

A Comparison of the Impact of an AI Teacher Professional Development Program on Science and non-Science Teacher AI Literacy

Abstract

In the face of the rising prevalence of artificial intelligence (AI) in daily life, there is a need to integrate lessons on AI literacy into K12 settings to equitably engage young adolescents in critical and ethical thinking about AI technologies. This exploratory study reports findings from a teacher professional development project designed to advance teacher AI literacy in preparation for teaching an AI curriculum in their inclusive middle school classrooms. Analysis compares the learning experiences of 30 participating teachers (including Computer Science, Science, Math, English, and Social Studies teachers). Results suggest Science teachers' understanding of AI concepts, particularly logic structures, is on average higher than their non-Science teacher counterparts. Teacher interviews reveal several thematic differences in Science teachers' learning from the AI PD as compared to their counterparts, namely learning from reflective discourse with diverse groups. Findings offer insights on the depth and quality of Science teacher AI literacy after participating in an AI teacher PD, with implications for future research in the integration of AI education into Science teachers' inclusive K12 classrooms.

1. Introduction

As artificial intelligence (AI) becomes increasingly commonplace in everyday life, particularly in social media, access to AI education has become a public need. Recognizing that AI education is no longer specialized training for the advanced sciences, a growing number of initiatives have developed tools, standards, and curriculum for teaching fundamental AI concepts and principles to K-12 learners, including Google's Teachable Machine, the AI4K12 working group, and MIT's Responsive AI initiative. Yet, the question remains as to how to integrate AI education into core subjects (ie., English Language Arts, Math, Science, Social Studies) so as to reach all students. Younger generations must understand the fundamentals of AI, not only to make informed decisions in civic lives and work (Knox, 2020; [2]), but to responsibly engage in social life. Young adolescents interact with AI technologies daily (e.g., social media, video games) often without realizing it (Williams, et al., 2022; National Science Board Vision 2030 Report). While AI-powered platforms provide youth with entertainment and educational opportunities, their algorithms have the potential to introduce misinformation (Fernández et al., 2021) and harm (Zhang et al., 2022; Coekelbergh, 2020). Thus, it is important that students are aware of how AI works in order to be responsible consumers, and ethical decision makers during their daily interactions with AI technologies.

In this paper, we report on the findings from a study, called *Everyday AI* (EdAI), of teacher learning of AI concepts, attitudes, perceptions, and self-efficacy in teaching AI. The aim of our analysis is to identify core subject areas in which teachers may be particularly well suited to integrate AI education into their core content area. We begin with Science teachers, as their background knowledge and training overlaps with several AI concepts (ie. abstractions of

complex processes). Our analysis focused on the hypothesis that science teachers may be relatively well suited to understand AI concepts when compared to their counterparts teaching other, non-Science, core subject areas (i.e., Math, English Language Arts, and Social Studies). This examination is guided by the following research questions: (1) Given an effective teacher professional development (PD) experience introducing AI concepts and pedagogy relevant for teaching AI, how do the AI literacy learning outcomes of Science teachers compare to those of non-Science teachers? (2a) After the PD, do Science teachers' attitudes and perceptions of AI differ from their non-Science counterparts? (2b) After the PD, does Science teachers' self-efficacy towards teaching AI differ from their non-Science counterparts? (3) Do Science teachers' descriptions of their learning experiences differ from their non-Science counterparts? How?

2. Relevance to Prior Work

This examination of the potential for Science teachers to integrate AI education into their science classrooms is based on research conducted during the *Everyday AI* (EdAI) teacher professional development program, which trained middle school teachers (from all subject areas) to implement the *Developing AI Literacy* (DAILy) curriculum.

EdAI Teacher Professional Development Program

The *Everyday AI* (EdAI) teacher professional development program (NSF Award #2048746) was developed to train in-service middle school teachers, new to AI concepts, to teach AI concepts, ethics, and careers (using the DAILy curriculum, see below for details) in their inclusive classrooms during the regular school day. To do this, EdAI introduces teachers to AI concepts, in the form of a book club, then creates opportunities for authentic practice teaching for the development of AI relevant pedagogy. Thus, the EdAI PD program includes a 20-hour AI Book Club (ABC) and a 30-hour summer Practicum. The ABC comprises weekly synchronous online meetings for 1.5 hours and 0.5 hours of asynchronous reflection assignments. Each week features an AI related topic, such as What is AI?, Algorithms as Opinions, Ethics in AI, Logic Systems, Perceptions and Machine Learning, Neural Networks and Deep Learning, and Generative AI. Readings are excerpts from *Artificial Intelligence: A Guide for Thinking Humans* (Mitchell, 2019), a book that interweaves AI concepts and the history of their development; examines AI hype versus reality; and discusses ethical concerns about AI. Summer Practicums are 2-week long (3-hours a day) summer camps for middle schoolers recruited by local youth serving organizations in which participating teachers implement the DAILy curriculum in co-teaching teams of 3-4 teachers. Team members are from the same school district, but not the same content area (e.g., Science teachers may co-teach with Math, English Language Arts, and Social Studies teachers).

DAILy Curriculum

Participating teachers (both Science and non-Science teachers) studied concepts from the *Developing AI Literacy* (DAILy) curriculum ((NSF Award #2022502). DAILy is a middle school curriculum that interweaves AI concepts, awareness of AI adoption in future jobs, and the

investigation of ethical issues in AI. The curriculum features: 1) an introduction to AI, 2) logic systems (e.g., decision trees constructed by humans), 3) supervised learning (e.g., concepts, processes, and bias), 4) neural networks (NN), 5) training and testing Machine Learning (ML) models, 6) Generative Adversarial Networks (GANs) and the ethical implications of GANs (ie., deep fakes), 7) the ethical design of everyday AI technologies (ie., Youtube), and 8) AI's impact on future careers. DAILY grounds a 30- hour AI curriculum in the literature on child development and ethics education. The curriculum was designed with an understanding of middle school students' capability to learn AI concepts, such as classification and prediction, to think more abstractly than younger students (Denham et al., 2009; Kaneshiro, 2023a), and incorporate new knowledge into existing schemas (Kaneshiro, 2023a, 2023b). DAILY also relies on research showing that even young children have a conception of ethical behavior (Payne, 2020).

3. Participants

In total, 30 in-service teachers from 3 US school districts participated in the EdAI PD program. Teachers taught a variety of disciplines: 60% (n=18) Computer Science/Technology, 17% (n=5) Science, 7% (n=2) English Language Arts, 7% (n=2) Math, 3% (n=1) Library/Media Literacy, 3% (n=1) Social Studies/Civics, and 3% (n=1) Art/Music. Their school districts served student populations that were largely from underrepresented groups in STEM and Computing (59%, 90% and 85% respectively). Ninety-percent of teacher participants (n=27) were from groups underrepresented in STEM and Computing education: 73% (n=22) being female, 7% (n=1) non-binary; 37% (n=11) being African American, 23% (n=7) being Hispanic Latino, and 3% (n=1) being Middle Eastern. Participating Science teachers (16%, n=5) were predominantly women of color: 100% female (n=5), 40% (n=2) African American, and 20% (n=1) Middle Eastern.

4. Method

Data collected in this study includes responses to the *teacher AI Concept Inventory* (AICI) comprising multiple scales, including teacher attitudes and perceptions of AI (A&P) and teacher self-efficacy (SE) towards teaching AI. The AICI was administered in two waves: (1) the AICI scales on AI concepts were administered before and after the ABC; and (2) the AICI scales on A&P and SE were administered before and after the full EdAI PD program (after the Practicum). The AICI includes (a) a concept inventory designed to assess teacher understanding of key AI concepts. This portion consists of 33 items with 3 sub-scales: AI general concepts, logic systems (ie., decision trees), ML general concepts (ie., supervised machine learning, and neural networks (NNs)). The reliability of these scales ranges from 0.75 to 0.88. The AICI also includes 5-point Likert scale items assessing teachers' (b) A&P (Cronbach's $\alpha=.85$) which consist of 25 questions examining teacher's interest in AI, anxiety toward AI, awareness of AI's impact on future jobs, and perceived relevance of AI to their lives; and (c) SE (Cronbach's $\alpha=.85$) which includes 19 items evaluating teachers' beliefs in middle school students' competency of

learning AI, confidence of teaching AI, and community support of teaching AI. To control for non-normal distributions, Welch's t-tests were used to compare results from surveys completed by participating Science teachers to non-Science teachers. Findings from the surveys were triangulated against teacher interview data.

Teacher interviews were conducted by a member of the research team once teachers finished implementation of the DAILY curriculum in their classrooms during the academic year (AY). Analysis focused on the following question: "Is there something that you think you've learned from the EdAI PD (including the AI Book Club and the summer Practicum) that you found useful for your teaching last fall?"

Table 1. Comparison of Science and non-Science Teachers' AI Literacy

	Science Teachers		Non-Science Teachers		<i>t-test</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
AICI total	15.80	0.84	14.50	2.41	2.101	0.05*	0.58
AI General	7.60	0.89	7.33	1.69	0.505	0.62	0.17
Logic Structures	3.00	0.00	2.16	0.87	4.703	< 0.01*	1.04
ML General	5.20	0.45	5.00	1.10	0.664	0.52	0.44

Table 2. Comparison of Science and non-Science Teachers' attitudes and perceptions (A&P), and self-efficacy (SE) towards teaching AI

	Science Teachers		Non-Science Teachers		<i>t-test</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Attitudes & Perceptions							
Interest in AI	4.80	0.45	4.71	0.46	0.414	0.693	0.20
Anxiety towards AI	2.80	0.45	2.45	0.66	1.418	0.193	0.54
Relevance of AI	4.40	0.55	4.50	0.51	-0.376	0.721	0.19
Career awareness	4.20	0.45	3.92	0.50	1.260	0.252	0.57
Self-Efficacy							
Belief in students' ability	4.20	0.84	4.13	0.61	0.190	0.859	0.12
Confidence in teaching AI	3.60	0.55	3.79	0.72	-0.670	0.523	0.27
Community support	4.00	0.00	4.08	0.65	-0.623	0.539	0.14

5. Results

On average, Science teachers performed slightly higher than their non-Science counterparts on the AICI post-survey of AI concepts (implemented after the ABC), $t(19.63) = 2.10$, $p = 0.05$, $d = 0.58$, with Welch's adjustment for non-normal distributions. Examination of survey responses by scale suggests that Science teachers outperformed their counterparts on items specific to logic structures (ie., decision trees), $t(23) = 4.20$, $p < 0.01$, $d = 1.04$. See Tables 1 and 2 for findings across all AICI scales, including scales on A&P and SE.

Science teachers' pre-survey AICI scores were not significantly different from their counterparts, although differences showed a medium effect size (Science, $M = 14.00$, $SD = 1.83$; non_Science, $M = 12.5$, $SD = 2.40$), $t(5.12) = 1.389$, $p = 0.22$, $d = 0.62$. Differences between Science teachers' pre- and post-survey scores were not-significant, although differences showed a large effect size (pre, $M = 14.00$, $SD = 1.83$; non_Science, $M = 15.80$, $SD = 0.84$), $t(4.00) = -1.825$, $p = 0.14$, $d = 1.33$. Significance may have been difficult to detect due to the small sample size. Figure 1 shows the distributions of the Science teacher AICI pre- and post-survey responses as compared to their non-Science teacher counterparts. Analysis of teacher A&P and SE responses, showed no significant differences between the pre- and post-survey (administered after the Practicum) and no significant differences between Science and non-Science teachers' post-survey responses.

Several themes emerged from among teacher interviews of Science and non-Science teachers. For instance, when reflecting on their learning from the EdAI PD, all Science teachers spoke of benefiting from conversations and collaborations with other teachers different from them in casual "non-pressured" contexts; one teacher shared, *"it's just amazing the collaboration that people get to have"* another shared, *"I think this type of professional development model really works. You know like teaching it in sort of a non-pressured situation that came you know in time for me to teach it in my class."* A third expanded on the merits of asynchronous, as well as synchronous, opportunities for reflection with other teachers on the curriculum, *"I felt like those reflective discussions that we have on Slack and stuff like that really just caused us to kind of have a different viewpoint."*

When reflecting on their learning, non-Science teachers described moments of applied learning such as (a) learning AI concepts because they were relevant to their everyday lives or (b) learning AI concepts right along with their students during the AY. For example, a Computer Science teacher shared that she, *"learned along with the kids,"* while another Computer Science teacher explained, *"I didn't know that AI was so into our lives as well I didn't even think about it like like I was shocked."* The majority of non-Science teachers responded by listing AI concepts learned during the EdAI PD, e.g., a Social Studies/Civics teacher shared, *"Like you know it makes sense once you know about it, but to really think okay get out the data. In order to get a more better prediction you need more data you need to include other perspectives,"* and a Computer Science teacher shared, *"I learned about GANs, general adversarial networks, I learned about teachable machines."* Interviews suggest that while their background knowledge may aid their learning of AI concepts, Science teachers may benefit from interacting with the learning experiences of other, non-Science content area teachers.

7. Conclusion

Science teachers may be relatively well suited to integrate AI literacy curriculum into their core content area. When provided with effective AI teacher PD, Science teachers tended to score higher on a post-survey of AI concepts than their non-Science teacher counterparts (even Computer Science teachers). In interviews, Science teachers shared an appreciation for learning

AI concepts in community through reflective discussion and collaboration with diverse groups of teachers. Further research is needed to replicate these results and to investigate how AI literate Science teachers integrate AI concepts into their classrooms.

References

- Coekelbergh, M. (2020). *AI ethics*. Bit Press.
- Denham, S. A., Wyatt, T. M., Bassett, H. H., Echeverria, D., & Knox, S. S. (2009). Assessing social-emotional development in children from a longitudinal perspective. *Journal of Epidemiology & Community Health*, 63(Suppl 1), i37-i52.
- Fernández, M., Bellogín, A., & Cantador, I. (2021). Analysing the effect of recommendation algorithms on the amplification of misinformation. *arXiv preprint arXiv:2103.14748*.
- Kaneshiro, N. K. (2023a). Adolescent development. *MedlinePlus*. <https://medlineplus.gov/ency/article/002003.htm>.
- Kaneshiro, N. K. (2023b). School-age children development. *MedlinePlus*. <https://medlineplus.gov/ency/article/002017.htm>.
- Knox, J. (2020). Artificial intelligence and education in China. *Learning, Media and Technology*, 45(3), 298-311.
- Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Penguin UK.
- National Science Board Vision 2030 Report: <https://www.nsf.gov/nsb/publications/2020/nsb202015.pdf>, last accessed 2023/08/14.
- Payne, B. H. (2020). Can my algorithm be my opinion?: an AI+ ethics curriculum for middle school students (Doctoral dissertation, Massachusetts Institute of Technology).
- Williams, R., Ali, S., Devasia, N., DiPaola, D., Hong, J., Kaputsos, S. P., Jordan, B. & Breazeal, C. (2022). AI+ ethics curricula for middle school youth: Lessons learned from three project-based curricula. *International Journal of Artificial Intelligence in Education*, 1-59.
- Zhang, H., Lee, I., Ali, S., DiPaola, D., Cheng, Y., & Breazeal, C. (2022). Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory study. *International Journal of Artificial Intelligence in Education*, 1-35.