HJSTEM
Hellenic Journal of STEM Education, 2026, 5(1), 1-18

http://www.hellenicstem.com/
ISSN 2832-840X

RESEARCH ARTICLE

The Interactive Design Process Framework (IDPF): Utilizing GenAI as a Collaborative Agent for Creating STEAM Projects

Dimitrios SOTIROPOULOS * University of Thessaly, GREECE
Apostolos XENAKIS University of Thessaly, GREECE
Michail KALOGIANNAKIS University of Thessaly, GREECE
Mehmet Fatih TAŞAR Georgia State University, USA

Abstract

Increasingly, teachers are considering using Generative AI (GenAI) and large language models (LLMs) as new tools for developing lessons and creating resources. Interdisciplinary STEAM project design specifically continues to be problematic because it requires many professionals to work together in integrative ways when developing successful STEAM project designs. Collaborating closely to develop authentic problems, establish a high level of pedagogical quality, use principles of inclusive design, and create valid and reliable assessment instruments often exceeds the individual teacher's professional expertise and available time. The purpose of this conceptual paper is to lay the foundation of what we call "the Interactive Design Process Framework (IDPF)" and propose its use in instruction. IDPF has four steps: Co-design, Analysis, Expert Review, and Synthesis. Using these four iterative (cyclical) steps, the IDPF framework positions GenAI as a collaborative designer to work with both classroom teachers and subject area expert reviewers. As a conceptual contribution, this paper does not present empirical data; rather, it articulates the theoretical rationale, structural logic, and design principles of the framework as a foundation for future empirical investigation. In addition to a primary education example of illustration, implications for practice, limitations of the study, and recommendations for future research are discussed.

Keywords

generative artificial intelligence; STEAM education; human-AI interaction; TPACK; universal design for learning; engineering design process; collaborative project design

• Received February 24 2026 • Revised 4 April 2026 • Accepted 8 April 2026

Introduction

Since the public launch of ChatGPT in late 2022, the use of Generative Artificial Intelligence (GenAI) in educational practice has increased at a geometrical rate. Teachers of all grade levels are using GenAI tools to support them in their work**, including lesson plan development and content creation as well as assessment creation and personalized feedback.** In a recent survey, 37% of lower secondary education teachers reported using AI as part of their job in 2024, and 57% stated that AI had helped them develop or refine their lesson plans (OECD, 2026). Overall, this trend indicates that GenAI is becoming increasingly accessible to educators and is perceived as a useful professional support tool.

In contrast to the routine task of lesson plan development, STEAM (Science, Technology, Engineering, Arts, and Mathematics) project development is holistic in its nature and presents significant challenges (Bryan & Guzey, 2020; Spyropoulou & Kameas, 2024). First, to develop a STEAM project, a teacher must integrate the various curricular and pedagogical approaches related to each discipline. Second, teachers must frame a project using an authentic, open-ended problem; determine how to blend various teaching methods to achieve successful outcomes; create inclusive learning opportunities; and develop assessments that accurately capture both discipline-specific and cross-disciplinary competencies. Finally, teachers' ability to develop STEAM projects are limited by several practical factors, including lack of expertise across all STEAM subject areas; lack of sufficient time; and lack of access to additional resources and to colleagues across disciplines who can review and refine their designs.

While GenAI tools can generate lesson plans, suggest activities, and produce assessment rubrics with impressive fluency, research consistently demonstrates that AI-generated educational content requires critical evaluation by domain-knowledgeable professionals (Clark & Van Kessel, 2024; Sakamoto et al., 2024; Sun et al., 2025). The quality of AI-assisted lesson planning depends heavily on the teacher's ability to clearly communicate pedagogical intentions, evaluate AI output critically, and integrate feedback from multiple sources into a coherent instructional design (Wardak et al., 2025; Karpouzis et al., 2024).

In response to the challenges outlined above, in this paper we propose an iterative, four-step framework, the Interactive Design Process Framework (IDPF) that involves three critical actors in the design of STEAM projects: the classroom teacher, a GenAI agent, and subject-area expert reviewers. Its iterative nature ensures improved design by repeated review and testing. All components of the framework are based on the principles of Universal Design for Learning (UDL) (CAST, 2024); UDL is used as the basis for the design of every step, tool and interaction so that unique aspects of each participant are taken into consideration from the outset; this inclusionary principle is highlighted within each of the framework's components.

https://doi.org/10.51724/hjstemed.v5i1.76
*Corresponding author

The Four Steps of IDPF (Figure 1)

Step A (Co-design): the teacher and GenAI produce an initial draft of the interdisciplinary STEAM project collaboratively.

Step B (Analysis): GenAI breaks the draft into review components that are tagged or labelled as belonging to specific STEAM disciplines and by the type of review (peer versus teacher or expert).

Step C (Expert Review): subject-area expert teachers review designated parts for the purpose of evaluating whether the content is accurate and the pedagogy is appropriate.

Step D (Synthesis): the teacher synthesizes the expert reviews with the GenAI-generated supporting data while retaining full professional judgement throughout the process.

In addition to UDL 3.0 being the foundational design principle for IDPF, we have utilized other design principles drawn from Technological Pedagogical Content Knowledge (TPACK model), PCK, Engineering Design Process (EDP), Design Thinking (DT), and Substitution, Augmentation, Modification, Redefinition (SAMR) model. All of these principle frameworks were chosen due to having previously documented positive research evidence for solving a specific documented issue. The individual components of IDPF are supported by research; however, there is currently no empirical validation of their collective effectiveness within this framework.

The TPACK Model offers a useful lens for examining the distribution of expertise across the three actors, since few individual teachers are likely to possess all the knowledge areas required to design an effective STEAM project. The PCK Model provides a two-dimensional analysis of the project component's content and pedagogy. Research shows that content accuracy and pedagogy are distinct dimensions of quality and require separate evaluation; therefore, this model is a valid tool for evaluating both dimensions of STEAM projects. EDP and DT methodologies were both used to provide the design logic for each of the student activities and as a meta-structure for the framework's own iterative cycle to help resolve teachers' lack of confidence in the Engineering component of STEAM projects.

The SAMR model (Puentedura, 2006) is applied here to support purposeful technology integration, helping to counter the tendency toward superficial tool use. We argue that while each component of the IDPF has been validated independently, their structured combination within a single framework may produce outcomes that exceed what any component achieves alone — though further systematic investigation is needed to confirm this proposition.

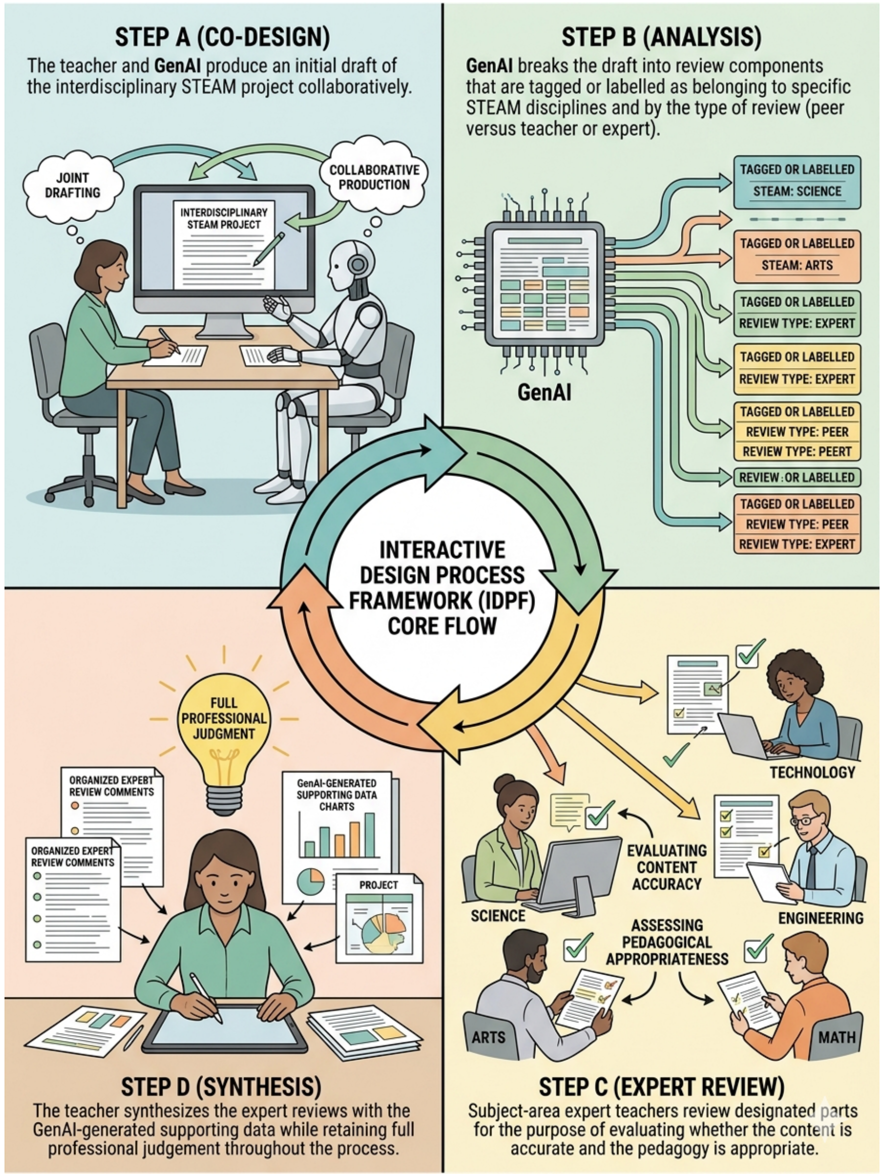

Figure 1. The four steps of IDPF. *Note.* Image generated using Google Gemini 3 from the prompt "create an image for the Interactive Design Process Framework - IDPF" on April 7, 2026, https://gemini.google.com/share/10b11c6eb5aa

The outline of the contents of this paper is as follows. Section 2 presents the theoretical basis and rationale for this study, outlining STEAM education, TPACK, PCK, Universal Design for Learning (UDL), Engineering Design Process/Design Thinking (EDP/DT), and Generative AI (GenAI). Section 3 provides a detailed description of the proposed Interactive Design Process Framework (IDPF) and where it fits within the literature on the subject. Section 4 presents the roles and responsibilities of the development team members responsible for creating these integrated frameworks. Section 5 gives an example of a project created through the use of an IDPF with primary school students. Section 6 discusses implications, limitations, and areas for future research concerning the use of IDPFs in future educational settings.

2. Background and Motivation

2.1 STEAM Education and Project Design

STEAM education expands on the traditional foundation of STEM by adding in the importance of the **arts** as an integral part of human creativity and design thinking (Land, 2013; Yakman, 2008). The STEAM epistemology correlates logic with creativity and science with art, promoting a holistic, integrative model of learning (Henriksen, 2014). It draws on constructionist principles — specifically, that knowledge is most effectively built through design, experimentation, iteration, and collaborative problem-solving (Papert, 1980). Recently, projects such as the EU-funded Road-STEAMer have helped define the characteristics of effective STEAM projects, such as equity, connections among disciplines, collaboration, real-world connections, and creativity (Chappell et al., 2025).

The issues with the overall design of STEAM projects have been still posing significant challenges for many teachers who expressed difficulty connecting multiple disciplines further than only at a surface level (Boice et al., 2024; Herro et al., 2019). Voicu et al. (2022) highlighted that different STEAM implementations across six countries have led to commonalities at both the preschool and primary school levels. Common reactions to implementing STEAM included a lack of teacher preparedness for interdisciplinary design, insufficient institutional support for this type of interdisciplinary design, and difficulty maintaining authentic interdisciplinary integration of a discipline over the course of the project. A particularly common concern noted by teachers was revealed by Mehddi et al. (2005) as the confidence level in their ability to incorporate engineering-related concepts into their projects, as it is a subject area that they are not as much familiar with as either science or mathematics.

It was also found that even teachers with extensive educational experience are having difficulties and experiencing confusion and conflict when attempting to create cross-disciplinary projects by focusing more on their discipline than the pedagogy behind making those cross-disciplinary projects work (Cohen & Bronshtein, 2025). Teachers support the belief that STEAM design which integrates disciplines will enhance their feelings of self-efficacy. However, teachers do not have sufficient time or resources to support collaborative design of cross-disciplinary pedagogical approaches (Vlasopoulou, Kalogiannakis & Sifaki, 2021). Moreover, there are ongoing challenges with providing teachers continued professional development, the complexity of assessing student work through different formats, and inadequate funding for materials as continuing barriers to teachers working collaboratively (Li, 2024). The consistent research finding across all these three studies cited above is that the ability to create a project that is authentic in combining the components of science, technology, engineering, the humanities, and mathematics in a pedagogically sound manner requires expertise that very few teachers posses individually and that accessing other teachers who are knowledgeable in different disciplines for the collaboration necessary to develop these types of projects and the foundational knowledge to do so is very limited (Saseendran & Thomas, 2025; Anderson & Tully, 2024).

2.2 TPACK: Distributing Expertise Across Actors

A common observation found in scholarly articles is that pedagogy, content and technology should integrate into STEAM learning, but the majority of teachers do not have sufficient expertise in how to use these elements in an integrated way (Celik et al., 2022; Mishra et al., 2023; Feldman-Maggor et al., 2025). The Technological Pedagogical Content Knowledge (TPACK) framework developed by Mishra & Koehler (2006) defines how to integrate these three separate forms of knowledge into one cohesive understanding through overlapping domains of knowledge (Content Knowledge (CK), Pedagogical Knowledge (PK) and Technological Knowledge (TK)). When combining the three domains through their overlapping areas of expertise (Pedagogical Content Knowledge (PCK), Technological Content Knowledge (TCK), Technological Pedagogical Knowledge (TPK), and Technological Pedagogical Content Knowledge (TPCK)), it is not possible for one teacher to possess all of these types of expertise in all STEAM disciplines and use all available technologies. To further complicate the things for teachers, there exists no general TPACK, it is inherently content and technology specific (Taşar & Yılmaz Ergül, 2023). It has to be re-created for every instance of teaching a topic.

To mitigate this issue, the IDPF proposes to share the burden of TPACK development by distributing responsibility for the development of TPACK across three stakeholders, rather than solely on the teachers. Experts' content (CK) and pedagogical (PK) feedback will assist in building capacity in CK and PCK. Research suggests that teachers who develop their CK and PCK in collaboration with subject-matter experts are more successful than those who attempt alone (Penuel et al., 2007), while PCK varies significantly across teaching contexts, requiring systematic rather than ad hoc support (Gözüm et al., 2022).

The use of the GenAI agent can help create TK, TCK, and TPK by suggesting the use of tools and technology and pedagogical pairings; research has suggested that this type of function can be performed by the GenAI agent when given a structured prompt and appropriate pedagogical intent (Peikos & Stavrou, 2025; Liu & Bates, 2025). The classroom teacher is at the intersection of TPACK and will take all inputs and combine them into an overall design; based upon their research, Feldman-Maggor et al. (2025) and Uğraş et al. (2025) have stated that this synthesis requires human judgment that cannot be performed by a GenAI agent. While there is no empirical research on the distributed model of use (as a whole), all individual elements of the distributed model have been supported through empirical research; therefore, by combining and using the structured components, it is possible that the outcome created from the combined use of those components will be greater than the individual outcomes; consequently, the combined outcome may exceed the sum of what each component produces independently.

2.3 PCK: Grounding the Analysis of Project Quality

Mishra and Koehler (2006) built upon Shulman's (1986) foundational work to develop the TPACK framework, which extends the concept of pedagogical content knowledge (PCK) to encompass technology's role in an educational setting.

In doing so, they added a third dimension to PCK, or the union of content knowledge and pedagogical knowledge, which places PCK in the larger context of pedagogical and technological knowledge bases.

Research has demonstrated that while the content produced by GenAI has a logical structure, much of it contains errors that only a professional with subject matter expertise could identify. Peikos & Stavrou (2025) discovered that when using ChatGPT without assistance, the generated material contained developmentally inappropriate content (e.g., using density formulas with primary age students), did not incorporate methods for addressing student misconceptions, and did not utilize differentiated instruction. Similar findings have been reported in multiple research contexts (Cooper, 2023; Clark & Van Kessel, 2024; Moundridou et al., 2024). The fundamental problem is that GenAI does not have PCK, or the integration of content knowledge and pedagogical knowledge, necessary for teachers to determine whether a representation, explanation, or activity will be effective for a group of students (Shulman, 1986; Gess-Newsome, 2015). Gemini AI firmly agrees with this statement and asserts that "As a GenAI, I cannot "create" PCK in the biological or professional sense that Shulman describes, but I can **simulate and support** the application of PCK through "Pedagogical Engineering"" (Google, 2026).

According to Carlson et al. (2019), the Refined Consensus Model of PCK includes classification of PCK into three components; collective PCK (knowledge that is shared amongst professional peers); personal PCK (knowledge that is individually owned by an individual/classroom teacher); enacted PCK (knowledge that is used in practice, including lesson planning). Otto and Everett (2013) developed a three-circle Venn Diagram to operationalize PCK for lesson planning that depicts Pedagogy, Content and Context as well as the pairwise intersections of each circle; thus providing a strong framework for lesson planning (Chaitidou et al., 2018). Gözüm et al. (2022) demonstrate that teachers' PCK may vary greatly amongst different cultures, suggesting that support should be provided in a structured way rather than only relying on individual experience. The IDPF plans to address the above-given challenge through two methods based on independently validated evidence-based practices. Firstly, Step B (Analysis) decomposes the project draft into Review Components (tagged with one or more descriptors) along two dimensions – content accuracy and pedagogical appropriateness – that directly reflect the PCK intersections identified by Otto and Everett (2013). This double-dimension approach was chosen based on the findings that content errors and pedagogical misalignments are two separate failure modes and require different areas of expertise for detection (Peikos & Stavrou, 2025; Feldman-Maggor et al., 2025). The iterative co-design dialogues of Step A, which emulate layer-prompting strategies that Peikos and Stavrou (2025) and Atlas (2023) demonstrate to be effective in creating quality improvements, are designed to create multiple opportunities to identify gaps and correct mistakes before the project reaches students. This is only a partial solution to closing quality gaps, but it does apply established practices from the literature, systematically providing ways for teachers to identify and address gaps before the project reaches students.

Blonder and Feldman-Maggor (2024) argue that GenAI is an excellent diagnostic and developmental resource for educators' PCK, while Celik (2023) created the Intelligent-TPACK model to extend TPACK to include an ethical dimension. These new models present a common point: the effective use of GenAI in education is determined by both the sophistication of the AI system and the level of human expertise available to guide, assess, and provide context for the output of the AI. The IDPF expands the teacher co-design model using ChatGPT previously proposed by Peikos and Stavrou (2025) by adding an expert-review dimension (Step C) and also providing additional scope to include interdisciplinary STEAM projects rather than single-discipline science lessons.

2.4 UDL: The Foundational Design Principle

Even though STEAM education aims to be equitable and inclusive for all learners, research findings indicate that many educators treat the principle of inclusive design more as a guiding value than as something they implement (Bequette & Bequette, 2012; Spyropoulou & Kameas, 2024). As a result, teachers typically develop instructional materials based on the expectation that all students can perform at the same level by using only one form of representation or expression and allowing very little choice for how students engage with the lesson. Teachers who requested lesson plans generated by GenAI have also found this same pattern in their materials; Peikos and Stavrou (2025) found that when they did not ask for differentiated strategies, those strategies did not appear in the generated lesson plans. The Universal Design for Learning (UDL) Framework (CAST, 2024) will be used as a foundational design principle for all components of the Interactive Design Process Framework (IDPF). This UDL 3.0 framework consists of three principles: Multiple Means of Engagement, Multiple Means of Representation, and Multiple Means of Action and Expression, which collectively provide more opportunities for students to establish identities and sense of belonging, to learn collaboratively and inclusively, and to eliminate historically-sustained biases and systemic barriers that have contributed to exclusion (CAST, 2024). UDL is not an additional framework; it is the base of all aspects of the IDPF, including all aspects of the IDPF design, all IDPF component selections, and all IDPF interactions between components. The IDPF process is designed for different levels of teachers to make it more accessible through the framework level, as an example. Each review brief has a UDL component included in it at the review level, which means that experts have evaluated each brief with a UDL perspective. Finally, at the activity level, each student-facing activity must include representation, expression and engagement options that provide students the opportunity to represent, express, and engage using different modes of representation, expression, and engagement. Research shows that explicit and systematic prompting of teachers to consider UDL principles during planning has resulted in activities that are significantly more accessible and differentiated (Tomlinson et al., 2003; Peikos & Stavrou, 2025). Initial applications in similar contexts suggest that embedding UDL as an intrinsic structural feature rather than an optional consideration results in a greater probability of inclusive design reaching the classroom; nonetheless, continued inquiry across diverse environments is required to provide evidence of the generality of this finding.

2.5 EDP and Design Thinking: Structuring Both Activities and the Design Process

Many educators indicate that they experience obstacles to incorporating Engineering content into the STEAM continuum (Mehddi et al., 2025) when compared to science or mathematics content. A systematic review conducted by Saseendran and Thomas (2025) concerning Design Thinking (DT) in STEM/STEAM found that teacher readiness, the rigid constraints of curriculum, and a lack of appropriate methods for assessing progress continue to be significant barriers to creating independent Design Challenges that are the focus of STEAM Instruction.

The Engineering Design Process (EDP) is one method for implementing a structured approach to design. EDP consists of the following iterative phases: problem identification, research, solution development, prototyping, testing, communicating the results of the test, and redesigning the solution (NRC, 2012; Psycharis et al., 2023). Chatzopoulos et al. (2024) noted a growing body of evidence supporting EDP as an effective interdisciplinary meta-framework for the foundation of project-based learning in higher education. DT has similar iterative phases to EDP but places more emphasis on empathy and ideation than EDP does (Saseendran & Thomas, 2025).

The IDPF utilizes a combination of the Engineering Design Process (EDP) and Design Thinking (DT) integrated together to create a comprehensive methodology for educators to help students bridge the gap between STEAM and the design challenges they face as they engage in STEAM project-based learning experiences. The GenAI Agent used in Step A in IDPF methodologically aligns student-centered STEAM activities to the EDP and DT phases of the process, allowing for the more substantive integration of the engineering component of the entire Design Thinking process versus just reference to it. While all components of IDPF are integrated through the use of EDP and DT, it is the substantive way that the engineering component is linked to the overall STEAM project-based learning experience that identifies the IDPF as an innovative methodology for educators to utilize. Researchers have documented this: Chatzopoulos et al. (2024) found increased engineering design outcomes in higher education students by participating in EDP structured projects. Psycharis et al. (2023) found that the use of explicit EDP helped to expand student's ability to develop computational thinking. Furthermore, the IDPF also applies the fundamental logic and circular/iterative sequences of design at the meta-level; however, there is no current amount of research that supports the meta-level application of this type of framework to create/design AI-enhanced/STEAM-based environments.

2.6 GenAI in Education: Capabilities and Persistent Limitations

The introduction of GenAI tools in educational settings has proven to have multiple applications related to the delivery of high-quality educational practice (i.e., lesson planning, content generation, assessment, differentiated instruction) (Kalogiannakis & Sotiropoulos; 2026a; 2026b; Kasneci et al., 2023; Xenakis A, 2005; Mollick & Mollick, 2024). The OECD Digital Education Outlook report of 2026 confirms that GenAI has the potential to support learning when used according to pedagogical principles, with 37% of educators indicating their use of AI-based tools in their work (OECD 2026). A systematic review performed by Spasopoulos et al. (2025) reported that there are many hopeful applications of GenAI; however, significant gaps still exist, most notably in laboratory/experimental learning contexts. Sotiropoulos and Kalogiannakis (2025) also identified opportunities for scaffolded inquiry and resource generation, while Uğraş et al. (2025) found that while teachers see the value of GenAI tools such as ChatGPT, they strongly believe they must be moderated by pedagogy. Liu and Bates (2025) proposed a collaborative model in which GenAI and teachers work together to enhance student learning through their respective roles in using the GenAI tool(s).

Despite the many positive aspects of how GenAI can help educators improve the quality of educational delivery (as referenced above), the validity of GenAI output remains an ongoing issue for educators. AI-generated lesson plans tend to lack context specificity, cultural responsiveness, and/or nuanced pedagogical decision-making (Clark & Van Kessel, 2024; Moundridou et al., 2024). Furthermore, pre-service teachers with a solid knowledge base in content identification were better positioned to analyze AI-generated lesson plans and identify the limitations of the AI (Kaplan-Rakowski et al., 2023).

Assessment design presents particular challenges: rubric-based approaches such as the LUPDA model (Lin et al., 2025), SSI-based STEAM rubrics (Kim & Martin, 2024), and NGSS-aligned rubrics (Achieve Inc., 2016; NGSS Lead States, 2013) offer validated instruments, but Li (2024) identified assessment innovation as a persistent challenge in STEAM implementation. These results support the promise of the IDPF, which suggests that GenAI is a very powerful generative tool. However, in order to satisfy the standards of quality demanded by educational practice, the output produced by GenAI (such as assessment design) will require a systematic form of human evaluation.

2.7 Conceptual Framework Development Methodology

The IDPF was created through the idea-based construction of a theoretical framework (IDPF) rather than from an empirical investigation. The framework's components were selected and integrated based on three criteria. Firstly, each component must support an identified documented problem in the design of STEAM projects found in research literature (e.g., insufficient interdisciplinary expertise, inconsistent quality of AI-generated content, and lack of adequate inclusive design). The second criterion is that there was independent empirical evidence supporting the effectiveness of each component for at least 1 of the characteristics of learning in an educational environment. Lastly, the third criterion was to ensure that the components were non-redundant, meaning all agree on their contribution to the design cycle.

The five frameworks that were chosen based on these criteria were TPACK for sharing expertise among stakeholders; PCK for evaluating the quality of the various aspects of the project; UDL 3.0 as the principle of inclusive design; EDP and DT for structuring student activities and the iterative nature of the framework itself; and SAMR for helping determine the selection of purposeful technology use. There were other frameworks considered to be potentially relevant (e.g., Community of Inquiry [CoI] Model, ADDIE instructional design methodology); however, these were not included: CoI focuses on the development of the community of inquiry process rather than how well a project was designed, and therefore would not apply here, nor would ADDIE's focus on the assessment of instructional design processes and does not incorporate multi-actor expert review or discipline-specific mapping.

The literature informing this synthesis was identified through targeted searches in Scopus, Google Scholar, and ERIC using combinations of keywords including "STEAM project design," "GenAI lesson planning," "TPACK," "UDL," "Engineering Design Process," and "AI education." The search prioritized peer-reviewed empirical studies published between 2020 and 2025, supplemented by foundational theoretical works. We acknowledge that this approach does not constitute a systematic review and that other framework combinations may warrant investigation; the rationale presented here is intended to make the construction logic transparent and open to scrutiny.

3. The Proposed IDPF Framework

The Interactive Design Process Framework (IDPF) is an iterative process with four steps for designing a STEAM project collaboratively, based on Universal Design for Learning (CAST, 2024). It consists of three actors: (1) a teacher in the classroom who leads and manages the project design and has final decision-making authority; (2) a Generative Artificial Intelligence agent temporarily embodied in general-purpose tools like ChatGPT, Claude, or Gemini that provides a collaborative partner in the design process, structural analysis for the project, and advice for the teacher about pedagogy; and (3) teacher experts in the content of the project who will be able to provide both content-specific and pedagogical feedback on specific aspects of the design of the project. The IDPF's four stages of design include: Step A: Co-design; Step B: Analysis; Step C: Expert Review; Step D: Synthesis (see Figure 1 above). After completion of Step D, if substantive changes are made, the teacher may continue iterating between Steps B–C–D. Table 1 compares the IDPF with other individual and established approaches for designing STEAM projects and school activities. Approaches such as Project-Based Learning, Engineering Design Process, and Design Thinking provide comprehensive student-facing methodologies but pertain only to the student learning process and not the quality assurance and design of the project. The IDPF operates at this earlier stage: how the teacher designs the project before students encounter it.

Table 1 illustrates the established approaches to STEAM project design and the positioning of the IDPF. The authors did not identify, in the literature reviewed, a single approach that addresses all five elements of the IDPF as a unified framework. In other words, an approach that incorporates the following five elements together:

(a) A structured, iterative, four-step cycle that has clearly defined roles;
(b) Three separate actors, including independent human expert reviewers;
(c) Dual-dimension analysis based on PCK grounded at the Review Component;
(d) Explicit STEAM discipline mapping that operationally drives the assignment of experts;
(e) UDL 3.0 as the foundational design principle of each project.

In addition, it is important to note that the comparison made here focuses on the project design phase (how the teacher creates the project and assures its quality for eventual implementation in the classroom) and does not consider the student-facing learning methodology at this time.

Recent research conducted on design in education supported by GenAI provides empirical evidence for the underlying mechanisms for each of the IDPF's steps and actors. While the IDPF is an original concept, its independent parts are grounded in previously validated principles of research/evidence.

Step A (Co-Design): The success of teacher-GenAI collaboration is largely determined by both the overall quality of the prompts, in addition to the teacher's pedagogical content knowledge (PCK). Based on research conducted by Peikos and Stavrou (2025) that considered various PCK specific elements included in initial prompts provided, as well as the use of layer prompt strategies used with ChatGPT generated lesson plans, demonstrated that the incorporation of additional PCK specific elements incorporated into initial prompts and then the use of layer prompt strategies improved the quality of ChatGPT generated science lesson plans.

Nyaaba et al. (2025) confirmed the idea that theory-based design frameworks provide better yields from the use of Culturally Responsive Lesson Planner (CRLP) than generic ones and also allow teachers to work together collaboratively. There are two case studies for chemistry provided by Blonder and Feldman-Maggor (2024) and Feldman-Maggor et al. (2025) showing teachers refining ChatGPT outputs iteratively to correct misconceptions through error analysis, which is an example of the collaborative design process defined in Step A.

Table 1. Established approaches to STEAM project design and the positioning of the IDPF

Approach	Core Logic	Actors	Iterative	Expert review	PCK	STEAM	UDL
Engineering Design Process (EDP)	8-phase problem-solving cycle: identify, research, develop, select, prototype, test, communicate, redesign	Teacher + Students	Yes (redesign loop)	No	No	Implicit (Engineering)	No
Project-Based Learning (PBL)	Students investigate authentic driving question over extended time	Teacher + Students	Partial (revision)	No	No	Possible	No
Design Thinking (DT)	Empathize, Define, Ideate, Prototype, Test; human-centered	Teacher + Students	Yes (non-linear)	No	No	Possible	No
5E Instructional Model	Engage, Explore, Explain, Elaborate, Evaluate	Teacher + Students	Partial (evaluate)	No	Implicit	No (science)	No
ADDIE	Analyze, Design, Develop, Implement, Evaluate	Instr. designer	Yes (evaluate)	Possible (SME)	No	No (generic)	No
Understand by Design (UbD)	Backward design: results, evidence, learning plan	Teacher (solo)	No (linear)	No	Implicit	No	No
GenAI-assisted planning	Teacher prompts GenAI to co-create lesson plans	Teacher + GenAI	Varies (ad hoc)	No	Varies	No	No
IDPF (this paper)	**4-step cycle: Co-Design, Analysis, Expert Review, Synthesis; RC decomposition**	**Teacher + GenAI + Experts**	**Yes (4-step, repeatable)**	**Yes (Step C, targeted RCs)**	**Yes (Step B, dual-dim.)**	**Yes (operational)**	**Yes (foundation)**

Step B (Analysis). Moundridou et al. (2024) took the initiative to categorize GenAI tools based on how they can be used throughout various phases of Inquiry-Based Learning Framework (IBL). They found that by matching specific requests to generative AI tools (rather than using general-purpose tools), they would see much better overall educational outcomes. The phase-specific match to Generative AI tools also aligns with what IDPF does with projects when componentizing them: when creating a review component, identify which STEAM discipline it belongs to and which review dimension it belongs to.

According to the OECD (2026), many teachers feel they are delivering lessons on material outside their subject areas. There needs to be a definitive way to identify which component of a project calls for a professional's expertise so that those experts can be involved.

Step C – Expert Review. As noted in the findings of Großmann et al. (2025), a study of 60 lessons generated by ChatGPT revealed varying degrees of quality and that the model was unable to provide appropriate responses to student inquiries. This provides a clear example of how expert validation is needed beyond the inputs of a single teacher generating the lesson. Karpouzis et al. (2024) described the Road-STEAMER project in which GenAI was designed to serve both as a teaching assistant and as an expert for subject matter expertise; however, as demonstrated by Abdulmajid et al. (2025), the expert that is simulated through AI can never replace the judgement of human expertise in a particular subject area, even when the use of AI enhances visual analytic skills but lacks the understanding of symbols and historical context. At the policy level, the European University Association (EUA, 2025) has endorsed the idea that human input based on academic expertise and peer review is critical to ensuring quality in higher education. Synthesis (Step D): Research on teacher agency collaborating with AI indicates that teachers are the ultimate arbiters of determining curriculum and evaluating AI generated content for integration into classroom instruction. Mishra et al. (2023) suggest that in the era of GenAI, that teachers must develop proficiency for engaging with AI, thus allowing them to play a key role as critical synthesizers.

Additionally, Liu and Bates (2025) recommend using a collaborative approach through structured processes when combining AI-generated lesson plans with the review of lesson plans completed by teachers for the purpose of increasing both the quality of the lesson plans as well as the teacher's trust in the AI-generated content thereby reinforcing the importance of teacher agency in the design process. Cooper's (2023) research study that determined while ChatGPT provides reasonably accurate science lesson plans; these lesson plans require significant teacher revisions in order to effectively use them within their classrooms. Other studies (Kasneci et al., 2023) identified use for GenAI by utilizing it as an agent for content generation and automated feedback; however, there are significant limitations associated with GenAI and its ability to understand contextually.

According to a major EEF-funded randomized controlled trial examining whether instructors would continue to use ChatGPT beyond the duration of the study (EEF, 2024), 31% of practitioners utilized ChatGPT to assist in reducing the overall time they spent planning their lessons. Further, while the study demonstrated the positive impact of the use of ChatGPT in the classroom by teachers, they also highlighted that quality assurance around the use of AI-created lessons in the classroom is needed. These findings reinforce the need for a quality assurance framework for AI-generated lessons. The IDPF is designed to establish quality assurance processes for the use of AI-generated lessons in the classroom setting, with at least a multi-actor structure for the IDPF to meet the requirements of quality assurance.

Subject-Area Expert Reviewers. Anderson and Tully (2024) found that STEM teacher collaboration in curriculum design significantly improves content accuracy and pedagogical alignment. Blonder and Feldman-Maggor (2024) showed that no single teacher possesses sufficient PCK across all STEAM domains, a gap that the IDPF addresses by distributing review across multiple specialists. This multi-expert approach is consistent with the consensus model of PCK (Carlson et al., 2019), which conceptualizes PCK as collectively held and refined through professional exchange.

Taken together, this evidence suggests that effective GenAI-assisted educational design benefits from: theory-informed prompting (Peikos & Stavrou, 2025; Nyaaba et al., 2025), iterative teacher-AI dialogue grounded in PCK (Blonder & Feldman-Maggor, 2024; Feldman-Maggor et al., 2025), systematic tool-function alignment (Moundridou et al., 2024), independent human expert verification (Großmann et al., 2025; EUA, 2025), and teacher agency as the synthesizing force (Mishra et al., 2023). The IDPF integrates these empirically supported mechanisms within a single, coherent four-step framework.

3.1 Step A: Co-Design

The first step involves creating an ongoing dialogue between the classroom teacher and a GenAI agent. In this case, as the first step in co-designing a STEAM project, the classroom teacher describes the educational context by providing information regarding the educational context (i.e., grade level, number of students participating, curricular guidelines/constraints, available resources), intended STEAM theme, and specific concerns/areas of uncertainty the classroom teacher may have related to the proposed STEAM project). Importantly, within the educational context (see example below), the classroom teacher is to share both their content-related goals and their uncertainty regarding methodology, assessment, and/or content (e.g. gaps).

The GenAI agent, in return, not only serves as a generator of content but also functions as a pedagogical advisor, generating recommendations related to multiple dimensions of pedagogy, including (but not limited to) project-based learning and inquiry-based learning; how to structure the learning activities (e.g., sequencing, timing) for the entire STEAM project; how to organize and group students based on assigned roles and when/how to rotate student groups; assessment method selection (e.g., rubric, formative assessment tools, and/or self-assessment) for each of the STEAM project learning activities; implementing differentiated instruction with respect to UDL principles (i.e., multiple means of representation, engagement, and action) and determining which technological tools to use and how to classify those tools based on the SAMR model (Puentedura, 2006). Following the GenAI agent's generation of project recommendation data, the teacher reviews the recommendations, asks clarifying questions, requests modifications, etc. These interactions may take place over several iterations (cyclical) until sufficient project materials are generated to create a draft of the project. At the conclusion of Step A, the draft contains a driving question/design challenge that has been articulated in EDP/DT format, a comprehensive outline of how the various activities indicated on the draft relate to the STEAM disciplines, explicit learning outcomes, timeline, student-facing instructional materials (worksheets/templates/card), an assessment rubric with clearly specified criteria and levels of performance, differentiation strategies, and a list of recommended tools along with brief instructions as to how to use each of those tools. All activities have been drafted using UDL, with the goal of providing multiple types of representation for information, allowing for multiple ways for students to express their learning, and engaging students through providing choices and making relevance for them.

Illustrative prompts (**Table 2**) are provided for teachers to use in engaging the GenAI agent in any of the three IDPF roles through each stage of the IDPF cycle; the examples given are not exhaustive but should serve as suggestions for teachers to modify to their own personal circumstances.

Table 2. Illustrative GenAI prompts by role

GenAI Role	Step	Illustrative Prompt
Co-Designer	A	I teach 5th grade, 25 students. I want a 3-week STEAM project about a vegetable garden. Suggest PBL with EDP phases.
Co-Designer	A	For Activity 3, suggest UDL alternatives for students who struggle with measurement.
Co-Designer	A	Classify the tools you recommended using the SAMR model.
Analyst	B	Decompose the project into Review Components. Tag each RC with the STEAM discipline, content, and pedagogy dimensions.
Analyst	B	For each RC, generate a review brief with context, checklist questions, and UDL considerations.
Synthesizer	D	The biologist suggested radishes instead of tomatoes. The math teacher recommended scale 1:10. Update the project.
Synthesizer	D	Present all expert feedback in a decision table: RC, expert, recommendation, impact, and suggested action.

3.2 Step B: Structural and Semantic Analysis

After producing a total draft of the project using a GenAI agent to create the required components, the project is deconstructed into smaller segments of the design called 'Review Components' (RCs). RCs are the smallest coherent parts of a project that can be left for expert review in a given content area. An RC can contain one activity, a worksheet, a rubric criterion, a group of learning objectives from one discipline, or a selection of a methodology. The breakdown follows three criteria of the RC: (a) Disciplinary Coherence – RCs should contain items from only one STEAM discipline, so an expert can be assigned accordingly; (b) Evaluative Independence – An expert will not need to read an entire project to complete the evaluation of each RC. They should be able to evaluate each RC independently; and (c) Pedagogical Completeness – RCs should contain enough context (target age, learning objectives, prior activities) related to that student and teacher so an expert can provide feedback on both the content and the method of instruction being used. Each RC is classified or 'tagged' in two different areas; content dimension (the STEAM discipline(s) that it is primarily associated with and what type of verification is needed in relation to the content) and pedagogical dimension (what specific aspects of what type of teaching method, activity design, scaffolding, or assessment require evaluation). Dual dimension tagging is unique to the instructional design process of the IDPF as can be seen in the intersections of the PCK framework described by Otto and Everett (2013).

Conventional review processes have traditionally only looked at the accuracy of the material. The IDPF understands that subject-area experts have specific knowledge (CK) about their particular area of study as well as pedagogical knowledge (PCK). An example would be a biologist can identify both that a plant-related claim is inaccurate, but also that a specific age group will have a common misconception about that claim and that an experiment, rather than a lecture, is the best method to address that misconception. This two-dimensional framework will assure that PCK is generated and retained.

Each Review Component (RC) will have a review brief generated by the GenAI agent. A brief is concise and gives the subject expert a sufficient understanding of the RC, such as what was developed, who the intended audience is, and what the purpose of the RC is. The review brief will also include questions related to the material's content (CK) and pedagogical aspects (PCK), as a checklist, as well as information on any UDL implications. Each subject expert will only be responsible for reviewing RC's specific to their area of expertise and it is estimated that it should take each expert 10-15 minutes to complete a review.

3.3 Step C: Expert Review

The subject area educators will only acquire the relevant review notes from their concentration and not from the entire copy of the project draft. Targeted engagement allows busy professionals to maximize their time constraints and provides the project with authentic disciplinary knowledge. Three criteria are used to identify expert reviewers: (a) the expert reviewer must have a subject area match; (b) the expert reviewer should have instructional experience with the age group of students; and (c) the amount of time required to complete a structured review brief is limited to 10 - 15 minutes per review. Expert reviewers will review the assigned review briefs, answer the questions, and may provide additional comments and/or suggestions.

If multiple subject area experts provide conflicting recommendations, the classroom educator becomes the ultimate arbitrator. The classroom educator's contextual insight into student, school culture, and resource availability provide the context to determine resolution of the conflicting recommendations. During Step D, the teacher will see the conflicting recommendations of the subject area experts represented on a decision table, enabling them to select the best course of action for their classroom.

The Teacher TPACK model maintains the teacher's professional autonomy while ensuring that expertise in a given field of study contributes to the final design, rather than controlling it.

Depending on the number of experts involved in the consultation process, there may be multiple disciplines represented in the design phase. For example, if there is an elementary school STEAM project related to a vegetable garden, the content related to plants and their needs, and seasonality, would be reviewed by a biologist; measurement and graphing activities would be reviewed by a mathematician; artistic design/aesthetic elements would be reviewed by an art educator; and overall methodology, group structure, and modes of assessment would be reviewed by a pedagogical consultant. Each expert reviewer has contributed to various parts of the TPACK model: the biologist will help to enhance the CK and PCK components while the pedagogical consultant enhances the PK and PCK components.

3.4 Step D: Synthesis and Finalization

The final step of the cycle involves the GenAI agent gathering the feedback received from all experts and presenting it to the classroom teacher in a structured decision table format. The teacher then reviews each piece of feedback to determine how (if at all) they will respond to each piece: accept, reject, or modify. After the teacher responds, the GenAI agent will make the necessary changes to the project draft, tracking all changes made to ensure cross-component coherence. At this point, the teacher retains complete professional judgement as to whether to accept the expert's recommendations based on their contextual knowledge of the specific classroom, student, and practical considerations. If the synthesis produces significant changes to the original draft, then any of the revisions may be sent back through the cycle for re-analysis at Step B, or RCs may be sent back to expert reviewers for a second review. However, in practical application, generally one or two iterations will be enough to develop a quality project.

3.5 STEAM Discipline Mapping

The IDPF requires every activity in a STEAM project to be explicitly connected to the STEAM disciplines it addresses. In frameworks such as EDP, discipline mapping serves primarily as a classification tool. In the IDPF, however, mapping performs a functional role: during Step B, when the GenAI agent decomposes the project into Review Components, the STEAM discipline tags assigned to each RC determine which expert is assigned to review it. This mechanism ensures that every content claim, pedagogical choice, and assessment element is reviewed by an individual with the appropriate disciplinary knowledge. The GenAI produces the STEAM Mapping Table as part of the Step A Draft, which is then validated during the Expert Review in Step C. An example is provided in **Table 3**.

Table 3. STEAM discipline mapping for vegetable garden project

Activity	S	T	E	A	M
What does a plant need?	✓				
Vegetable cards: research & choose	✓				
Measure & draw garden plan (1:10)		✓	✓		✓
Prepare beds & plant	✓		✓		
Design garden labels		✓		✓	
Record & graph plant growth	✓				✓
Presentation				✓	
Reflection (3 stars & 1 wish)					

Note: S = Science, T = Technology, E = Engineering, A = Arts, M = Mathematics.

3.6 TPACK Integration in the IDPF

According to its three actors, the IDPF allocates TPACK growth among several sources (see **Figure 2**). The classroom instructor starts with an existing level of PK and varies greatly in the amount of CK they possesses in relation to the STEAM disciplines. Each of the TPACK growth source entities contributes to each of the three zones of the TPACK model via the method of co-designing. The subject-area expert properties improve Content Knowledge (CK) by verifying facts and correcting erroneous perceptions held by teachers, and improve Pedagogical Content Knowledge (PCK) by recommending teaching methods and correcting how students work within their discipline. The generator AI source improves Technological Knowledge (TK) via introducing and scaffolding the use of various technology tools, improving Technological Content Knowledge (TCK) via recommending specific tools for specific application areas (e.g., Google Sheets for longitudinal student growth data), and improving Technological Pedagogical Knowledge (TPK) via recommending specific technology applications for pedagogical practices (e.g., Book Creator for multimodal, UDL aligned student portfolios). The classroom instructor creates TPACK at the intersection of the three sources to create a comprehensive, contextual design. To our knowledge, the conversions of the three source participants to the TPACK matrix appears to be novel in the existing literature; applications of TPACK within the context of teacher and AI tool typically do not have three sources verifying the development of TPACK across multiple sources via independent zones of TPACK (Blonder & Feldman-Maggor, 2024; Celik, 2023)

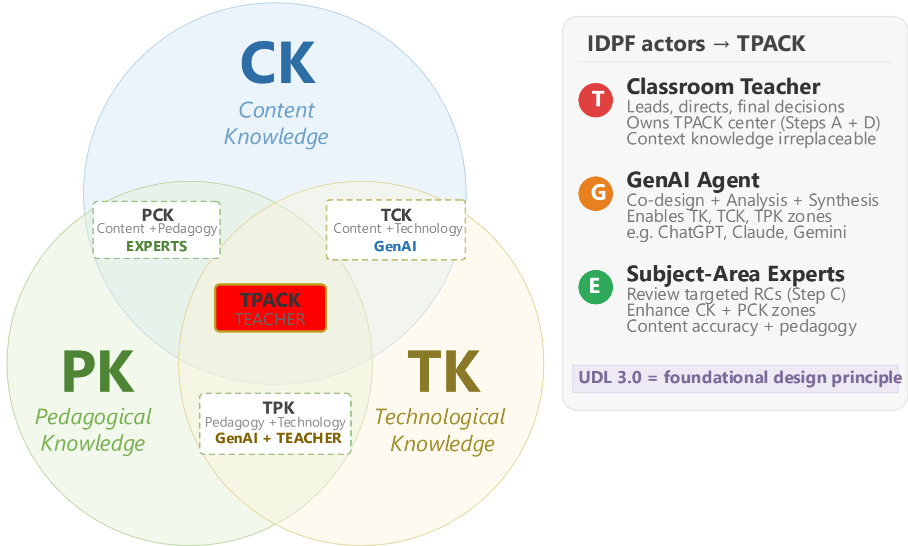

Figure 2. TPACK-IDPF actor integration. The three IDPF actors map to distinct TPACK zones: Experts enhance PCK (content + pedagogy); GenAI enables TCK and TPK (technology pairings); and the Teacher synthesizes all at the TPACK center. UDL 3.0 serves as the foundational design principle.

3.7 SAMR: Purposeful Technology Selection

An ongoing concern with incorporating technology into education is that educators often view technological resources as a direct replacement for static resources from other media (Boice et al., 2024; Mishra et al., 2023). All educators pointed out the lack of exploitation of the innovative possibilities created through technology in STEAM, and, thus, they utilized it at a basic level in terms of creating a possible learning experience. One way the IDPF is addressing this is to categorize each of the suggested tools identified by IDPF staff by level of the SAMR Taxonomy developed by Puentedura (2006), which describes four distinct levels: Substitution, Augmentation, Modification, and Redefinition. Research indicates that by understanding the level at which the educator is utilizing a technology, he or she will make more informed decisions regarding the integration of that technology (Mishra et al., 2023). Therefore, by explicitly referencing the SAMR classification of recommended tools in Step A, the IDPF will strengthen the rationale for selecting technology for educators' use; however, additional research is needed to determine whether an enhanced understanding will translate into classroom practice. It should be noted that the SAMR model (Puentedura, 2006) has not been published in a peer-reviewed journal and has received criticism for its linear hierarchy and limited empirical validation (Hamilton et al., 2016). Nevertheless, we include it here as a practical classification heuristic rather than a theoretical framework, acknowledging that its primary value lies in prompting educators to reflect on the level of technology integration rather than prescribing a progression. An example of SAMR classification is presented in **Table 4** below.

Table 4. SAMR classification of recommended IDPF tools

Tool	SAMR Level	Function in IDPF
Google Docs	Substitution	Draft writing (substitutes paper)
Canva	Augmentation	Enhanced worksheets, role cards
Google Sheets + charts	Modification	Auto graphing, real-time data
Book Creator	Modification	Multimedia student portfolios
GenAI (ChatGPT/Claude)	Redefinition	Co-design agent (new task)
Miro / FigJam	Redefinition	Visual RC decomposition
Google Forms	Augmentation	Structured review briefs

4. Roles and Responsibilities

Figure 3 shows the 'IDPF process flow' also showing each actor's participation in the four-step process.

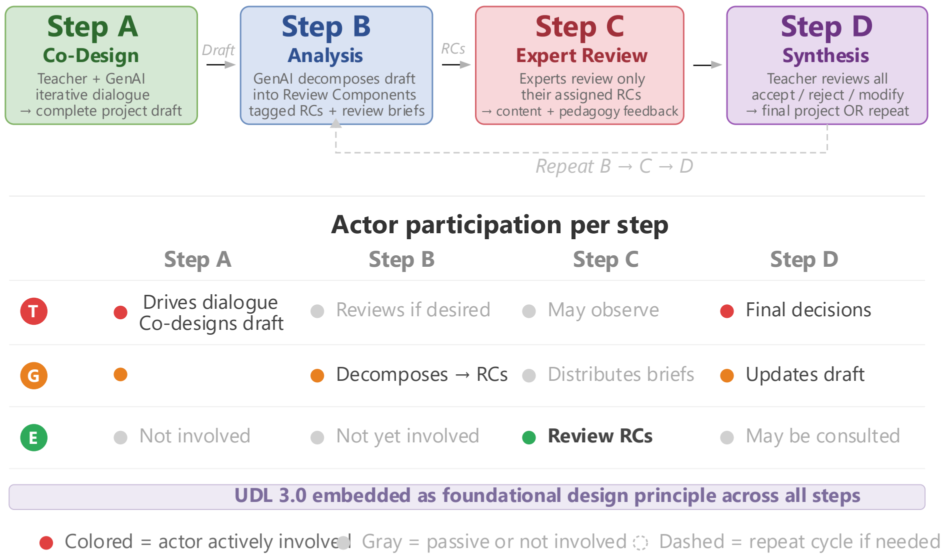

Figure 3. IDPF process flow and actor participation. Top: the four-step iterative cycle showing what passes between steps (draft, RCs, feedback). The dashed arrow indicates that Steps B–C–D may repeat. Bottom: actor participation matrix showing who is actively involved (colored) or passive (gray) at each step.

The arrows between steps identify what is transmitted - Step A identifies the production of a complete draft of the project, Step B decomposes the complete draft of the project to tagged Review Components (with associated structured Review Briefs for the Review Components), Step C generates expert feedback on the content accuracy and pedagogy of the materials (as noted in Step D but with different outcomes), and finally Step D synthesizes all input from the Reviewers to create a final project or repeats Steps B - C - D as required. The Actor Participation Matrix (located at the bottom of Figure 3) clarifies that only the Teacher and the GenAI are involved in both Steps A and D (Co-design and Synthesis), the GenAI leads Step B (Analysis) and Subject Area Experts are the Primary Actors in only Step C (Review). Table 5 below provides additional detail of the responsibilities of the respective actors and their respective TPACK contributions.

Table 5. Roles, responsibilities, and TPACK contributions per IDPF step

IDPF Step	Classroom Teacher	GenAI Agent	Subject-Area Experts
A: Co-Design	Describes context, goals, constraints, uncertainties; reviews and directs; makes decisions	Suggests methodology, activities, assessment, UDL, tools; generates draft; iterates with teacher	Not involved
B: Analysis	Not directly involved (reviews output if desired)	Decomposes draft into RCs; tags by STEAM discipline + review type; generates structured review briefs	Not yet involved
C: Expert Review	Not directly involved (may observe)	Distributes briefs; may provide context to experts	Review assigned RCs; provide content + pedagogical feedback; complete brief checklists
D: Synthesis	Reviews all feedback; accepts / rejects / modifies; makes final decisions	Compiles feedback into a decision table; updates the draft; ensures coherence	Not directly involved (may be consulted)
Primary TPACK contribution	***TPACK (synthesis center) — integrates all inputs into contextually appropriate design***	***TK, TCK, TPK — technology enabler + pedagogical advisor***	***CK, PCK — content accuracy + content-pedagogy enhancer***

5. Illustrative Example: A Primary School Vegetable Garden

To illustrate how the IDPF would be used in practice, we present the following hypothetical scenario. A 5th-grade teacher (Maria) who has been teaching for 8 years and wants to design a 3-week STEAM project with her 25 students based on the theme of Building a Vegetable Garden at the School. Maria has strong pedagogical knowledge and general classroom management skills, moderate knowledge of primary mathematics and environmental education, but limited knowledge of plant biology, soil science, and practical geometry (scale drawing). She has no prior experience with STEAM project design methodology or GenAI tools.

5.1 Step A: Co-Design

As Maria contemplates her own context of uncertainty within her assessment and methodology (on which to base her decision about whether to use the GenAI agent as her teaching consultant), she looks for ways to define her teaching practice: could Project-Based Learning (PBL) be her overall methodology, or can she find a more effective overall methodology? After talking with the GenAI agent, the GenAI agent recommended using Project-Based Learning methodology as a practice for this project; thus, the GenAI agent has suggested PBL's driving question for this project will be "Can we create a garden that will provide us with two different kinds of vegetables before the end of the school year?" This driving question meets EDP as it provides problem-based learning (PBL) opportunities, provides time and space limits, allows for researching and designing a solution to the problem, and provides a prototype for testing the solution once it is designed and developed (the product is the garden).

Through a continuous process of sharing in dialogue with one another, Maria and the GenAI agent co-created the following (in iteration): a three-phase project timeline by the phases of creating the project (research, developing, and observing); groups and rotating roles in groups; 8 activities provided at each phase of the STEAM disciplines; 3 worksheets (all designed following UDL principles); 1 assessment rubric for assessing each STEAM product/learning experience developed by the class provides a single set of criteria on which have been developed 3 levels of each criterion; differentiated teaching strategies for student supports for both students with learning difficulties and for advanced students; and a list of equipment/tools in which each piece (Canva for worksheets, Google Sheets for data recording,...) has been placed on the SAMR continuum.

5.2 Step B: Analysis

The GenAI agent takes the original draft and divides it into nine Review Components (RC) for the purposes of review and reconciliation. Each of these RCs has both a content and pedagogical dimension attached to them. For example, RC1 (activity entitled "What does a plant need?") has been coded with content related accuracy of biology, and for pedagogical appropriateness related to 40 minutes of direct instruction for 10-year-olds. Another example is RC3 (activity of creating scaled drawings), which has been coded for mathematical accuracy based on a 1:20 scale, and for scaffolding based on working individually or in pairs.

After the nine Review Components of the drafts are identified, each RC is assigned to the appropriate expert for review. Once the experts receive their RCs, structured review briefs are prepared for each expert containing specific checklist questions that pertain to both content and pedagogical dimensions of the assigned RCs.

5.3 Step C: Expert Review

The four experts completing the review of the materials each receive structured and targeted review briefs. There are a number of recommended changes made by the experts. The biologist has identified that the term "soil" is not accurate and that the term "nutrients" should be used. He/she suggests that tomatoes cannot produce fruit until after the school year ends (therefore growing fast growing alternatives such as radishes or rocket is preferred). The biologist also recommends changing the amount of direct instruction from 40 minutes to 15 minutes of elicitation followed by 25 minutes of hands-on experimentation with materials. The suggestions made by the mathematician include: Changing the scale of the drawings from 1:20 to 1:10; Using line graphs rather than bar graphs for presenting continuous data; Changing from individual to pair work when using graph/grid paper; Providing a pre-made data collection table. Developing a Labeling System for Fruit and Vegetable Gardens Using UDL Principles

From the previous activities, the expert notes that the timelines for the Label-making Activity need to be expanded from 1-2 hour sessions, the Art Educator should provide design templates for students to use (i.e., constrain creativity), and an Aesthetic Criterion should be included in the Rubric so all students will be able to use it when sending their respective labels home. The Pedagogical Advisor supports Changing Group Size from 6 Students Per Group to 5; Modifying the Driving Question to Include a Measurable Success Criterion; and Adding Alternative options for Expressing Ideas in the Rubric; to be compliant with UDL.

5.4 Step D: Synthesis

Maria synthesizes what the experts provided by reviewing all of the suggestions via a decision table from the GenAI Agent. Most of the suggestions are accepted with Professional Judgment applied to two in particular: Maria will continue using QR Codes on Garden Labels (because there are 5 tablets available at the school) even though the Art Educator Suggested Eliminating Them; and Maria will utilize 5 Vegetable Choices vs. 8 That Were Suggested by the Biologist (to reduce complexity for her students). The GenAI Agent updates the Project Draft to reflect the accepted Changes made to the original draft Project, while Still Maintaining Coherence between all components.

5.5 UDL Integration Across Activities

Table 6 illustrates how the use of UDL Principles will be integrated into all phases of this example Project, so all students, regardless of their Ability, Background or Learning Preference can have access to participate equitably in identifying their Labels.

Table 6. UDL integration across illustrative example activities

Activity	Representation (the what)	Action & Expression (the how)	Engagement (the why)
What does a plant need?	Speech + diagram + hands-on experiment	Worksheet OR drawing OR narration	Elicitation: 'What do you think?'
Vegetable cards: research & choose	Cards with image + text + difficulty symbols	Fill in card OR draw	Choice of vegetable (autonomy)
Measure & draw garden plan	Measurement outdoors then grid paper	In pairs, grid paper scaffold	Real-world measurement
Prepare beds & plant	Demonstration + instructions + hands-on	Physical action	'Gardener' role assignment
Design garden labels	Templates + examples + video	Template OR free design	Creative choice (colors, style)
Record & graph plant growth	Table + graph + photograph	Numbers OR sketch OR photo	'Scientist' role, routine
Presentation	N/A	Oral OR poster OR video OR Book Creator	Audience: other groups, parents

6. Discussion and Conclusions

6.1 Implications for Teacher Professional Practice

The IDPF has a range of important implications for teacher professional development. It positions the design of STEAM projects as an explicitly collaborative process (i.e., a collaborative endeavor involving both AI and human contributors) rather than something most teachers accomplish alone (or that would be happening in a vacuum). This mirrors existing theories about how teacher professional development is viewed by the literature as being fundamentally social and contextual (e.g., Penuel et al., 2007). One of the key requirements of the IDPF is for teachers to develop pedagogical prompting literacy – the capacity to articulate their pedagogical intent to an AI-based system, critically evaluate the output, and guide an iterative revision process based upon that evaluation. As there is already increasing usage of GenAI, this level of literacy will be a requirement for all teachers to acquire as a critical professional competency.

In addition, the model of distributed TPACK, as outlined in this framework, presents a principled approach to addressing an essential barrier to implementing STEAM projects: there is no teacher who possesses deep competence across all STEAM content domains and all technologies at the same time. Research independently supports the validity of each piece of the distribution; for example, subject-specific, collaborative engagement produces greater teacher PCK (Gözüm et al., 2022; Penuel et al., 2007), structured interaction with GenAI creates increased access to and ease of using technology (Liu & Bates, 2025; Spasopoulos et al., 2025), and the synthesis process employed by teachers ensures that projects are appropriate for their local contexts (Feldman-Maggor et al., 2025; Uğraş et al., 2025). Consequently, we hypothesize that the four validated approaches, when combined in a structured cycle, will result in a higher likelihood of producing quality STEAM projects than if the approaches were to be implemented by individual teachers. Nevertheless, the framework's effectiveness as an integrated whole awaits empirical confirmation.

The framework recommends a systematic way to address the quality gaps identified in recent research. According to Peikos and Stavrou (2025), ChatGPT has provided developmental inappropriate material when prompted through a variety of sophisticated prompters, as well as being unable to address student mischaracterizations of their work. The dual dimension Review Component (Step B) identified by the IDPF uses the relationship of the content knowledge (PCK) demonstrated by Otto and Everett (2013) as necessary for planning quality lessons. Structured expert feedback provided in Step C follows the same principle found by Großmann et al. (2025), which says human experts can increase the quality of AI-generated learning materials.

The integration of UDL at every level addresses the incongruity between inclusive intents and inclusive practices documented by Bequette & Bequette (2012) and Spyropoulou & Kameas (2024): evidence reveals that appropriately explicit prompting for differentiation results in designs that are more accessible (Peikos & Stavrou, 2025; Tomlinson et al., 2003). The IDPF does this systematically throughout the design process.

While many of these components are not guarantees against error in any project, they do provide multiple overlapping opportunities for quality assurance that are missing from unstructured, GenAI-assisted designs.

6.2 Limitations and Considerations

The Interactive Design Process Framework (IDPF) requires empirical validation before we can be confident in its efficacy as a conceptual framework. There are also several key questions that are yet to be determined: Is the quality of STEAM projects developed via the IDPF better than STEAM projects that are developed without using any form of AI assistance? Does using the IDPF streamline the amount of time that teachers spend designing STEAM projects without compromising project quality? What forms of professional development provide teachers with the most effective means of supporting them as they engage with the IDPF to create STEAM projects? In what way does the IDPF perform across different school contexts, especially in contexts where the teacher has limited expertise in their subject area or where access to Generative-AI (GenAI) tools is limited?

Lastly, there are critical considerations of equity and access. The IDPF assumes that schools will have access to a functional large language model (LLM), reliable internet access, and teachers familiar with technology in the classroom. These assumptions do not hold true for many schools. There may be ways to implement IDPF in schools with limited technology infrastructure moving forward. In practical terms, the IDPF requires the participation of multiple subject-area experts per project. In the illustrative example presented in Section 5, four experts reviewed nine Review Components, with each review estimated at 10–15 minutes per RC. This amounts to approximately 1.5–2.5 hours of total expert time per project, distributed across reviewers. While this represents a modest time investment for individual experts, identifying and coordinating willing professionals remains a logistical challenge, particularly in under-resourced settings. The surrogate GenAI pathway described in Section 6.4 offers one potential mitigation, but the framework's feasibility across diverse institutional contexts requires empirical investigation.

There are also some concerns regarding intellectual property, data privacy and ethics of using AI-generated materials in the classroom when looking at the IDPF. The IDPF's positioning of the teacher as the final decision-maker provides one response to the question of responsibility. Still, it may not fully resolve the ethical complexities of situations in which significant portions of a project's content are generated by an AI system.

6.3 Ethical, Privacy, and Intellectual Property Considerations

Concerns about intellectual property, ethical, and privacy-related issues arise with the introduction of GenAI into an educational design process. This requires explicit discussion on how data is shared with large language model (LLM) platforms during co-design, Step A. The classroom teacher will share context about their students, such as age, needs, and makeup of the classroom. The IDPF does not require that teachers provide specific identifying information about students; therefore, teachers must exercise caution to only provide context and not any student-level information that would allow anyone on other commercial LLM platforms to identify specific students; especially for those platforms that retain or otherwise use the data for model training. Schools and institutions that implement the IDPF should have protocols in place for the governance of student contextual data that comply with applicable laws, like the GDPR or other national regulatory frameworks.

Also still unresolved across jurisdictions is the copyright status of educational materials generated by GenAI; The IDPF identifies the classroom teacher as the author and decision-maker for the project; however, it's uncertain to what degree components developed by GenAI (e.g., activity descriptions, rubric elements, worksheet templates, etc.) will be afforded copyright protection. Teachers and schools must be aware of the fact that they may have generated materials using GenAI co-design, but they do not have the same protections under the law as those materials created solely by people.

Third, accountability within the IDPF is distributed but ultimately rests with the classroom teacher, who retains professional responsibility for all content delivered to students. The GenAI agent operates as a tool, not as a co-author with professional accountability, and subject-area expert reviewers provide advisory feedback without assuming responsibility for the final project. This accountability structure should be made explicit in any institutional implementation of the framework.

Finally, the potential for algorithmic bias in GenAI outputs–including cultural, linguistic, or disciplinary biases embedded in training data–should be acknowledged as an ongoing concern. The expert review process in Step C provides one mechanism for identifying such biases, but it cannot guarantee their complete detection. Future implementations of the IDPF should include explicit bias-awareness prompts during both the Co-Design and Expert Review stages.

6.4 Directions for Future Research

We recommend a few key areas for further exploration in the future. One option is for the implementation of a design-based research study including the IDPF (Interactive Design Process Framework) model across a range of different educational environments, with continued iterative improvement grounded in teacher input and the assessment of the quality of the IDPF-based projects generated by the teachers in each particular setting. Secondly, conducting a comparison study evaluating the outcomes of STEAM projects (science, technology, engineering, arts and mathematics) designed using the IDPF model versus STEAM projects designed without the support of an institutional framework, using independent expert panels to provide evaluations of each group of projects.

The most potentially impactful activity related to equitable access to education is the examination of alternative adaptations of the IDPF framework in settings where there are no content area experts available to work with. For example, many rural schools, low-enrolment / low-staffed educational institutions, or educational systems that lack the institutional support for interdisciplinary teaching and learning often lack the presence of a content area expert. In these settings, a tiered process could be implemented in which generative artificial intelligence (GenAI) serves as a surrogate reviewer in step C, producing an evidence-based evaluation of the work in question, with reference to source materials for the feedback provided. The alternative pathway represents the convergence of two emerging research patterns. The first is the research suggesting that providing GenAI with the necessary reference documents will significantly improve the quality and accuracy of GenAI responses, thus shifting from generic to domain-relevant (Peikos & Stavrou, 2025; Blonder & Feldman-Maggor, 2024).

GenAI literacy is being developed using multiple tools to compare outputs and cross-reference data as a way to manage hallucinations and better assure accuracy (Anh-Hoang et al., 2025). In this surrogate approach, there is a two-tiered validation process in which a GenAI tool is used to create the draft of the project (Step A). Then, a different tool with web search functionality (for example, Perplexity or a search-assisted LLM) would confirm the accuracy of the draft, identify areas where the drafts' accuracy may be in error, and provide factual citations for each review component in Step C. Triangulation across tools does not provide the contextual and pedagogical expertise and judgement as would occur in the judgment of an expert human reviewer.

However, it provides some estimate of expert-level verification of content and enables the identification of errors that an individual GenAI Tool might overlook. Research on strategies to minimize hallucinations supports the conclusion that the use of retrieval-augmented generation (RAG), multimodal approaches, and structured prompt engineering will materially reduce factual errors in LLM-generated outputs (Anh-Hoang et al., 2025; UNESCO, 2023). As has been the case, the teacher retains the ultimate authority and must use increased critical judgement for any decision in the absence of a human expert review.

The IDPF (Interactive Design Process Framework) is designed for tiered use (i.e., using human subject matter experts as a first option, with multi-instrument GenAI validation used as a potential fallback), allowing it to have broad applicability in different educational contexts while also maintaining its commitment to providing quality and accurate results. Future empirical research will investigate the types of conditions under which the surrogate pathway will generate results comparable to those produced by a thorough professional evaluation, and what types of training teachers require to be able to conduct a critical evaluation of the validity of the feedback they receive from GenAI.

In addition, longitudinal studies will look at whether or not using the IDPF over time will impact teachers' development of TPACK (Technology, Pedagogy, Content Knowledge) and whether or not teachers will be able to internalize the same expertise that was provided to them via GenAI's feedback, and via expert reviewers, initially. Additionally, further exploration will occur with regard to whether the framework can also be used to support STEM/STEAM student-facing learning activities through the integration of GenAI as a learning tool, while also creating a model for nested learning (i.e., GenAI supports the design of work and the work itself).

Finally, purpose-built GenAI agents will be developed that will be designed specifically to apply the concepts of the IDPF framework. At this time, the framework is being implemented through the use of general-purpose LLMs (i.e., ChatGPT, Claude, Gemini) through teacher-directed inputs.

In contrast to the traditional researchers' findings regarding the limited effectiveness of using AI-based GPT as a simple tool for teaching, emerging academic research has built on this idea. Research from Nyaaba et al. (2025) has detailed how a specific GPT model created using Culturally Responsive Teaching Theory, the Culturally Responsive Lesson Planner (CRLP) GPT, was able to demonstrate superior results in the areas of cultural competency, accuracy of assessment, and alignment with curriculum expectations compared to previous versions of GPT-4o.

Additionally, products like MagicSchool AI and Auto Classmate already offer pre-packaged (i.e., built on the GPT-4 architecture) lesson planning tools. If an IDPF-specific agent were created, it could embed within its architecture the four-step logic of the IDPF, as well as provide for generating (i.e., using prompts to create) a PCK-based review brief, tagging a STEAM discipline, and checking for compliance with UDL principles, thus reducing the number of prompts required by a teacher to generate lessons while increasing the consistency of quality from one use of the agent to the next (i.e., when scaffolding student learning across disciplines). Creating and validating this IDPF-specific agent represents an exciting future direction for translating the IDPF from a conceptual framework to a tool that is scalable and readily accessible to teachers.

Furthermore, researchers working with the IDPF could create a comprehensive case study showcasing its use from start to finish (i.e., starting with a teacher's prompt and documenting all four stages of implementation) through the creation of the final deliverable, which would include documentation of all GenAI interactions as well as review briefs, expert feedback, and teacher decisions made during this process. This companion paper would be lengthy and detailed in nature such that we envision it as a distinct resource from the original case study itself. Finally, we propose a comparative analysis that could evaluate STEAM projects completed using the IDPF and those created by teachers who have extensive experience creating STEAM projects without AI support. This evaluation would be conducted by independent expert panels using standardized rubrics for instructional quality (such as the NGSS-aligned rubric) and student learning outcomes (such as the LUPDA model) and would serve to compare the IDPF's effectiveness and document the areas in which it has added value as well as where traditional human-only design may remain superior in order to hold true to the IDPF's tentative nature regarding its claims while valuing the expertise of educators over all.

6.5 Concluding Remarks

The Interactive Design Process Framework (IDPF) proposed in this paper includes a structured, iterative, four-step approach to the collaborative design of STEAM projects in classrooms with a GenAI agent and subject area experts (reviewers). While other AI-assisted educational design approaches usually consist of a teacher and AI (the dyad) working within one discipline, the IDPF adds a third actor (subject area expert reviewers) into the iterative design cycle that is specifically created for interdisciplinary STEAM projects. Thus, the IDPF creates a three-actor structure with PCK-grounded analysis, UDL as the foundational principle, and EDP/DT as both an activity methodology and a meta-framework.

Additionally, the IDPF creates a theoretical framework that includes TPACK, UDL, and EDP/DT methodologies where each of the individual elements is supported by independent research. Furthermore, the IDPF creates a four-step cycle (Co-Design, Analysis, Expert Review, Synthesis) to allow for multiple and layered opportunities for quality assurance of products developed using AI in a systematic manner from the generation stage through to product evaluation by experts and the synthesis of the product through the professional judgment of the teacher.

A validated problem is addressed by each of the five components of the Framework: TPACK establishes distributed expertise; PCK provides a sound means for quality analysis; UDL ensures all students will be included; EDP/DT provides a structure for implementing the design; and SAMR will guide the selection of technology. The IDPF provides a vision of education as a cooperative venture between human teachers and AI, but is neither a techno-utopian nor techno-skeptical perspective. Instead, it is a pragmatic perspective that believes Generative AI should only be used for the activities for which it has been demonstrated to be effective (such as generating content, analyzing data, and organizing structured information) and that human expertise will continue to provide the essential elements that Generative AI cannot replace (such as contextual judgment, discipline depth, pedagogical wisdom, and the irreplaceable accountability of teachers). The framework does not guarantee that projects will be conducted without error; however, it provides a structured methodology that makes errors more likely to be identified and corrected through a methodical approach than through unstructured approaches. This framework is meant to provide a foundation for empirical research and to assist in developing a practice grounded in critical thought, evidence-informed methods, and inclusive design in STEAM projects.

Table of Abbreviations

Abbreviation	Full Term
CK	Content Knowledge
DT	Design Thinking
EDP	Engineering Design Process
GenAI	Generative Artificial Intelligence
IDPF	Interactive Design Process Framework
LLM	Large Language Model
PCK	Pedagogical Content Knowledge
PK	Pedagogical Knowledge
RC	Review Component
SAMR	Substitution, Augmentation, Modification, Redefinition
STEAM	Science, Technology, Engineering, Arts, Mathematics
TCK	Technological Content Knowledge
TK	Technological Knowledge
TPACK	Technological Pedagogical Content Knowledge
TPK	Technological Pedagogical Knowledge
UDL	Universal Design for Learning

References

Achieve, Inc. (2016). EQuIP rubric for science: Version 3.0. Next Generation Science Standards. https://www.nextgenscience.org/resources/equip-rubric-science

Abdulmajid, M., Alrashaidi, E., Alayar, M., & Alrashidi, N. (2025). AI as a pedagogical partner: Enhancing art history education and quality assurance through AI-generated critiques. *Quality Assurance in Education*. https://doi.org/10.1108/QAE-03-2025-0074

Anderson, J., & Tully, D. (2024). *STEM teacher collaboration in curriculum design*. In R. Tytler et al. (Eds.), Integrated STEM curriculum design. Springer.

Anh-Hoang, T., Tran, T., & Nguyen, N. (2025). Survey and analysis of hallucinations in large language models. *Frontiers in Artificial Intelligence, 8*, 1622292. https://doi.org/10.3389/frai.2025.1622292

Atlas, S. (2023). *ChatGPT for higher education and professional development: A guide to conversational AI*. University of Rhode Island. https://digitalcommons.uri.edu/cba_facpubs/548

Bequette, J. W., & Bequette, M. B. (2012). A place for art and design education in the STEM conversation. *Art Education, 65*(2), 40-47. https://doi.org/10.1080/00043125.2012.11519167

Blonder, R., & Feldman-Maggor, Y. (2024). AI for chemistry teaching: Responsible AI and ethical considerations. *Chemistry Teacher International, 6*(4), 385-395. https://doi.org/10.1515/cti-2024-0040

Boice, K. L., Alemdar, M., Jackson, J. R., et al. (2024). Exploring teachers' understanding and implementation of STEAM: One size does not fit all. *Frontiers in Education, 9*, 1401191. https://doi.org/10.3389/feduc.2024.1401191

Bryan, L., & Guzey, S. S. (2020). K-12 STEM Education: An Overview of Perspectives and Considerations. *Hellenic Journal of STEM Education*, *1*(1), 5–15. https://doi.org/10.51724/hjstemed.v1i1.5

Carlson, J., Daehler, K. R., Alonzo, A. C., et al. (2019). *The refined consensus model of pedagogical content knowledge in science education*. In A. Hume, R. Cooper, & A. Borowski (Eds.), Repositioning pedagogical content knowledge in teachers' knowledge for teaching science (pp. 77-94). Springer. https://doi.org/10.1007/978-981-13-5898-2_2

CAST (2024). Universal Design for Learning Guidelines version 3.0. https://udlguidelines.cast.org

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate AI-based tools into education. *Computers in Human Behavior, 138*, 107468. https://doi.org/10.1016/j.chb.2022.107468

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends, 66*, 616-630. https://doi.org/10.1007/s11528-022-00715-y

Chaitidou, M., Spyrtou, A., Kariotoglou, P., & Dimitriadou, C. (2018). Professional development in inquiry-oriented pedagogical content knowledge among primary school teachers. *The International Journal of Science, Mathematics and Technology Learning, 25*(2), 17-36.

Chappell, K., Hetherington, L., Juillard, S., Aguirre, C., & Duca, E. (2025). A framework for effective STEAM education: Pedagogy for responding to wicked problems. *International Journal of Educational Research Open, 9*, https://doi.org/10.1016/j.ijedro.2025.100474

Chatzopoulos, A., Xenakis, A., Papoutsidakis, M., Kalovrektis, K., Kalogiannakis, M., & Psycharis, S. (2024). *Proposing and testing an open-source and low-cost drone under the engineering design process for higher education*. In 2024 IEEE Global Engineering Education Conference (EDUCON) (pp. 1-7). IEEE. https://doi.org/10.1109/EDUCON60312.2024.10578677

Clark, C. H., & van Kessel, C. (2024). "I, for one, welcome our new computer overlords": Using artificial intelligence as a lesson planning resource for social studies. *Contemporary Issues in Technology and Teacher Education, 24*(2), 151- 183. https://citejournal.org/wp-content/uploads/2024/05/v24i2SocialStudies1.pdf

Cohen, N., & Bronshtein, O. (2025). Bridging disciplines: Exploring interdisciplinary curriculum development in STEM teacher education. *Education Sciences, 15*(8), 1064. https://doi.org/10.3390/educsci15081064

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. https://doi.org/10.1007/s10956-023-10039-y

Education Endowment Foundation (EEF). (2024). *Using Generative AI (ChatGPT) for KS3 Science Lesson Preparation: Teacher Choices Trial.* Education Endowment Foundation. https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/choices-in-edtech-using-generative-ai-chatgpt-for-ks3-science-lesson-preparation-2024-teacher-choices-trial

European University Association (EUA). (2025). *Assuring higher education's quality in the age of AI is no easy task.* https://www.eua.eu/our-work/expert-voices/

Feldman-Maggor, Y., Blonder, R., & Alexandron, G. (2025). Perspectives of generative AI in chemistry education within the TPACK framework. *Journal of Science Education and Technology, 34*, 1-12. https://doi.org/10.1007/s10956-024-10148-2

Google. (2026, April 7). *Do you agree with this claim: The fundamental problem is that GenAI does not have Pedagogical Content Knowledge (PCK), or the integration of content knowledge and pedagogical knowledge, necessary for teachers to determine whether a representation, explanation, or activity will be effective for a certain group of students (Lee Shulman, 1986 (Attached. It defines what PCK is)). Can you create PCK?* [Large language model]. https://gemini.google.com/share/1ef6d2a2cd4f

Gess-Newsome, J. (2015). *A model of teacher professional knowledge and skill including PCK.* In A. Berry, P. Friedrichsen, & J. Loughran (Eds.), Reexamining pedagogical content knowledge in science education (pp. 28-42). Routledge.

Gözüm, A. İ. C., Papadakis, S., & Kalogiannakis, M. (2022). Preschool teachers' STEM Pedagogical Content Knowledge: A comparative study of teachers in Greece and Turkey. *Frontiers in Psychology, 13*, 996338. https://doi.org/10.3389/fpsyg.2022.996338

Großmann, L., Koberstein-Schwarz, M., Krüger, D., & Krell, M. (2025). Lesson planning with ChatGPT for inquiry-based biology instruction – A(I) roll of the dice? International Journal of Science Education. https://doi.org/10.1080/09500693.2025.2567509

Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The Substitution Augmentation Modification Redefinition (SAMR) Model: A critical review and suggestions for its use. *TechTrends, 60*(5), 433–441. https://doi.org/10.1007/s11528-016-0091-y

Henriksen, D. (2014). Full STEAM ahead: Creativity in excellent STEM teaching practices. *The STEAM Journal, 1*(2), Article 15. https://doi.org/10.5642/steam.20140102.15

Herro, D., Quigley, C., & Cian, H. (2019). The challenges of STEAM instruction: Lessons from the field. *Action in Teacher Education, 41*(2), 172-190. https://doi.org/10.1080/01626620.2018.1551159

Kalogiannakis, M., & Sotiropoulos, D. (Eds.). (2026a). *Harnessing Generative Artificial Intelligence for Science Education.* IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-2913-0

Kalogiannakis, M., & Sotiropoulos, D. (Eds.). (2026b). *Generative AI-Enhanced Science Education From Early Childhood to Higher Education.* IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-9074-1

Kaplan-Rakowski, R., et al. (2023). *Generative AI and teachers' perspectives.* In Proceedings of AACE EdMedia.

Karpouzis, K., Pantazatos, D., Taouki, J., & Meli, K. (2024). Tailoring education with GenAI: A new horizon in lesson planning. In 2024 IEEE Global Engineering Education Conference (EDUCON) (pp. 1–10). IEEE. https://doi.org/10.1109/EDUCON60312.2024.10578690

Kasneci, E., Sessler, K., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kim, S., & Martin, S. N. (2024). Developing an evaluation rubric for planning and assessing SSI-based STEAM programs. *Research in Science Education, 54*, 457-482. https://doi.org/10.1007/s11165-023-10123-8

Land, M. H. (2013). Full STEAM ahead: The benefits of integrating the arts into STEM. *Procedia Computer Science, 20*, 547-552.

Li, J. (2024). Effective strategies for interdisciplinary integration in STEAM curriculum design. *Transactions on Social Science, Education and Humanities Research, 8*, 99-105. https://doi.org/10.62051/gvesha87

Lin, C.-L., & Tsai, C.-Y. (2025). LUPDA: A comprehensive rubrics-based assessment model for STEAM education. International *Journal of STEM Education, 12*, Article 570. https://doi.org/10.1186/s40594-025-00570-2

Liu, D., & Bates, T. (2025). A collaborative model for integrating teacher and GenAI into future education. *TechTrends*. https://doi.org/10.1007/s11528-025-01105-w

Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017-1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mehddi, F., Kazi, A. S., & Butt, A. I. (2025). From theory to practice: How STEAM professional development shapes teacher beliefs and perceptions about design thinking activities. *SAGE Open, 15*(1). https://doi.org/10.1177/21582440251355779

Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and Generative AI. *Journal of Digital Learning in Teacher Education, 39*(4), 235-251. https://doi.org/10.1080/21532974.2023.2247480

Mollick, E. R., & Mollick, L. (2024). Instructors as Innovators: A Future-focused Approach to New AI Learning Opportunities, With Prompts. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.4802463

Moundridou, M., Matzakos, N., & Doukakis, S. (2024). Generative AI tools as educators' assistants: Designing and implementing inquiry-based lesson plans. *Computers and Education: Artificial Intelligence, 7*, 100277. https://doi.org/10.1016/j.caeai.2024.100277

National Research Council. (2012). *A framework for K-12 science education.* National Academies Press. https://doi.org/10.17226/13165

Nyaaba, M., Zhai, X., & Aydin, H. (2025). Developing a theory-grounded AI tool for the generation of culturally responsive lesson plans. *Computers and Education: Artificial Intelligence, 8*, 100346. https://doi.org/10.1016/j.caeai.2025.100346

NGSS Lead States. (2013). Next Generation Science Standards. The National Academies Press. https://www.nextgenscience.org

OECD (2026). Digital Education Outlook 2026: Generative AI and Education. OECD Publishing.

Otto, C. A., & Everett, S. A. (2013). An instructional strategy to introduce pedagogical content knowledge using Venn diagrams. *Journal of Science Teacher Education, 24*(2), 391-403. https://doi.org/10.1007/s10972-011-9275-3

Papert, S. (1980). Mindstorms: Children, computers, and powerful ideas. Basic Books.

Peikos, G., & Stavrou, D. (2025). ChatGPT for science lesson planning: An exploratory study based on pedagogical content knowledge. *Education Sciences, 15*(3), 338. https://doi.org/10.3390/educsci15030338

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? *American Educational Research Journal, 44*(4), 921-958. https://doi.org/10.3102/0002831207308221

Psycharis, S., Iatrou, P., Kalovrektis, K., & Xenakis, A. (2023). The impact of the computational pedagogy STEAM model. In Lecture Notes in Networks and Systems (pp. 400-411). https://doi.org/10.1007/978-3-031-26190-9_41

Puentedura, R. R. (2006). Transformation, technology, and education. http://hippasus.com/resources/tte/

Sakamoto, Y., et al. (2024). *Cross-cultural evaluation of AI-generated math lesson plans.* In Proceedings of AIED 2024.

Saseendran, A., & Thomas, M. V. (2025). Design thinking in science and integrated STEM/STEAM education: A systematic review. *STEM Education, 5*(6), 1058-1101. https://doi.org/10.3934/steme.2025046

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14. https://doi.org/10.3102/0013189X015002004

Sotiropoulos, D., & Kalogiannakis, M. (2025). *Using GenAI tools to enhance teaching and learning in Science.* In Proceedings of the 14th International Conference New Perspectives in Science Education (pp. 293-301). Florence: Filodiritto-Pixel. https://doi.org/10.26352/L319_2420-9732

Spasopoulos, T., Sotiropoulos, D., & Kalogiannakis, M. (2025). Generative AI in pre-service science teacher education: A systematic review. *Advances in Mobile Learning Educational Research, 5*(2), 1501-1523. https://doi.org/10.25082/AMLER.2025.02.007

Spyropoulou, N., & Kameas, A. (2024). Augmenting the impact of STEAM education by developing a competence framework for educators. *Education Sciences, 14*(1), 25. https://doi.org/10.3390/educsci14010025

Sun, Y., et al. (2025). Exploring the role of GenAI in collaborative lesson planning. Journal of Computer Assisted Learning. https://doi.org/10.1111/jcal.70098

UNESCO. (2023). Guidance for generative AI in education and research. UNESCO. https://doi.org/10.54675/EWZM9535

Taşar, M. F. & Yılmaz Ergül, D. (2023). Technological pedagogical content knowledge (TPACK) in physics education. In M. F. Taşar & P. R. L. Heron (Eds.), *The International Handbook of Physics Education Research: Teaching Physics*, AIP Publishing, Melville, New York, pp. 1-1–1-30. https://doi.org/10.1063/9780735425712_001

Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted, 27*(2–3), 119–145. https://doi.org/10.1177/016235320302700203

Vlasopoulou, M., Kalogiannakis, M., & Sifaki, E. (2021). Investigating Teachers' Attitude and Behavioral Intentions for the Impending Integration of STEM Education in Primary School. In St. Papadakis and M. Kalogiannakis (Eds.), *Handbook of Research on Using Education Robotics to Facilitate Student Learning*, 235-256, Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-7998-6717-3.ch009

Uğraş, H., Uğraş, M., Papadakis, S., & Kalogiannakis, M. (2025). Innovative early childhood STEM education with ChatGPT: Teacher perspectives. *Technology, Knowledge and Learning, 30*, 809-831. https://doi.org/10.1007/s10758-024-09804-8

Voicu, C. D., Ampartzaki, M., Dogan, Z. Y., & Kalogiannakis, M. (2022). *STEAM implementation in preschool and primary school education: Experiences from six countries.* In M. Ampartzaki & M. Kalogiannakis (Eds.), Early childhood education: Innovative pedagogical approaches in the post-modern era (pp. 1-23). IntechOpen. https://doi.org/10.5772/intechopen.107886

Wardak, D., et al. (2025). Reshaping higher education designs and futures: Postdigital co-design with GenAI. Postdigital Science and Education. https://doi.org/10.1007/s42438-025-00595-4

Yakman, G. (2008). *STEAM education: An overview of creating a model of integrative education.* Purdue University STEM Colloquium.

Xenakis, A. (2025). AI-Based Models for Assessing STEAM Engineering Literacy for University Students: The Case of Digital Systems for Precision Agriculture. *Hellenic Journal of STEM Education, 4*(1), 1–9. https://doi.org/10.51724/hjstemed.v4i1.37