From Embeddings to Chatbots: Playful NLP Activities for Middle School AI Literacy

Jessica Vandenberg[1], Alex Goslen[2], Claire Aguiar[1], Wookhee Min[1], Veronica Catete[1], Bradford Mott[1]
[1]North Carolina State University, Raleigh, NC, USA
[2]Elon University, Elon, NC, USA
jvanden2@ncsu.edu, agoslen2@elon.edu, cmaguiar@ncsu.edu, wmin@ncsu.edu, vmcatete@ncsu.edu, bwmott@ncsu.edu

Abstract

As large language models (LLMs) and chatbots become increasingly prevalent, there is an urgent need to create engaging, age-appropriate learning activities that foster foundational AI literacy with a focus on natural language processing (NLP). This paper presents the iterative design and implementation of three instructional activities that introduce middle school learners (ages 11–14) to NLP concepts through playful, hands-on experiences aligned with the AI4K12 Big Idea of *Natural Interaction*. These activities include: (1) an unplugged card game that develops students' understanding of embeddings and similarity, (2) an unplugged collaborative sentence-generation challenge that illustrates how language models work, and (3) a web-based educational game in which students design and interact with chatbots. Each activity was implemented and refined across multiple educational contexts, including teacher professional development workshops, summer camps, and classroom implementations. All activities are designed to be easy to set up, requiring only commonly available classroom technology (e.g., laptops) and a few inexpensive materials (e.g., decks of cards), and are supported with facilitation guides and reflection prompts. Early implementations revealed areas for refinement, leading to clearer scaffolding that helped students connect gameplay to underlying NLP concepts, and post-refinement surveys indicated that students found the activities both enjoyable and educational. Findings suggest that blending unplugged and digital formats enhances comprehension, and that tailoring content to students' local contexts supports engagement. By making these activities openly available, this work contributes to the growing ecosystem of K–12 AI education resources and offers practical guidance for integrating NLP concepts into classroom instruction.

Introduction

Artificial intelligence (AI) is rapidly transforming the technological landscape and reshaping how individuals interact with the world (Floridi et al. 2018). From search engines to recommendation systems and digital assistants, AI technologies are now deeply embedded in the digital tools that shape modern experiences. As this integration accelerates, the ability to critically engage with AI systems is no longer a niche technical skill, it has become a foundational literacy necessary for all students (Long and Magerko 2020; Wang and Lester 2023). To support this need, efforts such as the AI4K12 initiative have proposed frameworks to guide educators in bringing core AI concepts to K–12 learners (Touretzky et al. 2019). Despite growing momentum, many K–12 educators lack the resources, training, and accessible curriculum materials necessary to meaningfully integrate AI concepts into classroom instruction, although efforts are beginning to bridge this gap (Gennari et al. 2023; Hinojosa et al. 2025; Katuka et al. 2023; Wang and Lester 2023).

Middle school is a particularly critical juncture for fostering STEM engagement and shaping students' long-term interest in computing (Faber et al. 2013). Research suggests that playful learning approaches are highly effective in this age group, supporting collaboration, iterative exploration, and sustained engagement (McLaren and Nguyen 2023; Wouters et al. 2013). Researchers have begun exploring the use of playful, game-based experiences to support AI education for K–12 students, particularly in areas such as search, planning, and machine learning (Gennari et al. 2023; Gupta et al. 2024; Hinojosa et al. 2025; Wang et al. 2022). However, far less attention has been devoted to the Big Idea of *Natural Interaction* from the AI4K12 framework, specifically to how AI systems understand and use human language (Dhama et al. 2023; Katuka et al. 2024). Given the growing prevalence of chatbots, voice assistants, and large language models (LLMs), introducing middle school students to the foundational concepts and limitations of natural language processing (NLP) is a vital step in fostering AI literacy.

This paper presents three NLP-focused learning activities designed to promote AI literacy in middle school students through playful, hands-on engagement. The activities introduce learners to key NLP concepts, including embeddings and similarity, sentence generation, and chatbot dialogue design. Each activity was iteratively refined through implementation in multiple educational contexts, including teacher professional development workshops, summer camps, and classroom settings. We describe the design and implementation of these activities, present learning outcomes and usability data, and highlight key design principles for integrating AI resources into middle school classrooms. This paper builds on a preliminary presentation of the work (Goslen et al. 2025) by providing expanded activity descriptions, new classroom implementation data, and substantially extended analysis. Together, these efforts provide instructional resources for the underexplored *Natural Interaction* Big Idea and contribute to the growing ecosystem of open, classroom-ready resources for teaching AI concepts in K–12 education.

Related Work

Game-based learning has been studied for many years, with computer games and simulations used for educational purposes as early as the 1950s (Wolfe and Crookall 1998). Contemporary research has focused on refining core principles and advancing theoretical frameworks to deepen our understanding of how games support learning across diverse educational domains (Adipat et al. 2021; Krath, Schürmann, and Von Korflesch 2021; Plass, Homer, and Kinzer 2015). Increasingly, studies have highlighted the potential of playful, game-based learning approaches to enhance subject-specific learning, including computer science and AI (Amzalag, Kadusi, and Peretz 2024; Min et al. 2020; Asbell-Clarke et al. 2021; Zhan et al. 2024).

In computer science education, game-based learning has been shown to foster engagement, problem-solving, and learning. Digital games provide interactive environments that motivate students and make computing concepts more approachable (Min et al. 2020; Denner et al. 2012). Many games require players to decompose complex problems, recognize patterns, and develop logical solutions, reinforcing core computer science skills (Grover and Pea 2013). Additionally, game-based learning supports iterative problem-solving, as students test hypotheses, receive immediate feedback, and refine their solutions, mirroring the debugging process central to programming (Shute, Sun, and Asbell-Clarke 2017). Multiplayer and cooperative games further promote collaboration, communication, and shared problem-solving, supporting teamwork skill development alongside computational thinking (Turchi, Fogli, and Malizia 2019). Finally, well-designed games also scaffold learning by gradually increasing complexity, ensuring that students build computing skills progressively without becoming overwhelmed (Cai et al. 2022). Together, this body of work highlights the potential of game-based learning as an effective strategy for developing computer science competencies, informing the design of our NLP-focused activities for AI literacy in middle school.

Researchers have also explored the benefits of playful, unplugged activities in teaching computing concepts, emphasizing their role in developing broader problem-solving skills without the need for a computer (Caeli and Yadav 2020; Pan et al. 2024). Caeli et al. underscored the importance of integrating unplugged approaches with digital ones to foster a deeper understanding of computational thinking (2020). Similarly, Brackmann et al. found that unplugged activities significantly enhance computational thinking skills (2017), with their effectiveness further amplified when combined with digital approaches. Additionally, research by del Olmo-Muñoz et al. suggests that unplugged learning not only supports computational thinking but also enhances student motivation, adding an element of engagement and enjoyment to the learning process (2020). More recently, Lim et al. introduced an unplugged board game centered on facial recognition to teach AI concepts, demonstrating the potential of unplugged formats for advancing AI literacy (2024).

As efforts to integrate AI literacy into K–12 classrooms expand, researchers have explored ways to extend the benefits of playful experiences and game-based learning to AI education. Touretzky et al. found that middle school students benefited from design-based lessons centered on hands-on activities that can be completed at their own pace (2022). Huber et al. offered a perspective on playful approaches for balancing the opportunities and challenges posed by the presence of LLMs in the classroom, further suggesting that LLMs themselves could be leveraged to create game-based activities that encourage cooperation between students and machines (2024).

In alignment with these findings, researchers have investigated how playful, game-based experiences can support AI literacy by fostering problem-solving skills, critical thinking, and engagement. Alam found that digital games offer a novel framework for classroom teaching of machine learning, artificial intelligence, and computational thinking in K–12 settings (2022). Henry et al. introduced an unplugged game-based approach to teaching ethical AI, in which students role-played as an AI, a developer, or a tester, promoting discussions on fairness, transparency, and bias (2021). Druga et al. examined how robots and conversational agents shaped young learners' conceptual understanding of human–AI interaction, demonstrating that playful engagement can deepen comprehension of how AI systems interpret and respond to human input (2017). Additionally, Wang et al. developed a game-based learning environment called ARIN-561 that teaches search algorithms through an engaging narrative (2022). Collectively, these studies highlight the potential of game-based learning as a powerful tool for AI education. By incorporating interactive, problem-solving, and design-based learning experiences, educators can enhance students' understanding of AI concepts while fostering engagement and critical thinking.

Playful NLP Activities

In response to the need for NLP-focused resources for middle school learners, we developed three playful, structured learning activities that introduce foundational NLP concepts: *SimilarityShuffle*, *NextWords*, and *DialogueCrafter*. Designed to be modular, accessible, and adaptable, the activities can be implemented in both formal classroom settings and informal learning environments. Each activity aligns with the AI4K12 *Natural Interaction* Big Idea, which emphasizes how intelligent agents interpret and use human language. The activities include two unplugged (no technology required) and one web-based digital game, providing multiple entry points for diverse learners and educational contexts. We next describe each of the activities in turn and then provide details about their target age group and learning contexts, setup and required resources, AI concepts addressed, and expected learning outcomes.

SimilarityShuffle

The first activity, *SimilarityShuffle*, is an unplugged card game designed to introduce students to the concepts of word embeddings and similarity scores, which are fundamental ideas in NLP where words are represented as vectors in a high-dimensional space that capture semantic and syntactic relationships. Instead of using actual word embeddings, the activity employs a standard French-suited deck of 52 playing cards, consisting of 13 ranks and 4 suits (diamonds, clubs, hearts, and spades). Students are also provided a *card embedding space* for the activity that organizes the cards on a 2D grid, allowing them to reason about the "similarity" of cards using spatial distance (Figure 1). Each axis of the grid corresponds to a distinct feature dimension (suit or rank), providing an analogy to how word embedding dimensions encode latent semantic properties. This structure reinforces the idea that distances in an embedding space reflect meaningful relationships.

Figure 1: Card embedding space for assessing similarity.

The *card embedding space* is arranged by suit and rank, where all cards of the same suit form a row and all of the same rank form a column. This physical layout mimics how word embeddings group similar words close together in a vector space. The arrangement allows students to visually assess how "close" (i.e., similar) different cards are to one another and determine distances between them. For example, in this space, the Jack of Diamonds is three steps away from the Jack of Clubs based on Manhattan distance. This also provides an opportunity to introduce other distance metrics such as Euclidean distance, when appropriate.

To play the game, students work in small groups of 4 to 8. At the start of the game, one student acts as the dealer, distributing five cards to each player and placing a single card face-up in the center of the group. Students take turns playing a card from their hand that they believe to be most similar to the center card, based on the *card embedding space*. They justify their reasoning aloud, facilitating both collaborative discussion and conceptual understanding. For example, if the center card is the Seven of Hearts and the students play the Seven of Spades, Two of Diamonds, Four of Clubs, and Five of Hearts, the student who played the Seven of Spades wins because it is one step away from the Seven of Hearts, while the other cards are at least two steps away. The game continues with the dealer placing another card face-up in the center until one player wins five rounds.

The process of counting steps using the *card embedding space* is designed to mimic how similarity scores are calculated in NLP models, reinforcing the concept of distance-based similarity in AI. This activity serves as a concrete entry point into abstract NLP concepts, making it especially effective for learners with no prior exposure to AI. It supports multiple rounds of play and requires only a standard deck of cards and a printable grid. It also scales easily for larger classes or camps and requires no technology infrastructure.

NextWords

The second activity, *NextWords*, is a collaborative unplugged activity that introduces students to the concept of probabilistic language generation. Inspired by how language models generate the "most likely" next word in a sentence based on training data, students participate in a structured sentence-generation challenge using a reference paragraph from a familiar story (Figure 2).

NextWords: Sentence Generation

While Sandy is hibernating for the winter, SpongeBob and Patrick are fascinated by the snow they see inside the Treedome and want to play in it. When they try to sneak inside, Sandy warns them in a pre-taped video not to disturb a mammal when she's hibernating. SpongeBob and Patrick, of course, ignore the warning, play in the snow, wake up Sandy and then have to contend with the sleepwalking and sleeptalking savage beast she's become. When they finally get her back to sleep, they realize the lock to the door is frozen shut and they will have to spend the winter there.

Figure 2: Example paragraph for sentence generation.

Working in small groups of 4 to 8, students are provided with a paragraph of text summarizing the plot of an episode from a popular television series, e.g., *SpongeBob SquarePants* (Springer, Greenblatt, and Williams 2001) and are then tasked with generating a new sentence that continues the story. They do so word by word, taking turns selecting the most likely next word based on the context of the given paragraph. For example, if the reference paragraph contains "SpongeBob and Patrick, of course, ignore the warning," and a student begins the next sentence with the word "SpongeBob," the next most probable word would be "and." This process continues until the group constructs a complete sentence. In the first round, students are restricted to only using words from the reference paragraph itself, mimicking a constrained model trained on a limited corpus. In a second round, this restriction is lifted.

The activity fosters an intuitive understanding of how training data influences language model behavior and introduces the concept of next-word prediction. It also opens up discussion about data bias, context dependence, and model limitations. Because it supports multiple rounds of play and requires only a reference story paragraph, the activity is accessible and easy to implement. By encouraging collaborative play and creativity, it helps students intuitively grasp how language models predict the next word based on likelihood (mirroring probabilistic choice) while critically reflecting on how such models function and their limitations.

DialogueCrafter

The final activity, *DialogueCrafter*, is a web-based educational game in which students explore principles of NLP and chatbot design through the refinement of interactive dialogue for non-player characters (NPCs). The primary focus of the activity is to guide students through the process of using, modifying, and creating dialogue for NPCs, helping them understand how conversational agents are structured and refined. To situate the activity in an authentic context, the game is set in a virtual farming community (Figure 3), reflecting the importance of agriculture in the region where many of our target learners live. Through interactions with multiple NPCs, students explore how AI-driven technologies are being applied to support agricultural practices.

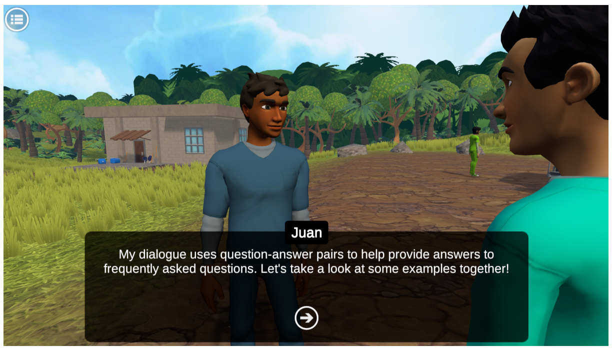

Figure 3: NPC interaction in *DialogueCrafter*.

The activity is scaffolded using a Use-Modify-Create framework (Lee et al. 2011; Lytle et al. 2019). In the *Use* phase, students interact with an NPC powered by a simple retrieval-based chatbot model. When a student types a question, the system compares it to the questions in the NPC's knowledge base using cosine distance between sentence embeddings generated by a pretrained lightweight model. It then returns the answer associated with the closest matching question. This helps students observe how similarity scores drive response selection. In the *Modify* phase, students are shown example question-answer pairs and invited to revise the responses. They can then re-ask questions to observe how their changes alter the dialogue. Finally, in the *Create* phase, students are given full control to design their own questions and answers, effectively building a custom chatbot. As they test their NPC's dialogue, they gain insight into both the power and limitations of AI systems, especially when faced with ambiguous or unexpected queries. This progression supports a deeper understanding of how chatbots could be developed using an NLP-based similarity metric, while the agricultural setting provides a meaningful application of AI in a domain relevant to students' lives.

A unique feature of this activity is the inclusion of an intentionally limited NPC that always gives the same response regardless of input. Students are asked to reflect on their experience with this rigid chatbot, prompting reflection on how fixed-response systems contrast with adaptive ones, and helping students articulate what makes dialogue feel natural and coherent. The game runs in web browsers on laptops and Chromebooks, making it accessible and easy for educators to adopt in classroom and camp settings.

Target Age Group and Learning Contexts

The learning activities are specifically designed for middle school students, typically ranging from ages 11 to 14 (grades 6 through 8). This period represents a key developmental stage when students begin forming lasting interests in STEM and develop more abstract reasoning skills, making it an opportune time to introduce foundational AI concepts (Faber et al. 2013). Importantly, middle school is also a time when students are frequently engaging with AI-powered technologies in their everyday lives, whether through smart speakers or voice assistants, yet often without understanding how these systems work or what their limitations are.

Our implementations of the activities prioritized broad reach by focusing on underrepresented and rural communities. Across our implementations held in the Southeastern United States, participants included racially and ethnically diverse students, many of whom had limited prior exposure to AI. For example, in a 2023 summer camp, 90% of students identified as Black and 10% as Hispanic or Latino. The activities were deliberately designed with low barriers to entry, emphasizing connections to students' lived experiences, such as using farming scenarios in the digital game and familiar media references in unplugged activities.

Educator engagement has also been central to our iterative design process. A PD workshop involving STEM teachers and media specialists provided feedback that directly informed refinements, particularly the value of situating AI content within contexts relevant to rural communities. This participatory design approach (Schuler and Namioka 1993) ensured that the resources were not only developmentally appropriate, but also resonant with learners' everyday realities and relevant to classroom teaching. Teachers noted strong potential for alignment with English Language Arts (ELA) and technology standards. For example, *NextWords* reinforces narrative structure and vocabulary development, while *DialogueCrafter* introduces revision practices. Future iterations will provide explicit mappings to standards to further ease classroom integration.

Setup and Resources Required

The three learning activities are designed to be flexible and easy to implement in both formal and informal learning environments, and can be customized to align with local curriculum standards. Educators do not need specialized technical backgrounds or programming expertise to facilitate them. Each activity includes a facilitation guide, setup instructions, discussion prompts, and reflection questions to support instructional use.

The required materials are minimal and familiar to both students and teachers (Table 1): standard decks of cards and a printable grid for *SimilarityShuffle*; printed story passages for *NextWords*; and access to a web browser for *DialogueCrafter*. All activities are accompanied by a shared set of introductory slides that educators can use to establish foundational knowledge about AI and NLP, connect the concepts to students' everyday experiences, and spark curiosity through open-ended questions.

These activities are also designed with time flexibility in mind. Each can be implemented as a stand-alone lesson in 30

Activity	NLP Concepts	Format	Description	Materials
SimilarityShuffle	Word embeddings, similarity metrics	Unplugged	Students use decks of playing cards and a *card embedding space* to reason about similarity based on spatial proximity, modeling embedding concepts.	Standard 52-card deck; printable 2D grid for *card embedding space*; optional worksheets for similarity estimation and justification.
NextWords	Language modeling, next-word prediction	Unplugged	Students collaboratively generate sentences word by word using constrained and unconstrained vocabularies, simulating how language models generate text.	Printed story passage (e.g., TV episode summary); paper and writing utensils; optional timer.
DialogueCrafter	Chatbot interaction, sentence embeddings, similarity scoring	Digital	Students interact with, modify, and create chatbot dialogues in a virtual farm setting to explore response generation and NLP system limitations.	Laptops or desktops with internet and modern web browser; optional headphones; browser-based access to game (no installation required).

Table 1: Summary of NLP activities: Concepts, formats, descriptions, and materials.

to 45 minutes, or combined into a longer instructional module. Optional extensions include ethical discussions, writing prompts, or exploratory design projects that deepen engagement with the underlying NLP concepts. All resources, including facilitation guides, printable materials, and the digital game, are freely available. They are distributed under a Creative Commons Attribution license (CC BY 4.0) to support widespread adoption and adaptation.

AI Concepts Addressed

Each activity targets core concepts in natural language processing (NLP) that are appropriate for middle school learners. All three activities were designed in alignment with AI4K12's *Natural Interaction* Big Idea, which emphasizes how intelligent agents engage with humans naturally, often through human language.

- *SimilarityShuffle* introduces embeddings by modeling how computers assess similarity. The card layout provides an intuitive analogy to word embedding spaces, reinforcing ideas like vector distance and similarity metrics (e.g., Manhattan or Euclidean distance).
- *NextWords* focuses on probabilistic language generation, illustrating how models predict likely next words based on prior context. It highlights concepts such as next-word prediction, training data limitations, and the role of probability in language modeling.
- *DialogueCrafter* presents chatbot design using a similarity-based question-answering model. Students explore how chatbots represent language and retrieve answers, evaluate similarity scores, and reflect on limitations in NLP-driven interactions. It supports iterative learning through a Use-Modify-Create progression.

Together, these activities help students conceptualize AI not just as an abstract algorithmic process but as a tool shaped by data, design decisions, and user interaction. All three activities include prompts for students to consider bias and fairness in language-based AI. For example, students reflect on how limited vocabulary in *NextWords* mirrors bias in training data, and the *DialogueCrafter* activity includes a "limited" chatbot to encourage reflection about how AI systems might fail to understand diverse or unexpected inputs.

Expected Learning Outcomes

Through participation in these activities, students are expected to develop both conceptual knowledge of NLP and practical skills for engaging with AI systems. The activities are designed to help learners build an intuitive understanding of how NLP models represent, process, and generate language, while also encouraging critical reflection on the strengths and limitations of these systems. In addition, students practice collaboration and creativity as they explore, modify, and design AI-driven content. Specifically, students are expected to:

- Explain how AI systems process and interpret human language.
- Describe how similarity can be computed using word embeddings as vector representations.
- Simulate the behavior of language models and identify factors that influence output.
- Demonstrate how chatbots can use similarity metrics to retrieve and generate responses.
- Reflect on the limitations of NLP systems and suggest ways to improve them.

- Collaborate with peers to explore, modify, and create AI-driven content.

The activities can also encourage metacognitive awareness as students reflect on their own interactions with NLP technologies in daily life, such as voice assistants or predictive text features.

Implementations

We implemented the activities across multiple contexts, including two summer camps (2023 and 2024), a teacher professional development workshop (2023), and a middle school classroom (2025). These settings provided opportunities to observe engagement, gather feedback, and conduct formative assessments and surveys. Before students participated in the implementations, parents were provided informed consent forms, and students completed assent forms, as approved by the university's ethics review board. Teachers also completed consent forms, prior to their involvement. Insights from these implementations directly informed iterative refinements to the activities, ensuring that they remained accessible, engaging, and developmentally appropriate.

Summer Camp 2023

In summer 2023, we implemented *NextWords* and *DialogueCrafter* as part of an AI-focused summer camp (Figure 4). The 10 student participants were made up of four girls and six boys, with 90% identifying as Black and 10% as Hispanic or Latino. Students responded positively to *NextWords*; however, confusion about similarity scores in *DialogueCrafter* prompted us to explore creating *SimilarityShuffle* as a new entry-level activity to introduce those concepts more explicitly. In the exit surveys for the activities, students demonstrated basic understanding of NLP-enabled technologies but struggled to articulate limitations of NLP. This prompted revisions to some of the introductory materials as well as *DialogueCrafter* to incorporate a fourth NPC that has the same response for every inquiry with a structured reflection afterwards. This feature was designed to emphasize limitations of NLP technologies, particularly in situations where systems fail to respond appropriately.

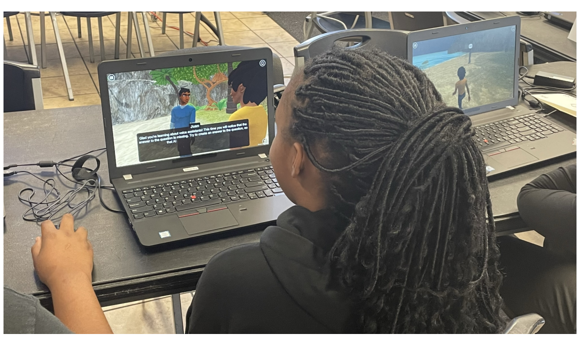

Figure 4: Student interacting with *DialogueCrafter*.

Professional Development Workshop 2023

In fall 2023, we conducted a professional development (PD) workshop with five middle school educators, including STEM teachers and media specialists (Figure 5). As part of the workshop, we reviewed introductory materials with the teachers and allowed them to play *DialogueCrafter* for approximately 30 minutes. Throughout the session, we prompted the teachers to provide feedback as they played the game, which was recorded through video and note-taking.

Figure 5: Educators participating in PD workshop.

The teachers emphasized the importance of aligning the activity's content with topics more familiar to their students, since the version they used focused on AI application areas that were less relevant to their local context. Because the participating rural student populations live in regions where agriculture plays a central role, teachers suggested that focusing the activity on AI technologies used in farming might improve student interest and focus. Based on this feedback, we refined *DialogueCrafter* so that each NPC highlights AI applications in agriculture. Teachers also highlighted the potential to align the activities with existing ELA and technology standards during the workshop.

Summer Camp 2024

In summer 2024, we held another AI-focused summer camp, where 23 middle school students participated and interacted with all three activities. Of these students, the following racial demographics were reported: 26% Black, 26% White, 4% American Indian, 31% South Asian or Indian, 4% Middle Eastern, 8% Multiracial. Students played each activity for around 30 minutes. Researchers observed the students as they engaged with the activities. Post-surveys showed that 83% correctly identified the definition of a similarity score and 79% articulated examples of challenges NLP systems face. Students frequently cited *SimilarityShuffle* and *NextWords* as their favorite activities. Observations and survey data indicated improved understanding of NLP concepts, increased engagement, and enthusiasm for AI topics.

Classroom Implementation 2025

In spring 2025, to examine the feasibility of deploying *DialogueCrafter* in formal educational settings, we conducted a classroom implementation with 40 middle school students. Project members introduced key ideas in NLP through a brief presentation, after which students played the game for approximately 20 minutes. The session concluded with a post-survey, including 15 Likert items focused on usability and perceptions of learning, and two open-ended questions.

Frustration was common during gameplay (42% of students agreed they felt frustrated while playing the game) but notably, only 27% reported frustration at the end of the game. This decline indicates that many students overcame early challenges and developed comfort with the game's mechanics over time. Regarding fairness, 38% of students strongly disagreed that the game was unfair, while only 13% agreed, indicating that students generally viewed the game as equitable.

Responses about rule clarity were mixed. While 43% of students reported understanding the rules, approximately one-third disagreed or were neutral. Similarly, 30% found the game's scenario interesting, while 40% were neutral. These findings highlight opportunities to strengthen the onboarding experience and enhance the relevance of the game's narrative context.

Open-ended responses revealed that several students were able to articulate key concepts related to NLP. The most common themes included understanding how AI uses similarity scores to match user input with predefined responses, and recognizing that AI requires ongoing development and human support. For example, one student explained, "AI gets certain answers, sees if the question is close or similar, and then gives an answer," while another noted, "AI will need some assistance throughout its development and won't always get everything right." Others emphasized that "there is no exact [answer]" in NLP and highlighted the "cool" potential of building AI into games. However, a notable portion of responses included either non-substantive answers, such as "idk," or expressions of confusion and disengagement.

When asked what questions they still had, many students responded with "none" or "nothing," which may indicate either a sense of closure or limited curiosity. A smaller subset posed thoughtful questions about AI learning and autonomy, such as "How does AI learn from its mistakes?" and "At what point will AI begin to think for itself?" A few students provided critical feedback about the game itself, pointing to opportunities for improvement in engagement and polish.

From a technical standpoint, the classroom implementation demonstrated that the game ran smoothly within the school's existing infrastructure. Students were able to access and play *DialogueCrafter* on a school computer without any installation or compatibility issues. No network connectivity problems were reported, and the browser-based format ensured that the activity was accessible across devices. These findings indicate that the game can be reliably deployed in typical classroom settings without placing additional demands on school technology resources.

Taken together, these results suggest that while the *DialogueCrafter* provided a useful entry point for many students to engage with core NLP concepts, there is variability in how students internalize and articulate what they learn. The game succeeded in conveying the mechanics of NLP to some students, particularly around similarity scoring and AI reasoning, and the smooth technical deployment demonstrated that it can be reliably integrated into classroom environments. However, the small sample size, limited exposure time, and the absence of concrete learning measures restrict our ability to fully assess its impact. Follow-up interviews and planned pre/post measures will explore how students' conceptual understanding of embeddings and prediction evolves when supported with more explicit reflection prompts. At the same time, more structured reflection and differentiated supports may be needed to deepen understanding and foster broader engagement.

Conclusion and Future Work

In this work, we introduced three NLP-focused instructional activities designed to foster foundational AI literacy in middle school learners. Developed through iterative design and implementation in informal and formal contexts, these activities leverage both unplugged and digital formats to introduce students to foundational NLP concepts that serve as conceptual anchors for later engagement with more advanced AI topics. Emphasizing accessibility, contextual relevance, and playful learning, the resources align with the AI4K12 *Natural Interaction* Big Idea and are designed to be well-suited for underserved student populations. By offering low-barrier, classroom-ready activities, the resources provide educators with practical tools for introducing AI concepts without requiring specialized technical expertise.

Results from implementations suggest that students found the activities engaging and educational, although initial confusion and frustration highlight the need for additional scaffolding and clearer instructions. Early indications showed notable conceptual growth by students, particularly in understanding similarity scores and the challenges NLP systems face. Open-ended survey responses reflected a range of engagement, from specific insights about NLP mechanisms to thoughtful questions about AI learning and sentience, while also revealing the need for more structured support to guide reflection and deepen understanding for all learners. These findings point to the value of embedding AI education in experiential and reflective learning formats, especially when aligned with learners' lived experiences.

Future work will expand classroom-based implementations to assess how these resources align with formal curriculum goals and constraints. We also aim to explore integrating large language models into the *DialogueCrafter* game to offer students a more authentic interaction with modern NLP tools while simultaneously introducing critical perspectives on bias, fairness, and transparency. Additional development will focus on enhancing narrative engagement, expanding ethical discussion prompts, and refining onboarding materials to reduce cognitive load for first-time users. By continuing to refine these resources in collaboration with educators, we aim to support a growing community of practice around K–12 AI education.

Acknowledgements

This research was supported by the National Science Foundation through Grant DRL-2148680. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Adipat, S.; Laksana, K.; Busayanon, K.; Asawasowan, A.; and Adipat, B. 2021. Engaging students in the learning process with game-based learning: The fundamental concepts. *International Journal of Technology in Education*, 4(3): 542–552.

Alam, A. 2022. A digital game based learning approach for effective curriculum transaction for teaching-learning of artificial intelligence and machine learning. In *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 69–74. IEEE.

Amzalag, M.; Kadusi, D.; and Peretz, S. 2024. Enhancing academic achievement and engagement through digital game-based learning: An empirical study on middle school students. *Journal of Educational Computing Research*, 62(5): 989–1013.

Asbell-Clarke, J.; Rowe, E.; Almeda, V.; Edwards, T.; Bardar, E.; Gasca, S.; Baker, R. S.; and Scruggs, R. 2021. The development of students' computational thinking practices in elementary- and middle-school classes using the learning game, Zoombinis. *Computers in Human Behavior*, 115: 106587.

Brackmann, C. P.; Román-González, M.; Robles, G.; Moreno-León, J.; Casali, A.; and Barone, D. 2017. Development of computational thinking skills through unplugged activities in primary school. In *Proceedings of the 12th Workshop on Primary and Secondary Computing Education*, 65–72.

Caeli, E. N.; and Yadav, A. 2020. Unplugged approaches to computational thinking: A historical perspective. *TechTrends*, 64(1): 29–36.

Cai, Z.; Mao, P.; Wang, D.; He, J.; Chen, X.; and Fan, X. 2022. Effects of scaffolding in digital game-based learning on student's achievement: A three-level meta-analysis. *Educational Psychology Review*, 34(2): 537–574.

del Olmo-Muñoz, J.; Cózar-Gutiérrez, R.; and González-Calero, J. A. 2020. Computational thinking through unplugged activities in early years of primary education. *Computers & Education*, 150: 103832.

Denner, J.; Werner, L.; Campe, S.; and Ortiz, E. 2012. Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts? *Computers & Education*, 58(1): 240–249.

Dhama, S.; Katuka, G.; Celepkolu, M.; Boyer, K. E.; Glazewski, K.; and Hmelo-Silver, C. 2023. NLP4Science: Designing a platform for integrating natural language processing in middle school science classrooms. In *Proceedings of the 2023 IEEE Symposium on Visual Languages and Human-Centric Computing*, 269–273. IEEE.

Druga, S.; Williams, R.; Breazeal, C.; and Resnick, M. 2017. "Hey Google is it ok if I eat you?" Initial explorations in child-agent interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children*, 595–600.

Faber, M.; Unfried, A.; Wiebe, E. N.; Corn, J.; Townsend, L. W.; and Collins, T. L. 2013. Student attitudes toward STEM: The development of upper elementary school and middle/high school student surveys. In *2013 ASEE Annual Conference & Exposition*, 23–1094.

Floridi, L.; Cowls, J.; Beltrametti, M.; Chatila, R.; Chazerand, P.; Dignum, V.; Luetge, C.; Madelin, R.; Pagallo, U.; Rossi, F.; et al. 2018. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4): 689–707.

Gennari, R.; Melonio, A.; Pellegrino, M. A.; and D'Angelo, M. 2023. How to playfully teach AI to young learners: A systematic literature review. In *Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter*, 1–9.

Goslen, A.; Aguiar, C.; Vandenberg, J.; Cateté, V.; Min, W.; and Mott, B. 2025. Playful pathways to AI literacy: Designing NLP activities for middle school AI education. In *International Conference on Artificial Intelligence in Education*, 60–67. Springer.

Grover, S.; and Pea, R. 2013. Computational thinking in K-12: A review of the state of the field. *Educational Researcher*, 42(1): 38–43.

Gupta, A.; Lee, S.; Mott, B.; Chakraburty, S.; Glazewski, K.; Ottenbreit-Leftwich, A.; Scribner, A.; Hmelo-Silver, C. E.; and Lester, J. 2024. Supporting upper elementary students in learning AI concepts with story-driven game-based learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 23092–23100.

Henry, J.; Hernalesteen, A.; and Collard, A.-S. 2021. Teaching artificial intelligence to K-12 through a role-playing game questioning the intelligence concept. *KI-Künstliche Intelligenz*, 35(2): 171–179.

Hinojosa, C.; Kumar, P.; Rajarajan, P. D.; and Martin, F. 2025. TrainYourSnakeAI: A novel tool to teach reinforcement learning to middle school students. In *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1*, 506–512.

Huber, S. E.; Kiili, K.; Nebel, S.; Ryan, R. M.; Sailer, M.; and Ninaus, M. 2024. Leveraging the potential of large language models in education through playful and game-based learning. *Educational Psychology Review*, 36(1): 25.

Katuka, G. A.; Auguste, Y.; Song, Y.; Tian, X.; Kumar, A.; Celepkolu, M.; Boyer, K. E.; Barrett, J.; Israel, M.; and McKlin, T. 2023. A summer camp experience to engage middle school learners in AI through conversational app development. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, 813–819.

Katuka, G. A.; Chakraburty, S.; Lee, H.; Dhama, S.; Earle-Randell, T.; Celepkolu, M.; Boyer, K. E.; Glazewski, K.; Hmelo-Silver, C.; and Mcklin, T. 2024. Integrating natural language processing in middle school science classrooms: An experience report. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, 639–645.

Krath, J.; Schürmann, L.; and Von Korflesch, H. F. 2021. Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, 125: 106963.

Lee, I.; Martin, F.; Denner, J.; Coulter, B.; Allan, W.; Erickson, J.; Malyn-Smith, J.; and Werner, L. 2011. Computational thinking for youth in practice. *ACM Inroads*, 2(1): 32–37.

Lim, H.; Min, W.; Vandenberg, J.; Cateté, V.; and Mott, B. 2024. Unplugged K-12 AI learning: Exploring representation and reasoning with a facial recognition game. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 23285–23293.

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.

Lytle, N.; Cateté, V.; Boulden, D.; Dong, Y.; Houchins, J.; Milliken, A.; Isvik, A.; Bounajim, D.; Wiebe, E.; and Barnes, T. 2019. Use, Modify, Create: Comparing computational thinking lesson progressions for STEM classes. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, 395–401.

McLaren, B. M.; and Nguyen, H. A. 2023. Digital learning games in Artificial Intelligence in Education (AIED): A review. *Handbook of Artificial Intelligence in Education*, 440–484.

Min, W.; Mott, B.; Park, K.; Taylor, S.; Akram, B.; Wiebe, E.; Boyer, K. E.; and Lester, J. 2020. Promoting computer science learning with block-based programming and narrative-centered gameplay. In *2020 IEEE Conference on Games*, 654–657. IEEE.

Pan, Y.; Adams, E. L.; Ketterlin-Geller, L. R.; Larson, E. C.; and Clark, C. 2024. Enhancing middle school students' computational thinking competency through game-based learning. *Educational Technology Research and Development*, 72(6): 3391–3419.

Plass, J. L.; Homer, B. D.; and Kinzer, C. K. 2015. Foundations of game-based learning. *Educational Psychologist*, 50(4): 258–283.

Schuler, D.; and Namioka, A. 1993. *Participatory design: Principles and practices*. CRC press.

Shute, V. J.; Sun, C.; and Asbell-Clarke, J. 2017. Demystifying computational thinking. *Educational Research Review*, 22: 142–158.

Springer, A.; Greenblatt, C.; and Williams, M. 2001. SpongeBob SquarePants. Survival of the idiots. Television series episode. Created by Stephen Hillenburg.

Touretzky, D.; Gardner-McCune, C.; Cox, B.; Uchidiuno, J.; Kolodner, J.; and Stapleton, P. 2022. Lessons learned from teaching artificial intelligence to middle school students. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 2*, 1371–1371.

Touretzky, D.; Gardner-McCune, C.; Martin, F.; and Seehorn, D. 2019. 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.

Turchi, T.; Fogli, D.; and Malizia, A. 2019. Fostering computational thinking through collaborative game-based learning. *Multimedia Tools and Applications*, 78.

Wang, N.; Greenwald, E.; Montgomery, R.; and Leitner, M. 2022. ARIN-561: An educational game for learning artificial intelligence for high-school students. In *Proceedings of the International Conference on Artificial Intelligence in Education*, 528–531. Springer.

Wang, N.; and Lester, J. 2023. K-12 Education in the Age of AI: A Call to Action for K-12 AI Literacy. *International Journal of Artificial Intelligence in Education*, 33(2): 228–232.

Wolfe, J.; and Crookall, D. 1998. Developing a scientific knowledge of simulation/gaming. *Simulation & Gaming*, 29(1): 7–19.

Wouters, P.; Van Nimwegen, C.; Van Oostendorp, H.; and Van Der Spek, E. D. 2013. A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2): 249.

Zhan, Z.; Tong, Y.; Lan, X.; and Zhong, B. 2024. A systematic literature review of game-based learning in Artificial Intelligence education. *Interactive Learning Environments*, 32(3): 1137–1158.