Check for updates

OPEN ACCESS

EDITED BY Joana Carneiro Pinto, Catholic University of Portugal, Portugal

REVIEWED BY Ozden Sengul, Boğaziçi University, Türkiye Christopher Nakamura, Saginaw Valley State University, United States

*correspondence Xian Wu ⊠ xian.wu@uconn.edu

RECEIVED 12 February 2025 ACCEPTED 28 April 2025 PUBLISHED 16 May 2025

CITATION

Wu X, Li Y, Zu T, Hutson J, Loschky LC and Rebello NS (2025) Using multimedia hints to facilitate conceptual problem solving in physics: investigating the effects of multiple modalities. *Front. Educ.* 10:1568406.

doi: 10.3389/feduc.2025.1568406

COPYRIGHT

© 2025 Wu, Li, Zu, Hutson, Loschky and Rebello. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Using multimedia hints to facilitate conceptual problem solving in physics: investigating the effects of multiple modalities

Xian Wu¹*, Yaoguang Li^{2,3}, Tianlong Zu⁴, John Hutson⁵, Lester C. Loschky⁶ and N. Sanjay Rebello^{7,8}

¹Department of Physics, University of Connecticut, Storrs, CT, United States, ²AGQ Solutions, South Windsor, CT, United States, ³UConn Library, University of Connecticut, Storrs, CT, United States, ⁴Department of Physics and Astronomy, Northwestern University, Evanston, IL, United States, ⁵Educational Testing Service, Princeton, NJ, United States, ⁶Department of Psychological Sciences, Kansas State University, Manhattan, KS, United States, ⁷Department of Physics and Astronomy, Purdue University, West Lafayette, IN, United States, ⁸Department of Curriculum & Instruction, Purdue University, West Lafayette, IN, United States

Multimedia hints are widely used in educational materials to support conceptual learning, yet their comparative effectiveness across modalities remains underexplored. Prior studies suggest that graphical hints can enhance learners' performance on physics problems, but it is unclear how they interact with other modalities such as text and voice. Understanding these interactions is essential for designing effective instructional tools. In this study, we investigated the effects of graphical, typographic, and vocal hints, individually and in combination, on students' problemsolving performance. A total of 162 students from a conceptual physics course participated in individual interviews and solved four sets of isomorphic problems. Each set included an initial problem (pretest), six training problems, a near transfer problem, and a far transfer problem. We employed a $2 \times 2 \times 2$ between-subject quasi-experimental design to examine the effects of the three hint modalities. Results from paired-sample t-tests showed significant performance gains from pretest to both near and far transfer tasks, indicating that solving isomorphic problems with hints promotes learning. Among modalities, graphical hints led to better training performance than typographic or vocal hints. Notably, combining typographic and vocal hints produced worse outcomes than using either modality alone, contradicting the auditory superiority effect and suggesting potential cognitive overload. These findings highlight the effectiveness of visual support and caution against indiscriminate integration of multiple hint modalities. We provide evidencebased recommendations for designing multimedia instructional materials that optimize cognitive processing and support conceptual problem solving in physics.

KEYWORDS

physics, problem-solving, multimedia hints, auditory superiority effect, modality

1 Introduction

In today's educational landscape, multimedia has become an essential tool for delivering instructional content, particularly in STEM fields where abstract concepts often require visual and interactive aids to support student understanding. In disciplines like physics, where students must grasp complex, concept-heavy material, multimedia resources provide unique value. However, designing effective multimedia for physics education requires more than visual and auditory appeal; it demands careful consideration of the cognitive demands placed on learners, especially in problem-solving contexts.

Physics education researchers have long recognized that problemsolving is central to developing deep understanding in the discipline, though it presents unique challenges, particularly for novices (Docktor et al., 2016; Burkholder et al., 2020). Effective problem-solving relies on leveraging prior knowledge to overcome cognitive challenges, a process that can be enhanced through well-designed instructional materials. Hints, multimedia elements that guide learners toward relevant information, play a pivotal role in facilitating problemsolving. However, the relative effectiveness of different hint modalities remains underexplored.

Research on multimedia learning emphasizes the importance of aligning instructional designs with cognitive theories, such as Mayer's (2017) Cognitive Theory of Multimedia Learning (CTML) and Wickens (2002) Multiple Resources Theory. These frameworks highlight the benefits of using dual-channel processing and multimodal information to reduce cognitive load. Graphical hints have been shown to direct attention effectively and aid comprehension by leveraging visual-spatial processing pathways. Typographic and vocal hints, on the other hand, rely on linguistic processing and may vary in effectiveness based on their presentation format and interaction with graphical elements.

Despite these theoretical insights, prior studies have rarely compared the effectiveness of graphical, typographic, and vocal hints in a systematic way. This study addresses this gap by examining how these modalities influence problem-solving performance in conceptual physics tasks. By integrating insights from CTML and Multiple Resources Theory, we aim to provide a comprehensive understanding of how multimedia hints can optimize cognitive resources and enhance problem-solving success.

2 Theoretical background

To guide our study design, we developed a theoretical framework (Figure 1) by synthesizing insights from several well-established cognitive theories. We grounded our framework in Ohlsson's (1992) modified Representational Change Theory to explain the cognitive mechanisms underlying impasse resolution. To account for how multimedia hints influence these mechanisms, we incorporated Mayer's (2017) Cognitive Theory of Multimedia Learning (CTML) and Wickens (2002) Multiple Resources Theory. These theories, when combined, help explain how different modalities of hints (i.e., graphical, typographic, vocal) may support representational change by leveraging distinct perceptual channels and cognitive resources. The resulting framework serves to bridge problem-solving processes and multimedia design principles in a unified model.



shared processing resources, while response generation is treated separately

Because our framework focuses on the role of hints in supporting conceptual problem-solving, it is important to first clarify the cognitive nature of problem-solving itself and how it differs from learning. Although the two processes are related, they operate differently (Schnotz and Kürschner, 2007). Learning involves associating newly gathered information with prior knowledge to create schemas, enabling the encoding of new information from working memory into long-term memory. In contrast, problemsolving relies on retrieving prior knowledge from long-term memory to address novel situations and may not directly result in learning. For example, to understand Faraday's Law, learners might be tasked with solving for the current induced in a rod moving in a magnetic field or determining the electric potential on a spinning rod. These tasks require learners to transform the given state (i.e., a moving or spinning rod) into the goal state (i.e., the current or potential on the rod). However, understanding Faraday's Law, which might emerge as a byproduct of this transformation process, is not necessarily encoded in long-term memory during problem-solving. Since this study focuses on conceptual physics problem-solving, clarifying the definitions of "problem" and "problem-solving" is essential (Lestari Syafril et al., 2021).

Jonassen (2010) defined problem-solving as "a question or issue that is uncertain and so must be examined and solved," emphasizing the cognitive challenges inherent in solving problems. Insight problems, a specific class of problems characterized by an impasse, exemplify these challenges. Dow and Mayer (2004) defined insight problems as "a special type of non-routine problem in which the problem primes an inappropriate solution procedure that is familiar to the problem solver" (p. 389). An impasse occurs when solvers apply familiar but unsuitable strategies to a problem. Breaking through an impasse often results in an "Aha" moment, a sudden realization of how to proceed. Insight problem-solving, therefore, involves overcoming initial failure to achieve eventual success.

Ohlsson's (1992) modified Representational Change Theory provides a framework for understanding how solvers encounter and resolve impasses in insight problem-solving. According to this theory, impasses arise when a solver's mental representation of a problem limits the activation of necessary prior knowledge. To break the impasse, the unproductive mental representation must be altered through one of three mechanisms: elaboration, re-encoding, or constraint relaxation. Elaboration involves adding information internally (e.g., recalling relevant knowledge) or externally (e.g., receiving hints). Re-encoding entails restructuring the problem's mental representation, while constraint relaxation involves loosening perceived restrictions on potential solutions. These mechanisms help solvers shift focus from irrelevant to relevant information, facilitating progress.

Problem-solving hints play a critical role in this process. Research has shown that learners often encounter impasses by focusing on thematically irrelevant information within problems (Madsen et al., 2012; Rouinfar et al., 2014). Well-designed graphical hints can help learners restructure their problem representation, redirect attention to relevant information, and activate more effective prior knowledge, enabling them to overcome impasses (see the left side of Figure 1). For example, Thomas and Lleras (2007) demonstrated that visual hints embedded in a seemingly unrelated task sequence could effectively guide attention and significantly improve performance on insight problems. In more recent educational technology contexts, a systematic review by Albus et al. (2021) found that collaborative learning in virtual reality environments can enhance problem-solving by leveraging spatial, embodied, and social cues. However, questions remain regarding the relative effectiveness of different modalities of multimedia hints, graphical, typographic, or vocal, particularly in visually complex problems involving figures or graphs. While educational materials often blend these modalities, the optimal approach for maximizing effectiveness is still unclear (Girwidz and Kohnle, 2021).

Cognitive psychologists often evaluate instructional designs based on their ability to reduce cognitive load or minimize resource demands (Kirschner, 2002; Paas et al., 2004, 2010). Mayer's CTML offers insights into the cognitive loads imposed by different modalities (2017). CTML posits two information-processing channels: auditory and visual. Vocal hints follow the auditory pathway, while graphical hints engage the visual pathway. Typographic hints, though processed visually, require conversion to a phonological format due to their linguistic nature. Given the limited cognitive resources of each channel, CTML suggests that leveraging both channels simultaneously can reduce cognitive overload and improve learners' ability to process external information. This underlies the auditory superiority effect, which suggests that combining spoken text with visuals should be more effective than combining written text with visuals. For instance, pairing graphical and vocal hints may optimize dual-channel use, whereas pairing graphical and typographic hints could overload the visual channel. The right side of Figure 1 illustrates this dualchannel system, offering a theoretical foundation for understanding how multimedia facilitates learning across different modalities.

Problem-solving is inherently complex, requiring the simultaneous execution of multiple cognitive processes. To understand the cognitive resources demanded at different stages of problemsolving, we incorporate Wickens' Multiple Resources Theory (Wickens, 2002). The theory offers a four-dimensional model to explain cognitive load and predict task performance, focusing on processing codes (analog vs. categorical), visual channels (focal vs. ambient), perceptual modalities (auditory vs. visual), and stages (perception, cognition, and response). While processing codes and visual channels are not relevant to this study, the dimensions of perceptual modalities and stages align well with CTML's dual-channel framework. The stages dimension, in particular, clarifies the cognitive resources needed when hints are perceived, integrated with prior knowledge to alter unproductive mental representations, and ultimately used to generate solutions. According to Multiple Resources Theory, the cognitive resources required for perceiving and processing hints are distinct from those needed for generating answers. This separation suggests that processing hints to trigger representational changes should not interfere with answering questions. On the other hand, more recent studies showed that perception and cognition draw on shared processing resources (Cichy et al., 2014; Fast and McGann, 2017; Phillips, 2019), which is why they are represented as an integrated block in our theoretical framework (see the center of Figure 1).

3 Significance of study

Integrating Cognitive Theory of Multimedia Learning (Mayer, 2017) and Multiple Resources Theory (Wickens, 2002) provides a robust framework for understanding problem-solving with multimedia hints. However, several critical gaps remain in educational research regarding how multimedia materials can enhance

problem-solving performance. While many studies have examined the perception of printed text (Van Orden and Goldinger, 1994; Tzeng and Singer, 1981), comprehension of digital and printed text (Tanner, 2014; Ross et al., 2017), supportive interactions between typographic and vocal information (Sohoglu et al., 2014), and the redundancy effects-both positive and negative-between graphical, typographic, and vocal information (Trypke et al., 2023), there has been little direct comparison of these modalities in the context of problem-solving. For example, Klingner et al. (2011) found that visually presented numbers imposed less cognitive load than verbally presented numbers during numerical tasks. However, no study has systematically compared graphical, typographic, and vocal presentations of the same information. Addressing this gap, our study investigates the effects of these modalities in conceptual physics problem-solving, offering new insights into how multimedia hints can optimize cognitive resources and enhance performance in complex tasks. Guided by our theoretical framework, we formulated two hypotheses:

Hypothesis 1: Participants will perform better on near and far transfer problems than on the pretest problems, and the extent of this improvement will vary depending on the type of hint received during training. This suggests that working through training problems supports learning, and that different hint modalities may influence how effectively that learning transfers to new contexts.

Hypothesis 2: Participants' performance on training problems will be influenced by the modality and combination of multimedia hints they receive. Based on the *auditory superiority effect* embedded in the Cognitive Theory of Multimedia Learning, we expect vocal hints to be more effective than typographic hints. Furthermore, we anticipate interactions between modalities, such that combining vocal and graphical hints will be more effective than combining typographic and graphical hints. Finally, we expect that exploring these effects across different problem sets will help identify which hint modalities or combinations most consistently support conceptual physics problem-solving.

4 Method

4.1 Participants

Participants (N = 162) were recruited from conceptual physics courses at a midwestern university in the United States and received course credit for their participation. The majority were sophomores and juniors, with over 80% being future elementary teachers. Fewer than 10% had taken a physics course in high school, and none had prior experience with college-level physics. All participants gave informed consent in accordance with IRB-approved procedures, and each received \$10 compensation for their time.

4.2 Materials

Each participant solved four sets of conceptual problems in the interview. These problems were adapted from those used in our previous study (Rouinfar et al., 2014). We selected these problem sets

because they are well-suited for investigating students' conceptual understanding in physics and have been shown to elicit distinct patterns of visual attention. We named each of these after the main object in the problem -- "Ball," "Graph," "Roller Coaster (RC)" and "Skier" (see Figure 2). Each set had one initial problem, six training problems, one near transfer problem and one far transfer problem. Each of the training problems differed from the initial one only in terms of surface features. They had the same physics concept and the same representation, only a minor change in the details of the situation (see Figure 3). The problems were presented to participants with multimedia hints discussed in detail below. The near transfer problem was designed based on the same physics concept and representation but in a different context. The far transfer problem was again based on the same physics concept and representation, but the context was substantially different from the context of the training and near transfer problems. The topics relevant to the problems were kinematics and energy conservation, which had been covered in lectures prior to the recruitment of participants. As a part of our experiment design, we randomized conditions, the sequence of sets, and the sequence of training problems within each set. A complete list of all problems used in the study is provided as Supplementary material.

4.3 Experiment procedure

Each participant in this study completed an individual interview session lasting about 45 min on average. A short oral explanation of the interview was given to each participant before the interview started. The explanation included the goal of this study, the procedure of interview, a request for informed consent, and information regarding extra credit the participant would receive for their participation in the study.

We used a full factorial design: 2 (graphical hint / no graphical hint) \times 2 (typographic hint/no typographic hint) \times 2 (vocal hint/no vocal hint) with eight conditions in total. Participants were randomly assigned to one condition: no hint (N = 20), graphical hint (N = 20), typographic hint (N = 22), vocal hint (N = 21), graphical + typographic hint (N = 18), graphical + vocal hint (N = 19), typographic + vocal hint (N = 20), and graphical + typographic + vocal hint (N = 22). All problems were presented on a computer screen. Participants were instructed to read problems carefully, view hints when they were available, and then verbally provide their answers and reasons to the interviewer when they were ready. In all the seven hint conditions, participants were asked to wait at least 10 s after the problem appeared on the screen to prevent participants from rushing through problems and hints without carefully reading them. Participants were instructed that they could view hints as many times as they wanted. Participants asked follow-up questions only for clarification purposes. The interviewer took notes on participants' answers and reasons during the interview. The entire interview session was audio and video recorded.

4.4 Multimedia hint design

Except for the condition with no hint, participants received hints with different modalities when they solved the training problems.



Participants were not provided with any hints on the initial, near transfer, or far transfer problems, in any of the conditions.

We adopted graphical hints from our previous study and the more detailed explanation of these hints can be found there (Rouinfar et al., 2014). The graphical hint for each training problem was eight-seconds long, highlighting the area of the diagram that was related to the correct answer conceptually. The highlighting patterns of Figure 2 are examples of a graphical hint. The design of our typographic hints and vocal hints was done to ensure that they conveyed the same amount of information as the corresponding graphical hint. An example of a typographic hint for another task set can be seen in Figure 2. We invited the instructor (one of the co-authors) of the course that we recruited participants from, to record vocal hints since participants

should be familiar with his voice. The length of vocal hints for all the problems was between 7 and 8 s long. This was the same duration as the graphical hints and the typographic hints.

5 Experimental data and analysis

5.1 Scoring procedure

The correctness of participants' responses was determined after all interviews were finished. Four raters completed the rating. Each of them was assigned to one set to maximize consistency. To be coded as correct, a participant's response needed to have both the



correct answer and correct reason. Table 1 provides the rubric used to determine whether participants' explanations met the criteria for a correct reason in each problem set. Each rater graded 10 participants' interview notes with the help of videotapes. Afterward they discussed their ratings with the first author to have an agreement on grading rubric for each set. Then they graded all participants' responses for one set separately. They marked the ambiguity responses for the first author to review with the videotapes. The inter-rater reliability for the four task sets was above 95%.

On some occasions, participants who were assigned conditions with hints accidentally gave the answers and reasons before they were presented with the hints, or the interviewer did not remind the participant to access the hints. All these responses were excluded from our data analysis, resulting in less than 3% of missing data for each set. Each interview contained four sets, and each set had one problem as a pretest, six isomorphic problems as a training process and then two problems as a transfer test. Pretest performance was calculated as the average correctness rate across the four pretest problems. Training performance was measured by averaging the correctness rates of the 24 training problems. Near transfer performance was defined as the average correctness rate across four near transfer problems, and far transfer performance was similarly calculated using the four far transfer performance was similarly calculated the assumptions of ANOVA (normality and homogeneity of variance), Welch's ANOVA was used to assess differences among cross types. Statistical significance was set at p < 0.05. When multiple comparisons were conducted simultaneously, the Bonferroni correction was applied. All statistical analyses were conducted in R version 4.4.2 (R Core Team, 2013), and

Problem topic	Criteria for a correct reason
Ball	Interprets the spacing between consecutive snapshots as an indicator of speed.
Graph	Relates the slope of the graph at a specific time to the instantaneous velocity.
Roller coaster	Compares the height change from the initial to the final position to infer speed.
Skier	Relates the height change of each track section to the resulting speed at the end of that section.

TABLE 1 Grading rubric for correct reasons across the four problem sets.

A response was considered correct only if it included both the correct answer and an explanation that matched the criteria below for each topic.

visualizations were produced using the ggplot2 package (Wickham, 2016).

5.2 Pretest and transfer performances

This section addresses Hypothesis 1, to evaluate this hypothesis, we first assessed baseline equivalence across conditions, then tested for overall performance improvement, and finally examined whether the extent of improvement was moderated by hint condition.

As a first step, we examined whether participants across all eight experimental conditions started with equivalent levels of performance. This baseline check ensures that any observed differences in later performance are not due to pre-existing differences among groups. A one-way analysis of variance (ANOVA) was conducted. The results revealed no statistically significant difference between the conditions [F(7, 154) = 1.00, p = 0.432]. The pretest performance for each of the eight conditions is detailed in Table 2.

With group equivalence established, we then tested whether participants' performance improved following the training session. Paired-sample *t*-tests comparing pretest and near transfer performances indicated a significant improvement in near transfer performance (M = 0.29, SD = 0.25) compared to pretest performance (M = 0.10, SD = 0.16), t(161) = 9.78, p < 0.01. Similarly, a paired-sample *t*-test comparing pretest and far transfer performances showed a significant increase in far transfer performance (M = 0.32, SD = 0.23) compared to pretest performance (M = 0.10, SD = 0.23) compared to pretest performance (M = 0.10, SD = 0.16), t(161) = 12.16, p < 0.01. The improvements observed in near and far transfer performances suggested that the training session positively influenced participants' understanding of the relevant physics topics.

Apart from the control condition, all seven other conditions provided hints before participants answered the training problems. Given that the statistical results demonstrated superior near and far transfer performances compared to the pretest, we sought to examine the effect of the hint condition on performance improvement. A two-way repeated measures ANOVA was conducted to assess the impact of hint condition on pretest/near transfer performance improvement, revealing no statistically significant interaction [*F*(7, 292) = 1.33, *p* = 0.237]. Similarly, another two-way repeated measures ANOVA was conducted to assess the effect of hint condition on pretest/far transfer performance improvement, with no statistically significant interaction found [*F*(7, 292) = 0.50, *p* = 0.835]. Therefore, there was no statistical evidence indicating that the performance improvement was influenced by hint conditions.

These findings support the first part of Hypothesis 1: participants showed significant improvement from pretest to both near and far transfer problems, suggesting that working through training problems with hints enhanced their conceptual understanding. However, the data did not support the second part of the hypothesis, as the type of hint received during training did not significantly affect the degree of improvement.

5.3 The effects of hint modalities on training performance

To evaluate Hypothesis 2, we first analyzed the overall effects and interactions among graphical, typographic, and vocal hints, then explored whether these effects were consistent across different problem sets. This two-step approach allowed us to test both the predicted auditory superiority effect and broader patterns of modality interaction.

We conducted a three-way (graphical hint/no graphical hint × typographic hint/no typographic hint × vocal hint/no vocal hint) ANOVA for examining the effects of different hint modalities and their combinations. No statistically significant three-way interaction was observed [F(1, 154) = 0.07, p = 0.799]. However, a significant two-way interaction surfaced between typographic and vocal hints [F(1, 154) = 4.97, p = 0.027]. Simultaneously, there was a significant simple main effect of graphical hints [F(1, 154) = 40.09,p < 0.001] and a significant simple main effect of typographic hints [F(1, 154) = 4.71, p = 0.032]. Post-hoc Tukey's multiple comparisons were employed to explore the interaction between typographic and vocal hints. The analysis revealed that typographic hints significantly improved training performance in the absence of vocal hints [M (typographic without vocal) – M (no typographic no vocal) = 0.15, p = 0.012]. However, typographic hints did not show a significant improvement when presented with vocal hints [M (typographic with vocal) – M (no typographic with vocal) = -0.001, p = 1.000]. Figure 4 presents the interactions between hint modalities.

To further investigate the effects of hint modality on problem sets, we conducted a two-way ANOVA comparing training performance across four problem sets [Ball, Graph, Roller Coaster (RC), and Skier] and eight hint conditions. A significant interaction was found between problem set and condition [F(21, 160) = 10.01, p < 0.001], along with significant main effects of the problem set [F(3, 160) = 247.22, p < 0.001] and condition [F(7, 160) = 73.33, p < 0.001]. To further explore these interactions, we ran separate one-way ANOVAs for each problem set, followed by Tukey's Honest Significant Difference (HSD) post-hoc comparisons.

The detailed results are presented in Figure 5. In this figure, scenarios (i.e., combinations of conditions and problem sets) were denoted by "a," "b," "c," and so forth, representing the comparison outcomes and indicating the statistical significance of differences between two scenarios. Specifically, a scenario labeled with "a" signified its superiority within the problem set. It was statistically superior to conditions lacking the "a" label, such as "b," "bc," "c," etc. Furthermore, it exhibited marginal superiority compared to conditions labeled with "a" in combination with other labels, including "ab," "abc," "abcd," and so on.

Among the four problem sets, "Ball" is the only one where the no-hint scenario performed significantly worse than all hint scenarios.

Condition	Participant num.	Pretest	Training	Near transfer	Far transfer
No Hint	20	0.10 ± 0.15	0.12 ± 0.14	0.19 ± 0.20	0.25 ± 0.16
Graphical	20	0.11 ± 0.15	0.42 ± 0.20	0.33 ± 0.25	0.31 ± 0.28
Typographic	22	0.10 ± 0.17	0.33 ± 0.24	0.33 ± 0.26	0.40 ± 0.25
Vocal	21	0.11 ± 0.19	0.26 ± 0.21	0.26 ± 0.27	0.36 ± 0.23
Graph. + Type.	18	0.04 ± 0.10	0.52 ± 0.23	0.31 ± 0.22	0.33 ± 0.17
Graph. + Vocal	19	0.16 ± 0.19	0.51 ± 0.24	0.34 ± 0.25	0.32 ± 0.25
Type. + Vocal	20	0.14 ± 0.17	0.30 ± 0.23	0.28 ± 0.28	0.26 ± 0.25
Graph. + Type. + Vocal	22	0.07 ± 0.14	0.46 ± 0.26	0.31 ± 0.27	0.34 ± 0.24

TABLE 2 The pretest, training, near and far transfer performances (1 = 100%) with standard deviations of eight conditions.

This suggests that the hint design for the "Ball" problems was more effective than for the other problem sets. Scenarios with graphical hints were consistently labeled as "a," while no scenarios without graphical hints were labeled as "a," indicating that graphical hints were more effective than typographic or vocal hints. However, combining graphical hints with typographic and/or vocal hints did not further enhance their effectiveness. The scenario with only typographic hints was labeled as "b," and the one with only vocal hints was labeled as "c," showing that typographic hints were more effective than vocal hints. The scenario with typographic and vocal hints was labeled as "bc," suggesting that combining typographic and vocal hints resulted in an intermediate effect.

For the "Graph" problem set, the scenario with only vocal hints and the one with no hints were labeled as "c," indicating that vocal hints alone did not improve performance. Scenarios with graphical and/or typographic hints were labeled as "a" or "ab," showing that graphical and typographic hints together was not significantly better than presenting graphical or typographic hints alone. The scenario with graphical and typographic hints was labeled as "a," while the scenario with typographic and vocal hints was labeled as "b." This suggests that vocal hints were less effective than graphical hints for this problem set.

In the "RC" problem set, the no-hint scenario was labeled as "d." Scenarios with typographic and/or graphical hints also received labels containing "d," indicating that neither graphical nor typographic hints improved performance for this problem set. However, all scenarios with vocal hints included the label "a," showing that vocal hints were more effective than graphical or typographic hints. Combining vocal hints with other modalities did not result in a significant improvement over vocal hints alone.

For the "Skier" problem set, the no-hint scenario and the scenario with typographic and vocal hints were labeled as "e" and "de," respectively, showing that combining these two hints did not improve performance compared to the no-hint scenario. Scenarios with single-modality hints (graphical, typographic, or vocal) were labeled as "d" or "cd," indicating that these modalities were equally effective on their own. The scenario with graphical and typographic hints was labeled as "a," while the scenario with graphical and typographic hints was labeled as "b," and the scenario with all three hints was labeled as "bc." This suggests that graphical and typographic hints together were the most effective combination for this problem set, but adding vocal hints reduced their effectiveness. Table 3 summarizes the statistical analyses conducted in this study, highlighting their purposes and the key findings that support our interpretations.

In summary, these results indicate that graphical hints were the most consistently effective modality across problem sets. In contrast, vocal hints were generally less effective and combining vocal and typographic hints often led to reduced performance, suggesting a redundancy effect. These findings run counter to the *auditory superiority effect*. While participants' training performance was clearly shaped by the modality of hints they received, the overall pattern challenges the assumption that auditory presentation necessarily supports more efficient learning when paired with visual content. We examine the implications of these findings in the following discussion section.

6 Discussion and applications

6.1 Are graphical hints the best?

This section and the one that follows reflect on the findings through the lens of our two guiding hypotheses: examining both whether participants improved after training and which types of hints most effectively supported conceptual problem-solving. Here, we focus on the advantages of graphical hints over typographic and vocal alternatives across most problem sets.

The superior performance associated with graphical hints aligns with prior research emphasizing the cognitive efficiency of visual representations. Larkin and Simon (1987) argued that graphical information is more computationally efficient than linguistic information, making it easier to process and manipulate. Nesbit and Adesope (2006) demonstrated that graphical information enhances memory and recall, while Harris (2021) found that visual representations reduce cognitive load and facilitate shared understanding among healthcare professionals. Moreover, Dansereau and Simpson (2009) found that graphical information aligns with the brain's natural ability to recognize patterns, relationships, and overall structure. Supported by gestalt principles, graphical representations are processed as cohesive wholes, unlike linguistic information, which requires analytical decomposition and imposes higher cognitive demands. Graphical representations, therefore, allow for faster and more efficient comprehension.

Our findings also refine current theoretical models. While Mayer's (2017) Cognitive Theory of Multimedia Learning and Wickens (2002) Multiple Resources Theory emphasize a dual-channel system, visual and auditory, they do not account for differences in processing efficiency within a single channel. Our results highlight that complete, self-contained graphical information may be more easily processed

Analysis type	Analysis factor	Purpose	Key findings
One-way ANOVA	Pretest performance across eight conditions	To assess baseline differences among conditions	No significant difference in pretest performance across conditions
Paired-sample <i>t</i> -tests (×2)	Pretest vs. Near Transfer Pretest vs. Far Transfer	To evaluate performance improvement after training	Significant improvement from pretest to both near and far transfer performance
Two-way repeated measures ANOVA (×2)	Conditions × (Pretest vs. Near Transfer) Conditions × (Pretest vs. Far Transfer)	To examine interaction between condition and transfer performance improvement	No significant interactions for either near or far transfer
Three-way ANOVA	Graphical × Typographic × Vocal hints	To investigate effects of hint modality combinations on training performance	No three-way interaction; significant two-way interaction between typographic and vocal hints; significant main effects for graphical and typographic hints
Two-way ANOVA	Problem Set × Condition	To analyze training performance differences across problem sets and hint conditions	Significant interaction between problem set and condition; significant main effects for both problem sets and conditions

TABLE 3 Summary of statistical analyses conducted in the study.

than visually presented text. This distinction explains why combining multiple hint modalities did not generally enhance performance: when graphical hints were sufficient, adding more information offered little additional benefit. The "Skier" problem set was the exception, where combining graphical and typographic hints improved performance, possibly due to the added clarity needed to support understanding in a more complex context.

6.2 Typographic vs. vocal hints

This section continues the discussion of Hypothesis 2, particularly the component grounded in the *auditory superiority effect*. Prior studies have found that auditory information serves as a more effective companion to graphical content than written text (Rias and Zaman, 2010; Dousay, 2016), an effect attributed to the lower cognitive load imposed when information is distributed across auditory and visual channels. A more recent study by Haavisto et al. (2023) found that cognitive load did not predict learning outcomes, and instead proposed that ecological factors, such as learners' familiarity with media formats, control over pacing, and expectations for interaction, may offer a more accurate explanation. This evolving perspective suggests that while the *auditory superiority effect* may still be observed, its underlying mechanisms are likely more contextual and learnerdependent than previously assumed.

However, our results did not support these predictions. Typographic hints outperformed vocal hints in the "Ball" and "Graph" problem sets, performed similarly in the "Skier" set, and were only outperformed by vocal hints in the "RC" set. In fact, the "RC" set also showed the highest number of hint scenarios that were no better than the no-hint scenario, suggesting a potential issue with the hint design for this specific problem set. Overall, typographic hints were more effective than vocal hints across most problem sets.

While Haavisto et al.'s (2023) work offers valuable insight, it does not account for our findings. In our study, all hints were presented through a uniform interface: graphical and typographic hints were displayed on screen for the same duration, and vocal hints were delivered through speakers. No modality offered greater user control or familiarity. Instead, our findings align more closely with those of Reinwein and Tassé (2022), who also found that written text outperformed spoken text in a sentence-picture comparison task. Like us, they questioned the generalizability of the *auditory superiority effect* and highlighted the importance of task complexity. However, their results showed a written-over-oral advantage only in low-complexity tasks, with no modality effect observed under higher complexity. This contrasts with our findings, where typographic hints were more effective in the context of conceptual physics problems. Solving those problems posed a substantial challenge to participants, as reflected in their low pretest performance (see Table 2). Our theoretical framework offers an explanation rooted in the idea that perception and cognition share common processing resources, as represented by the integrated perception-cognition block in the center of Figure 1. In this view, problem-solving is not only cognitively demanding but also perceptually intensive, requiring careful allocation of limited mental resources.

Under such high-load conditions, the efficiency of perceptual input becomes critical. Unlike vocal hints, which are temporally transient and processed sequentially, typographic hints persist visually on screen for longer time, providing learners continuous access to the full message. This visual persistence may have eased working memory demands by enabling participants to re-read and re-process information as needed—offering stronger support for complex reasoning. The observed advantage of typographic hints, therefore, may stem not from their modality alone, but from how their delivery structure aligns with the joint cognitive and perceptual demands of conceptual problem-solving.

6.3 Suggestions for instructional multimedia design

Graphical hints should be prioritized when designing multimedia instructional materials to support students in solving physics problems. These hints can direct attention to relevant parts of the problem or illustrate key concept structures or patterns. When the hint design and content are effective, graphical hints are better than linguistic hints in any modality.

However, designing effective graphical hints can be challenging, as they cannot explicitly tell students what to do. In such cases, typographic hints can serve as complementary support to make graphical hints more explicit. For example, presenting graphical and typographic hints together was significantly better for the "Skier"





represents typographic hints only, "V" signifies vocal hints only, "CT" involves graphical and typographic hints, "GV" combines graphical and vocal hints, "TV" includes typographic and vocal hints, and "GTV" encompasses graphical, typographic, and vocal hints. Results are presented for each problem set.

problem set than any single-modality hint scenario. In contrast, combining typographic and vocal hints is not recommended, as their interaction can reduce the effectiveness of typographic hints due to

redundancy. This finding aligns with multimedia design principles based on Mayer's CTML, which caution against redundant information in instructional design.

7 Limitations and future work

One limitation of this study is the lack of observed performance differences on near-transfer or far-transfer problems across all eight conditions. These findings partially contradict those of our previous work (Rouinfar et al., 2014), which demonstrated that visual hints alone significantly improved near-transfer performance compared to no hints. We suspect that this discrepancy may be due to differences in the physics backgrounds of participants. In Rouinfar et al.'s (2014) study, participants were enrolled in an algebra-based physics course for life science majors, with most having completed high school physics. In contrast, participants in this study were recruited from a conceptual physics course, where few had prior high school physics experience, and none had completed a college-level physics course. This lack of problem-solving experience may have hindered participants' ability to recognize the connections between training problems and transfer problems, given their differences in representation and content. Future studies should consider replicating this work with participants who have comparable physics backgrounds to those in Rouinfar et al.'s (2014) study.

Another important finding of our study is that typographic hints were more effective than vocal hints for helping students solve physics problems, likely due to the visual persistence of typographic information. To better understand this effect, future research could investigate typographic hints presented sequentially, word by word, to mimic the temporal nature of vocal hints. Such a study would provide deeper insights into the differential effects of these modalities on problem-solving performance.

Finally, we identified potential issues with the hint design for the "Roller Coaster" problem set in this study, as many scenarios did not perform better than the no-hint condition. Revising the hint design for this problem set and conducting follow-up studies could provide additional evidence to clarify the comparative effectiveness of graphical, typographic, and vocal hints.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Institutional Review Board at Kansas State University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

References

Albus, P., Vogt, A., and Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Comput. Educ.* 166:104154. doi: 10.1016/j.compedu.2021.104154

Burkholder, E., Blackmon, L., and Wieman, C. (2020). Characterizing the mathematical problem-solving strategies of transitioning novice physics students. *Phys. Rev. Phys. Educ. Res.* 16:020134. doi: 10.1103/PhysRevPhysEducRes.16.020134

Author contributions

XW: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. YL: Data curation, Formal analysis, Software, Writing – review & editing. TZ: Data curation, Writing – review & editing. JH: Writing – review & editing. LL: Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing. NR: Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work is supported in part by the U.S. National Science Foundation under Grant No. 1348857. Opinions expressed are of the authors and not of the Foundation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1568406/ full#supplementary-material

Cichy, R. M., Pantazis, D., and Oliva, A. (2014). Resolving human object recognition in space and time. *Nat. Neurosci.* 17, 455–462. doi: 10.1038/nn.3635

Dansereau, D. F., and Simpson, D. D. (2009). A picture is worth a thousand words: The case for graphic representations. *Prof. Psychol. Res. Pr.* 40, 104–110. doi: 10.1037/a0011827

Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., Hsu, L., Jackson, K. A., et al. (2016). Assessing student written problem solutions: A problem-solving rubric with

application to introductory physics. *Phys. Rev. Phys. Educ. Res.* 12:010130. doi: 10.1103/PhysRevPhysEducRes.12.010130

Dousay, T. A. (2016). Effects of redundancy and modality on the situational interest of adult learners in multimedia learning. *Educ. Technol. Res. Dev.* 64, 1251–1271. doi: 10.1007/s11423-016-9456-3

Dow, G. T., and Mayer, R. E. (2004). Teaching students to solve insight problems: Evidence for domain specificity in creativity training. *Creat. Res. J.* 16, 389–398. doi: 10.1080/10400410409534550

Fast, C. D., and McGann, J. P. (2017). Amygdalar gating of early sensory processing through interactions with locus coeruleus. *J. Neurosci.* 37, 3085–3101. doi: 10.1523/JNEUROSCI.2797-16.2017

Girwidz, R., and Kohnle, A. (2021). "Multimedia and digital media in physics instruction" in *Physics education*. challenges in physics education. eds. H. E. Fischer and R. Girwidz (Cham: Springer).

Haavisto, M., Jaakkola, T., and Lepola, J. (2023). Video outperforms illustrated text: Do old explanations for the modality effect apply in a learner-paced fifth-grade classroom context? *Comput. Educ.* 199:104775. doi: 10.1016/j.compedu.2023.104775

Harris, L., (2021). Designing a visual grammar to enable more effective stakeholder participation in scoping organizational change: a physics of notations approach. PhD diss., Memorial University of Newfoundland.

Jonassen, D. H. (2010). Learning to solve problems: a handbook for designing problem-solving learning environments. New York: Routledge.

Kirschner, P. A. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learn. Instr.* 12, 1–10. doi: 10.1016/S0959-4752(01)00014-7

Klingner, J., Tversky, B., and Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology* 48, 323–332. doi: 10.1111/j.1469-8986.2010.01069.x

Larkin, J. H., and Simon, H. A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cogn. Sci.* 11, 65–100. doi: 10.1111/j.1551-6708.1987.tb00863.x

Lestari Syafril, S., Latifah, S., Engkizar, E., Damri, D., Asril, Z., and Yaumas, N. E. (2021). Hybrid learning on problem-solving abiities in physics learning: A literature review. *J. Phys. Conf. Ser.* 1796:012021. doi: 10.1088/1742-6596/1796/1/012021

Madsen, A. M., Larson, A. M., Loschky, L. C., and Rebello, N. S. (2012). Differences in visual attention between those who correctly and incorrectly answer physics problems. *Phys. Rev. ST Phys. Educ. Res.* 8:010122. doi: 10.1103/PhysRevSTPER.8.010122

Mayer, R. E. (2017). Using multimedia for e-learning. J. Comput. Assist. Learn. 33, 403-423. doi: 10.1111/jcal.12197

Nesbit, J. C., and Adesope, O. O. (2006). Learning with concept and knowledge maps: a meta-analysis. *Rev. Educ. Res.* 76, 413–448. doi: 10.3102/00346543076003413

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. *Advances Psychol. Thinking*. (eds.) M. T. Keane and K. J. Gilhooly (Hemel Hempstead, Hertfordshire, UK: Harvester Wheatsheaf) 1, 1–44.

Paas, F., Renkl, A., and Sweller, J. (2004). Cognitive load theory: instructional implications of the interaction between information structures and cognitive architecture. *Instr. Sci.* 32, 1–8. doi: 10.1023/B:TRUC.0000021806.17516.d0

Paas, F., van Gog, T., and Sweller, J. (2010). Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. *Educ. Psychol. Rev.* 22, 115–121. doi: 10.1007/s10648-010-9133-8

Phillips, B. (2019). The shifting border between perception and cognition. *Nous* 53, 316–346. doi: 10.1111/nous.12218

R Core Team, (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/

Reinwein, J., and Tassé, S. (2022). Modality Effects Examined by Means of an Online Sentence-Picture Comparison Task. *J. Psycholinguist. Res.* 51, 521–542. doi: 10.1007/s10936-022-09849-9

Rias, R.M., and Zaman, H.B., (2010). Investigating the redundancy effect in multimedia learning on a computer science domain. In Proceedings 2010 International Symposium on Information Technology - Visual Informatics, ITSim'10 1, 1-4

Ross, B., Pechenkina, E., Aeschliman, C., and Chase, A.-M. (2017). Print versus digital texts: understanding the experimental research and challenging the dichotomies. *Res. Learn. Technol.* 25. doi: 10.25304/rlt.v25.1976

Rouinfar, A., Agra, E., Murray, J., Larson, A. M., Loschky, L. C., and Rebello, N. S. (2014). "Can visual cues and correctness feedback influence students reasoning?" in 2013 Physics Education Research Conference Proceedings. Presented at the 2013 Physics Education Research Conference (Portland, OR: American Association of Physics Teachers), 305–308.

Schnotz, W., and Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educ. Psychol. Rev.* 19, 469–508. doi: 10.1007/s10648-007-9053-4

Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2014). Top-down influences of written text on perceived clarity of degraded speech. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 186–199. doi: 10.1037/a0033206

Tanner, M. J. (2014). Digital vs. Print: Reading Comprehension and the Future of the Book. *Student School Inform. Student Res. J.* 4:12. doi: 10.31979/2575-2499.040206

Thomas, L. E., and Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychon. Bull. Rev.* 14, 663–668. doi: 10.3758/BF03196818

Trypke, M., Stebner, F., and Wirth, J. (2023). Two types of redundancy in multimedia learning: a literature review. *Front. Psychol.* 14:1148035. doi: 10.3389/fpsyg.2023.1148035

Tzeng, O. J. L., and Singer, H. (1981). Perception of print: reading research in experimental psychology. *1st* Edn. London, UK: Routledge. doi: 10.4324/9781315454375

Van Orden, G. C., and Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 1269–1291. doi: 10.1037/0096-1523.20.6.1269

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theor. Issues Ergon. Sci.* 3, 159–177. doi: 10.1080/14639220210123806

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis (2nd Edn. 2016). AG, Switzerland: Springer International Publishing. doi: 10.1007/978-3-319-24277-4