existing AI and math knowledge, gathered via pre-survey, and the number of errors made in BFS, DFS, and Greedy modules, and in in-game quizzes, search tree expansions, and transfer puzzles.

Alternatively, we performed correlations of student errors made in different game modules. In particular, we are interested in whether student errors in the tutorial modules (e.g., BFS in-game quiz and search tree expansion) is related to errors in transfer modules (e.g., BFS transfer puzzle). Results show that the number of errors made in the BFS puzzle is significantly but weakly correlated with those in BFS in-game quiz ($r = .1; p = .009$), and not with those in BFS tree expansion activities ($r = .027, p = .519$). The number of student errors made in the DFS puzzle is significantly correlated with those in DFS in-game quiz ($r = .183, p < .001$) and in DFS tree expansion ($r = .314; p < .001$). The errors made in the greedy transfer puzzle are significantly but weakly correlated with those in greedy in-game quiz ($r = .125, p = .002$) and in greedy tree expansion ($r = .161; p < .001$). This indicates that in general the performances in the transfer modules are positively correlated with performance in the tutorial modules.

4.3 In-game Error and AI Learning

We conducted pairwise correlations on outcome variables, such as AI knowledge gained and changes in AI interest (pre-post), and the variables on student errors. Overall, we did not find any statistically significant correlation between changes in AI interest and student errors made in any module in the game. The AI knowledge gained is negatively but weakly correlated with total errors made in DFS ($r = -.137, p < .001$) and greedy algorithm modules ($r = -.123, p = .002$), but not in BFS modules ($r = .006, p = .879$). The AI knowledge scale includes 3 subscales for each of the search algorithms covered in the game (BFS, DFS, and Greedy search). We then conducted pairwise correlations of AI knowledge gained in a specific game module (e.g., BFS) and student errors made in that module (e.g., quiz in BFS tutorial, BFS tree expansion, and BFS puzzle). Again, we did not find any significant correlation between any AI knowledge subscale and errors made in corresponding game modules.

The pre- and post-survey included identical measures of AI knowledge. We performed a median-split on AI knowledge gained (pre-post) and group students into High/Low AI knowledge gained. Independent Sample t-test showed that the group of students in high AI knowledge gain group made fewer mistakes in all modules combined in DFS ($M_{high} = .603, M_{low} = .644, p < .001$) and greedy ($M_{high} = .538, M_{low} = .590, p = .002$), but not in BFS ($M_{high} = .513, M_{low} = .503, p = .568$). Additional t-test showed significant differences between high/low AI knowledge gain group on total mistakes made in in-game quizzes ($M_{high} = .404, M_{low} = .428, p = .006$) and search tree expansions ($M_{high} = .732, M_{low} = .765, p = .009$), but not in transfer puzzles ($M_{high} = .420, M_{low} = .439, p = .132$, BFS, DFS, greedy combined for all three measures). This indicates that students who had higher AI knowledge gain made fewer mistakes in DFS and Greedy, and in tutorial modules (quizzes and search tree expansions). Performances in these modules could be indicative of which group students may fall into on AI knowledge gains.

5 Discussion

In this paper, we discussed a large-scale evaluation study to assess the efficacy of ARIN-561. Inspired by student feedback, we conducted an analysis of errors students committed in the game. Results show that in-game performances varied a great deal between students and between game modules. Overall, students struggled when learning DFS and in the activities of search tree expansion, which occurs right after the game "removes" instructional support and asks the students to continue expanding the search tree to demonstrate mastery of the knowledge. This indicates the potential need for delaying the removal of the instructional support in search tree expansions, and the additional support to help students learn DFS.

The significant (negative) relationships between prior AI and mathematical knowledge and the observed in-game student errors suggest an educational game that is optimized for youth who already enter with a strong mathematical foundation. This would challenge efforts at using the current iteration of the game for a broad high-school population with a wide range of prior math competencies. This also suggests that in-game mini lessons to strengthen students' math knowledge related to AI learning can potentially improve student performances in the game.

The result also shed light on the relationship between activities in tutorial modules and those in transfer modules. In particular, student performances in tutorial modules are weakly related to performances in transfer tasks. While the tutorial-transfer learning module placements are by design, the results highlighted students' need to better master the learning material while still in the tutorials. Additional and more fine-grained analysis is needed to uncover where inside the tutorial additional support is needed.

One of the limitations of the study is that it was dependent on a researcher-developed measure of AI knowledge, with limited evidence available of its validity with the population sampled. This speaks to the current dearth of AI knowledge measures developed for pre-college-aged youth, a challenge that our research team, and others, are working to address through ongoing research and measurement development.

6 Acknowledgement

This research was supported by the National Science Foundation (NSF) under Grant #1842385. Any opinions and finding expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

1. AAAI: Aaai launches "ai for k-12" initiative in collaboration with the computer science teachers association (csta) and ai4all (2018), https://aaai.org/Pressroom/Releases/release-18-0515.php
2. AI+: Developing an artificial intelligence curriculum adapted to European high schools (2021), https://aiplus.udc.es/
3. Curcio, F.R.: Comprehension of mathematical relationships expressed in graphs. Journal for Research in Mathematics Education **18**(5), 382–393 (1987)
4. (Ed.), U.: K-12 ai curricula: A mapping of government-endorsed ai curricula (2022)
5. Forsyth, S., Dalton, B., Foster, E.H., Walsh, B., Smilack, J., Yeh, T.: Imagine a more ethical AI: Using stories to develop teens' awareness and understanding of AI and its societal impacts. In: Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology. pp. 1–2 (2021)
6. Friel, S.N., Curcio, F.R., Bright, G.W.: Making sense of graphs: Critical factors influencing comprehension and instructional implications. Journal for Research in Mathematics Education **32**(2), 124–158 (2001)
7. Gardner-McCune, C., Touretzky, D., Martin, F., Seehorn, D.: AI for K–12: Making room for AI in K–12 CS curricula. In: ACM Technical Symposium on Computer Science Education. pp. 1244–1244 (2019)
8. Greenwald, E., Leitner, M., Wang, N.: Learning artificial intelligence: Insights into how youth encounter and build understanding of ai concepts. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 35, pp. 15526–15533 (2021)
9. Ho, J.W., Scadding, M., Kong, S., Andone, D., Biswas, G., Hoppe, H., Hsu, T.: Classroom activities for teaching artificial intelligence to primary school students. In: Conference on Computational Thinking Education. pp. 157–159 (2019)
10. Kandlhofer, M., Steinbauer, G., Laßnig, J.P., Baumann, W., Plomer, S., Ballagi, A., Alfoldi, I.: Enabling the creation of intelligent things: Bringing artificial intelligence and robotics to schools. In: IEEE Frontiers in Education Conference. pp. 1–5 (2019)
11. Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., Werner, L.: Computational thinking for youth in practice. ACM Inroads **2**(1), 32–37 (2011)
12. Lee, S., Mott, B., Ottenbreit-Leftwich, A., Scribner, J., Taylor, S., Park, K., Rowe, J., Glazewski, K., Hmelo-Silver, C.E., Lester, J.: AI-infused collaborative inquiry in upper elementary school: A game-based learning approach. In: Symposium on Education Advances in Artificial Intelligence. vol. 35, pp. 15591–15599 (2021)
13. Leitner, M., Greenwald, E., Montgomery, R., Wang, N.: Design and evaluation of arin-561: An educational game for youth artificial intelligence education. In: Proceedings of the 30th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education (2022)

14. Lin, P., Van Brummelen, J., Lukin, G., Williams, R., Breazeal, C.: Zhorai: Designing a conversational agent for children to explore machine learning concepts. In: AAAI Conference on Artificial Intelligence. vol. 34, pp. 13381–13388 (2020)
15. Liu, X., Koirala, H.: The effect of mathematics self-efficacy on mathematics achievement of high school students. In: Proceedings of NERA (2009)
16. Long, D., Magerko, B.: What is AI literacy? Competencies and design considerations. In: CHI conference on human factors in computing systems. pp. 1–16 (2020)
17. MIT AI Education Initiative: What's the DAILy curriculum? (2021), https://raise.mit.edu/daily/index.html
18. Payne, W.C., Bergner, Y., West, M.E., Charp, C., Shapiro, R.B.B., Szafir, D.A., Taylor, E.V., DesPortes, K.: danceON: Culturally responsive creative computing. In: CHI Conference on Human Factors in Computing Systems. pp. 1–16 (2021)
19. Peterson, D., Goode, K., Gehlhaus, D.: AI education in China and the United States: A comparative assessment. Tech. rep., Center for Security and Emerging Technology (2021)
20. PLAN, S.: The national artificial intelligence research and development strategic plan (2016)
21. Plass, J.L., Mayer, R.E., Homer, B.D.: Handbook of game-based learning. MIT Press (2020)
22. Ritter, S., Aglio, J., Stetzer, R., Wilson, G., Navta, N., Lippl, C.: Teaching AI as computational thinking for middle school students. In: Workshop on K-12 AI Education at the 20th International Conference on AI in Education (2019)
23. Rodríguez-García, J.D., Moreno-León, J., Román-González, M., Robles, G.: Evaluation of an online intervention to teach AI with LearningML to 10–16 year old students. In: ACM Symposium on Computer Science Education. pp. 177–183 (2021)
24. Sabourin, J., Rowe, J.P., Mott, B.W., Lester, J.C.: When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In: Proceedings of AIED. pp. 534–536. Springer (2011)
25. Sakulkueakulsuk, B., Witoon, S., Ngarmkajornwiwat, P., Pataranutaporn, P., Surareungchai, W., Pataranutaporn, P., Subsoontorn, P.: Kids making AI: Integrating machine learning, gamification, and social context in STEM education. In: IEEE international conference on teaching, assessment, and learning for engineering (TALE). pp. 1005–1010 (2018)
26. Shamir, G., Levin, I.: Transformations of computational thinking practices in elementary school on the base of artificial intelligence technologies. In: Proceedings of EDULEARN20. vol. 6, pp. 1596–1605 (2020)
27. Wang, N., Johnson, M.: AI education for K–12: Connecting AI concepts to high school math curriculum. In: Proceedings of the IJCAI Workshop on Education in Artificial Intelligence K-12 (2019)
28. Youjun, X., Jiqing, W., Jinson, H.: Textbook series on Artificial Intelligence for Elementary and Middle Schools. East China Normal University Press (2018)
29. Yukun, C., Tang, X.: Foundamentals of Artificial Intelligence for High Schools. East China Normal University Press (2018)
30. Zhou, X., Tang, J., Daley, M., Ahmad, S., Bai, Z.: "Now, I want to teach it for real!": Introducing machine learning as a scientific discovery tool for K–12 teachers. In: Proceedings of AIED. pp. 486–499. Springer (2021)