International Conference
The Future of Education

Dewey Meets the Machine: Guided Inquiry Using Generative AI for Students Ages 7-11

Timothy Mattison[1], Jill M. Raisor[2]

University of Southern Indiana, United States[1]
University of Southern Indiana, United States[2]

Abstract

When Open AI made ChatGPT publicly available in 2022, teachers scrambled in the absence of theory-driven research to figure out how to use it and subsequent generative artificial intelligence (GenAI) programs to drive their students' learning. After almost three years and several studies on the use of GenAI in education, none have yet systematically engaged the philosophical and theoretical foundations of learning. This manuscript helps to fill that gap. Engagement with philosophies and theories on learning reveals how GenAI can be used, especially in primary education, to help children develop habits that facilitate critical thinking and prevent overreliance on technology. The researchers used the qualitative content analysis approach to explore the specific capabilities of GenAI use in education, and how the technology is currently being used in early childhood and primary education. They then present the foundational philosophy of John Dewey, focusing on his description of how children engage in inquiry to solve problems and how GenAI could enhance the process. The manuscript then provides a framework for teacher modeling, guidance, and supervision of students' use of GenAI during the inquiry process. It concludes with concrete examples of students' use of GenAI in the inquiry process, in addition to the limitations of GenAI and the safeguards required to protect students as they use the technology. The ultimate goal of presenting this framework to other researchers is to engage them in working hypotheses they can test through empirical studies.

Keywords: *Primary Education, Artificial Intelligence (AI), Generative Artificial Intelligence (GenAI)*

Author Note
Timothy Mattison https://orcid.org/0000-0002-5292-0760
Jill Raisor https://orcid.org/0000-0002-3167-3289
Correspondence concerning this manuscript proposal should be addressed to Timothy Mattison, tmattison@usi.edu, and Jill Raisor, jmraisor@usi.edu, University of Southern Indiana, Teacher Education Department, 8600 University Blvd, Evansville, IN, 47712.

1. Introduction

In November of 2022, OpenAI released its GPT-3.5 model of ChatGPT. This new program was powered by artificial intelligence (AI) that, unlike online search engines such as Google or Bing, gave users the ability to get direct responses to questions, followed by a sequence of logical follow-up questions. GPT-3.5 also remembered and learned from the user's previous questions and responses, making the program even more attuned to the user's needs the longer the interaction continued. Thus, interacting with GPT-3.5 was like conversing with an attentive human assistant who had committed much of the world's knowledge to memory. Furthermore, if GPT-3.5 detected any false premise or ambiguity in a question, it would provide the user advice for rephrasing and even offer tutorials if that initial advice was confusing. In this sense, GPT-3.5 was also a teacher. Therefore, the program was, in many respects, a sophisticated realization of the science-fiction AI android named Data from the popular 1990s television program Star Trek: The Next Generation. However, what made GPT-3.5 most groundbreaking was its ability to respond to user commands to create new text. For instance, the program could quickly generate a unique email, research report, or any other textual deliverable that, in many cases, appeared to be written by a human being. It was this new generative capability that spawned the designation *generative artificial intelligence* or *GenAI*.

According to Su and Yang (2022) [51], before the release of GPT-3.5, primary teachers knew very little about how to use non-generative artificial intelligence programs like Apple's *Siri* to support the learning of young students. Therefore, these teachers likely knew even less about using a more sophisticated generative program like GPT-3.5 for teaching purposes. To complicate matters, GenAI technology advanced more quickly than primary teachers' adoption of it, making it difficult for teachers to keep up with new and changing functionalities [6]. Consider how quickly OpenAI replaced GPT-3.5 with ChatGPT 4.0 in 2023. The speed of GenAI advancement is not slowing down, as OpenAI plans to release GPT versions 4.5 and 5.0 within one year after this publication. Furthermore, in 2023, companies like Google and Microsoft launched their own GenAI programs Gemini and Copilot, respectively [37]. Google and Microsoft continuously update their GenAI programs and add functionalities [37]. Even more companies are releasing GenAI programs on platforms not explored here. Consequently, the gap in primary teachers' knowledge of how to use GenAI to support their students' learning will only continue to grow.

GenAI programs released after GPT-3.5 offered primary teachers most of the same capabilities. For example, they still allowed teachers to create lesson plans, activities, assessments aligned to state educational standards, and students' grade levels, lexiles, and personal interests. However, programs like Google's Gemini 2.0 Flash also allowed teachers to create unique and fantastical images to indulge children's imagination, such as futuristic versions of cities or even make-believe animals [23]. There are now many other programs such as Invideo, Canva, and DeepAI, which give users the ability to create engaging educational videos with scripts, narrators, images, and background music [20]. IGI Global Scientific Publishing. The primary teacher simply types in a prompt telling the GenAI program what kind of video they want to create and the teacher can download an mp4 file. Thus, teachers no longer need to search a vast library of videos to supplement their lessons; they can instead create their own in mere minutes. Similarly, primary students can create new and creative video content with no cinematography experience.

GenAI programs are more than additional tools in primary teachers' tool belts. They have the potential to radically change the very nature of primary teachers' instructional methods, and students' acquisition of knowledge. Primary teachers could potentially shift some of their time away from lesson or activity planning and grading to acting as their students' learning thought partner and motivator. Meanwhile, primary students could possibly use GenAI chatbots to engage in a constant and personalized process of inquiry involving real-world problems, questions, hypotheses, and testing. For instance, a third grader studying an informational text could use a speech interface with a GenAI program to check whether they have correctly identified the main idea. They could then ask the GenAI program to help them explore why their main idea is right or wrong. Finally, the third grader could use the GenAI explanation to guess again and receive more feedback from the GenAI program.

John Dewey's piece entitled *The School and Society*, which was the first to explore this process of inquiry in education, is foundational to most teacher preparation programs and inspired Jerome Bruner's famed discovery approach to learning [14]. Dewey envisioned inquiry as an ideal process that would condition children to think scientifically to solve real-world problems, continuously drive their curiosity, and ultimately grow intellectually. Therefore, introducing GenAI programs in primary school classrooms puts Dewey's well-established *philosophy on learning through inquiry* face to face in a potential partnership with "the machine." If this partnership were a locomotive, Dewey's process of inquiry would be the engine and GenAI would be the fuel.

The exploration in this manuscript could not be timelier. Primary teachers have broad access to free GenAI programs, but there is relatively little guidance in existing literature on how to use these applications within existing learning theories like Dewey's description of children's process of inquiry. Furthermore, unless parents, teachers, school leaders, and policymakers decide how they envision the use of GenAI in schools, primary students will simply create their own system of GenAI use. This is precisely what happened two decades ago during the proliferation of social media. Many primary students gained access to programs like Facebook, Instagram, and Twitter, and it took many teachers years to realize the need to learn how their students were using them [33]. It took education researchers even longer to study these social media platforms. For example, it was over a decade after Facebook went public before the seminal work entitled *Education and Social Media: Toward a Digital Future* reached library shelves. Meanwhile, young people created an online ecosystem of social interaction that previously occurred only face-to-face [54]. As a result, social media created unmitigated social problems such as anxiety, depression, and even cyberbullying, which have been difficult for primary teachers to manage in the classroom [58]. Unmitigated use of GenAI could lead to similar social problems.

Unmitigated use could also lead to interruptions to the process of learning for primary students. Imagine, for instance, a primary student using a GenAI program to generate images for a class project on the habitats of animals in dense forests. Then imagine that the GenAI program produces several images of animals that do not live in forests. Such errors are not merely theoretical. For instance, the journal Frontiers in Cell and Developmental Biology had to retract an article due to its nonsensical illustrations of mammalian reproductive systems in the manuscript [12]. Without a critical framework like Dewey's process of inquiry, these students might not recognize obvious errors in the images and thus unwittingly spread false information to other students. It is therefore in the interest of all those responsible for primary education to habituate primary students in the positive and productive uses of GenAI, especially in using critical frameworks like Dewey's process of inquiry.

2. Methods

When conducting the analysis for this manuscript, we used qualitative content analysis, as described by White and Marsh (2006) [59]. These authors define content analysis as a systematic and rigorous qualitative or quantitative approach to analyzing resources a researcher identifies as part of the literature search [59]. More specifically, content analysis is a method that enables the researcher to make valid inferences from resources to the contexts where these inferences can be applied [59]. These inferences are part of an analytical process that uses existing theories or resources [59]. According to Krippendorff (2004) [31], as cited by White and Marsh (2006) [59], qualitative content analysis involves the following four components: creating research questions, collecting a sample of texts, distinguishing parts of a text through excerpts and quotes, and contextualizing these excerpts.

For the study in this manuscript, we created the following research question: What do prominent developmental learning theories have to say about the use of GenAI in primary school? To answer this question, we collected a sample of articles on the ways in which researchers propose to use GenAI in education. We used EBSCOhost to search for articles through a variety of search terms. Our initial searches uncovered just a few resources directly related to GenAI capabilities in education. This is likely because GenAI is such a new technology, and research on it is in its early stages. Therefore, we expanded our search to resources on AI capabilities in education. This search unearthed multiple articles, most of which were published from 2021 to 2024.

Next, we extracted excerpts from these multiple texts and used open coding when identifying specific capabilities. These capabilities constituted the themes that emerged from our coding process. After that, we chose the major educational theorists that we present to our students at a 4-year institution of higher education. Finally, we contextualized GenAI capabilities within the themes that emerged from the works of the educational theorists by using the process of logical inference.

3. GenAI Capabilities in Education

3.1 Background

GenAI emerged in education after decades of artificial intelligence (AI) development. Definitions of AI vary, but most characterize the technology by its performance on tasks that mirror human intelligence, such as categorization, inference, error analysis, and problem-solving [7]. AI was meant to use these abilities to improve the efficiency of work tasks in areas such as instruction and learning [7]. The most critical point in the development of AI in education was the creation of the personal computer in the 1970s, which programmers used to design predetermined instructions in AI through a system of coding [7]. This AI development culminated in computer-aided instruction, learning, and, more recently, the internet [7]. Within just a few decades, AI has evolved with the introduction of machine learning into programs such as adaptive standardized tests and sophisticated search engines [34]. Machine learning models analyze patterns of data generated by the user's actions to predict, using numerous coexistent parameters, which the user needs next [34]. These parameters are values akin to coefficients in an equation that enable GenAI models to make their predictions.

GenAI builds upon earlier versions of AI known as large language models, which are part of an umbrella of different types of machine learning [36]. However, the current generative models of artificial intelligence (i.e., GenAI) are far more sophisticated than AI programs of the past, now often incorporating as many as 175 billion parameters [36]. These parameters operate within a system of sophisticated algorithms that generate content using data from all existing textual sources and other input data [25]; [36]. This level of sophistication in GenAI models has empowered programmers to design applications that generate new information and content by leveraging the stochastic properties of the data, rather than merely analyzing the patterns within existing content [36]. Programmers *train* large language models with close to a trillion words that form sequential patterns to ultimately predict subsequent words in a sequence [61]. Predictions from these models generate outputs, such as text, images, and more, which users of GenAI receive as responses to the prompts they enter into programs like ChatGPT.

GenAI currently gives users a variety of ways to interface with it. These interactive modalities include text, image, video, code, sound, and other produced content such as molecules and 3D renderings, which can then be translated from one modality to another [5]. For example, the user can currently translate media as follows: Text to Text; Text to Image; Text to Audio/Speech; Image to Text; Image to Image; Text-to-Video; Text+Video to Video; Video to Video; Image+Text to Image; Text-Driven 3D Content Generation; Text to 3D Image; and Text to 3D Animation [4].

3.2 GenAI Use by Teachers

The scholarly literature on GenAI in education can be divided into two broad categories. The first focuses on teachers' use of GenAI programs to provide high-quality and innovative instruction. Su and Yang (2023) [52] and Kanders et al. (2024) [26] were among the limited sources located that focus on GenAI use specifically in younger students. Furthermore, both researchers only explicitly addressed the use of GenAI by teachers. For instance, Kanders et al. (2024) [26] lists multiple ways in which teachers can use GenAI to organize resources that enrich students' learning such as drafting lesson plans, supporting teacher training, and supporting speech and language therapy. Kanders et al. (2024) [26] also listed ways to use GenAI to bolster creative content for teachers.

Another potential role for GenAI is as the teacher's assistant [8]; [21]. For example, GenAI could help teachers assist students in a variety of inquiry and problem-solving tasks, such as information searches, answering questions, and enhancing writing in a variety of languages [8]; [13]; [52]. It is important to note that by using the phrase "teacher's assistant," those researchers are not suggesting GenAI be used to replace humans in classrooms. The researchers instead suggest that GenAI can assist by extending a teacher's capabilities. GenAI could also help teachers provide students with differentiated feedback, tailored support, and quick answers to their questions [9]; [13]; [21]; [27].

GenAI helps provide feedback and support, which could be especially impactful for special populations such as students with special needs and English learners. For instance, GenAI could help students with speech recognition for the hearing impaired [1]. The technology could additionally summarize text in different writing styles and in more digestible language [9]; [41]. Furthermore, GenAI could provide translations for class resources, such as videos and scaffolding for writing, to support different levels of language proficiency [9]; [41]; [56].

3.3 GenAI Use by Students

The other category of GenAI use in education targets student uses of GenAI programs for guided and independent learning. Many researchers that study the current capabilities of GenAI point to its potential to provide students with independent or assisted personal tutoring, individual feedback to correct errors, dialectic exercises, and practice responding to ethical quandaries with a chatbot [1]; [7]; [8]; [9]; [13]; [21]; [22]; [46]; [52]; [56]; [62]. GenAI can also work through a chatbot to adapt its explanations, teaching methods, and materials to student misconceptions in addition to their interests and learning levels [1]; [9]; [41]; [46]. This adaptation is possible, because GenAI collects and analyzes performance data while it interacts with the student [8]; [13]; [27]; [46].

GenAI can also facilitate the brainstorming and creative writing process and generate new content such as text (e.g. research papers), images, audio, videos, computer code, and 3D models by typing requests [2]; [1]; [8]; [9]; [13]; [26]; [50]; [56]; [62]. GenAI can also help create unique children's books with the child reader as the main character, and voice narration [41]. In addition, the GenAI in a program called *Midjourney* can help guide students through visual and graphic art creation [8]; [11]; [13]; [26]. Other GenAI programs can produce high-quality concept art, music, and animation [11]; [13]; [62]. The GenAI in the program *Dream Studio* by Stability AI can even enable students to construct and visualize their dreams in narrative text and visual images [2].

3.4 Limitations

GenAI offers exciting opportunities to improve student learning in the field of education. However, it is crucial to recognize the technological limitations. Perhaps the most concerning limitation currently is the persistence of what GenAI researchers call *hallucinations*. Hallucinations in GenAI refer to the fabrication and random parroting of false "facts" [35]. GenAI hallucinations stem from two primary, generalizable issues: the complexity of models and algorithms for processing data and the use of incomplete or biased training data [15]. GenAI hallucinations also result from bias in model design, as well as the matching of heuristic data collection, incorrect encoding of input data, and decoding of output data [49]. Studies such as Kim et al. (2025) [28] in Bioengineering, have developed tools for addressing the causes of hallucinations in GenAI. Meanwhile, other studies explore the utility of artificial intelligence literacy, to identify GenAI hallucinations and mitigate their negative effects, which places more responsibility on the user than on fixing the technology [18]; Walter, 2024). Nevertheless, the problem of hallucinations in GenAI remains, so more time and research are required to resolve this issue.

4. GenAI Within Dewey's Model for Inquiry

When GenAI programs enter the primary school classroom, they come face to face with traditional educational practices inspired by more than a century of theory and research on learning and cognitive development. According to Wu (2023) [60], GenAI programs such as ChatGPT do not make traditional learning theories or past research obsolete. GenAI programs simply give teachers more opportunities to engage students in the learning processes promoted by educational theory and research [60]. Therefore, deep consultation with educational theory and research can help teachers make reasoned decisions on how to incorporate GenAI technology into their classrooms.

4.1 Pragmatist Paradigm

Much of past and present theory, research, and practice in education draw on the seminal work by John Dewey *The School and Society*, which was published in 1899. As the founder of pragmatism, Dewey stressed the need for students' use of practical tools and hands-on experiences in the learning process to solve practical problems [39]. Dewey likely never imagined there would be the kinds of tools that exist in GenAI programs today. Nevertheless, given his belief in the importance of using all available tools for students' learning, he may well have supported the use of GenAI programs in primary education. Furthermore, he bucked against classical metaphysical philosophies in a similar way that GenAI challenges the paradigms behind traditional educational practices. For instance, Dewey argued against the standard educational model of his day which viewed children as passive receivers of knowledge. Therefore, he might have seen GenAI programs as helpful for providing children with an inexhaustible treasure trove of new learning experiences. An endless supply of new learning experiences would promote what Dewey believed was the purpose of education–never-ending growth.

4.2 Model for Inquiry

According to Dewey, growth involves the seeking of truth. However, he rejected the concept of metaphysical truth and instead defined truth as a series of *warranted assertions*, i.e., hypotheses. Students make warranted assertions during a process of inquiry. He argued that inquiry is a process of thinking and problem solving, which involves identifying a problem, exploring and generating hypotheses, planning to test hypotheses, creating alternative hypotheses, testing hypotheses, and finally, reflecting on the results.

On its face, inquiry involving hypothesis testing might seem too complex a framework for students in primary education, and far more suitable during what Jean Piaget labeled the formal operational stage, starting at the age of 11, where students begin to engage in abstract thought [44]. However, to understand the suitability of Dewey's model for inquiry for younger age students, we must look to Vygotsky's theory of learning development. Vygotsky argued that the cognition involved in students' problem solving can be categorized as elementary processes and higher psychological functions [57]. The difference between elementary and higher-level processes lies in the complexity and sophistication of children's planning behavior [57]. In other words, children with elementary-level processes can divide the planning to solve a problem into fewer and simpler steps. Evidence of children's elementary processes abounds in empirical studies like Koksal-Tunser and Sodian (2018) [30], which revealed that primary school students often have the ability to make connections between causal hypotheses and evidence. Therefore, for Dewey's model of inquiry involving hypothesis testing to be appropriate for primary education, it simply needs to align to the level of sophistication in students' planning behavior.

4.2.1 Identifying a Problem

The first step of inquiry in Dewey's model is identifying a problem. In mathematics education, a *problem* is an unknown and important piece of information. As evidence, researchers of mathematics education such as Cooper (1986) [10] define problem-solving as the attempt to find an unknown piece of information to achieve one central goal–to answer a question. Therefore, identifying a problem means naming an unknown and important piece of information. In primary mathematics, students look to the question in an assessment item to identify the unknown. Imagine application questions such as, "if Olivia had 10 grapes, and she ate four, how many does she have left?" The unknown here is how many grapes Olivia has left, which is a discrete piece of unknown information.

The definition of a problem in mathematics applies well to other subjects too. Consider the third-grade student whose teacher asks them to determine the meaning of a word based on how it is used in a reading passage, or a fourth-grade student who must write down their thoughts and figure out what linking words are appropriate to string them together. Each of these examples involves an unknown that the student must find to achieve specific goals. However, the potential success of identifying these unknowns depends heavily on the requirements for student analysis. For instance, in the above examples, the third grader would need to first realize that the word in question has multiple meanings and that the unknown is which meaning fits the word in the context of the passage. Likewise, the fourth grader would need to first recognize that linking words can show when thoughts complement or contradict each other, and that the unknown is whether a string of thoughts is complementary or contradictory. This type of analytical thinking is not easy for many young students. Peltier and Vannest (2017) [42] refer to this difficulty of analytical thought, in the context of mathematics, as an impediment to students' ability to develop problem representation. Problem representation involves the understanding of the text of a problem and the potential solution pathways. This is why many students require intervention by teachers to facilitate the thought process [42].

Rubenstein et al. (2020) [47] explored one helpful model for how students think through problem identification. These researchers identified three broad approaches, including inductive, deductive, and a hybrid of both approaches [47]. In inductive reasoning, students generalize from a set of cases and in deductive reasoning students use the process of elimination [16]. To illustrate, recall the third-grade student identifying the meaning of a word in context. The student could use inductive reasoning by examining uses of the word in various contexts and then create rules for using the word in certain types of contexts. By contrast, the student could use deductive reasoning by listing all the possible meanings of the word in question and then test each meaning of the word in the context of their reading passage. Each type of reasoning helps the student specify and better understand the nature of the unknown.

4.2.2 Exploring a Problem

After identifying the problem, students engage in an exploration of the problem, which involves a process of investigation and discovery. Jerome Bruner was the first contemporary theorist to explore discovery learning as it is known today, and he described it as a process of induction where teachers guide young students through a series of examples and cases or even vague questions [38]. The purpose of this guided process of induction is to have students arrive at an understanding of an abstract concept by detecting patterns and creating rules that explain the patterns [38]. For instance, the third grader determining the meaning of a word based on how it is used in a specific reading passage would examine multiple uses of the word in different contexts and create a rule from patterns of its usage. They could then apply this rule to the passage assigned by the teacher. Similarly, the fourth grader figuring out what linking words will appropriately string their thoughts together would analyze multiple uses of linking words, and then state a rule based on the patterns they see. That rule could help them choose linking words that would best string their thoughts together. The teacher’s role during this process of discovery is to ask students probing questions that direct them closer to important details in the examples and cases under investigation, and to give them feedback when they identify incorrect patterns and rules [19].

4.2.3 Generating and Testing Hypotheses

Hypotheses follow an exploration of the problem. Hypotheses are constructs based on logic, usually involving a cause-and-effect relationship, which propose possible answers to questions during inquiry and guide the thinking process [53]. The ability to engage in hypothesis testing typically occurs at around the age of 6 for most children [48]. Students can generate and test hypotheses during their development when the details of a new experience, e.g., an example or case taught in the classroom, involve dynamics that contradict their prior knowledge [32]. Take, for instance, when students begin to learn the basics of telling time, and they find out that a day contains a certain number of hours, which correspond to the position of the sun. This learning can create disequilibrium between the students’ prior belief and the new information. For example, they might have believed that a day ends when their parents put them to bed, which does not necessarily correspond to the position of the sun. Legare (2012) [32] argues that this inconsistency generates the impetus for students’ exploration, which their teachers facilitate. To help students test for the relationship between the length of a day and the position of the sun, teachers might teach them how to use a sundial. When students use the sundial for a week, they can confirm that the position of the sun is consistent at each time of day. However, children in primary school struggle to engage in hypothesis creation and testing independently, and therefore, they require cues from teachers [24].

Research is lacking on how teachers provide cues to engage students in hypothesis creation and testing, and the scant studies focus on science instruction in primary school [45]. In light of this, Peterson and French (2008) [43] suggest guiding experiments using hypothesis testing for existing activities in many schools, such as arts and crafts. For example, a teacher could guide students to predict the result of mixing two colors together and then mix the colors to test their predictions.

4.2.4 Reflecting on Results of Hypothesis Testing

Dewey’s model for inquiry finishes with a reflection on the results of hypothesis tests and potentially plans to conduct other hypothesis tests, if the results do not solve the problem and reveal the unknown. Reflection is a purposive process of analysis and evaluation to reach a deeper understanding of what is known and what still needs to be learned [40]. Therefore, reflection on the results of hypothesis testing involves analysis and evaluation of the beliefs and evidence that undergird the hypothesis, as well as what information is required to reformulate it. Amsterlaw and Wellman (2006) [3] suggest that the impetus during reflection to reformulate and retest hypotheses comes from student’s realization of what these researchers call *false belief.* A false belief in this case refers to the original hypothesis that is not confirmed by a test. Students begin to reformulate their hypotheses when teachers prompt them to explain why they think their initial hypothesis test failed [32]. This ability to reformulate hypotheses typically develops as early as 6 years old and becomes more sophisticated over time [29]. Student's explanation of failed hypothesis tests involves an evaluation of the cause-and-effect relationship they hypothesized initially, which then sparks ideas for other cause/effect hypotheses [32]. Once they have created other hypotheses, students engage in more hypothesis testing until they reveal the unknown and solve the problem.

4.3 Role for GenAI in Students' Inquiry Process

Dewey argued that teachers need to support students by modeling the inquiry process. This is important because students actively select behaviors to emulate from teachers to find solutions for real-world problems. Furthermore, teacher support ensures the development of students' cognitive processes to solve practical problems, which ultimately increases students' levels of personal self-efficacy [57]. However, Dewey suggested teachers should not prevent students from experiencing failure when engaging in inquiry, because failure is the mechanism that encourages students to continue creating and testing hypotheses.

As previously discussed, GenAI has the potential to act as a teacher's assistant and help facilitate students' thought processes in problem identification. However, it is important for teachers to first *train* GenAI programs on how to facilitate this process. In programs like ChatGPT, Microsoft Copilot, and Google Gemini, training requires teachers to choose a model for students to think through problem identification, upload text explaining the model to the GenAI program, and direct the program to follow the model as it interacts with the student.

5. Modeling, Guidance, and Supervision for GenAI Use in Primary Education

Many existing sources that contemplate the various uses of GenAI assume its positive utility in the field of education. These sources suggest that students could use GenAI in unsupervised, semi-supervised, and fully supervised learning environments [17]. However, these same sources promote caution in the use of GenAI by young students. Students can lack the critical thinking skills necessary to protect themselves from inappropriate and false information. Therefore, young students are more susceptible to deception, and it is important to teach them to think critically about the veracity of information from GenAI [56]. Of particular concern, as AI takes on more humanistic qualities such as appearance and voice, a younger student is more likely to believe misinformation. For instance, students could fall prey to future versions of generative AI chatbots designed to convincingly impersonate humans and convince them to do immoral and dangerous things for their age levels [55]. They argue that teachers should take an active role in supporting their young students' use of all forms of technology in the classroom [8]. Teachers should take an even more active role in guiding and supporting young students when using GenAI to protect their safety and well-being [52].

Su and Yang (2023) [52] offer a helpful framework called IDEE that could be helpful as a general guide for modeling the use of GenAI programs such as ChatGPT in education. IDEE is an acronym that stands for *Identify the Desired outcomes*, *Determine the Appropriate Level of Automation*, *Ensure Ethical Considerations*, and *Evaluate the Effectiveness* [52]. The analysis in this manuscript of what the major learning theories suggest for the use of GenAI relates to the *Determine the Appropriate Level of Automation* and *Ensure Ethical Considerations* parts of Su and Yang's (2023) [52] framework. The major learning theories suggest tempering the level of automation of GenAI use based on students' developmental levels to protect their safety and to teach them how to use the technology ethically as well as resist an overreliance of technology use with students.

The IDEE framework also includes six broad considerations that can guide teachers' modeling of GenAI use to young students. These considerations include the following: start with the basics, use age-appropriate language, demonstrate AI in action, explain the benefits of AI, discuss the ethical implications of AI, and encourage curiosity [52]. Teachers should address each consideration through the lens of Dewey's philosophy analyzed in this manuscript. For example, when demonstrating GenAI in action, teachers should model the standard process for inquiry. The standard process of inquiry involves posing and testing hypotheses. Modeling the standard process for inquiry will prepare students to engage in inquiry independently.

6. Examples of Inquiry with GenAI

Common Core K.MD.1. Describe measurable attributes of objects, such as length or weight.
Describe several measurable attributes of a single object. 2. Directly compare two objects with a measurable attribute in common, to see which object has “more of”/“less of” the attribute, and describe the difference. For example, directly compare the heights of two children and describe one child as taller/shorter.
Activity: Let’s go to the zoo!
GenAI Tool: Microsoft Designer
Engage: What animals can we see at the zoo?
Description: What does a [choose animal] look like? Which animals are tall, short, heavy, light, small, big, wide, narrow?
Problem: The animals get out and the zoo needs our help to get them back in their cages. There are cages of different sizes. How can we make sure all the animals fit into their cages?
Prompts: Which animals are taller, shorter, heavier, lighter, smaller, bigger, wider, narrower?
Hypotheses: Examples - The giraffe is taller than the hippo. The hippo is wider than the lion.
Test Hypotheses: Ask GenAI tool to show - e.g., a giraffe next to a hippo, a hippo next to a lion, etc.
Reflect: Were you correct? Is the giraffe taller than the hippo? What other animals are also taller than the hippo? Is the hippo wider than the lion? What other animals are also wider than the lion?
Common Core 2.MD.8. Solve word problems involving dollar bills, quarters, dimes, nickels, and pennies, using $ and ¢ symbols appropriately. Example: If you have 2 dimes and 3 pennies, how many cents do you have?
Activity: It’s a bake sale!
GenAI Tool: Microsoft Designer
Engage: What are your favorite cookies?
Description: How much does a small cookie cost? How many quarters, dimes, and nickels would we need to buy a cookie? (e.g., 65 cents)
Problem: Ten customers want to buy cookies from us, two of each. If you are the cashier, how much money should we have when we count it?
Prompts: Ask GenAI to create a picture of four different kinds of cookies, and another picture of 10 different customers, each with a name.
Hypotheses: Draw circles on your paper to show the number of each coin we will get from each customer. Then count the total of each coin we have altogether.
Test Hypotheses: Ask GenAI to create a picture of the total quarters, dimes, and nickels we will have from 10 customers, who buy 4 cookies each at 65 cents per cookie.
Reflect: How does your picture compare to the one from GenAI? How could you change your picture to make it the correct number of each coin?

7. Concluding Thoughts

Teachers make decisions daily for their classrooms based on what is in the best interest of students. If a teacher should want to use GenAI, first the teacher must be open and willing to seek ways in which GenAI can be of benefit to the classroom. They must be willing to model best practices and pace gradual independence with technology. The above examples can be used as guidance for best practice. For young students, human interaction in the inquiry process is critical. Rashel et al. (2024) [46] posed concerns about the overuse of GenAI therefore limiting human interaction. Van der Berg (2024) expressed concerns such as accuracy, reliability, professional and moral use, as well as overdependence. GenAI can never replace the vital role a human teacher has in a child’s development. GenAI should be viewed as supplemental to learning and teaching, not a replacement. Teaching students how to think, problem solve, be creative, etc. is all uniquely human. Wu (2023) [60] comments on how human learning involves learning from mistakes and refining our understanding. Safeguards for AI include having guided then supervised use while working toward more independence. The teacher should act as an authority in the classroom and be aware of all interactions with a student and technology.

REFERENCES

[1] Ahmad, N., Murugesan, S., & Kshetri, N. (2023). Generative artificial intelligence and the education sector. *Computer, 56*(6), 72-76.

[2] Ali, S., Ravi, P., Williams, R., DiPaola, D., & Breazeal, C. (2024, March). Constructing dreams using generative AI. In *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(21), pp. 23268-23275).

[3] Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of cognition and development, 7*(2), 139-172.

[4] Bandi, A., Adapa, P. V. S. R., & Kuchi, Y. E. V. P. K. (2023). The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges. *Future Internet, 15*(8), 260.

[5] Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets, 33*(1), 63. https://doi.org/10.1007/s12525-023-00680-1

[6] Chan, C. K. Y., & Tsi, L. H. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation, 83*, 101395.

[7] Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *Ieee Access, 8*, 75264-75278.

[8] Chiu, T. K. (2023). The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 1-17.

[9] Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology, 32*(3), 444-452.

[10] Cooper, T. (1986). *Problem solving.* Queensland: Mathematics Education, Brisbane College of Advanced Education.

[11] Epstein, Z., Hertzmann, A., Investigators of Human Creativity, Akten, M., Farid, H., Fjeld, J., ... & Smith, A. (2023). Art and the science of generative AI. *Science, 380*(6650), 1110-1111.

[12] Franzen, C. (2024, February 19). *Science Journal retracts peer-reviewed article containing AI generated "nonsensical" images*. VentureBeat. Retrieved from https://venturebeat.com/ai/science-journal-retracts-peer-reviewed-article-containing-ai-generated-nonsensical-images/

[13] Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research, 25*(3), 277-304.

[14] Garrett, L. (1997). Dewey, Dale, and Bruner: Educational philosophy, experiential learning, and library school cataloging instruction. *Journal of education for library and Information science,* 129-136.

[15] Gondode, P., Duggal, S., & Mahor, V. (2024). Artificial intelligence hallucinations in anaesthesia: Causes, consequences and countermeasures. *Indian Journal of Anaesthesia, 68*(7), 658-661. https://doi.org/10.4103/ija.ija_203_24

[16] Goswami, U. (2010). Inductive and deductive reasoning. *The Wiley‐Blackwell handbook of childhood cognitive development,* 399-419.

[17] Gozalo-Brizuela, R., & Garrido-Merchan, E. C. (2023). ChatGPT is not all you need. A State of the Art Review of large Generative AI models. *arXiv preprint arXiv:2301.04655*.

[18] Hien, L. T. & Nga, L. P. (2024). Current Status and Measures to Develop Critical Thinking for Pedagogical Students in Using Generative Artificial Intelligence. *Vinh University Journal of Science, Educational Science and Technology, 53.* https://doi.org/10.56824/vujs.2024.htkhgd180

[19] Honomichl, R. D., & Chen, Z. (2012). The role of guidance in children's discovery learning. Wiley Interdisciplinary Reviews: *Cognitive Science, 3*(6), 615-622.

[20] Hossain, M., Amin, S. Z., & Hossain, F. (2025). Empowering Education Harnessing AI Tools for Innovative Content Development. In *Strengthening Human Relations in Organizations With AI* (pp. 295-326). IGI Global Scientific Publishing.

[21] Hu, Y. H. (2024). Implementing generative AI chatbots as a decision aid for enhanced values clarification exercises in online business ethics education. *Educational Technology & Society, 27*(3), 356-373.

[22] Hwang, G. J., & Chen, N. S. (2023). Exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions. *Journal of Educational Technology & Society, 26*(2).
[23] Imran, M., & Almusharraf, N. (2024). Google Gemini as a next generation AI educational tool: a review of emerging educational technology. *Smart Learning Environments, 11*(1), 22.
[24] Jones, A., Markant, D. B., Pachur, T., Gopnik, A., & Ruggeri, A. (2021). How is the hypothesis space represented? Evidence from young children's active search and predictions in a multiple-cue inference task. *Developmental psychology, 57*(7), 1080.
[25] Kalota, F. (2024). A Primer on Generative Artificial Intelligence. *Education Sciences, 14*(2), 172. https://doi.org/10.3390/educsci14020172
[26] Kanders, K., Stupple- Harris, L., Smith, L., & Gibson, J. L. (2024). Perspectives on the impact of generative AI on early- childhood development and education. *Infant and Child Development*, e2514.
[27] Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research, 34*(2), 313-338.
[28] Kim, S., Park, C., Jeon, G., Kim, S., & Kim, J. H. (2025). Automated Audit and Self-Correction Algorithm for Seg-Hallucination Using MeshCNN-Based On-Demand Generative AI. *Bioengineering, 12*(1), 81.
[29] Köksal, Ö., & Sodian, B. (2023). The origins of hypothesis testing: Young children's developing understanding of information seeking compared to effect production. *European Journal of Developmental Psychology, 20*(4), 694-708.
[30] Köksal-Tuncer, Ö., & Sodian, B. (2018). The development of scientific reasoning: Hypothesis testing and argumentation from evidence in young children. *Cognitive development, 48*, 135-145.
[31] Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human communication research, 30*(3), 411-433.
[32] Legare, C. H. (2012). Exploring explanation: Explaining inconsistent evidence informs exploratory, hypothesis- testing behavior in young children. *Child development, 83*(1), 173-185.
[33] Lu, J., Hao, Q., & Jing, M. (2016). Consuming, sharing, and creating content: How young students use new social media in and outside school. *Computers in Human Behavior, 64*, 55-64.
[34] Luan, H., & Tsai, C. C. (2021). A review of using machine learning approaches for precision education. Educational Technology & Society, 24(1), 250-266.
[35] Maleki, N., Padmanabhan, B., & Dutta, K. (2024, June). AI hallucinations: a misnomer worth clarifying. In 2024 IEEE conference on artificial intelligence (CAI) (pp. 133-138). IEEE.
[36] Megahed, F. M., Chen, Y. J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2024). How generative AI models such as ChatGPT can be (mis) used in SPC practice, education, and research? An exploratory study. *Quality Engineering, 36*(2), 287-315. https://doi.org/10.1080/08982112.2023.2206479
[37] Narayanaswamy, A. (2024). Evolution of microsoft copilot. In *Microsoft Copilot for Windows 11: Understanding the AI-Powered Features in Windows 11* (pp. 1-24). Berkeley, CA: Apress.
[38] Neber, H. (2012). Discovery learning. *Encyclopedia of the sciences of learning, 1009*, 1012.
[39] Noddings, N. (2018). *Philosophy of education*. Routledge.
[40] Nuraini, N. L. S., Cholifah, P. S., Mahanani, P., & Meidina, A. M. (2020, November). Critical thinking and reflective thinking skills in elementary school learning. In *2nd Early Childhood and Primary Childhood Education (ECPE 2020)* (pp. 1-5). Atlantis Press.
[41] Pack, A., & Maloney, J. (2023). Potential affordances of generative ai in language education: Demonstrations and an evaluative framework. *Teaching English with Technology, 23*(2), 4-24.
[42] Peltier, C., & Vannest, K. J. (2017). A meta-analysis of schema instruction on the problem-solving performance of elementary school students. Review of Educational Research, 87(5), 899-920.
[43] Peterson, S. M., & French, L. (2008). Supporting young children's explanations through inquiry science in preschool. *Early childhood research quarterly, 23*(3), 395-408.
[44] Piaget, J., & Inhelder, B. (2008). *The psychology of the child*. Basic books.
[45] Ramanathan, G., Carter, D., & Wenner, J. (2022). A framework for scientific inquiry in preschool. *Early Childhood Education Journal, 50*(7), 1263-1277.

[46] Rashel, M. M., Khandakar, S., Hossain, K., Shahid, A., Kawabata, T., Batool, W., ... & Rafique, T. (2024). AI in Education: Unveiling the Merits and Applications of Chat-GPT for Effective Teaching Environments. *Revista de Gestão Social e Ambiental, 18*(10), e09110-e09110.
[47] Rubenstein, L. D., Callan, G. L., Neumeister, K. S., & Ridgley, L. M. (2020). Finding the problem: How students approach problem identification. *Thinking Skills and Creativity, 35,* 100635.
[48] Ruffman, T., Perner, J., Olson, D. R., & Doherty, M. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis- evidence relation. *Child Development, 64*(6), 1617-1636.
[49] Siontis, K. C., Attia, Z. I., Asirvatham, S. J., & Friedman, P. A. (2024). ChatGPT hallucinating: can it get any more humanlike?. *European Heart Journal 45*(5), 321-323. https://doi.org/10.1093/eurheartj/ehad766
[50] Stokel-Walker, C. (2023). CHATGPT LISTED AS AUTHOR ON. *Nature, 613*.
[51] Su, J., & Yang, W. (2022). Artificial intelligence in early childhood education: A scoping review. *Computers and Education: Artificial Intelligence, 3*, 100049.
[52] Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education, 6*(3), 355-366.
[53] Supino, P. G. (2012). The Research Hypothesis: Role and Construction. In *Principles of Research Methodology: A Guide for Clinical Investigators* (pp. 31-53). New York, NY: Springer New York.
[54] Tsaliki, L. (2022). Constructing young selves in a digital media ecology: Youth cultures, practices and identity. *Information, Communication & Society, 25*(4), 477-484.
[55] UNICEF. (n.d.). Generative AI: Risks and opportunities for children. *Innocenti Global Office of Research and Foresight*. https://www.unicef.org/innocenti/generative-ai-risks-and-opportunities-children
[56] Van den Berg, G. (2024). Generative AI and educators: Partnering in using open digital content for transforming education. *Open Praxis, 16*(2), 130-141.
[57] Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes (Vol. 86)*. Harvard university press.
[58] Weigle, P. E., & Shafi, R. M. (2024). Social media and youth mental health. *Current psychiatry reports, 26*(1), 1-8.
[59] White, M. D., & Marsh, E. E. (2006). Content analysis: A flexible methodology. *Library trends, 55*(1), 22-45.
[60] Wu, Y. (2023). Integrating generative AI in education: how ChatGPT brings challenges for future learning and teaching. *Journal of Advanced Research in Education, 2*(4), 6-10.
[61] Wu, Y., Jiang, A. Q., Li, W., Rabe, M., Staats, C., Jamnik, M., & Szegedy, C. (2022). Autoformalization with large language models. *Advances in Neural Information Processing Systems, 35*, 32353-32368.
[62] Yu, H., & Guo, Y. (2023, June). Generative artificial intelligence empowers educational reform: current status, issues, and prospects. In *Frontiers in Education* (Vol. 8, p. 1183162). Frontiers Media SA.