An Effectiveness Study of Teacher-led AI Literacy Curriculum in K-12 Classrooms

Anonymous submission

Abstract

Artificial intelligence (AI) has rapidly pervaded and reshaped almost all walks of life, but efforts to promote AI literacy in K-12 schools remain limited. There is a knowledge gap in how to prepare teachers to teach AI literacy in inclusive classrooms and how teacher-led classroom implementations can impact students. This paper reports a comparison study to investigate the effectiveness of an AI literacy curriculum when taught by classroom teachers. The experimental group included 89 middle school students who learned an AI literacy curriculum during regular school hours. The comparison group consisted of 69 students who did not learn the curriculum. Both groups completed the same pre and post-test. The results show that students in the experimental group developed a deeper understanding of AI concepts and more positive attitudes toward AI and its impact on future careers after the curriculum than those in the comparison group. This shows that the teacher-led classroom implementation successfully equipped students with a conceptual understanding of AI. Students achieved significant gains in recognizing how AI is relevant to their lives and felt empowered to thrive in the age of AI. Overall this study confirms the potential of preparing K-12 classroom teachers to offer AI education in classrooms in order to reach learners of diverse backgrounds and broaden participation in AI literacy education among young learners.

Introduction

In the past decade, artificial intelligence (AI) is reshaping the industrial landscape and bringing tremendous economic benefits: 50% of companies that embrace AI over the next five to seven years have the potential to double their cash flow (Bughin et al. 2018). At the same time, various AI-enabled technologies have been widely adopted in daily lives, raising a variety of ethical concerns over bias, privacy, and security. This has produced an urgent need for cultivating AI literacy among all learners to better prepare everyone for a future embedded with AI. Preparation includes gaining an awareness of bias, privacy concerns, as well as potential for regional prosperity, and workforce development. The need has attracted educators' attention and led to the establishment of various educational initiatives and development of curricular materials such as AI4K12's K-12 Guidelines for Artificial Intelligence (Touretzky et al. 2019b), code.org's AI for Oceans curriculum, and MIT's Responsible AI for Social Empowerment and Education (RAISE) initiative.

Many of the current AI curricular activities, however, were taught by curriculum developers or AI content experts in out-of-school settings (Casal-Otero et al. 2023). There is very limited understanding of how teachers teach AI literacy lessons in classrooms and as important, if not more, whether/how teachers' teaching can impact student learning of AI literacy. Research is urgently needed to fill in this knowledge gap because teaching AI literacy in school has the most potential for reaching diverse populations and meeting the need of preparing the next generation of AI literate citizens to thrive in life and work with AI.

In this paper we report an effectiveness study that investigates the impact of an AI literacy curriculum called [anonymized] curriculum when taught by classroom teachers in classrooms. The research employed a quasi-experiment design and included students from two consecutive academic years: students in Year 1 were in the comparison group who did not learn the AI literacy curriculum; students in Year 2 were in the experimental group who learned the AI literacy curriculum taught by their teachers after they participated in an AI literacy professional development (PD) program. The [anonymized] curriculum was found to help students develop their AI literacy when taught by curriculum developers in prior summer's AI workshop settings (Authors 2023). This study aims to further investigate the impact of the curriculum through the comparison of student learning outcomes of the experimental and comparison groups. The research questions are:

- To what extent can first year AI literacy teachers help students develop AI literacy by teaching an AI literacy curriculum in classroom settings? Do students of different characteristics perform differently?
- What implementation strategies were employed by the teachers? Do different curriculum implementation strategies impact students differently?

Rationale

The need of fostering AI literacy among K-12 students has drawn more and more educators' attention. However, many of the current definitions of AI literacy focus on college students and the workforce. For instance, (Aoun 2017) defined

AI literacy in higher education as the ability to realize and utilize AI by understanding its concepts and usage. (Kong, Cheung, and Zhang 2021) provided an operational definition of AI literacy that comprises understanding AI concepts, competencies in using AI concepts for evaluation and using AI concepts to understand the real world. (Long and Magerko 2020), based on a synthesis of the literature on interdisciplinary learning, defined AI literacy as a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace. While these definitions aim to be comprehensive, they may not sufficiently reflect the cognitive capabilities and limitations of young learners due to the dearth of research on youth's capabilities of AI.

In recent years educators started shifting their attention to AI literacy education at the K-12 level. The AI4K12 (Touretzky et al. 2019a) published the five "big ideas" of AI as a framework for guiding curriculum development to foster AI literacy: perceptions, representation and reasoning, learning, natural interaction, and societal impact. (Wong et al. 2020) perceived AI concepts, applications, ethics, and safety as sub-elements of AI literacy. (Ng et al. 2021) explicated four aspects of AI literacy (i.e., know and understand, use and apply, evaluate and create, and ethical issues). They further proposed AI literacy as a part of digital literacy, a fundamental skill for everyone, not just for computer scientists. These definitions of AI literacy are typically designed to set goals for AI education. To better promote AI literacy education in K-12 schools, we need a definition that has been tested and validated in real classroom settings.

Fostering AI Literacy at the K-12 Level: Developing Conceptual Understanding and Empowerment

In this paper we define AI literacy as the necessary knowledge and skills for K-12 students to become informed citizens, critical consumers of AI, and potentially, future developers of AI. We posit that fostering AI literacy at the K-12 level needs to not only develop students' understanding of AI concepts and related ethics issues, but also support them with new abilities and ways to feel empowered to thrive in the era of AI. Our definition of AI literacy encompasses three dimensions: technical concepts, related ethics, and attitudes toward AI's impact on future careers.

The first two dimensions fall into the domain of conceptual learning, meaning that students need to understand fundamental concepts of AI and AI's potential for bias. For instance, the conceptual understanding of Supervised Learning (SL) at the middle school level involves understanding what the training and testing phases are, what happens during each phase and why each phase proceeds in this way, why it needs labeled data, what may cause bias, and how to mitigate potential bias. Given that teenage students typically do not possess sufficient technical background to understand the mathematics underlying AI, it is unreasonable to expect them to develop their own AI algorithms. Yet, it is reasonable and age appropriate for them to get an intuitive sense of how AI model work and what processes are used in training and testing AI models. Understanding how AI works is also essential for helping learners develop an understanding of how the systems may be biased and subsequent ethical issues with their use. Teenagers nowadays interact with AI enabled tools and are exposed to various AI generated media on a daily basis. They may be targeted by fake information, believe it is real, and act accordingly. By gaining an understanding of how AI works, students will be able to judge the ethical considerations involved and the value and impact of AI tools while engaged with the technology. This competency is also essential for every future worker because they must be able to evaluate whether an AI solution is ethical and decide whether to use it in work and life (Eaton et al. 2018).

The third dimension falls into the domain of empowerment, i.e., giving people the tools they need to better control their lives and expand their coping skills (Mäkinen 2006). In the context of AI literacy, (Kong, Cheung, and Zhang 2021) proposed that AI empowerment at the higher education level or the public level should include AI self-efficacy (how well learners believe they are doing when they engage with AI), meaningfulness (the perceived value and significance that AI has for learners in their everyday lives), impact (the degree to which interacting with AI is perceived as making a difference or having societal impacts), and creative self-efficacy (learners' conviction that he or she can come up with new ideas and solutions). Informed by their work and considering the characteristics of K-12 students (e.g., many do not recognize that they are interacting with AI), we argue that fostering AI empowerment among younger learners should focus on motivating and sustaining their interest in AI, enabling them to recognize that AI is relevant to their daily lives (not just computer scientists), and helping them develop a critical stance toward AI applications. Further, research has suggested that adolescence is an important time for youth to develop future career ideas (Porfeli and Lee 2012). AI is and will inevitably be changing the landscape of the job market. It will replace many jobs and will likely transform almost all occupations at least to some degree (Frank et al. 2019). This has raised public fears of mass unemployment and uncertainty of how to adapt to the technological change. Developing AI empowerment among youth should also foster their career awareness (teaching them AI's effect on jobs and how future jobs may change) and adaptability (helping them develop the ability and coping skills to adapt to changes in future jobs) (Johnston 2018) so that they can feel confident that they are flexible and adaptable in future careers.

Promoting AI Literacy Education in K-12 Schools

One of the most documented factors that hinder the adoption of curriculum innovations in the classroom is teachers, including teachers' understandings of the innovation, teacher training, lack of guidance for teachers, as well as other challenges teachers face such as insufficient resources and lack of administrative support (Carless 2003). Amid these difficulties, teachers' training is particularly important for bringing AI literacy education into schools. Almost all classroom teachers have not dealt with AI during their prior education or training and are, therefore, facing a stiff challenge when it comes to teaching it in classrooms. Surveys of teachers' perceptions of AI show that their knowledge has been broadly influenced by media coverage and includes many misconceptions and misunderstandings that would easily reflect in their teaching (Lindner et al. 2019). Teachers seem to have almost no knowledge and even no concrete preconcepts concerning technical aspects of AI (Lindner and Berges 2020). This lack of background content knowledge leads to a need for teacher PD program that emphasizes basic concepts and technical details of AI as well as AI's potential for social and ethical consequences.

Yet, simply offering a teacher PD is not enough. Research has found that teachers' adoption and implementation of any novel approach in classrooms is a gradual process and allowing teachers to conduct local implementations and critique the innovation can accelerate the process (Sedova 2017). Luehmann (Luehmann 2002) found across five teacher case studies that teachers went through a non-linear appraisal-customization process before they finalized a new curriculum. The teachers need to first gain evidence of improvements in student performances before they can decide whether or not to incorporate the change in their curriculum. Similarly, Guskey's Model of Teacher Change (Guskey 2002) posits that the experience of a successful implementation is much more effective in changing teachers' beliefs in adopting a curriculum innovation than any PD. These findings coincide with Roger's a five-stage diffusion of innovation process (Rogers 2003), where the teacher needs to first become aware of the innovation (Stage One: Knowledge), develop an opinion or attitude towards it with respect to its value and possible use (Stage Two: Persuasion), and decide whether to try the new change (Stage Three: Decision). Afterwards, the teacher can implement the innovation to a varying degree depending on the situation to determine its usefulness (Stage Four: Implementation), and finalizes his/her decision to continue using and customizing it (Stage Five: Confirmation).

To promote AI literacy education into K-12 schools, teachers should be allowed to carry out classroom implementations of AI literacy curriculum, see their students' reactions to the curriculum, evaluate the effect, and make plans for further modifications. This paper reports the findings of two middle school teachers who taught an AI literacy curriculum in their classrooms the first time to evaluate its effectiveness. The two teachers completed a PD program that prepared them with content and pedagogical knowledge to teach AI literacy. See below for details of the curriculum and the PD program.

Background

[Anonymized] curriculum

The [anonymized] curriculum aims to foster middle and early high school students' AI literacy and encourage them to think about AI as a system with personal, career, and societal implications. The design of the curriculum was based on our definition of AI literacy wherein students must learn three core domains to become AI literate citizens: technical concepts, ethical and societal implications, and AI's impact on careers.

The [anonymized] curriculum features activities that teach five key AI concepts: (1) a general introduction to AI, (2) logic systems, (3) Supervised learning, (4) Neural Networks (NN), and (5) Generative Adversarial Networks (GANs). Within each topic, students learn the key concepts (e.g., processes, steps), investigate potential for bias in the datasets and algorithms (and potential mitigation strategies), discuss the societal and ethical impacts of biased AI systems, connect to their daily lives and future selves, and engage in "AI and my future" career exploration activities that support students to become aware of AI's impact on jobs, recognize their own strengths and interests, and realize the importance of technical skills development and the ongoing nature of change and adaptation in today's job world. For instance, in the topic of Supervise Learning, students first learned how to use Google's Teachable Machine to train AI models to detect faces, then reflected on their experiences to understand why labeled data are needed, and what takes place during the training and testing phases. They then explored bias in their trained AI models and experimented whether an increasing amount of data can mitigate bias. Afterwards students explored algorithmic bias from tools that may be used in everyday life (e.g., the Quick Draw), and explained why certain items in a dataset, the size of the dataset, and the diversity of items in the dataset might introduce algorithmic bias. Then they brainstormed, justified, and tested a method of re-curating a dataset to solve an algorithmic bias problem in their own trained AI models. In the career training, students learned by watching and discussing a video about how autonomous driving may impact long haul truck drivers' work and what new job opportunities can emerge.

In total the curriculum includes 22 lessons and lasts 30 hours. Our previous implementations of the curriculum in summer camps have shown that middle schoolers were highly engaged and achieved significant gains in their AI literacy knowledge and skills (Authors 2023). See the supplementary materials for detailed information of the curriculum.

[Anonymized] Teacher Professional Development

The teacher participants in this study completed an [anonymized] PD program to prepare for their classroom teaching of the [anonymized] AI curriculum. The PD program (Authors 2022) features three components:

- a 20-hour AI Book Club (ABC), which included synchronous online meeting for 1.5 hours and asynchronous "homework" assignment of approximately 0.5 hour every week. Each week's homework engaged teachers in reading and reviewing curricular materials and readings about an AI related topic, such as What is AI?, Ethics in AI, Perceptions and Machine Learning, Neural Networks and Deep Learning, and Generative AI. During the online ABC meetings, teachers first reviewed what they learned from the assignment and then experienced the AI curriculum activities as learners where the PD instructors and facilitators provided teaching demonstra-

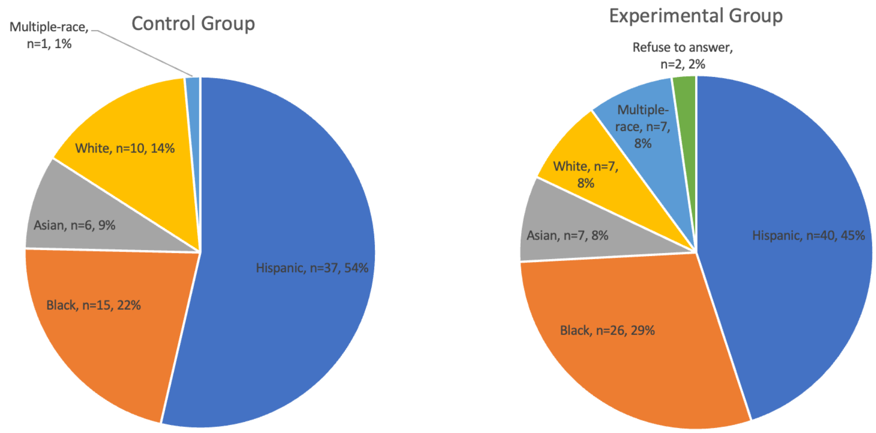

Figure 1: Demographics of control and experimental group participants

tions and pedagogical tips. They also reviewed student work from previous implementations of the AI curriculum and discussed key points, common misconceptions, and how to assess student understanding.

- a 30-hour summer practicum, during which teachers co-taught the AI curriculum in camps with two other peer teachers who attended the ABC together. Each teacher needed to lead eight or nine hours of activities. At the end of each day, the teachers discussed their experiences of leading the activities.
- monthly webinars during the school year for teachers to share their experiences of ongoing classroom implementations, discuss challenges, and brainstorm solutions.

Methods

Participants and Research Design

This study employed a quasi-experimental research design (Gopalan, Rosinger, and Ahn 2020) and involved two teachers and their students from two consecutive academic years. Students of Academic Year 2021-2022 participated as the comparison group and did not teach any AI related curriculum in school. Students of Academic Year 2022-2023 were included as an experimental group whose teachers taught the [anonymized] AI curriculum in classrooms. To establish a baseline for the comparison, both groups completed the same pre and post-test.

Student participants were middle school students enrolled in technology classes (comparison group: 6th graders: n=41, 59%, 7th graders: n=20, 29%, 8th graders: n=8, 12%; experimental group: 6th graders: n=51, 57%, 8th graders: n=38, 43%). Sixty-one percent of students in the comparison group were male (n=42) and the remaining 39% students were female. Fifty-eight percent of students in the experimental group were male (n=56) and the remaining 42% students were female. The percentages of students from racial and ethnic minority groups were high, with 77% in the comparison group and 84% in the experimental group. Figure 1 shows detailed students' race and ethnicity information in each group.

The two teachers (Ms. JT and Ms. BK) were experienced teachers who had multiple years of experience teaching technology classes at the middle school level. Neither of them had any prior experience of teaching AI literacy in the classroom. Both teachers participated in the [anonymized] PD program in the spring of 2022 and completed the ABC and the summer practicum. During the academic year of 2022-2023, they decided to try the [anonymized] curriculum in the classrooms. Due to their schedule constraints, the two teachers chose different strategies of implementation: Ms. JT chose to implement the curriculum once per week and the implementation lasted five months; Mr. BK implemented the curriculum everyday in class and completed the classroom implementation in two and half months. Both teachers managed to implement the whole curriculum.

Instruments

To compare student learning outcomes, we administered the same pre and post-test in the two groups. The instruments included two sets of questions. Both sets of questions have been validated in our prior work with satisfactory internal consistency (AI-CI: Cronbach's alpha=.73, Attitudes toward AI: Cronbach's alpha ranged from .71 to .94 across the subscales). See supplementary materials for the instrument.

- AI Literacy Concept Inventory (AI-CI), which measures student understanding of AI literacy concepts including AI general concepts, logic systems, machine learning general concepts, supervised learning, neural networks, and GANs through multiple-choice questions. The development of the AI-CI was based on the literature of concept inventory for Computer Science (Taylor et al. 2014) and prior work that shows a set of AI concepts that middle school students are capable of learning (Payne 2020).
- AI attitudes and career futures survey, which includes 5-Likert scale questions examining participants' interest in AI, anxiety about AI, perceived relevance of AI to their life, career awareness, and career adaptability (a subscale that includes multidimensional adaptive behavior for future careers such as identifying educational choices to join future jobs). We utilized student performance on the

	experimental group				Control group				ANCOVA results
	Pretest		Posttest		Pretest		Posttest		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
All items	.52	.15	.64	.12	.51	.19	.56	.16	F(1, 155)=16.51, *p*<.001
AI general	.48	.17	.60	.16	.45	.18	.48	.19	F(1, 155)=19.08, *p*<.0001
Logic systems	.69	.33	.84	.27	.78	.35	.66	.38	F(1, 154)=14.78, *p*<.001
ML general	.55	.33	.58	.25	.58	.37	.64	.30	F(1, 154)=2.20, *p*=.14
Supervised learning	.44	.23	.59	.21	.44	.27	.51	.25	F(1, 154)=5.44, *p*<.05
GANs	.55	.28	.71	.24	.49	.29	.59	.23	F(1, 155)=7.83, *p*<.01

Figure 2: Control and experimental groups' performance on AI-CI

AI attitudes and career futures survey as an indicator of whether and to what extent they feel they are empowered for the AI era.

This paper mainly reports findings from student data to determine the effectiveness of the curriculum. We conducted an analysis of covariance (ANCOVA) to compare pre and post-test performance of the two groups. We also compared student learning outcomes of the teachers to examine whether or how teachers' different implementation strategies may impact student learning of AI literacy differently. Teachers' post-lesson reflection interviews were used to triangulate the findings and reveal more information about teachers' classroom implementations.

Findings

Effectiveness of the Curriculum on Conceptual Understanding

Overall, the ANCOVA results (see Figure 2) shows a significant effect of the curriculum on student AI literacy as measured by the AI CI after controlling for their pretest scores on AI CI, [F(1, 155)=19.08, $p < .0001$]. Students of the two groups started with the same level of prior knowledge on the post-test. On the post-test the experimental group achieved much higher scores (Mean=.64, SD=.12) with an effect size (Cohen's *d*) of .88. The comparison group did not achieve as much (Mean=.56, SD=.16, Cohen's *d* =.28).

We also conducted ANCOVA using students' characteristics (grades, gender, and race and ethnicity information) as independent variables. None of these variables were found to have a statistically significant impact on students' learning outcomes. This suggests that there was no significant difference in student conceptual learning of AI literacy across students of different characteristics. Students of different characteristics benefited similarly from the curriculum.

A further examination of student performance on the subscales shows that students in the experimental group demonstrated a much better understanding on all subscales than the comparison group, except for the Machine Learning (ML) general concepts (i.e., distinguishing between supervised and unsupervised learning, distinguishing between AI that are used for classification or for generation). Students in both groups started with high prior knowledge on this subscale: their average scores on the pretest were over .50, meaning that they answered over half of the questions correctly. On the posttest, the comparison group even had a higher average score (.64) than the experimental group (.58), but the difference was not statistically significant. This indicates that middle school students through their daily interactions with AI may have established some vague ideas of different types and uses of AI. The significant gains achieved by the experimental group on all the other subscales (AI general concepts, logic systems, supervised learning and GANs) showed that the [anonymized] curriculum can foster students' understanding of AI literacy concepts in formal educational settings. These findings also demonstrate that teachers without an AI background were able to successfully teach AI literacy lessons in classrooms after being trained by the PD program, suggesting the success of the [anonymized] PD program.

Effectiveness of the Curriculum on AI Empowerment

ANCOVA on students' pre/post-test performances on the AI attitudes and career futures survey shows that students in the experimental group achieved greater gains in their perceived relevance of AI to their own life, anxiety about AI, career awareness, and adaptability of adjusting to future careers than the comparison group (see Figure 3). This suggests that students in the experimental group felt that they were more

	experimental group				Control group				ANCOVA results
	Pretest		Posttest		Pretest		Posttest		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Interest	3.45	.75	3.31	.68	3.29	.82	3.28	.82	F(1, 153)=.73, *p*>.05
Perceived relevance	3.56	.53	3.59	.51	3.38	.56	3.34	.59	F(1, 153)=4.34, *p*<.05
Anxiety*	1.80	1.41	3.06	.59	1.64	1.56	2.82	.72	F(1, 155)=4.66, *p*<.05
Career awareness	3.13	.79	3.28	.63	3.00	.74	2.95	.68	F(1, 146)=8.74, *p*<.01
Career adaptability	3.88	.71	3.94	.58	3.52	.71	3.44	.71	F(1, 134)=16.34, *p*<.001

*Note. All the Likert-scale questions were coded from 0 to 5. The anxiety items were reverse coded, i.e., the higher the score, the lower the anxiety.

Figure 3: Control and experimental groups' performance on Likert scale questions

prepared for the AI's impact on future jobs and that they can deal with changes in future careers than those in the comparison group.

Students in the experimental group did not achieve as much gains on the interest scale. Students started with high interest scores on the pretest (average mean score on the pretest was 3.45 out of 5), indicating that they were curious about AI before learning the curriculum. Their average interest score, however, dropped after learning the curriculum (Pretest Mean= 3.45, Posttest Mean=3.31), although not statistically significant (p=.09). One reason for the decreased interest might be the introduction of the related negative impact (e.g., bias, discrimination) and ethical issues of AI. Students were astonished and even disappointed to find that a lot of AI enabled tools are biased, *"I think [Investigating Bias activity] opened their eyes to more, like the implications of bias, and how we have to be careful with it. A lot of them [students] didn't like it, they felt they were offended…"* (Ms. JT's reflective interview after classroom implementation). While it is critical to prepare students to become knowledgeable of potential biases of AI, introducing such topics could challenge students' misconceptions that technologies are always correct and fair. The realization that machines can be biased or incorrect may lower students' interest in and enthusiasm for AI. To counteract this, teachers suggested incorporating more examples of how AI can benefit people in the curriculum.

Further ANCOVA using students' characteristics as independent variables shows that students in different grades achieved different gains on the subscale of career adaptability [F(2, 131)=3.85, $p < .05$] and that there was an interaction between grade and student groups [F(1, 131)=4.90, $p < .05$]. Grade 8 students in the experimental group achieved the most gains. Their posttest score of the career adaptability scale increased to 3.97, much higher than Grade 6 or 7 students in the comparison group whose average score was 3.49 on the post-test. Given that early adolescence is an important time for students to ideate their future career ideas, it is crucial to prepare 8th graders and early high schoolers to become confident and comfortable with adjusting to future jobs.

Impact of Different Implementation Strategies

As noted earlier, Ms. JT and Mr. KB chose different implementation strategies to teach AI literacy lessons the first time in their classrooms. Ms. JT taught the AI curriculum once every week and Mr. KB taught the curriculum everyday. First, paired t-tests showed that students of both teachers significantly improved their AI-CI and the AI empowerment scores on the posttest. The effect size of Ms. JT's students was .78 and the effect size of Mr. KB's students was 1.00 for the AI-CI. This suggests that either implementation strategy successfully helped students improve their understanding of AI literacy concepts.

To further compare the impact of implementation strategies on student AI-CI and survey performances, we conducted ANCOVA using teacher as an independent variable. The results showed no significant differences in all survey subscales. With regard to the conceptual learning outcomes, the ANCOVA showed a statistically significant difference between the performances of students of the two teachers [F(1, 86)=5.14, $p < .05$]. Mr. KB's students outperformed Ms. JT's students on the post-test after controlling for their pre-test scores [Mr. KB's students posttest average=.68 (.12), Ms. JT's students posttest average=.61 (.11)].

Further examinations (see Figure 4) revealed that Mr. KB's students achieved significant gains on items assessing their conceptual understanding of the processes involved in supervised learning and GANs [SL: F(1, 86)=5.11, $p <$ .05; GANs: F(1, 86)=6.97, $p < .01$]. On the posttest, Mr. KB's students on average scored .66 on SL questions and .79 on GANs questions, compared to Ms. JT's students' mean

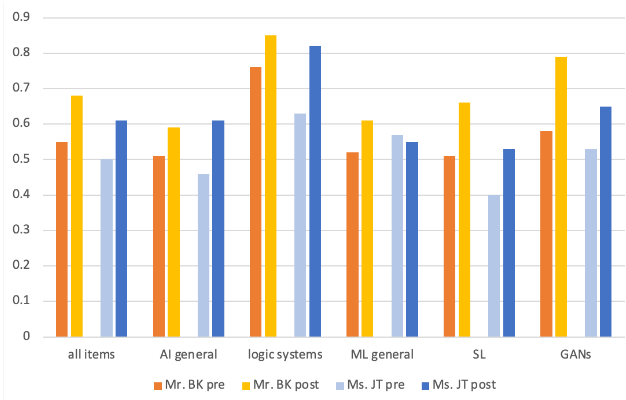

Figure 4: Pre and posttest AI-CI performances of Mr. BK and Ms. JT's students

scores of .53 on SL questions and .65 on GANs questions. This suggests that Mr. KB's students on average developed a deeper understanding of the processes of SL and GANs than Ms. JT's students (see Figure 4). This finding indicates that higher frequency of classroom implementations of AI lessons may be more effective in terms of supporting students to develop a deeper understanding of AI literacy. This also coincides with teachers' views and impressions of student learning. For instance, Mr. KB, in his reflective interview, talked about his impression of student learning,

> *"the fact that I was able to do it every day for basically two months, and we could feed off the previous lesson, the curriculum seemed to flow well. I didn't really have any distractions from teaching the curriculum. The engagement was great from the students. And I think part of that engagement is because they see how one lesson flows to the next lesson flows to the next lesson. Whereas I feel like if I was just doing it one or two days a week or when the curriculum fit in, that flow wouldn't be there and it would feel very choppy..."*

Conclusions

Promoting AI literacy education in K-12 schools is not an easy task. Besides the complexity of formal education that hinder its widespread implementation, the field of AI literacy education has few curricula that have been tested in classrooms. Having validated curricula is crucial because teachers must see evidence of student learning before adopting any new curriculum (Guskey 2002). To motivate teachers to offer AI literacy education in the classroom, research is needed to demonstrate to teachers the effectiveness of AI literacy education. This paper reports an effort to investigate the effectiveness of an [anonymized] curriculum taught by two classroom teachers through the comparison of the learning of the experimental group with the comparison group (i.e., students who did not engage with the curriculum). Overall, the results showed that students in the experimental group developed a deeper understanding of AI literacy concepts and felt more empowered to deal with the changes AI will bring to future jobs. Students, regardless of their backgrounds (race, ethnicity, gender), all benefited from the curriculum. This confirms that the [anonymized] curriculum can successfully foster students' AI literacy skills and knowledge when taught by in-service teachers in inclusive classrooms. The large gains students in the experimental group achieved also reflect the success of the [anonymized] PD program. Although both teachers had no prior experience of teaching AI literacy, they learned sufficient content and pedagogical knowledge and skills through the PD to teach the curriculum in their own classrooms. Both teachers felt they learned a lot from the PD and from the classroom implementation.

Further, the comparison between the two teachers' implementations showed that teaching the curriculum with higher frequency could lead to more student learning gains. Mr. BK's students, who learned the curriculum everyday, developed a more solid understanding of AI concepts of supervised learning processes and GANs than Ms. JT's students, who learned the curriculum once a week. Teachers' interviews provide more evidence of this. Mr. BK reflected that

> *"doing this [teaching the curriculum] every day, where you can see the connections from lesson to lesson, made the teaching and learning smoother. It went very well, the kids were very engaged. They were interactive, they loved it."*

In contrast, Ms. JT reflected that

> *"I have to try to slow myself down and check what they're feeling... I have to keep telling myself, okay, they're kids, they haven't really been learning this for a week, they have forgotten, so it takes time... And I'm doing it again, and it's just like that... So, I'm learning to have a lot more patience, more repeating myself, and more understanding."*

This evidence indicates that when teaching AI literacy in the classroom, it may be more effective if the lessons are taught with shorter time intervals to support students to develop a more coherent understanding of AI.

Finally, we are cognizant that there are a few limitations of this study. For instance, it is quasi-experimental study that measured immediate effects of the intervention using a pre-post test comparison design whereas some complex affective constructs such as interest may require more time for students to reflect and internalize. Conducting a delayed post-test and interviews could clarify the long-term effects of the curriculum experience on student attitudes toward AI. Another limitation is that although the teachers were trained together in our PD program, they are likely to use different teaching practices in their classrooms, which could significantly impact student learning. As our next step, we plan to analyze and compare teachers' self-reported teaching practices and investigate if and how these teaching practices correlate with student learning outcomes.

Acknowledgments

We are grateful to the teachers and students involved in this study. We are also grateful for our funding agency.

References

Aoun, J. E. 2017. *Robot-proof: higher education in the age of artificial intelligence*. MIT press.

Authors. 2022. XXX. In *XXX*, XXX.

Authors. 2023. XXXX. *XXX*, XX(XX): XX.

Bughin, J.; Hazan, E.; Lund, S.; Dahlström, P.; Wiesinger, A.; and Subramaniam, A. 2018. Skill shift: Automation and the future of the workforce. *McKinsey Global Institute*, 1: 3–84.

Carless, D. R. 2003. Factors in the implementation of task-based teaching in primary schools. *System*, 31(4): 485–500.

Casal-Otero, L.; Catala, A.; Fernández-Morante, C.; Taboada, M.; Cebreiro, B.; and Barro, S. 2023. AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, 10(1): 29.

Eaton, E.; Koenig, S.; Schulz, C.; Maurelli, F.; Lee, J.; Eckroth, J.; Crowley, M.; Freedman, R. G.; Cardona-Rivera, R. E.; Machado, T.; et al. 2018. Blue sky ideas in artificial intelligence education from the EAAI 2017 new and future AI educator program. *AI Matters*, 3(4): 23–31.

Frank, M. R.; Autor, D.; Bessen, J. E.; Brynjolfsson, E.; Cebrian, M.; Deming, D. J.; Feldman, M.; Groh, M.; Lobo, J.; Moro, E.; et al. 2019. Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences*, 116(14): 6531–6539.

Gopalan, M.; Rosinger, K.; and Ahn, J. B. 2020. Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, 44(1): 218–243.

Guskey, T. R. 2002. Professional development and teacher change. *Teachers and teaching*, 8(3): 381–391.

Johnston, C. S. 2018. A systematic review of the career adaptability literature and future outlook. *Journal of Career Assessment*, 26(1): 3–30.

Kong, S.-C.; Cheung, W. M.-Y.; and Zhang, G. 2021. Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds. *Computers and Education: Artificial Intelligence*, 2: 100026.

Lindner, A.; and Berges, M. 2020. Can you explain AI to me? Teachers' pre-concepts about Artificial Intelligence. In *2020 IEEE Frontiers in education conference (FIE)*, 1–9. IEEE.

Lindner, A.; Romeike, R.; Jasute, E.; and Pozdniakov, S. 2019. Teachers' perspectives on artificial intelligence. In *12th International conference on informatics in schools,"Situation, evaluation and perspectives", ISSEP*.

Long, D.; and Magerko, B. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, 1–16.

Luehmann, A. L. 2002. Understanding the appraisal and customization process of secondary science teachers. In *annual meeting of the American Educational Research Association: New Orleans, LA*.

Mäkinen, M. 2006. Digital empowerment as a process for enhancing citizens' participation. *E-learning and Digital Media*, 3(3): 381–395.

Ng, D. T. K.; Leung, J. K. L.; Chu, K. W. S.; and Qiao, M. S. 2021. AI literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology*, 58(1): 504–509.

Payne, B. H. 2020. *Can my algorithm be my opinion?: an AI+ ethics curriculum for middle school students*. Ph.D. thesis, Massachusetts Institute of Technology.

Porfeli, E. J.; and Lee, B. 2012. Career development during childhood and adolescence. *New directions for youth development*, 2012(134): 11–22.

Rogers, E. 2003. Diffusion of Innovations fifth Ed Free Press. New York. *Rezvani, Z., Jansson, J. & Bodin*.

Sedova, K. 2017. A case study of a transition to dialogic teaching as a process of gradual change. *Teaching and Teacher Education*, 67: 278–290.

Taylor, C.; Zingaro, D.; Porter, L.; Webb, K. C.; Lee, C. B.; and Clancy, M. 2014. Computer science concept inventories: past and future. *Computer Science Education*, 24(4): 253–276.

Touretzky, D.; Gardner-McCune, C.; Martin, F.; and Seehorn, D. 2019a. Envisioning AI for K-12: What should every child know about AI? In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, 9795–9799.

Touretzky, D. S.; Gardner-McCune, C.; Martin, F.; and Seehorn, D. 2019b. K-12 guidelines for artificial intelligence: what students should know. In *Proc. of the ISTE Conference*, volume 53.

Wong, G. K.; Ma, X.; Dillenbourg, P.; and Huan, J. 2020. Broadening artificial intelligence education in K-12: where to start? *ACM Inroads*, 11(1): 20–29.