education sciences

Article

The Impact of AI-Generated Instructional Videos on Problem-Based Learning in Science Teacher Education

Nikolaos Pellas

Department of Primary Education, University of Western Macedonia, 53100 Florina, Greece; aff00192@uowm.gr

Abstract: Artificial Intelligence (AI) has gained significant prominence in science education, yet its practical applications, particularly in teacher training, remain underexplored. Specifically, there is a lack of research on AI's potential to support personalized professional development through automated analysis of classroom interactions and tailored feedback. As science teacher education requires skill development in complex scientific concepts within problem-based learning (PBL) contexts, there is a growing need for innovative, technology-driven instructional tools. AI-generated instructional videos are increasingly recognized as powerful tools for enhancing educational experiences. This study investigates the impact of AI-generated instructional videos, designed using established instructional design principles, on self-efficacy, task performance, and learning outcomes in science teacher education. Employing a within-subjects design, the current study included pre-test, post-test, and transfer assessments to evaluate learning durability and transferability, consistent with design-based research methodology. Moreover, this study compares the effectiveness of two AI-generated instructional video formats: one with an embedded preview feature allowing learners to preview key concepts before detailed instruction (video-with-preview condition) and another without this feature (video-without-preview condition). It specifically examines the role of preview features in enhancing these outcomes during training on scientific concepts with 55 Greek pre-service science teachers (n = 55; mean age 27.3 years; range 22–35). The results demonstrated that the videos effectively supported self-efficacy, task performance, and knowledge retention. However, no significant differences were observed between videos with and without preview features across all assessed metrics and tests. These findings also indicate that AI-generated instructional videos can effectively enhance knowledge retention, transfer, and self-efficacy, positioning them as promising assets in science teacher education. The limited impact of the preview feature highlights the need for careful design and evaluation of instructional elements, such as interactivity and adaptive learning algorithms, to fully realize their potential.

Keywords: artificial intelligence; instructional videos; narration; problem-based learning; science teacher education

check for updates

Academic Editors: Laurinda Leite and Luís Dourado

Received: 6 December 2024
Revised: 2 January 2025
Accepted: 13 January 2025
Published: 18 January 2025

Citation: Pellas, N. (2025). The Impact of AI-Generated Instructional Videos on Problem-Based Learning in Science Teacher Education. *Education Sciences*, *15*(1), 102. https://doi.org/10.3390/educsci15010102

1. Introduction

The digital transformation of education has redefined how knowledge is delivered, absorbed, and applied. In science teacher education, the challenge is imparting theoretical knowledge and cultivating essential skills, reflecting on critical thinking, collaborative problem-solving, and the ability to translate complex scientific concepts into engaging teaching practices (Koumi, 2013). Among the most impactful pedagogical strategies, problem-based learning (PBL) is still of great importance for its student-centered, inquiry-driven approach, encouraging learners to engage deeply in real-world problems (Akçay, 2009).

PBL empowers future (pre-service) science teachers to experience firsthand the dynamic interplay between scientific inquiry and practical problem-solving, making it a cornerstone of effective science teacher training (Gumisirizah et al., 2024).

However, the shift to digital contexts, driven by technological advancements and global circumstances, has introduced complexities in implementing PBL effectively. Maintaining student engagement, fostering learning within problem-solving contexts, and providing contextualized problem scenarios are significant hurdles in digital settings (Lange & Costley, 2020). Traditional methods of instruction often fall short of offering the interactive and immersive experiences necessary for meaningful learning, particularly in science education. This gap necessitates innovative approaches that harness the power of emerging technologies to transform learning environments. Integrating instructional videos into the PBL context enhances learning by presenting real-world, information-rich scenarios that foster critical thinking and engagement. Kumar (2010) emphasized the value of interactive video anchors in contextualizing learning through authentic problem-solving activities, such as investigating environmental or historical challenges. These examples highlight the importance of tailoring video content to specific educational objectives. Rasi and Poikela (2016) demonstrated the multimodal potential of video triggers in PBL, aligning with modern students' participatory learning preferences. Their review confirmed that video triggers promote collaborative inquiry and self-directed learning when supported by appropriate infrastructure and educational alignment. Aidoo (2023) found that technology-integrated PBL empowers teacher educators to foster both independent and collaborative learning. Conversely, challenges such as resource constraints, time limitations, and varying competence levels among educators point to the need for professional development to optimize video-based PBL contexts. Gumisirizah et al. (2024) reported that supplementing PBL with video resources in physics education significantly improved student performance, particularly in private schools. These findings suggest that videos enhance PBL's effectiveness by providing accessible and contextually relevant learning aids. Son (2024) highlighted the role of interactive video simulations in developing preservice teachers' adaptive teaching skills, showcasing the potential of responsive, video-based instructional tools. Similarly, Taningrum et al. (2024) demonstrated that animated videos in sports education improved students' critical thinking and understanding of complex concepts. Therefore, studies underline the transformative potential of instructional videos in PBL contexts, emphasizing strategies such as curating authentic content, promoting collaborative learning, addressing implementation challenges, and leveraging narration to simplify complex topics and engage learners effectively.

One such transformative innovation is nowadays the integration of Artificial Intelligence (AI)-generated instructional videos into PBL contexts. These AI-generated videos, created using advanced natural language processing and machine learning algorithms, can dynamically adapt to educational needs. They offer a visually rich and interactive medium to present complex scientific phenomena, simulate real-world scenarios, and provide guided support during problem-solving activities (Pellas, 2024). Unlike static resources, AI-generated videos can be tailored to specific educational contexts, ensuring relevance and accessibility. For future science teachers, such videos can serve as both instructional tools and an example of how technology can enhance classroom learning (Pellas, 2023b; Saqr & López-Pernas, 2024). Integrating AI-generated instructional videos into PBL contexts offers transformative potential by enriching learning experiences through personalized, interactive, and visually engaging content. Leveraging Large Language Models (LLMs), these videos enable adaptive scaffolding and support collaborative problem-solving, making them particularly valuable in science teacher education (Pellas, 2023a).

Generative AI technologies also facilitate scalable, high-quality content creation by converting slides or images into hyper-realistic videos, expanding accessibility across diverse educational contexts (Bewersdorff et al., 2024; Shahzad et al., 2024). Research (e.g., Li et al., 2024; Pellas, 2023b) shows that AI-generated videos achieve comparable outcomes to traditional recordings, with enhanced retention rates and effective support for transfer tasks, which are associated with self-efficacy. Moreover, these tools foster self-directed learning and autonomy, equipping future science educators with the skills to explore and implement advanced strategies independently. A balanced approach, where AI complements human educators rather than replaces them, preserves the relational aspects of teaching while maximizing AI's potential (Arkün-Kocadere & Çağlar-Özhan, 2024). For science teacher training, AI-driven tools not only enhance content delivery but also model innovative instructional practices, fostering interdisciplinary engagement and future-ready teaching skills. By balancing innovation with ethical considerations, AI-generated instructional videos can redefine PBL settings, transforming science education and equipping educators with innovative capabilities (Bewersdorff et al., 2024; Pellas, 2023a; Taningrum et al., 2024). In particular, preview conditions embedded within AI-generated instructional videos can serve as essential scaffolding tools (Lim, 2024). Preview conditions assist in establishing clear goals, activating prior knowledge, and reducing cognitive load by familiarizing learners with key concepts before delving into complex tasks (Van Der Meij, 2017). By drawing attention to critical elements and stimulating curiosity, preview features enhance engagement and facilitate deeper processing, ultimately improving retention and task performance (Van Der Meij et al., 2018a).

As indicated above, the rapid integration of AI in education presents transformative opportunities, yet its practical applications in teacher training programs remain underexplored (Blummer & Kritskaya, 2009; Shahzad et al., 2024). Particularly, science teacher education demands innovative methods to prepare educators for technology-rich classrooms and equip them with the skills necessary to engage diverse learners. PBL settings offer promising frameworks for fostering critical thinking, collaboration, and real-world problem-solving skills (Pellas, 2024; Son, 2024). However, the effectiveness of AI-generated instructional videos within these environments, particularly in enhancing self-efficacy, task performance, and learning outcomes for pre-service science educators, is yet to be fully understood. Moreover, key instructional features—such as conversational narratives, visual demonstrations, and interactive elements—are crucial for maximizing the impact of AI-generated videos. These features align with established multimedia design principles, ensuring that the content is accessible, engaging, and tailored to diverse learning needs (Gumisirizah et al., 2024). By integrating such elements thoughtfully, instructional videos can act as a bridge between theoretical knowledge and practical application, fostering both confidence and competence among pre-service science teachers (Taningrum et al., 2024).

Among the most impactful pedagogical strategies, PBL is one of the most prevalent for its student-centered, inquiry-driven approach, encouraging learners to engage deeply in real-world problems (Akçay, 2009). PBL's emphasis on authentic problem-solving aligns particularly well with the challenges faced by science teachers in preparing students to engage with complex scientific concepts and real-world issues. PBL can also empower future (pre-service) science teachers to experience firsthand the dynamic interplay between scientific inquiry and practical problem-solving, making it a cornerstone of effective science teacher training (Gumisirizah et al., 2024). However, the shift to digital contexts, driven by technological advancements and global circumstances, has introduced complexities in implementing PBL effectively (Li et al., 2024; Magaji et al., 2024). Therefore, there is a need for innovative tools and strategies that can effectively support PBL in digital environments, particularly in science teacher education.

This study investigates the potential of AI-generated instructional videos to address these challenges. AI-generated videos can provide visually engaging and contextually rich scenarios that enhance student engagement in online PBL activities. The use of AI allows for the creation of dynamic and adaptable video content, enabling the presentation of complex scientific phenomena in an accessible and engaging manner. AI's capacity for natural language processing and avatar animation enables the creation of conversational and relatable learning experiences, especially in science education. In the context of complex PBL scenarios, preview conditions within AI-generated videos can be particularly valuable in reducing cognitive load by providing an initial framework for understanding the problem and its related scientific concepts. Preview can help activate prior knowledge relevant to the problem, facilitating more effective engagement in the initial stages of the PBL process, such as problem definition and hypothesis generation. This study aims to investigate the impact of AI-generated instructional videos, with and without preview features, on pre-service science teachers' ability to apply scientific concepts within a PBL context, their problem-solving performance, and their self-efficacy in tackling complex scientific problems. It not only contributes to the growing discourse on AI in education but also offers practical solutions to enhance science teacher training, ensuring educators are equipped with innovative capabilities of AI-generated instructional video to meet the demands of this contemporary education.

2. Background

2.1. Instructional Videos into PBL Contexts

The integration of instructional videos into PBL contexts has been shown to significantly enhance learning by providing real-world, information-rich scenarios that stimulate critical thinking and engagement. Kumar (2010) demonstrated that interactive video anchors, derived from cognitive theories, help students to contextualize learning by embedding students in immersive and authentic problem-solving situations. Examples such as analyzing pollution in a river ecosystem or investigating historical illnesses showcased how videos can effectively bridge theoretical knowledge with practical application. These examples underline the need for educators to thoughtfully curate or create video content tailored to specific learning objectives. In addition, Rasi and Poikela (2016) highlighted the multimodal and participatory affordances of video triggers and production in PBL, emphasizing their alignment with modern students' communication and content creation preferences. Their review of higher education contexts confirmed that video triggers support collaborative learning and self-directed inquiry when implemented under optimal conditions, such as proper technological infrastructure and alignment with educational goals. Instructional videos offer effective support within PBL contexts, particularly for science teachers tackling complex, open-ended problems requiring interdisciplinary knowledge integration. Without sufficient guidance, students may find the problem-solving process challenging. Instructional videos can provide targeted support at crucial junctures by prior works (Lange & Costley, 2020; Pellas, 2024):

- Introduction: Videos can introduce essential background knowledge and concepts necessary for understanding the problem context.
- During problem-solving: Videos can offer just-in-time support by demonstrating specific skills or procedures relevant to particular stages of the problem-solving process.
- Post-problem reflection: Videos can provide expert commentary or alternative solutions, facilitating deeper reflection and learning.

Despite these benefits, challenges such as limited resources, time constraints, and varying levels of teacher competence were noted, underscoring the need for targeted professional development to enhance the adoption of video-based PBL strategies. The effectiveness of video-enhanced PBL in physics education was investigated by Gumisirizah et al. (2024), who reported significant improvements in academic performance across various school settings. Students engaging with video-based PBL performed better than those in traditional or PBL-only settings, particularly in private schools. These findings suggest that video resources can amplify the benefits of PBL by providing accessible and contextually relevant visual aids. Son (2024) demonstrated that AI chatbot simulations enhanced preservice teachers' noticing expertise and movements, reinforcing the value of responsive teaching. Lastly, Taningrum et al. (2024) extended the application of animated videos in sports education, finding that problem-based learning videos significantly improved students' critical thinking skills and understanding of complex football tactics. The study concluded that animated videos can effectively simplify challenging content while encouraging analytical thinking.

The findings from the above studies emphasize the transformative potential of integrating instructional videos into PBL settings. Key strategies for maximizing impact include:

1. Curating authentic and relevant content: Instructional videos should simulate real-world scenarios that resonate with learners' experiences and challenge their problem-solving skills.
2. Promoting multimedia-supported learning: Video triggers and productions foster participatory and interactive learning environments that cater to diverse student preferences.
3. Addressing challenges: Science teacher education in programs and adequate technological support are essential to overcome barriers such as resource limitations and educator readiness.
4. Leveraging AI and animation: Advanced technologies like AI chatbots and animated videos enhance adaptability, interactivity, and engagement, especially in addressing complex or abstract topics.

By aligning pedagogical strategies with technological innovations, educators can create impactful, student-centered experiences within PBL contexts that prepare learners for real-world challenges. For instance, AI-generated instructional videos offer unique opportunities to enhance understanding of Newton's laws of motion by visualizing abstract principles through dynamic animations and real-world scenarios (Savinainen et al., 2004; Thornton & Sokoloff, 1998). For instance, a video could illustrate the first law (inertia) by showing an object at rest and an object in motion until acted upon by an external force. For the second law (F = ma), the video might depict varying forces applied to objects of different masses, demonstrating how acceleration changes in proportion. Finally, for the third law (action-reaction), an animation could simulate a rocket launch, highlighting the expulsion of gas and the equal and opposite force propelling the rocket upward. To promote active learning, the video could include interactive segments where students predict outcomes before animations play or solve related physics problems during pause points. These elements can make AI-generated videos a powerful tool for demystifying physics concepts and improving learners' performance.

2.2. Observational and Demonstration Learning in Video Training

Instructional videos have gained momentum due to their ability to engage learners through multimedia representation and interactive features (Simanjuntak et al., 2019). These types of videos offer notable advantages, such as multimedia representation, alignment between screen animation and task execution, and user-friendly modeling. Research (Lim, 2024; Semeraro & Vidal, 2022) has demonstrated their superiority as an instructional medium in software training, although limitations remain, such as passive information processing, lack of structural overview, and pace inflexibility. Nevertheless, the effectiveness of videos can vary significantly depending on their design and implementation.

The theoretical foundation for videos lies in Bandura's social cognitive theory, which emphasizes the importance of observational learning. By observing and imitating the actions of others, learners can acquire new knowledge and skills. Demonstration-Based Training (DBT) builds upon this theory by systematically designing instructional features to enhance learning outcomes. This theoretical underpinning posits that much of human knowledge is acquired through observing others' actions. DBT builds on these principles, employing systematically designed instructional features to enhance knowledge acquisition, skills development, and attitudes (Van Der Meij et al., 2018b). DBT aligns instructional features with Bandura's four observational learning processes: attention, retention, production, and motivation. To support attentional processes, which are critical for observational learning, features such as narration and highlighting were utilized (Hurzlmeier et al., 2021). Narration is synchronized speech with animation for clarity in video. Highlighting techniques, including color coding, and zooming, directed learners' focus on essential interface elements, improving readability and recognition (Newbold & Gillam, 2010). Empirical evidence supports the effectiveness of such visual cues in improving learning outcomes (Zhang et al., 2024) and examines the functions and designs of added cues in instructional videos (Wang et al., 2020). User control, allowing learners to pause, replay, or skip sections, is another vital feature. It enables self-paced learning, overcoming the passivity often associated with videos and fostering active engagement. Studies indicate that user control enhances knowledge transfer, particularly for low-ability learners or complex tasks (Pi et al., 2019). The pacing of videos also influences learning outcomes. This can pass too quickly risk cognitive overload, while overly slow pacing may reduce attention (Martin & Martin, 2015). Previous studies (Van Der Meij, 2017) suggested an optimal pacing strategy that balances these concerns, with shorter videos under three minutes proving most engaging. In this study, video lengths ranged from approximately 3 to 5 min, aligning with these findings.

Lastly, previews serve as essential tools in video-based training, functioning as advanced organizers that structure and guide learners' understanding of the subsequent demonstration (Van Der Meij & Van Der Meij, 2014). By providing a concise overview of the upcoming content, previews can significantly enhance learning outcomes. While research (Blummer & Kritskaya, 2009; Van Der Meij, 2017) on the specific impact of previews in multimedia learning is limited, studies on text comprehension offer valuable insights. Some studies have shown that initial summaries, similar to previews, can improve information retention, particularly when they actively engage learners. This suggests that a well-designed preview can capture learners' attention, motivate them, and reduce cognitive load.

Prior works (Lim, 2024; Hurzlmeier et al., 2021; Van Der Meij et al., 2018a) indicate some key guidelines for effective preview design:

1. Clear objectives: The preview should clearly state the instructional video's learning goals.
2. Conversational style: The narrative should be engaging and easy to follow, resembling a conversation.
3. Introduction of key concepts: Critical or novel concepts should be briefly introduced, priming learners for deeper understanding during the demonstration.

The preview condition in this study adhered to the above principles, providing a brief overview of the task, and highlighting key elements. This approach aimed to boost learners' self-efficacy and reduce cognitive load by familiarizing them with essential terminology and concepts before the detailed demonstration. The preview process can ensure that learners are well-prepared to absorb the information presented in video (Van Der Meij & Van Der Meij, 2014). The preview in this study adhered to the aforementioned principles, offering learners a concise yet structured introduction to the task at hand. By providing a brief overview and emphasizing critical components, the preview served as a cognitive scaffold, equipping learners with the foundational knowledge needed to engage effectively with the detailed demonstration. This preparatory phase is essential in establishing learners' self-efficacy, as it assures them that the task is approachable and within their capabilities (Rismark & Sølvberg, 2019). According to Bandura's theory of self-efficacy (2006), individuals who perceive a task as manageable are more likely to persist, perform effectively, and achieve positive outcomes.

In this context, the preview fulfills multiple functions. It familiarizes learners with essential terminology and key concepts, reducing the cognitive load associated with processing new information during the demonstration. By doing so, it allows learners to focus their mental resources on understanding task-specific actions rather than struggling with unfamiliar jargon or concepts (Costley et al., 2020). Additionally, the preview acts as an anticipatory guide, helping learners organize incoming information, establish a mental framework, and activate relevant prior knowledge. This alignment between the preview and the learner's cognitive schema fosters a sense of preparedness and confidence, both critical components of self-efficacy (Marshall, 2024). Furthermore, the preview's ability to simulate a clear and achievable task pathway reinforces learners' belief in their capacity to succeed. For instance, by introducing screen objects, procedural steps, and task goals in an accessible manner, the preview transforms potentially intimidating content into manageable chunks (Brame, 2016). This breakdown not only supports cognitive processing but also nurtures a positive mindset, encouraging learners to approach the subsequent videos with motivation and readiness (Ester et al., 2023). In essence, the preview primes learners for success by creating a low-stakes environment for initial exposure to enhance their attention and memory, ensuring they are psychologically and cognitively equipped to engage with the full demonstration (Schacter & Szpunar, 2015). The integration of such previews aligns closely with the pre-training principle advocated by Kulgemeyer (2018), who posits that teaching essential components before a task enhances understanding and retention. By proactively addressing potential barriers to comprehension, the preview creates a smoother learning curve and supports the overarching goals of self-efficacy development, cognitive load reduction, and instructional effectiveness, when it comes to instructional explanation of scientific topics.

However, the integration of AI in videos provides additional ethical considerations. AI tools were employed to streamline the design process and generate dynamic visuals, but these tools were evaluated to ensure the produced content was accurate, culturally inclusive, and pedagogically effective. Potential biases in AI algorithms were assessed to avoid perpetuating stereotypes or misinformation. To enhance learner engagement and accessibility, the videos were designed with natural, relatable visuals and conversational narration (Al-Zahrani & Alasmari, 2024). The AI-generated components were developed with transparency, emphasizing fairness and accountability. Measures were taken to provide learners with control over video interaction and to ensure the content's adaptability to diverse learning needs. Additionally, the ethical implications of AI tools in education, including data privacy and algorithmic decision-making, were carefully considered to align with both General Data Protection Regulation (GDPR) regulations and best practices for educational fairness (Hamad et al., 2024).

2.3. AI-Generated Instructional Videos into PBL Contexts

One of the primary advantages of integrating AI-generated instructional videos into PBL settings is their potential to enrich learning experiences. By leveraging LLMs, these videos can deliver personalized, visually engaging, and interactive content that resonates with diverse learners. This capability enhances engagement with scientific concepts and provides dynamic scaffolding for collaborative problem-solving tasks. Such tools are particularly valuable in science teacher education, where contextualized and adaptive learning resources are essential for cultivating practical teaching skills (Bewersdorff et al., 2024). Moreover, generative AI pipelines that transform slides or images into hyper-realistic videos enable scalable, high-quality content creation, making innovative educational tools accessible across varied contexts (Saqr & López-Pernas, 2024).

Studies have shown that AI-generated videos can achieve comparable learning outcomes to traditional recorded videos. While retention rates with AI-driven resources are often higher, their performance on transfer tasks aligns closely with that of human-instructor-led videos, demonstrating their viability as an alternative instructional tool. Li et al. (2024) advocated that AI-generated videos support self-directed learning by fostering autonomy and facilitating personalized guidance, particularly valuable for training science educators to independently explore and implement advanced teaching strategies. These capabilities position AI-generated content as a significant enhancement to traditional PBL methodologies, addressing the limitations of static instructional resources.

However, challenges remain in achieving optimal engagement and a sense of social presence in AI-generated instructional videos. For instance, Shahzad et al. (2024) indicated that learners often perceive these videos as less engaging or relatable compared to human instructors. Concerns such as distraction, discomfort, and a lack of connectedness highlight the importance of refining the human-like qualities of AI avatars and videos. Despite these limitations, Arkün-Kocadere and Çağlar-Özhan (2024), in their study pointed out that learners tend to overlook these issues when the content is intrinsically engaging or aligned with their learning goals. As generative AI continues to advance, these differences are expected to diminish, further normalizing the use of its content in education.

Ethical and pedagogical considerations also play a critical role in the integration of AI-generated instructional videos into education. A study conducted by Pellas (2024) revealed that creativity mediates the relationship between academic performance and attitudes toward machine learning, though its moderating effect was not significant. Problem-solving and critical thinking, while not mediators, significantly moderates the link between academic performance and machine learning attitudes. Pellas (2023b) also showed that challenges such as data privacy, algorithmic bias, and the potential over-reliance on technology necessitate robust frameworks to ensure responsible use in AI-generated video content creation. This study examines the impact of digital storytelling and generative AI platforms on narrative intelligence and writing self-efficacy in undergraduate students. A pretest–posttest study, Pellas (2023a) compared traditional platforms (e.g., Storybird, Storyjumper) with generative AI tools (e.g., Sudowrite, Jasper). Results showed that generative AI platforms significantly enhance narrative intelligence and writing self-efficacy, though no notable differences were found in creative identity. These findings underscore the potential of AI tools to support students' story creation and development skills. Lin et al. (2025) admitted that a balanced approach is essential, where AI complements the unique relational and adaptive capabilities of human educators rather than replacing them. This equilibrium preserves the educator's role as a mediator of learning experiences while maximizing the potential of AI-generated tools.

For science teacher education, the integration of AI-generated videos offers a dual benefit: enhancing content delivery and modeling innovative instructional practices. By engaging in AI-driven resources during their training, Qu et al. (2024) admitted that future educators gain insight into how technology can be leveraged to address diverse learning needs effectively. Additionally, the multimodal capabilities open new possibilities for delivering immersive, interdisciplinary learning scenarios, fostering deeper engagement in scientific practices and equipping teachers with adaptable, future-ready skills (Pellas, 2024).

While the above research studies were focused on instructional videos, observational learning, and PBL are extensive, a significant gap exists at the intersection of these domains, particularly concerning the use of AI-generated videos. Although previews have been shown to be beneficial in traditional video-based learning by activating prior knowledge and reducing cognitive load, their specific role within AI-generated instructional videos used in PBL contexts remains largely unexplored. Existing research has not adequately addressed how the unique characteristics of AI-generated videos, such as their dynamic content and potential for personalization, might interact with preview strategies to influence learning outcomes. Furthermore, there is a lack of empirical evidence examining the combined effect of previews and AI-generated videos on pre-service science teachers' self-efficacy and their ability to transfer learned concepts to novel problem-solving situations within a PBL activity. This lack of research is particularly relevant in science teacher education, where preparing future educators to effectively use technology in the classroom is crucial. Although AI offers promising tools for creating engaging and accessible learning resources, it is essential to understand how established instructional design principles, such as the use of previews, can be optimally integrated into these new technologies. Investigating this intersection can provide valuable insights into maximizing the effectiveness of AI-generated videos in PBL settings, ultimately contributing to improved science teaching practices and enhanced student learning. This study aims to address this critical gap by specifically examining the impact of previews in AI-generated instructional videos on pre-service science teachers' learning, problem-solving performance, and self-efficacy within PBL contexts.

3. Materials and Methods

3.1. Research Purpose and Questions

The purpose of this study is to explore the role of preview features in enhancing the effectiveness of AI-generated instructional videos, particularly within PBL contexts for science teacher education. As the integration of AI in educational technology grows, this research seeks to examine how AI-generated videos—especially those with preview segments—can revolutionize the learning process in teacher training environments. By comparing videos with and without preview features, this study provides a unique contribution to the field of AI-driven multimedia learning and instructional design, offering insights into how subtle design elements can impact educational outcomes. In addition, research (Bandura, 2006; Shaukat et al., 2024) has shown that higher levels of self-efficacy are associated with increased persistence, better problem-solving strategies, and improved academic performance. By enhancing students' self-efficacy through instructional videos users will be more likely to engage deeply with the material and approach physics problems with greater confidence and competence. Therefore, any increase in self-efficacy measured in our study would suggest not only an improvement in students' beliefs about their abilities but also a potential improvement in their actual academic performance.

This research employs a within-subjects design (Greenwald, 1976), a research methodology grounded in well-established educational frameworks, allowing for a more direct comparison of the effects of preview features on learners' (pre-service teachers) performance (Johnson et al., 2017; Reeves et al., 2005). Through this rigorous evaluation, the current study aims to generate evidence that can inform the design of more engaging, interactive, and effective learning experiences. Specifically, it investigates how the inclusion of preview features in AI-generated videos influences key factors such as self-efficacy, task performance, and overall learning outcomes. Incorporating principles from design guidelines for software training videos, such as the importance of contextualized previews to enhance understanding and retention (Arkün-Kocadere & Çağlar-Özhan, 2024; Bewersdorff et al., 2024; Gumisirizah et al., 2024), this study anticipates that AI-generated videos will improve participants' confidence and skill acquisition in teaching science concepts. Moreover, it is expected that preview features would further enhance these outcomes by providing learners with a brief but informative overview of the content, thus improving both the retention and transferability of the learned material. The findings of this study may inform future instructional video design, particularly in the context of science teacher education, and contribute to the broader understanding of how AI can be leveraged to support effective, technology-enhanced learning.

To investigate the impact of AI-generated instructional videos on pre-service science teachers' learning within PBL contexts, the following research questions (RQs) guided this study:

- RQ1: Do AI-generated instructional videos, with and without a preview feature, affect pre-service science teachers' ability to apply scientific concepts?
- RQ2: Do AI-generated instructional videos, with and without a preview feature, influence pre-service science teachers' immediate and delayed problem-solving performance?
- RQ3: Do AI-generated instructional videos, with and without a preview feature, impact pre-service science teachers' self-efficacy in addressing complex scientific problems?

This study intends to provide evidence-based insights into how AI-generated instructional videos, particularly those enhanced with preview features, can effectively enhance science teacher education through PBL settings.

3.2. Research Design

While AI has garnered significant attention in education, its application within teacher training, particularly concerning the effective design and utilization of digital learning resources, remains underexplored. Though AI-generated content for student use is gaining traction, the potential of AI to support teachers in creating, adapting, and integrating these resources into their instruction is an area ripe for investigation. For example, AI could be used to: (a) automatically generate diverse examples and practice problems for instructional videos; (b) provide real-time feedback to teachers on the design and delivery of video lessons; or (c) create personal recommendations for teachers on how to best utilize video resources within PBL contexts.

This study investigates the comparative effectiveness of AI-generated instructional videos with and without preview features on learning outcomes, self-efficacy, and knowledge transferability. Using a within-subjects design, where all participants experienced both video conditions, minimizes variability due to participant differences and aligns with design-based research principles emphasizing iterative refinement of educational interventions (Reeves et al., 2005). Following Greenwald's (1976) principles, learning was assessed using pre-test, post-test, and transfer assessments to evaluate learning durability and transferability, directly addressing the need for practical evaluations of design principles in instructional media (Lin et al., 2025). This approach bridges theoretical insights from instructional design with actionable recommendations for educators. Participants' learning of scientific concepts, such as Newton's laws of motion, was assessed as follows:

- Pre-test: A 10-item multiple-choice and 2-item short-answer test assessed baseline knowledge (e.g., 'Which of Newton's laws explains why an object in motion stays in motion unless acted upon by an external force?').
- Post-test: Administered immediately after the intervention, the post-test used a parallel format with reworded or scenario-based questions assessing the same concepts.
- Transfer test: Transfer was assessed by requiring participants to apply Newton's laws to a novel, real-world context—analyzing a video of a complex mechanical system (e.g., a bicycle braking system)—distinct from the training examples (e.g., simple projectile

motion problems). This ensured measurement of genuine knowledge application rather than simple recall.

Participants completed a post-test immediately after finishing the online learning module. A delayed post-test was administered two weeks later to assess retention. This two-week interval was chosen to provide a more realistic measure of long-term retention, allowing sufficient time for potential memory decay and the influence of everyday experiences (Figure 1).

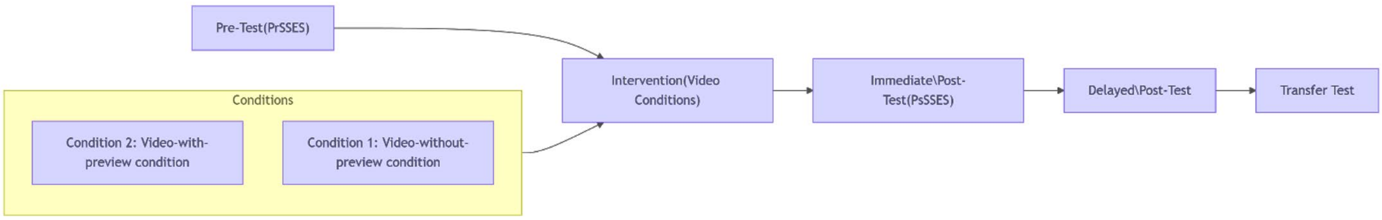

Figure 1. Flowchart of the research design.

3.3. Participants

Fifty-five Greek pre-service teachers participated in the study, comprising 31 females (56.36%) and 24 males (43.64%). Their average age was 27.3 years (SD = 3.56, range: 22–35). Most participants had 4–5 years of teaching experience (42.77%, n = 23), followed by less than 4 years (31.29%, n = 17), and first-year teachers (12.25%, n = 7). Their specialization areas included chemistry (32.73%, n = 18), biology (27.27%, n = 15), geology (21.82%, n = 12), and physics (18.18%, n = 10). All participants were experienced in using AI platforms, particularly in the context of media and information literacy.

The participants were selected based on their strong background in science education and prior exposure to AI-driven learning environments, ensuring their familiarity with instructional technology and innovative teaching methods. The majority (82.91%, n = 46) had completed traditional teacher training in Greece, involving a 4-year bachelor's degree or a 2-year master's degree, followed by a mandatory induction phase.

Most participants had 6 or more years of experience (53.98%, n = 29), with fewer having 4–5 years (42.77%, n = 23). The majority worked full-time (56.74%, n = 31). Approximately 22.51% (n = 12) were adjunct science teachers in public secondary schools, with the remaining 77.49% (n = 41) employed in private secondary schools. This sample also represented diverse educational contexts, with 37% (n = 20) teaching in rural areas and 63% (n = 35) in urban schools. Participants' motivation to participate was driven by professional development opportunities (78%, n = 43) and a shared interest in integrating AI into science education.

Participants were recruited from pre-service teacher education programs at three prominent Greek universities. Recruitment was conducted through a multi-pronged approach to maximize reach and ensure a representative sample. These emails provided a concise overview of the study's purpose, the estimated time commitment (approximately 2–3 h), information about data confidentiality, and a clear statement of voluntary participation. The email also included a link to an online platform (e.g., Google forms) where interested participants could access the informed consent form.

Inclusion criteria for participation were as follows: (a) current enrollment in a pre-service science teacher education program at one of the participating universities; (b) access to a computer with reliable internet access; and (c) willingness to participate in both data collection sessions. Participants were not offered any direct compensation, but their participation was framed as a valuable opportunity for professional development and contributing to the advancement of research in educational technology.

A total of approximately 80 email invitations were sent, resulting in 55 completed responses and a response rate of approximately 69%. This response rate is considered moderate and suggests a reasonable level of engagement with the study. While we acknowledge that any non-response may introduce some selection bias, the achieved sample size is sufficient for the planned statistical analyses. We address the potential limitations related to response rate in the discussion section.

3.4. Instruments

3.4.1. Training Materials

Participants were trained in key concepts reflecting Newtonian mechanics. The instructional videos covered specific concepts, including Newton's three laws of motion—inertia, force and acceleration, and action-reaction pairs. These concepts were chosen because they are fundamental to understanding basic physics principles and are essential for teaching introductory physics at the secondary level. The learning objective was to enable pre-service teachers to visualize Newton's laws in terms of analyzing and explaining the motion of objects in various real-world scenarios (Savinainen et al., 2004; Thornton & Sokoloff, 1998). This objective directly informed the design of the post-test and transfer test, ensuring alignment between instruction and assessment.

A. Instructional AI-generated videos: This study integrates AI-generated instructional videos into PBL for science teacher education, leveraging the IDEA (Interpret, Design, Evaluate, Articulate) framework by Lin et al. (2025) to enhance the effectiveness of training materials. Each of the four treatment phases has been restructured to align with the IDEA principles, ensuring a comprehensive, pedagogically grounded approach to developing pre-service science teachers' skills.

Treatment phase 1: Observing and explaining natural phenomena

- Interpret: Participants are guided to interpret and understand natural phenomena by observing and analyzing scientific events such as the phases of the moon, changes in shadow length, and the motion of objects under various forces. Teachers provide context for these tasks, helping students identify key observational goals and comprehend the significance of the phenomena. Newton's laws are introduced through AI-generated scenarios, such as a ball rolling on different surfaces to demonstrate inertia and force. AI tools offer culturally and contextually relevant visual prompts and scenarios to scaffold understanding. This focus on interpretation laid the groundwork for the transfer task, which required participants to interpret a complex mechanical system.
- Design: Participants collaborate with AI tools to develop detailed observation plans. For instance, they may design templates to measure and record motion data, such as acceleration or force diagrams. These tools support multi-lingual and multimodal interactions, allowing students to generate text-based or visual representations of their observations.
- Evaluate: Participants critically analyze their observation data against Newtonian principles. AI feedback systems help highlight inaccuracies or areas of improvement, promoting reflection. Peer discussion sessions, guided by AI-generated insights, enable collaborative evaluation of findings, fostering deeper understanding of concepts such as inertia and action-reaction pairs.
- Articulate: Participants provide their interpreted observations and reasoning to peers through presentations or reports. AI assistance ensures clarity and precision, offering real-time feedback and generating culturally relevant explanations to support their arguments, particularly around Newton's laws.

Treatment phase 2: Conducting Experiments and Analyzing Results

- Interpret: Participants are introduced to experimental scenarios, such as testing the effects of mass on acceleration or analyzing the forces acting on an inclined plane. Teachers guide students to understand the purpose, methodology, and significance of these experiments. AI-generated videos include dynamic animations and visual overlays to enhance comprehension of experimental setups and protocols involving Newton's laws.
- Design: Using AI tools, each participant designs experimental workflows, including hypotheses, variables, and data collection methods. AI offers real-time assistance by suggesting modifications to ensure safety and experimental validity, tailoring feedback to students' cultural and educational contexts. Experiments emphasize key principles such as force, mass, and acceleration.
- Evaluate: Participants evaluate their experimental results by comparing their findings with expected outcomes derived from Newton's laws. AI-powered analytics tools guide this process, identifying trends and highlighting discrepancies. Peer assessment sessions, supported by AI-generated discussion prompts, further refine their understanding of concepts such as force interactions and equilibrium.
- Articulate: Participants can articulate their experimental findings through interactive presentations, supported by AI-generated visual aids such as graphs, tables, and infographics. These outputs enable effective communication of results to diverse audiences, fostering scientific literacy with a focus on Newtonian mechanics.

Treatment phase 3: Advanced PBL in science education contexts

- Interpret: The main researcher introduces complex scientific challenges, such as designing a water filtration system or analyzing the dynamics of a pendulum. AI helps students interpret these problems by breaking them into manageable components, using simulations to visualize underlying principles, including Newton's laws.
- Design: Participants design solutions using AI-generated tools to create and refine prototypes or simulate outcomes. For example, they might model a pendulum's motion to study forces and acceleration or simulate vehicle collisions to understand action-reaction pairs. AI tools provide iterative feedback, ensuring designs are scientifically sound and feasible.
- Evaluate: Participants test their solutions using AI-driven simulations and evaluate the results. They reflect on the success of their approaches and identify areas for improvement. AI-generated comparisons with best-practice solutions help them understand gaps in their designs and how Newton's principles apply.
- Articulate: Participants can present their problem-solving processes and outcomes to peers and instructors, supported by AI-generated explanatory content. Real-time AI translation and contextual feedback help them effectively communicate complex ideas, particularly those related to Newtonian mechanics.

Treatment phase 4: Interactive features and technical considerations

- Interpret: The main researcher guides students to understand the importance of using digital tools for effective learning. The interactive features of the AI-generated videos, such as pausing, rewinding, and adjusting playback, are demonstrated to help students optimize their learning experience. Videos include real-time examples of Newton's laws in action to reinforce understanding.
- Design: Participants design their learning schedules and practices around these interactive features. AI provides personalized suggestions, such as recommending specific playback settings for segments illustrating complex Newtonian principles.

- Evaluate: Participants reflect on the utility of the interactive features in improving their comprehension and retention of Newton's laws. Peer feedback and self-assessment help identify which features are most effective for different learning styles.
- Articulate: Participants articulate their experiences with the interactive features through written or verbal feedback, supported by AI prompts. This phase encourages how these tools enhance their learning of fundamental physics concepts.

By integrating the IDEA framework into the chapters, this approach emphasizes a holistic and student-centered methodology. It ensures that students not only acquire scientific knowledge and skills but also develop critical thinking, collaboration, and communication abilities essential for effective science education. The key examples implemented focus on providing Newton-specific examples that explicitly relate to his laws of motion. In the preview videos, examples now address friction and inertia (Newton's First Law) and action-reaction pairs (Newton's Third Law). Demonstration videos showcase concepts like force and acceleration (Newton's Second Law) and the conservation of momentum and collisions (Newton's Third Law). The practice files offer opportunities to apply these laws by calculating force and acceleration (Newton's Second Law), analyzing forces on an inclined plane, and exploring projectile motion (incorporating multiple laws). Furthermore, instructional content videos and practice files connect to Newton's laws, providing more detailed examples to give a better sense of their content.

A. *Preview videos:* The preview videos provide concise overviews of the scientific phenomena or experimental techniques covered in each chapter. The narration begins conversationally, posing a problem or question to engage learners (e.g., *"Have you ever wondered why a rolling ball eventually stops?"* or *"What happens when you push a stationary object?"*). These videos include brief demonstrations of critical steps, enhanced by animations and zooming to emphasize key details. For example, a preview video might show a ball rolling across different surfaces (carpet vs. smooth floor) to illustrate the concept of friction and how it relates to Newton's First Law (inertia). Another example could show a brief animation of a rocket launching to introduce the concept of action-reaction pairs (Newton's Third Law). The average duration of preview videos is 1.30 min, with a five-second pause at the end to help students process the information (range 1.15−1.60). These previews served as initial exposure to the concepts tested in the pre- and post-tests.

B. *Demonstration videos*: Demonstration videos deliver detailed step-by-step instructions for conducting experiments or solving scientific problems. Narration is structured to explain both actions and outcomes (e.g., *"We will now measure the acceleration of a cart as we apply different forces, demonstrating Newton's Second Law. First, we'll measure the mass of the cart..."* or *"Observe how the force applied to the cart directly affects its acceleration."*). These videos incorporate AI-enhanced visuals, such as showing force vectors acting on an object or graphing the relationship between force and acceleration, to reinforce learning. For instance, a video could demonstrate how changing the mass of an object affects its acceleration when a constant force is applied. Another example could showcase the collision of two objects, demonstrating the conservation of momentum and Newton's Third Law. Each step is followed by a brief pause to allow reflection, and the entire video concludes with a five-second pause. Demonstration videos have an average length of 1.77 min. These demonstrations directly addressed the content assessed in the post-test and provided concrete examples that aided in conceptual understanding relevant to the transfer task.

C. *Final video files:* Accompanying the video tutorials are practice files, designed to replicate the scenarios presented in the videos. These files include step-by-step instructions for conducting experiments, paired with before-and-after images or data tables to guide learners through each task. For example, a file might provide instructions for calculating the force required to accelerate an object at a given rate (applying Newton's Second Law) or analyzing the forces acting on an object on an inclined plane. Another example could involve analyzing the motion of a projectile, calculating its trajectory based on initial velocity and launch angle (incorporating concepts from Newton's laws). These tasks allow students to apply what they have learned from the videos, with the option of re-visiting the segments for guidance. These practice files offered opportunities for applying learned concepts, reinforcing learning and preparing participants for the application-focused aspects of the post-test and the transfer task.

The above framework aligns with the study's focus on science teacher education, emphasizing active engagement with scientific phenomena and problem-solving through AI-generated instructional content.

3.4.2. Self-Efficacy Scale

To evaluate self-efficacy, the validated *Standards Self-Efficacy Scale (SSES)* was utilized (Shaukat et al., 2024). This study explored the potential of the SSES to assess pre-service teachers' self-efficacy about teaching standards, separating into other two: the Pre-Self-Efficacy Questionnaire (PrSSES) and the Post-Post-Self-Efficacy Questionnaire (PsSSES). The psychometric properties of the SSES were examined to evaluate its validity and reliability. It consists of 29 items and is selected due to its individual item's internal consistency. This was designed to capture participants' confidence levels in completing specific tasks, both before and after engaging with a video. The former (PrSEQ) was administered before the video to assess baseline self-efficacy. Based on Bandura's recommendations for designing self-efficacy measures, the questionnaire included seven items. Each item provided a brief description of a task along with corresponding "before-and-after" screenshots. Participants were asked to rate their perceived competence in completing the task using a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree."

This scale is widely used to assess teachers' beliefs in their ability to meet the demands of their profession. For the purpose of this study, we adapted the SSES to specifically measure pre-service teachers' self-efficacy related to applying Newtonian mechanics concepts within the context of a roller coaster design PBL activity. This adaptation involved rephrasing items to refer to specific tasks related to applying Newton's motion laws during the roller coaster design PBL activity. The adapted scale used a 5-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5).

Pre-self-efficacy questionnaire (PrSSES): The PrSSES, administered before the intervention, consisted of seven items adapted from the original SSES. These items were designed to assess baseline self-efficacy related to basic understanding and application of Newton's Laws, such as interpreting motion graphs, understanding force vectors, and calculating basic forces. The PrSSES demonstrated strong reliability with a Cronbach's alpha of 0.83. These items established a baseline for participants' confidence in applying the foundational concepts of Newtonian mechanics. Examples of PrSSES items (relating to Newton's Laws) are the following:

1. "*I interpret a graph of velocity versus time to determine the acceleration of an object.*" (Relates to understanding motion and Newton's First Law/Inertia if the velocity is constant, and to Newton's Second Law if the velocity changes).
2. "*I identify the forces acting on a stationary object.*" (Relates to Newton's First Law and the concept of balanced forces).
3. "*I describe the force required to accelerate an object given its mass and acceleration.*" (Directly relates to Newton's Second Law: F = ma).

Post-self-efficacy questionnaire (PsSSES): The PsSSES, administered after the intervention, also consisted of seven items adapted from the original SSES. These items were designed to assess self-efficacy related to more complex applications of Newton's Laws, such as analyzing forces on an inclined plane, predicting projectile motion, and understanding the relationship between force, mass, and acceleration *within the context of designing a roller coaster*. The PsSSES also demonstrated strong reliability with a Cronbach's alpha of 0.81. These items measured participants' confidence in applying Newton's Laws in more complex scenarios, similar to those encountered in the PBL activity. Examples of PsSSES items (relating to Newton's Laws and the Roller Coaster Context) are the following:

1. "*I analyze the forces acting on a roller coaster car as it moves along an inclined track, including gravity, normal force, and friction.*" (Combines Newton's First and Second Laws in a realistic scenario).
2. "*I predict the speed of a roller coaster car at different points on the track using the principles of energy conservation and Newton's Laws.*" (Integrates multiple concepts and laws within the PBL context).
3. "*I explain how Newton's Third Law (action-reaction) affects the interaction between the roller coaster car and the track, especially during turns.*" (Directly relates to a key aspect of roller coaster design).

To comprehensively analyze self-efficacy changes, this study employed tools to capture participants' perceived abilities at two distinct stages. Given the Greek context of this research, the scales were rigorously translated into Greek using Brislin's (1970) back-translation method to maintain cultural and linguistic equivalence. Online questionnaires were distributed to assess participants' perceptions.

3.4.3. Task Performance Tests

The task performance tests were developed to evaluate the practical skills introduced in the AI-generated instructional video. Four distinct tests were administered: the pre-test, immediate post-test, delayed post-test, and transfer test, following the guidelines by Van Der Meij et al. (2018b). Each task was assessed on a binary scoring system: tasks completed entirely and correctly were awarded one point, while incomplete or incorrect tasks received zero points. The maximum achievable scores were nine points for the pre-test, immediate post-test, and delayed post-test, and four points for the transfer test.

A. Pre-Test, Immediate Post-Test, and Delayed Post-Test: The pre-test, conducted before the tutorial, consisted of seven tasks designed to evaluate participants' initial understanding and skills related to science phenomena. These tasks were carefully aligned with the content of the AI-generated instructional videos, focusing on key scientific concepts and processes, including:
    1. Adjusting variables in a simulation to observe changes in a scientific phenomenon.
    2. Identifying relationships between dependent and independent variables in an experiment.
    3. Interpreting data from graphs or tables representing scientific phenomena.
    4. Organizing observations into structured formats, such as diagrams or flowcharts.
    5. Creating and labeling a model to represent a specific science concept.

These tasks established a baseline for evaluating participants' ability to comprehend and apply scientific concepts, as facilitated by the AI-generated instructional videos.

The pre-test included an instructional guide and seven science-related tasks. Each task was introduced with a brief description, supported by before-and-after visualizations illustrating the expected results. Participants were instructed to open corresponding simulation files, stored in their personalized computer folders, to perform the tasks. For instance, one task required participants to analyze a graph of velocity vs. time and determine the acceleration of an object. This task assessed their understanding of Newton's First Law (inertia) by requiring them to recognize that a constant velocity indicates zero net force and therefore no acceleration (connecting to Newton's first law). Another task asked participants to draw force vectors acting on a block sliding down an inclined plane without friction to connect with Newton's second law. This task assessed their understanding of Newton's Second Law (F = ma) by requiring them to identify the forces acting on the block (gravity, normal force) and their relationship to the block's acceleration. The pre-test was designed to have a maximum score of nine points.

The immediate post-test, administered directly after the tutorial, and the delayed post-test, conducted seven days later, followed the same structure as the pre-test. While the tasks and instructions remained similar, the content of the associated science simulations was adjusted slightly to ensure learners demonstrated conceptual understanding without repetition. While the pre-test asked participants to analyze a graph with constant acceleration, the post-test presented a graph with varying acceleration (Connecting to Newton's Second Law). This required a deeper understanding of Newton's Second Law and the relationship between force and changing acceleration. Connecting to Newton's Laws in combination, similarly, the inclined plane problem in the post-test included friction. This required participants to apply both Newton's First and Second Laws to account for the additional force of friction and its effect on the block's motion. These tests were also designed to have a maximum score of nine points.

B. Transfer test: The transfer test evaluated participants' ability to apply their knowledge and skills to science phenomena that were not explicitly addressed in the AI-generated instructional videos. This test aimed to measure adaptability, critical thinking, and independent problem-solving skills. The transfer test included a separate instruction file and four new tasks:

   1. Adjusting environmental variables in a simulation to observe changes in system behavior.
   2. Reorganizing data within a visual representation, such as a graph or chart.
   3. Identifying patterns in new experimental results.
   4. Creating a hypothesis and explaining its implications using unfamiliar scientific data.

Each task in the transfer test was paired with uniquely designed simulation files to ensure participants faced challenges extending beyond the tutorial's scope. This approach encouraged deeper cognitive engagement and the application of foundational scientific principles in new contexts. An appropriate example of connecting to Newton's Third Law and conservation of momentum was required. One task presented a scenario involving two colliding objects with different masses and asked participants to calculate the final velocities of each object. This task assessed their understanding of Newton's Third Law (action-reaction) and the principle of conservation of momentum. Connecting to multiple laws in a complex system. Another task asked participants to predict the trajectory of a projectile launched at an angle, considering the effects of air resistance. This required participants to integrate their understanding of all three of Newton's Laws to account for gravity, air resistance, and the resulting motion of the projectile.

By integrating these structured assessments, the study systematically evaluated the effectiveness of the AI-generated instructional videos in fostering understanding, retention, and transferable skills within the domain of science teacher education.

3.5. Procedure

The present study was conducted in a middle school computer lab, equipped with headphones for each participant. Experiments spanned two sessions over six days. During Session 1, participants were assigned ID cards and completed the PrSSES and a pre-test after a brief introduction. Following a short break, participants watched the video tutorial, practiced tasks, and completed the PsSSES, and an immediate post-test. This session lasted approximately 90 min (Figure 2).

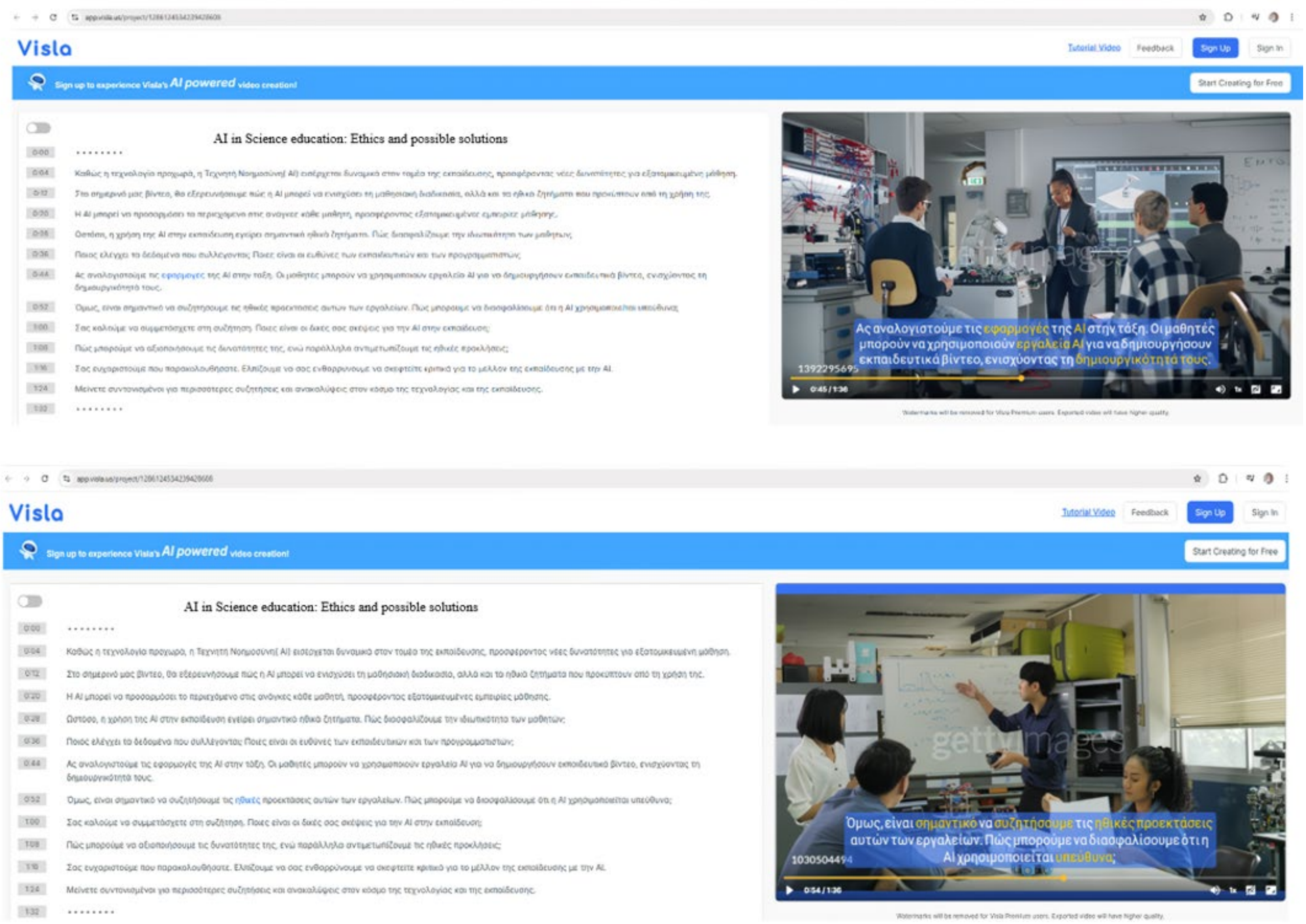

Figure 2. Participants create AI-generated instructional videos using Visla.

In Session 2, conducted a week later, participants completed a delayed post-test and a transfer test without access to the tutorial. This session lasted 45 min. Participants interested in pursuing careers within public or private schools can gain valuable research experience by exploring the potential of generative AI chatbots for education. This could involve:

- Evaluating existing tools: Analyzing the effectiveness of current AI-powered educational applications like ChatGPT, Claude, and Gemini to identify areas for improvement.
- Designing integration frameworks: Developing frameworks for seamlessly integrating AI into existing digital literacy curriculums.

Furthermore, all participants can enhance their practical skills by creating AI-supported educational resources tailored to diverse age groups and learning objectives. This process encompassed the following:

- Interactive multimedia content: Designing and developing educational videos and images enriched with interactive digital elements to enhance learners' digital literacy across various subjects.
- AI-generated video production: Utilizing platforms like Sudowrite (https://www.sudowrite.com, accessed on 8 December 2024), Visla (https://www.visla.us, accessed on 8 December 2024), or Jasper (https://www.jasper.ai, accessed on 8 December 2024) for creating engaging educational video content through AI assistance.

This study employed a within-subjects design, where all participants experienced both conditions: AI-generated instructional videos with a preview feature and AI-generated instructional videos without a preview feature. This design was chosen to minimize the influence of individual differences on the results, allowing a more direct comparison of the effects of the preview feature within the same participants (Greenwald, 1976). However, the lack of a separate control group receiving a different type of instruction limits the ability to contextualize findings against alternative instructional approaches. Therefore, the current study consisted of the following phases:

- Pre-test: Participants completed the pre-test, which assessed their baseline knowledge of Newtonian mechanics and related problem-solving skills. They also completed the PrSSES, measuring their perceived self-efficacy in problem-solving.
- Intervention (Video conditions): Participants were exposed randomly to two conditions of AI-generated instructional videos:
    - Condition 1: Video with preview: Participants watched AI-generated instructional videos on Newtonian mechanics with an embedded preview feature (video-with-preview condition).
    - Condition 2: Video without preview: Participants watched the same AI-generated instructional videos on Newtonian mechanics without the embedded preview feature (video-without-preview condition).

  A 15-min break was given between the two conditions to minimize carryover effects.
- Immediate post-test: After viewing both video conditions, participants completed the immediate post-test, which assessed their understanding and application of the concepts presented in the videos. They also completed the PrSSES again to measure any changes in self-efficacy.
- Delayed post-test: Seven days after the intervention, participants completed the delayed post-test to assess retention of the learned material over time.
- Transfer test: Immediately following the delayed post-test, participants completed the transfer test. This assessed their ability to apply the learned concepts to novel problems, specifically a roller coaster design PBL activity.

This counterbalancing ensured that any observed differences between the conditions were not due to the order in which they were presented. It is important to emphasize that the core instructional content of the AI-generated videos was identical in both conditions. Both videos covered the same concepts of Newtonian mechanics and provided the same demonstrations and explanations. The sole difference between the conditions was the inclusion or exclusion of the short preview segment at the beginning of the video. The instructional videos were generated using a variety of platforms for AI-generated video production. Each one was selected based on its ability to create dynamic, engaging content suitable for educational purposes. AI with detailed text scripts that outlined the key concepts of Newtonian mechanics, such as Newton's First and Second Laws of Motion, along with specific descriptions of the animations we wanted to accompany the explanations. For example, the script for Newton's Second Law included a description of a moving car and a force diagram to visually demonstrate the relationship between force, mass, and acceleration. Additionally, we specified visual elements such as color schemes, motion graphics, and the inclusion of text annotations to reinforce the concepts. The AI then generated video segments that included animated explanations, synthesized voiceovers, and text captions. The voiceovers were created using text-to-speech technology, ensuring clear and accurate pronunciation of the key terms and concepts. The animations were designed to be visually engaging and pedagogically appropriate, aimed at enhancing learners' understanding of the physics concepts. Once the videos were generated, they were thoroughly reviewed and edited by two physics educators to ensure both accuracy and pedagogical soundness. The educators verified that the content was correct from a scientific standpoint, ensuring that the animations effectively represented the physics concepts and that the voiceovers were coherent and appropriate for the target audience. Any errors in terminology or conceptual explanations were corrected, and adjustments were made to the animations to improve clarity and enhance educational value.

All tests were administered online using Google Forms. The videos were hosted on an individual server and were accessible through unique links provided to each participant. Participants were instructed to complete all phases of the study individually and without external assistance. The entire procedure, including all tests and video viewings, took approximately two hours to complete.

3.6. Data Collection and Data Analysis

Data were gathered through structured tasks aligned with the instructional videos. The tasks assessed participants' ability to comprehend and apply the concepts presented in the videos. Each task was scored based on accuracy, with zero points for incorrect or incomplete tasks and full points for correct completions.

Data analysis was performed using IBM SPSS 23. Baseline equivalence was assessed through independent-sample t-tests (for age, PrSSES, and pre-test scores) and Chi-Square tests (for gender distribution). Paired-sample t-tests were used to evaluate the effect of the tutorial on self-efficacy and test performance across time points. ANCOVAs were conducted to examine condition effects, with PrSSES and pre-test scores as covariates. The homogeneity of variance was checked beforehand. Significance was set at $\alpha = 0.05$, with effect sizes interpreted per Cohen's guidelines (small: $d = 0.2$, medium: $d = 0.5$, large: $d = 0.8$) (Cortina, 1993).

Data was analyzed using Paired-Sample t-tests to compare pre- and post-test scores within each condition. Additionally, ANCOVAs were conducted for pre-existing differences in participants' knowledge of the scientific concepts. Specifically, pre-test scores on the knowledge assessment were used as covariates. This was done because pre-test scores are a strong predictor of post-test performance and allow us to account for individual differences in baseline knowledge. This approach allowed us to isolate the effect of the instructional videos with and without preview features on learning outcomes, independent of participants' initial knowledge levels

The video tutorial (final video file) designed for this study serves as a key component in integrating AI-generated instructional content into PBL contexts for science teacher education. This process was specifically developed to align with Lin et al.'s (2025) IDEA framework for effective instructional video design, adapting these guidelines for science education. The treatment was organized into five phases, each progressively addressing more complex scientific phenomena and laboratory practices (Table 1).

Table 1. Summarizing the phases and data collection.

Phases	Activity	Data Collected	Time Point
Pre-Intervention	Completion of pre-test and Pre-Self-Efficacy Questionnaire (PrSSES)	Pre-test scores, PrSSES scores	Before Intervention
Intervention	Viewing AI-generated videos (with/without preview—counterbalanced)	Video viewing data (completion, time spent)	During Intervention
Immediate Post-Test	Completion of immediate post-test and Post-Self-Efficacy Questionnaire (PsSSES)	Immediate post-test scores, PsSSES scores	Immediately after Intervention
Delayed Post-Test	Completion of delayed post-test	Delayed post-test scores	7 Days after Intervention
Transfer Test	Completion of transfer test	Transfer test scores	After Delayed Post-Test

3.7. Ethical Considerations

The lead researcher (author) prioritized participant well-being and rights throughout this study, adhering to rigorous ethical guidelines. Ethical protocols encompassed obtaining informed consent, ensuring confidentiality and anonymity, and safeguarding participants' privacy and overall welfare while addressing the ethical, social, and educational implications of utilizing AI technologies in research and instructional design (Al-Zahrani & Alasmari, 2024). Participation was entirely voluntary, with all individuals providing informed consent before data collection and being informed of their right to withdraw from the study at any time without repercussion.

Specific to the development and use of AI-generated video tutorials, ethical considerations extended to the transparency and accountability of AI algorithms, ensuring the generated content was accurate, unbiased, and aligned with educational standards. Potential biases in AI training data were carefully evaluated to prevent unintended reinforcement of stereotypes or misinformation. Specific to the development and use of AI-generated video tutorials, ethical considerations extended to the transparency and accountability of AI algorithms, ensuring the generated content was accurate, unbiased, and aligned with educational standards. Potential biases in AI training data were carefully evaluated to prevent unintended reinforcement of stereotypes or misinformation. Additionally, the main researcher took measures to enhance the naturalness and relatability of AI-generated visuals and narratives, ensuring the content was engaging, culturally sensitive, and supportive of a positive learning environment (Al-Zahrani & Alasmari, 2024).

Before initiating the instructional intervention, participants were thoroughly briefed on the study's purpose (Hamad et al., 2024). They also signed a consent form detailing: (a) the potential implications of using the assessment platforms and AI-generated content, (b) the collection and handling of their data under GDPR, and (c) their unconditional right to withdraw from the study at any point without negative consequences. By addressing both traditional and AI-specific ethical considerations, the study aimed to uphold the highest standards of integrity and respect for participant autonomy. This study complied with the ethical guidelines outlined in the Declaration of Helsinki. Necessary permissions were secured from all participating institutions. Moreover, all subjects were fully informed of their right to withdraw from this study at any time without repercussions. Both verbal and written consent were obtained from each participant (Bush & Grotjohann, 2020).

4. Results

4.1. Descriptive Analysis

A total of 55 participants completed their AI-generated instructional video tasks. A chi-square test indicated no statistically significant difference in gender distribution between the two groups, $\chi^2(1, n = 55) = 0.49$, $p = 0.64$. A p-value of 0.64 is greater than the typical significance level of 0.05, indicating no statistically significant difference in gender distribution between the two groups. The participants' average age was 25.44 years (SD = 0.91), with ages ranging from 22 to 36. Similarly, an independent sample t-test revealed no significant difference in age between the two groups, $t(53) = 0.84$, $p = 0.58$. A p-value of 0.58 is also greater than 0.05, indicating no statistically significant difference in age between the two groups.

The PrSSES scores were comparable across the two conditions. Participants scored an average of 5.74 (SD = 1.41) in the video-with-preview condition and 5.81 (SD = 0.95) in the video-without-preview condition, $t(53) = 0.81$, $p = 0.45$. However, a significant difference emerged in the pre-test scores when participants were first assigned to different starting conditions. Participants who viewed the video-without-preview condition first achieved a higher mean pre-test score (M = 0.52, SD = 0.21) compared to those who viewed the video-with-preview condition first (M = 0.42, SD = 0.22), $t(55) = 3.44$, $p = 0.003$. This significant difference suggests that participants demonstrated greater initial ability in developing AI-generated video tutorials when starting with the video-without-preview condition compared to the video-with-preview condition. Consequently, the baseline task performance was not equivalent across conditions at the outset. The first plot in Figure 3 compares the pre-test scores between the video-with-preview and video-without-preview conditions, with error bars indicating the standard deviations. The second plot in the same figure shows the PrSSES scores for both conditions, again with error bars to represent the variability. The first plot highlights the higher initial scores when participants started with the video-without-preview condition, compared to the video-with-preview condition. The second plot illustrates the close similarity in self-efficacy perceptions (PrSSES scores) between the two conditions, with error bars representing standard deviations.

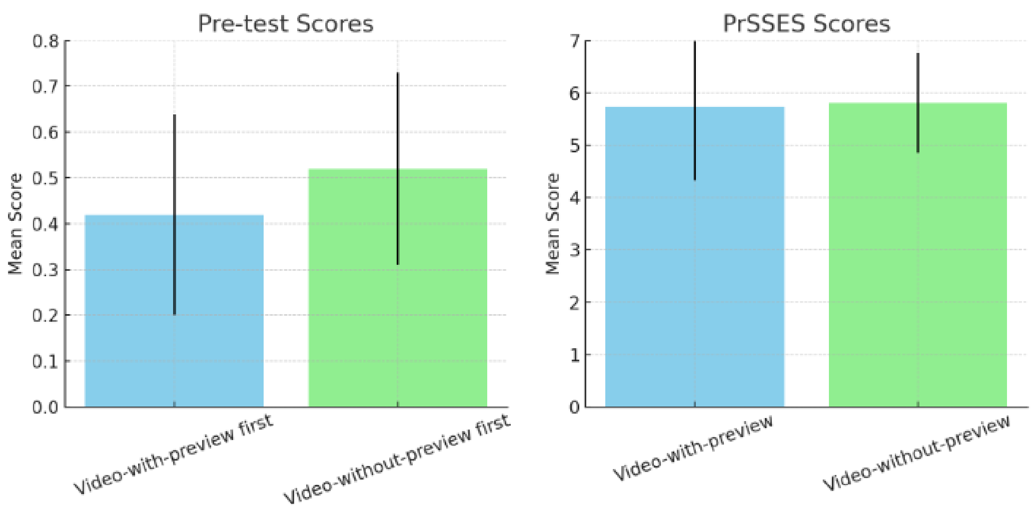

Figure 3. The Pre-test and PrSSES scores by condition.

Some key points of view based on the above analysis are the following:

1. Baseline differences in performance: Participants exhibited significantly higher pre-test scores when starting with the video-without-preview condition compared to the video-with-preview condition, indicating a disparity in initial task performance

capabilities. This suggests that further investigation is needed to address such baseline differences and ensure equivalent starting points for all participants, regardless of the order of conditions.

2. Consistency in self-efficacy perceptions: Despite differences in baseline performance, the PrSSES scores were comparable across the video-with-preview and video-without-preview conditions. This finding suggests that the AI-generated instructional videos were perceived similarly in terms of their ability to support self-efficacy across conditions, regardless of the order in which participants experienced the videos.
3. Implications for instructional design: The significant difference in pre-test scores suggests that the preview feature alone may not be sufficient to ensure equivalent learning conditions. It is important, therefore, to incorporate additional scaffolding strategies or tailored previews to accommodate the diverse starting points of learners, as observed in the baseline performance differences.
4. Generalizability of AI-generated tutorials: Participants showed similar levels of self-efficacy across both conditions, highlighting the broad potential of AI-generated instructional tools to enhance participant confidence, regardless of specific instructional design elements, such as the inclusion or exclusion of a preview feature.

4.2. Top of Form

Self-Efficacy Before and After Training

As illustrated in Figure 4, participants' self-efficacy significantly improved after completing the training (M_post = 6.66, SD_post = 0.72) compared to their self-efficacy before the training (M_pre = 5.72, SD_pre = 1.23), with t(53) = 5.37, *p* = 0.002 t(53) = 5.37, *p* = 0.002 t(53) = 5.37, *p* = 0.002. The effect size, calculated as Cohen's d = 0.96 indicates a large impact of the intervention. These findings confirm the hypothesis that the video tutorial positively influenced participants' self-efficacy. Additionally, the relatively high initial self-efficacy score (M_pre = 5.88), which exceeds the scale's midpoint, suggests that the word formatting tasks were perceived as relatively straightforward by most participants.

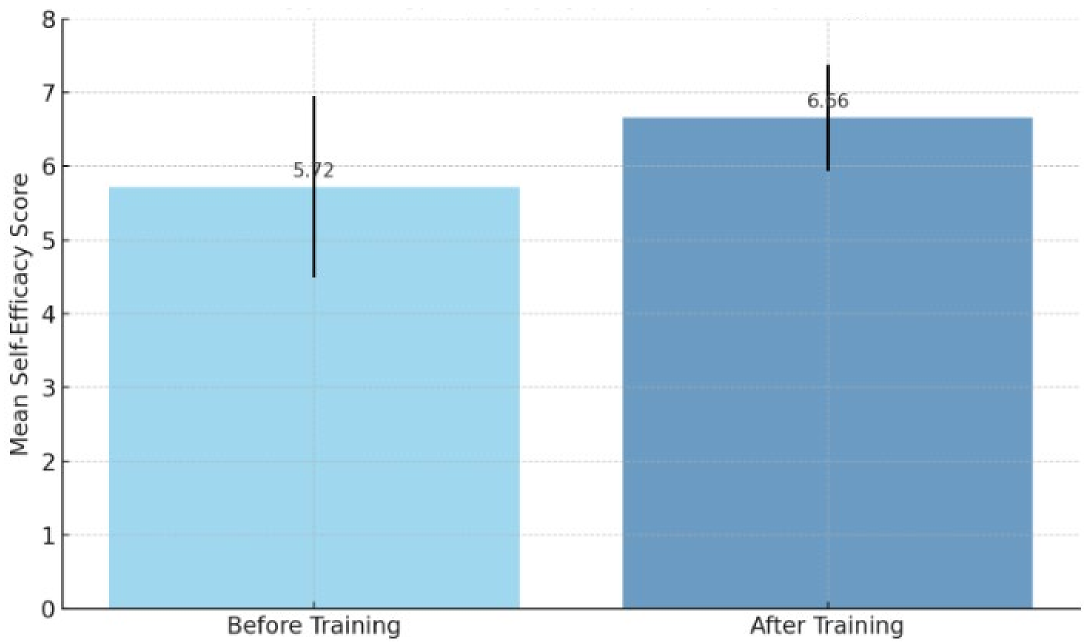

Figure 4. Self-efficacy results before and after training.

While the training significantly improved self-efficacy overall, analyses comparing the two video conditions (with and without the preview feature) revealed no statistically significant differences. A post-hoc power analysis conducted using G*Power with an alpha level of 0.05, an observed effect size of Cohen's d = 0.15, and a sample size of 55, indicated low achieved power (0.35). This suggests a low probability of detecting a statistically significant difference in self-efficacy between the two conditions, even if a true effect of that magnitude existed. Therefore, while the training effectively increased self-efficacy, the non-significant findings regarding the preview feature should be interpreted cautiously, as the study may have been underpowered to detect a small effect.

Regarding the impact of the experimental condition, ANCOVA tests revealed no significant difference between the participants in video-with-preview and video-without-preview conditions, $F(1, 54) = 1.73$, $p = 0.19$. This result does not support the hypothesis that self-efficacy improvements would be significantly greater in the video-with-preview condition compared to their counterparts in video-without-preview condition. Figure 4 illustrates the mean self-efficacy scores of participants before and after treatment, with error bars representing standard deviation. The increase in mean scores suggests a positive effect of the training on self-efficacy across both groups, although this increase was not significantly different between the conditions.

Based on the analysis above, three key points of view should be highlighted:

1. Overall self-efficacy improvement: Participants significantly improved their self-efficacy after training, with a large effect size.
2. Initial confidence: The pre-training self-efficacy scores were already above average, suggesting that the task was perceived as relatively manageable.
3. Condition comparison: No significant difference between instructional video-without-preview and video-with-preview conditions, implying both benefited similarly from the training.

4.3. Task Performances and Learning

Across the four assessments (pre-test, immediate post-test, delayed post-test, and transfer test), participants achieved the highest scores on the delayed post-test (M = 0.75, SD = 0.33), while the pre-test recorded the lowest scores (M = 0.41, SD = 0.22). A detailed descriptive analysis of the results for each test is presented in Figure 5, offering a clear visual comparison of the performance across conditions and test types. The visual and descriptive analysis in Figure 5 highlights the importance of different instructional conditions in enhancing learning outcomes and retention over time.

Some of the most notable points of view are the following:

1. Delayed post-test performance: The delayed post-test scores are the highest across all conditions, indicating a strong retention of the material over time.
2. Pre-test comparison of participants: (a) the pre-test scores are the lowest overall, showing significant room for improvement before the intervention and (b) the participants in video-without-preview conditions scored higher than the video-with-preview group in the pre-test, indicating an initial disparity in baseline performance.
3. Post-test insights: Participants in both conditions show improvement from the pre-test to the immediate post-test, but those in video-with-preview conditions' delayed post-test scores surpass their counterparts, suggesting a longer-term benefit of the preview condition.
4. Transfer test: Scores in the transfer test are consistent across conditions, suggesting that the intervention equally prepared both groups for applying learned concepts to new contexts.

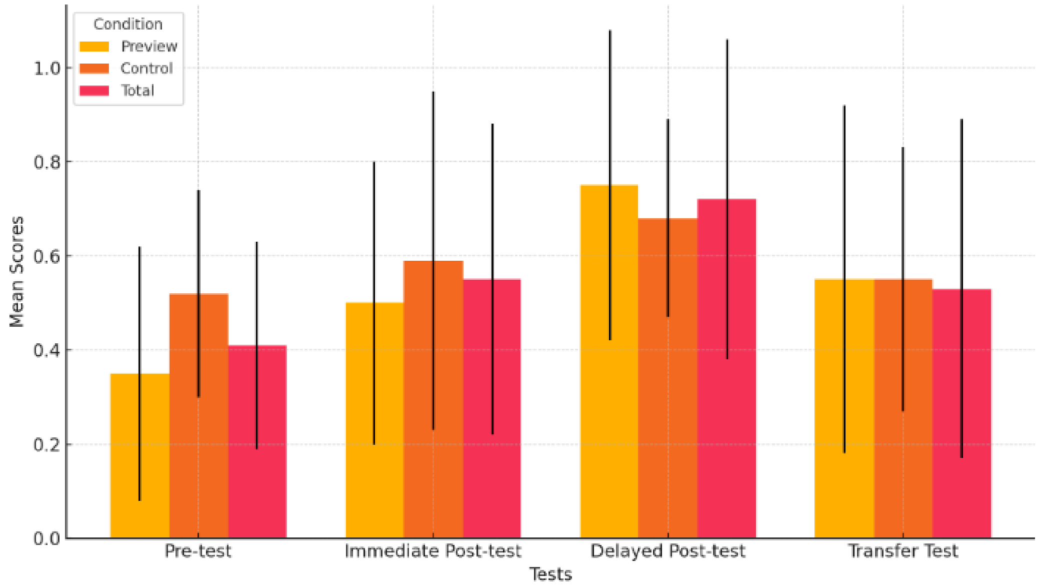

Figure 5. Performance across tests by condition.

One-sided paired-samples t-tests revealed that immediate post-test scores (M = 0.65, SD = 0.42) were significantly higher than pre-test scores (M = 0.41, SD = 0.33), t(53) = 4.58, $p < 0.001$, d = 0.65. This represents a moderate effect size. Similarly, delayed post-test scores (M = 0.69, SD = 0.24), t(53) = 7.95, $p < 0.001$, d = 1.65, and transfer test scores (M = 0.67, SD = 0.32), t(53) = 2.97, $p = 0.002$, d = 0.67, were also significantly higher than pre-test scores. These results confirm that the video tutorial positively impacted participants' task performance on the immediate post-test, delayed post-test, and transfer test (Figure 6).

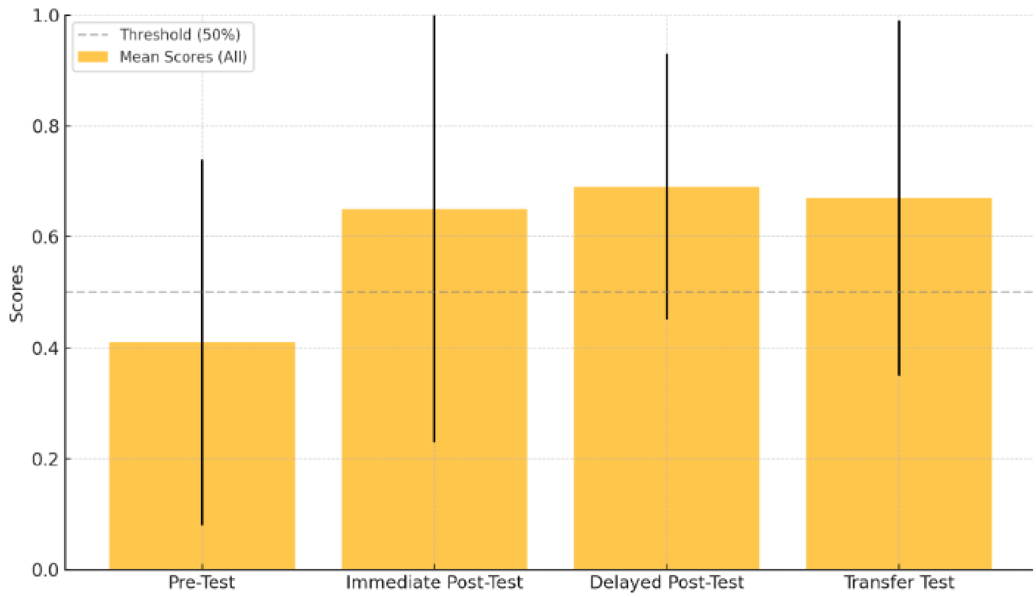

Figure 6. Overall task performance across tests.

The bar charts in Figure 6 show the mean scores for all participants across the four tests (pre-test, immediate post-test, delayed post-test, and transfer test), with error bars representing standard deviations. The performance increases steadily from the pre-test to the delayed post-test, emphasizing the effectiveness of the video tutorials.

An ANCOVA was conducted to examine the effects of conditions, with the pre-test scores included as a covariate. First, the assumption of homogeneity of variance was checked across all tests, and no violations were detected. A one-sided ANCOVA revealed the following:

- For the immediate post-test, the preview condition (M = 0.68, SD = 0.36) did not score significantly higher than the video-without-preview condition (M = 0.62, SD = 0.41), $F(1, 54) = 0.003$, $p = 0.54$.
- For the delayed post-test, the preview condition (M = 0.70, SD = 0.21) did not score significantly higher than the video-without-preview condition (M = 0.71, SD = 0.22), $F(1, 54) = 0.04$, $p = 0.51$.
- For the transfer test, the preview condition (M = 0.53, SD = 0.28) also did not score significantly higher than the video-without-preview condition (M = 0.51, SD = 0.29), $F(1, 54) = 0.22$, $p = 0.73$.

The bar charts compare the performance between conditions for the Immediate post-test, delayed post-test, and transfer test. Error bars show the standard deviations. While the preview consistently shows slightly higher means, the differences are not statistically significant. Therefore, the assumption that the preview feature in the AI-generated instructional video tutorial would lead to participants' performance improvement on the immediate post-test, delayed post-test, and transfer test cannot be supported (Figure 7).

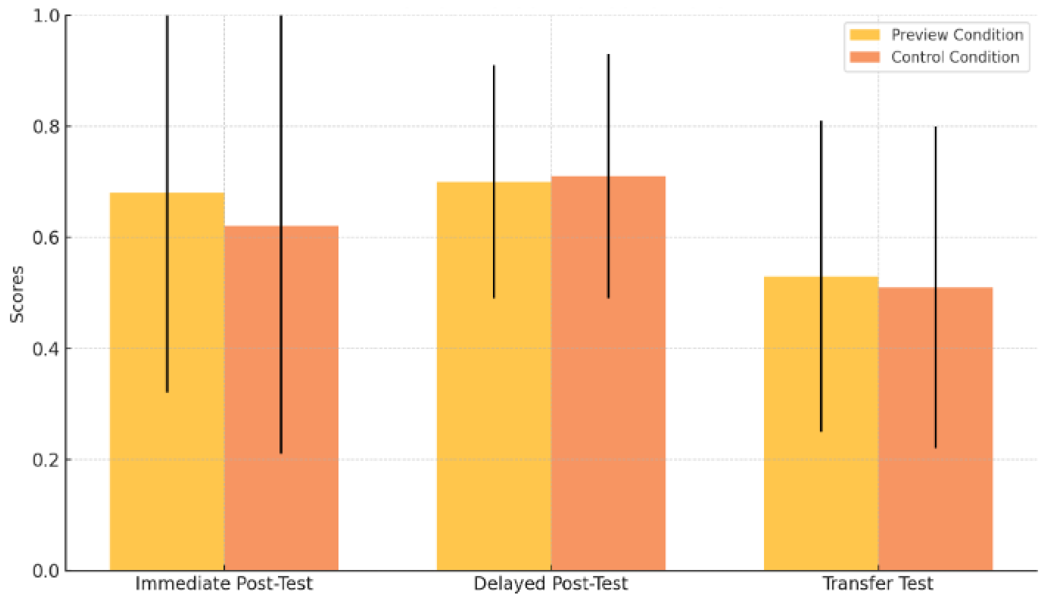

Figure 7. Results from the preview and conditions.

Based on the results presented above, the following key points should be considered:

- Improvement across all tests: The immediate post-test, delayed post-test, and transfer test scores were all significantly higher than the pre-test scores. This suggests that the intervention (AI-generated video tutorial) had a positive impact on participant task performance, improving their ability to complete both tasks immediately and overtime. Specifically, the large effect sizes in the delayed post-test (d = 1.65) and the moderate effect size in the transfer test (d = 0.67) emphasize the sustained benefits of the intervention. These results confirm that the video tutorial facilitated learning and retention of knowledge, extending beyond the immediate learning phase.
- No significant difference between conditions: Although the preview condition showed slight improvements in scores across the immediate post-test, delayed post-test, and transfer test, these improvements were not statistically significant when compared to video-without-preview condition. This suggests that the inclusion of a preview in the AI-generated instructional video did not have a notable additional effect on

performance relative to the video-without-preview condition. This indicates that other factors, such as the video content itself or individual differences, might be influencing task performance more than the preview alone.

- AI-generated instructional video 's effectiveness: Despite the lack of significant differences between conditions, the overall improvement in task performance from the pre-test to the subsequent tests (immediate, delayed, and transferred) supports the effectiveness of the AI-generated video tutorials in enhancing participants' understanding and task execution. The video tutorials, whether with or without a preview, appear to be a useful tool for promoting learning outcomes.
- Potential limitations of the preview effect: The absence of significant findings regarding the preview conditions may point to potential limitations in how the preview was implemented. While previews are generally helpful in setting up context and expectations, it seems that the format, content, or delivery of the preview might not have been optimized in this case to enhance task performance. Future research could explore different approaches to delivering previews or investigate other factors that could interact with the preview to better support learning.

In summary, the key aspects of these results are that AI-generated video tutorials have a positive effect on learning outcomes, but the specific role of a preview in enhancing performance remains unclear and warrants further exploration. These findings suggest that while AI-generated video tutorials are effective at enhancing participant performance, further investigation is needed into the precise mechanisms that make them successful. Research can further focus on refining the preview strategies, examining the role of other instructional features (such as interactive elements), or exploring the impact of learners' prior knowledge and motivation on tutorial effectiveness.

5. Discussion

This study aimed to examine the impact of AI-generated instructional videos on the learning outcomes of participants (pre-service science teachers) in PBL contexts, with particular emphasis on the role of preview content and its effect on pre-service science teachers' self-efficacy, task performance, and retention. The results presented here provide valuable insights into the effectiveness of AI-generated videos and the associated instructional features, such as the preview, in enhancing learning for future science educators.

Regarding RQ1, this study's results demonstrate that the use of AI-generated instructional videos significantly enhanced participants' task performance, as evidenced by the improvements observed in the immediate post-test, delayed post-test, and transfer test scores. These results are also consistent with those from previous studies (Arkün-Kocadere & Çağlar-Özhan, 2024; Pi et al., 2019; Rasi & Poikela, 2016), showing a clear increase in performance from the pre-test (M = 0.41) to the subsequent tests, with the delayed post-test (M = 0.75) achieving the highest scores. This pattern is consistent with the notion that AI-generated videos provide learners with a clear and engaging medium for acquiring and retaining knowledge, reinforcing the positive impact of multimedia learning (Gumisirizah et al., 2024; Son, 2024). The fact that the transfer test scores were also significantly higher than the pre-test suggests that the participants were not only able to remember information but were also able to apply it to novel situations. These results aligned with a portion of researchers (Arkün-Kocadere & Çağlar-Özhan, 2024; Li et al., 2024; Semeraro & Vidal, 2022) and showed that multimedia learning environments, particularly those employing AI technology, can enhance both the retention and application of learned material. Importantly, the improvement in transfer tasks emphasizes the potential of AI-generated videos to foster critical thinking and real-world problem-solving skills, which are essential for future educators.

The preview, as an instructional feature, was expected to play a significant role in setting up the learners' cognitive framework and promoting higher self-efficacy and performance (Bandura, 2006). However, the results of the answer RQ2 did not support the hypothesis that the preview condition would lead to a significantly higher task performance compared to the video-without-preview condition. Both the immediate post-test, delayed post-test, and transfer test results showed no significant differences between the preview and video-without-preview conditions, suggesting that the preview did not have a substantial effect on task performance, which comes in line with findings from previous works (Al-Zahrani & Alasmari, 2024; Brame, 2016). This finding is somewhat surprising given the theoretical backing for the preview's role in enhancing learning outcomes. In line with Lin et al.'s (2025) IDEA framework, the preview was designed to introduce participants to key concepts and reduce cognitive load. While the preview may have reduced cognitive load and provided context for the learners, it seems that its impact was not strong enough to create significant differences in performance. One possible explanation is that the format or content of the preview did not effectively engage participants in active cognitive processing, as suggested by previous research (Aidoo, 2023; Kumar, 2010). Alternatively, the effectiveness of the preview might depend on other factors, such as learners' prior knowledge or the instructional design of the tutorial itself. The null effect of the preview condition also highlights the complexity of instructional design in AI-generated video tutorials. Although the preview provided some context and clarity, it may not have been sufficiently engaging or tailored to the individual needs of the participants. Future research could explore variations in how previews are presented—perhaps incorporating interactive elements or asking participants to predict outcomes based on the preview, which might foster more active engagement and, consequently, better learning outcomes (Lange & Costley, 2020; Pellas, 2024; Saqr & López-Pernas, 2024).

Another key aspect of this study was the impact of the AI-generated video tutorials on participants' self-efficacy regarding RQ3. The results indicated that participants' self-efficacy scores significantly improved from the pre-test (M = 5.72) to post-training (M = 6.66), with a large effect size (d = 0.96). This finding underscores the positive influence of video tutorials on participants' confidence in their abilities to perform the tasks associated with the tutorial. The improvement in self-efficacy is consistent with previous research (Li et al., 2024; Shahzad et al., 2024), which has shown that multimedia instructional tools can enhance participants' perceived competence and motivation (Van Der Meij et al., 2018b). Interestingly, the comparison between conditions did not reveal significant differences in self-efficacy scores, suggesting that while video tutorials overall contributed to an increase in self-efficacy, the preview did not appear to provide additional benefits over the video-without-preview condition. This could suggest that the inherent benefits of the instructional videos themselves, such as clear instructions and demonstrations, were sufficient to boost participants' confidence without the need for a preview. Therefore, this finding stresses the importance of refining instructional elements to maximize their impact on participants' outcomes.

While the training itself led to significant improvements in self-efficacy and overall task performance, no statistically significant differences were observed between the AI-generated instructional videos with and without the preview feature across any of the assessed metrics (self-efficacy, immediate and delayed task performance, and transfer). The effect sizes for these comparisons were small to negligible: Cohen's d = 0.15 for self-efficacy, Cohen's d = 0.65 for immediate task performance, Cohen's d = 1.65 for delayed task performance, and Cohen's d = 0.67 for transfer. These small effect sizes suggest that the preview feature had a minimal practical impact on learning outcomes, as Sullivan and Feinn (2012) suggest.

Several factors may explain these non-significant findings. One possibility is that the preview did not provide sufficiently novel or crucial information beyond what was already presented in the main video content. Furthermore, participants may not have effectively utilized the preview, perhaps skipping it or not fully engaging with its content. The design of the preview itself may have also played a role. Future iterations of the preview could be improved by optimizing its timing (e.g., presenting it before complex sections), focusing its content on the most critical information, and incorporating interactive elements to promote active processing. An adaptive preview that tailors its content to individual learner needs could also be explored. These changes could enhance the preview's effectiveness and potentially lead to more significant learning gains.

6. Implications

The results of this study have several important implications for the design, implementation, and evaluation of AI-generated instructional tools in educational contexts, particularly in science and PBL settings. First, the findings confirm that AI-generated videos are a viable and effective medium for supporting learning, particularly in fostering retention and transfer of knowledge. The improvement demonstrated in participants' delayed post-test and transfer test performance highlights the potential of AI-driven multimedia tools to promote deeper learning outcomes, including the application of knowledge in science teacher education. By leveraging AI technologies such as adaptive avatars and natural language processing, these tools can provide a dynamic, engaging, and potentially personalized approach to content delivery that aligns with diverse learner needs. AI-generated instructional tools can also play a transformative role in bridging gaps in resource-limited educational settings. Their scalability and ability to deliver high-quality, consistent content make them a valuable asset for institutions seeking to expand access to high-quality education. This is particularly relevant in teacher training programs, where such tools can enhance pedagogical practices and prepare future educators for integrating technology into their classrooms.

The mixed findings regarding the preview feature offer valuable lessons for instructional design. While previews are theorized to set expectations and reduce cognitive load by providing an overview of key concepts, this study suggests that their efficacy depends on additional contextual factors. Instructional designers must carefully evaluate how features like previews are implemented. For example, interactive previews that prompt learners to make predictions or actively engage with the material may yield stronger learning outcomes compared to passive content delivery. Similarly, scaffolding strategies, such as linking preview content to specific tasks or goals, could enhance their relevance and impact. This underscores the importance of iterative design and testing in AI-generated instructional tools. Designers should consider learner feedback, content complexity, and pacing when integrating such features. Moreover, incorporating adaptive learning algorithms that adjust preview content based on individual learner profiles or prior knowledge could make these tools more responsive and effective.

On the one hand, the substantial improvement in self-efficacy observed across conditions brings to light a crucial advantage of AI-generated instructional tools, which is their ability to boost learners' confidence in their abilities. This finding is particularly significant in PBL contexts, where self-efficacy is a key predictor of success. AI tools that integrate motivational elements, such as encouraging feedback from avatars or personalized progress tracking, can further enhance learners' confidence and engagement. On the other, the lack of differential effects between the video-with-preview and video-without-preview conditions suggests that the core design of the video tutorials, including clear demonstrations and step-by-step guidance, was sufficient to support self-efficacy. Future developments could explore how additional features, such as reflective questions or gamified elements, might further amplify this effect.

7. Limitations and Directions for Future Research

While this study provides valuable insights into the use of AI-generated instructional videos, several limitations should be noted. The sample size, though adequate for this analysis, may limit the generalizability of the findings, and future research with a larger and more diverse sample could provide more robust conclusions. Additionally, while the study focused on the impact of video tutorials on self-efficacy and task performance, future research could explore other variables, such as participants' motivation, engagement, and perceptions of the technology.

Another potential avenue for future research is the examination of the long-term impact of AI-generated video tutorials. While the delayed post-test indicated sustained learning benefits, the retention and application of skills over a longer period remain an important area to explore. Longitudinal studies could assess whether the positive effects of AI-generated tutorials persist over time or if additional instructional interventions are required to maintain high levels of engagement and performance.

This study's findings open up avenues for future research. For instance, the role of the preview warrants further investigation. Future studies could experiment with different types of previews, such as interactive previews or those that involve active participation, to determine whether these approaches can more effectively support learning outcomes. Additionally, examining how individual learner characteristics, such as prior knowledge or learning styles, interact with the preview and other instructional features could provide further insights into optimizing AI-generated tutorials. Studies should also explore the role of interactivity, adaptive features, and learner agency in AI-generated instructional tools. For example, incorporating real-time feedback mechanisms or allowing learners to navigate content non-linearly could lead to more personalized and impactful learning experiences. Investigating how these tools can be adapted for diverse educational contexts, such as cross-disciplinary PBL programs or culturally specific learning environments, would further broaden their applicability. Lastly, this study's findings underline the need for longitudinal research to evaluate the long-term impact of AI-generated instructional tools on learners' knowledge retention, skill application, and overall academic performance. Such studies can provide deeper insights into how these tools shape learning trajectories over time and inform best practices for their integration into curricula.

8. Conclusions

The present study aspires to bridge the gap between the theoretical potential of AI-generated instructional video tutorials and its practical application in science training programs. The findings aim to provide a foundation for designing innovative pedagogical strategies that integrate AI-generated videos into PBL settings, enhancing both the learning experience and the professional preparedness of future science educators. The current study contributes to the evolving narrative of how technology reshapes education, empowering teachers to inspire the next generation of scientists and critical thinkers. It also aligns with broader educational goals of fostering digital competence among future educators. As schools increasingly adopt technology-enhanced learning environments, science teachers must understand and utilize advanced digital tools effectively. By experiencing AI-driven instructional resources in their training, future teachers gain not only subject-specific knowledge but also a deeper appreciation of how technology can be leveraged to address diverse learning needs. AI-generated instructional tools hold tremendous promises for transforming education, but their success hinges on thoughtful design and evidence-based implementation. By addressing the nuanced challenges highlighted in this study, educators and designers can harness the full potential of these tools to foster meaningful, sustained learning outcomes.

In conclusion, this study contributes to the growing body of research on the use of AI-generated instructional videos in science teacher education. The results indicate that these tools are effective in enhancing participants' learning outcomes, particularly in terms of knowledge retention and transfer tasks. While the preview did not have a significant impact on task performance, the overall positive effects of the AI tutorials for science education themselves suggest that they can be a valuable resource in future teacher training programs. This study also contributes to broader educational goals of fostering digital competence among future educators. Digital competence for educators encompasses a range of skills, knowledge, and attitudes related to the effective and responsible use of digital technologies in teaching and learning. This includes not only the ability to use digital tools but also the capacity to critically evaluate digital resources, integrate technology into pedagogical practices, and foster students' digital literacy.

The current work directly contributes to pre-service teachers' digital competence by providing them with hands-on experience using and evaluating AI-generated instructional videos. This experience enhances their ability to integrate digital resources into their pedagogical practices (Göltl et al., 2024). By comparing different versions of the instructional videos, participants are encouraged to consider the impact of specific design features, such as the preview, on learning outcomes, thus developing their evaluative skills related to digital content. Furthermore, this study contributes to a broader understanding of the pedagogical potential of AI in education. By experiencing AI-driven instruction, pre-service teachers gain insights into how these technologies can be leveraged to enhance learning and support different learning styles.

While this study provides valuable insights into the effectiveness of AI-generated instructional videos in teacher training, several avenues for future research remain. Future studies should investigate the effectiveness of interactive and adaptive previews, incorporating elements such as short quizzes or personalized content based on learner needs. Employing eye-tracking techniques could also provide valuable insights into how learners interact with the preview feature. Furthermore, exploring the use of gamification and personalized feedback within the video tutorials could further enhance learner engagement and motivation. Finally, future research could explore broader applications of AI in teacher training, such as using AI to provide automated feedback on teacher practice or developing AI-powered training simulations for classroom management.

Funding: This research was not funded by any source.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to federal data protection regulations (no personal data was collected during the study).

Informed Consent Statement: Participation was completely voluntary, and informed consent was obtained from all participants before data collection.

Data Availability Statement: Data is available upon request to the corresponding author.

Acknowledgments: The author would like to thank all participants who took part in this study.

Conflicts of Interest: The author declares no conflict of interest.

References

Aidoo, B. (2023). Teacher educators' experience adopting problem-based learning in science education. *Education Sciences*, *13*(11), 1113. [CrossRef]

Akçay, B. (2009). Problem-based learning in science education. *Journal of Turkish Science Education*, *6*(2), 26–36. Available online: https://www.tused.org/index.php/tused/article/view/104 (accessed on 3 December 2024).

Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities and Social Sciences Communications*, *11*, 912. [CrossRef]

Arkün-Kocadere, S., & Çağlar-Özhan, Ş. (2024). Video lectures with AI-generated instructors: Low video engagement, same performance as human instructors. *The International Review of Research in Open and Distributed Learning*, *25*(3), 350–369. [CrossRef]

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares, & T. C. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Information Age Publishing.

Bewersdorff, A., Hartmann, C., Hornberger, M., Seßler, K., Bannert, M., Kasneci, E., & Nerdel, C. (2024). Taking the next step with generative artificial intelligence: The transformative role of multimodal large language models in science education. *arXiv*, arXiv:2401.00832.

Blummer, B. A., & Kritskaya, O. (2009). Best practices for creating an online tutorial: A literature review. *Journal of Web Librarianship*, *3*(3), 199–216. [CrossRef]

Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE Life Sciences Education*, *15*, es6. [CrossRef] [PubMed]

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. [CrossRef]

Bush, A., & Grotjohann, N. (2020). Collaboration in teacher education: A cross-sectional study on future teachers' attitudes towards collaboration, their intentions to collaborate and their performance of collaboration. *Teaching and Teacher Education*, *88*, 102968. [CrossRef]

Cortina, J. M. (1993). What is the coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98. [CrossRef]

Costley, J., Fanguy, M., Lange, C., & Baldwin, M. (2020). The effects of video lecture viewing strategies on cognitive load. *Journal of Computing in Higher Education*, *33*, 19–38. [CrossRef]

Ester, P., Morales, I., & Herrero, L. (2023). Micro-videos as a learning tool for professional practice during the post-COVID era: An educational experience. *Sustainability*, *15*, 5596. [CrossRef]

Göltl, K., Ambros, R., Dolezal, D., & Motschnig, R. (2024). Pre-service teachers' perceptions of their digital competencies and ways to acquire those through their studies and self-organized learning. *Education Sciences*, *14*, 951. [CrossRef]

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological Bulletin*, *83*(2), 314. [CrossRef]

Gumisirizah, N., Nzabahimana, J., & Muwonge, C. M. (2024). Supplementing problem-based learning approach with video resources on students' academic achievement in physics: A comparative study between government and private schools. *Education and Information Technologies*, *29*, 13133–13153. [CrossRef]

Hamad, Z. T., Jamil, N., & Belkacem, A. N. (2024). ChatGPT's impact on education and healthcare: Insights, challenges, and ethical considerations. *IEEE Access*. [CrossRef]

Hurzlmeier, M., Watzka, B., Hoyer, C., Girwidz, R., & Ertl, B. (2021). Visual cues in a video-based learning environment: The role of prior knowledge and its effects on eye movement measures. In E. De Vries, Y. Hod, & J. Ahn (Eds.), *Proceedings of the 15th International Conference of the Learning Sciences—ICLS 2021* (pp. 3–10). International Society of the Learning Sciences. [CrossRef]

Johnson, C., Hill, L., Lock, J., Altowairiki, N., Ostrowski, C., Da Rosa Dos Santos, L., & Liu, Y. (2017). Using design-based research to develop meaningful online discussions in undergraduate field experience courses. *International Review of Research in Open and Distributed Learning*, *18*(6), 36–53. [CrossRef]

Koumi, J. (2013). Pedagogic design guidelines for multimedia materials: A call for collaboration between practitioners and researchers. *Journal of Visual Literacy*, *32*(2), 85–114. [CrossRef]

Kulgemeyer, C. (2018). A framework of effective science explanation videos informed by criteria for instructional explanations. *Research in Science Education*, *50*, 2441–2462. [CrossRef]

Kumar, D. D. (2010). Approaches to interactive video anchors in problem-based science learning. *Journal of Science Education and Technology*, *19*(1), 13–19. [CrossRef]

Lange, C., & Costley, J. (2020). Improving online video lectures: Learning challenges created by media. *International Journal of Educational Technology in Higher Education*, *17*(16), 1–18. [CrossRef]

Li, B., Wang, C., Bonk, C. J., & Kou, X. (2024). Exploring inventions in self-directed language learning with generative AI: Implementations and perspectives of YouTube content creators. *TechTrends*, *68*, 803–819. [CrossRef]

Lim, J. (2024). The potential of learning with AI-generated pedagogical agents in instructional videos. In *Extended abstracts of the CHI conference on human factors in computing systems* (pp. 1–6). ACM: USA. [CrossRef]

Lin, C., Zhou, K., Li, L., & Sun, L. (2025). Integrating generative AI into digital multimodal composition: A study of multicultural second-language classrooms. *Computers and Composition*, *75*, 102895. [CrossRef]

Magaji, A., Adjani, M., & Coombes, S. (2024). A systematic review of preservice science teachers' experience of problem-based learning and implementing it in the classroom. *Education Sciences*, *14*(3), 301. [CrossRef]

Marshall, E. (2024). Examining the effectiveness of high-quality lecture videos in an asynchronous online criminal justice course. *Journal of Criminal Justice Education*, 1–18. [CrossRef]

Martin, N. A., & Martin, R. (2015). Would you watch it? Creating effective and engaging video tutorials. *Journal of Library & Information Services in Distance Learning*, *9*(1–2), 40–56. [CrossRef]

Newbold, N., & Gillam, L. (2010). Text readability within video retrieval applications: A study on CCTV analysis. *Journal of Multimedia*, *5*(2), 123–141. [CrossRef]

Pellas, N. (2023a). The effects of generative AI platforms on undergraduates' narrative intelligence and writing self-efficacy. *Education Sciences*, *13*(11), 1155. [CrossRef]

Pellas, N. (2023b). The influence of sociodemographic factors on students' attitudes toward AI-generated video content creation. *Smart Learning Environments*, *10*, 276. [CrossRef]

Pellas, N. (2024). The role of students' higher-order thinking skills in the relationship between academic achievements and machine learning using generative AI chatbots. *Research and Practice in Technology Enhanced Learning*, *20*, 36. [CrossRef]

Pi, Z., Zhang, Y., Zhu, F., Xu, K., Yang, J., & Hu, W. (2019). Instructors' pointing gestures improve learning regardless of their use of directed gaze in video lectures. *Computers & Education*, *128*(1), 345–352. [CrossRef]

Qu, Y., Tan, M. X. Y., & Wang, J. (2024). Disciplinary differences in undergraduate students' engagement with generative artificial intelligence. *Smart Learning Environments*, *11*, 51. [CrossRef]

Rasi, P. M., & Poikela, S. (2016). A review of video triggers and video production in higher education and continuing education PBL settings. *Interdisciplinary Journal of Problem-Based Learning*, *10*(1), 7. [CrossRef]

Reeves, T. C., Herrington, J., & Oliver, R. (2005). Design research: A socially responsible approach to instructional technology research in higher education. *Journal of Computing in Higher Education*, *16*(2), 96–115. [CrossRef]

Rismark, M., & Sølvberg, A. M. (2019). Video as a learner scaffolding tool. *International Journal of Learning, Teaching and Educational Research*, *18*(1), 62–75. [CrossRef]

Saqr, M., & López-Pernas, S. (2024). Why explainable AI may not be enough: Predictions and mispredictions in decision making in education. *Smart Learning Environments*, *11*, 343. [CrossRef]

Savinainen, A., Scott, P., & Viiri, J. (2004). Using a bridging representation and social interactions to foster conceptual change: Designing and evaluating an instructional sequence for Newton's third law. *Science Education*, *89*(2), 175–195. [CrossRef]

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 60–71. [CrossRef]

Semeraro, A., & Vidal, L. T. (2022, April 30–May 5). *Visualizing instructions for physical training: Exploring visual cues to support movement learning from instructional videos*. CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA. [CrossRef]

Shahzad, M. F., Xu, S., & Asif, M. (2024). Factors affecting generative artificial intelligence, such as ChatGPT, use in higher education: An application of technology acceptance model. *British Educational Research Journal*. [CrossRef]

Shaukat, S., Wiens, P., & Garza, T. (2024). Development and validation of the Standards Self-Efficacy Scale (SSES) for US pre-service teachers. *Educational Research: Theory and Practice*, *35*, 71–87.

Simanjuntak, L. Y. A., Perangin-Angin, R. B., & Saragi, D. (2019). Development of scientific-based learning video media using problem-based learning (PBL) model to improve student learning outcomes in fourth grade students of elementary school Parmaksian, Kab. Sambair Toba. *Budapest International Research and Critics in Linguistics and Education Journal*, *2*(4), 297–304. [CrossRef]

Son, T. (2024). Noticing classes of preservice teachers: Relations to teaching moves through AI chatbot simulation. *Education and Information Technologies*. [CrossRef]

Sullivan, G. M., & Feinn, R. (2012). Using effect size—Or why the p value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. [CrossRef]

Taningrum, N., Kriswanto, E. S., Pambudi, A. F., & Yulianto, W. D. (2024). Improving critical thinking skills using animated videos based on problem-based learning. *Retos*, *57*, 692–696. [CrossRef]

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics*, *66*(4), 338–352. [CrossRef]

Van Der Meij, H. (2017). Reviews in instructional video. *Computers & Education*, *114*, 164–174. [CrossRef]

Van Der Meij, H., Rensink, I., & Van Der Meij, J. (2018a). Effects of practice with videos for software training. *Computers in Human Behavior*, *89*, 439–445. [CrossRef]

Van Der Meij, H., & Van Der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & Education*, *78*, 150–159. [CrossRef]

Van Der Meij, H., Van Der Meij, J., & Voerman, T. (2018b). Supporting motivation, task performance, and retention in video tutorials for software training. *Educational Technology Research and Development*, *66*(3), 597–614. [CrossRef]

Wang, X., Lin, L., Han, M., & Spector, J. M. (2020). Impacts of cues on learning: Using eye-tracking technologies to examine the functions and designs of added cues in short instructional videos. *Computers in Human Behavior*, *107*, 106279. [CrossRef]

Zhang, C., Wang, Z., Fang, Z., & Xiao, X. (2024). Guiding student learning in video lectures: Effects of instructors' emotional expressions and visual cues. *Computers & Education*, *218*, 105062. [CrossRef]