Can short duration visual cues influence students' reasoning and eye movements in physics problems?

Adrian Madsen,[1] Amy Rouinfar,[1] Adam M. Larson,[2] Lester C. Loschky,[3] and N. Sanjay Rebello[1]

[1]*Department of Physics, Kansas State University, Manhattan, Kansas 66506, USA*
[2]*Department of Psychology, University of Findlay, Findlay, Ohio 45840, USA*
[3]*Department of Psychology, Kansas State University, Manhattan, Kansas 66506, USA*

We investigate the effects of visual cueing on students' eye movements and reasoning on introductory physics problems with diagrams. Participants in our study were randomly assigned to either the cued or noncued conditions, which differed by whether the participants saw conceptual physics problems overlaid with dynamic visual cues. Students in the cued condition were shown an initial problem, and if they answered that incorrectly, they were shown a series of problems each with *selection* and *integration* cues overlaid on the problem diagrams. Students in the noncued condition were also provided a series of problems, but without any visual cues. We found that significantly more participants in the cued condition answered the problems overlaid with visual cues correctly on one of the four problem sets used and a subsequent uncued problem (the transfer problem) on a different problem set. We also found that those in the cued condition spent significantly less time looking at ''novicelike'' areas of the diagram in the transfer problem on three of the four problem sets and significantly more time looking at the ''expertlike'' areas of the diagram in the transfer problem on one problem set. Thus, the use of visual cues to influence reasoning and visual attention in physics problems is promising.

I. INTRODUCTION

When learners attempt to solve physics problems, in our case those containing text and an image, various parts of the visual information compete for the learners' attention. This is because our brains can only process some of the information received by our retinas at a given time [1] and we generally only become aware of that part of the retinal information that has been attended to and entered into working memory [2–4]. There are a large number of visual learning or assessment environments in physics, which contain information that is both relevant and irrelevant to the task at hand, competing for the learner's attention. To facilitate learning, one can help the learner to focus their attention, and thus their cognitive resources, on relevant information in instructional problems and to avoid focusing on irrelevant information. Mayer's cognitive theory of multimedia learning [5] explains that learning occurs when relevant information is successfully selected and organized into a coherent representation and integrated into the existing knowledge base. All of these processes occur in the learner's working memory, which is a limited cognitive resource that can be used up by focusing on irrelevant information. To help alleviate this problem, visual cues can help the learner attend to and notice relevant information in the problem, which

they may have previously ignored. Having attended to relevant information in the problem, the learner may be more likely to retrieve relevant information from long-term memory, which will then be available for problem solving processes occurring within the constraints of their working memory limits. Providing visual cues by no means guarantees that the learner will reach a correct solution and understanding of the problem. However, such cues could at least facilitate correct problem solving and learning by helping the learner avoid "red herrings" based on common alternate conceptions about physics.

Visual cues have been studied in a large range of contexts and have been found to be useful in many of these contexts. Grant and Spivey [6] studied how shifts in visual attention to critical diagram features can facilitate correct problem solving. The researchers studied Duncker's radiation problem[1] [7] and manipulated the diagram so that either the relevant or irrelevant area of the diagram pulsed or the diagram remained static. They found that the group of participants who viewed the relevant area pulsing (expanding by six pixels repeatedly) spent more time

---

[1]Duncker's radiation problem (adapted by Grant and Spivey [6]) states, "Given a human being with an inoperable stomach tumor, and lasers which destroy organic tissue at sufficient intensity, how can one cure the person with these lasers and, at the same time, avoid harming the healthy tissue that surrounds the tumor?" The solution is to fire several low intensity lasers at the tumor from different locations so that, although each laser is not powerful enough to damage the tissue surrounding the tumor, the combined power of several laser beams is sufficient to destroy the tumor.

looking at the relevant area and were significantly more likely to produce a correct solution than those who saw the irrelevant area pulsing or a static diagram. They argued that drawing visual attention to the critical area of a diagram can influence cognitive processing in ways that lead to correct problem solving. Thomas and Lleras [8] conducted a follow-up study on the work of Grant and Spivey [6] to determine the existence and nature of an implicit connection between eye movements and cognition. To do this, they overlaid visual cues on Duncker's radiation problem diagram for 4 sec at the end of a 26 sec free viewing period. This was repeated 20 times or until the participant answered correctly. These visual cues moved in four different patterns. One group of participants saw a pattern of random characters that mimicked the solution to the problem.[2] Participants in this ''embodied solution'' group were significantly more likely to solve the problem correctly. They concluded that manipulating eye movements can serve as an implicit guide to influence thinking on spatial reasoning tasks.

Thomas and Lleras conducted another study to determine if the greater rate of correct problem solving by those in the embodied solution group was a result of shifts in attention or actual eye movements [9]. The ''eye-movement'' group saw random digits appear in a pattern that embodied the solution and followed these digits with their eyes. The ''attention-shift'' group saw the same string of random digits as the eye-movement group but was instructed to follow the digits with their attention but to keep their eyes fixated at the center of the screen. The eye-movement and attention-shift groups were more likely than other groups to answer the problem correctly. However, no significant difference was found between the eye-movement and attention-shift groups. The results of this study suggest that the primary mechanism behind the increased correct solution rates in the study by Thomas and Lleras [8] for the embodied solution group is the shift of attention that immediately preceded the directed eye movements. Thus, directed shifts of attention have been found to influence cognitive processing on spatial insight problems and increase rates of correct solutions. Based on successful problem solving presented in this previous work, we apply visual cueing to static physics problems to explore whether such cues serve as an implicit guide to improve problem solving performance.

A. Theoretical background

There are two relevant theoretical frameworks which help us interpret the functions and mechanisms of visual cueing. The first, representational change theory [10], is

related to the cognitive mechanisms involved in solving problems that require insight. The second, the cognitive theory of multimedia learning [5], pertains to the use of multimodal information in learning, and more specifically for our current purposes, cueing.

1. Representational change theory

Representational change theory [10] provides a framework to understand the cognitive mechanisms involved in solving problems—particularly, problems that require insight, as opposed to merely algorithmic problems. This theory is relevant to our work on visual cues, as the problems we are interested in require conceptual insight and are not merely algorithmic in nature.

Representational change theory explains that the way a problem is represented in a solver's mind mediates the knowledge that the solver retrieves from long-term memory. The retrieval process is based on spreading activation among concepts or pieces of knowledge in long-term memory. An impasse or block occurs when the way a problem is represented does not permit retrieval of necessary operators or possible actions. Breaking the impasse requires changing the problem representation. A new mental representation acts as a retrieval cue for relevant operators in long-term memory, extending the information available to the problem solver. Insight is achieved when the impasse is broken and the retrieved knowledge operators are sufficient to solve the problem.

According to representational change theory, there are three mechanisms by which an impasse to solving a problem is broken: (i) adding information to the problem to enrich and extend the existing representation (i.e., elaboration), (ii) replacing the existing representation with a different more productive representation (i.e., reencoding), or (iii) removing unnecessary constraints often self-imposed by the problem solver (i.e., constraint relaxation). Once the impasse is broken, the new mental representation of the problem can activate relevant concepts in long-term memory, extending the information available to the problem solver. When the relevant concept or pieces of knowledge are available to the solver, she can apply the concept to answer the question correctly.

2. Cognitive theory of multimedia learning

Many physics problems involve the use of information presented with text and diagrams. Thus, learning to solve physics problems involves coordinating information provided in multiple modalities with the learner's prior knowledge. The cognitive theory of multimedia learning [5] identifies three distinct processes—*selection*, *organization*, and *integration*—involved in learning from information presented in multiple modalities. Based on Mayer's cognitive theory of multimedia learning, de Koning *et al.* [11] proposed a framework for attention cueing which suggests that visual cues, if designed appropriately, can

---

[2]Participants in this embodied solution group saw random characters that moved from outside the skin to inside the tumor at several locations around the diagram. This embodied the solution of using several weak lasers incident at different angles.

facilitate all three processes involved in multimedia learning.

Mayer describes selection as the process of attending to certain pieces of sensory information from each modality. de Koning extends selection to visual cueing and explains that selection cues help the learner attend to relevant information in a visual representation. Organization is using the selected information in each modality to create a coherent internal representation in that modality. Organization involves structuring information to facilitate comparison, classification, enumeration, generalization, and cause-effect relationships. Organizational cues emphasize structure and facilitate the identification and the subsequent representation of the material's organization. Integration is combining internal representations from different modalities with activated prior knowledge. Two kinds of integration processes are important for multimedia learning and problem solving: integrating elements (i) within a single representation that are widely spatially separated [12] or (ii) across multiple representations or modalities such as coordinating graphs and pictures with text to create an operational situational mental model [13] to solve the problem. Cueing learners to relate elements within a single representation is especially important if the elements they need to integrate are widely spatially separated [12]. Cueing can also be useful when the problem is complex and could have more than one method for solution and schema construction, imposing a high cognitive load on the learner. Cues that make implicit causal or functional relations between elements more explicit can potentially improve problem solving ability. Integration cues can be particularly helpful for learners when they must integrate textual and graphical information to create a situation model in order to solve a math or science problem [13].

The distinctions among cue types proposed by de Koning *et al.* [11] are useful in guiding one's thinking about cues. However, in practice it seems that those cues that would be useful in solving a problem often involve combinations of the three types. More specifically, virtually any organization or integration cue will simultaneously serve as a selection cue, since both must attract attention to a problem element in order to be effective. Thus, in our current study, we utilized visual cues that served to both select relevant information and integrate related elements in a problem diagram. However, the current study did not investigate organization cues.

B. Research questions

In the current study, we aim to answer the following research questions.

(1) Do short duration, dynamic visual integration and selection cues improve students' problem solving performance on introductory conceptual physics problems?

(2) How do dynamic visual cues influence participants' eye movements on current cued problems and subsequent uncued problems?

(3) Does students' problem solving performance on a subsequent uncued problem improve after seeing visual cues on similar problems?

II. METHOD

A. Participants

We conducted individual sessions with 63 participants concurrently enrolled in either a first or second semester introductory physics course. These students were enrolled in either a traditional course, which included lectures, a recitation, and a lab, or a reformed course, which included studios [14] and lectures. There were separate instructors for the traditional first and second semester physics courses as well as for the reformed first semester course. All courses were at the same large, Midwestern public university. Students were invited via an Email sent to all students enrolled in the course and were paid $10.00 for participation. We collected data over the duration of two semesters, but ensured that students had covered relevant topics in their physics course before recruiting them to volunteer in our study. We invited students from the same courses to participate both semesters. We also ensured that each student participated only once.
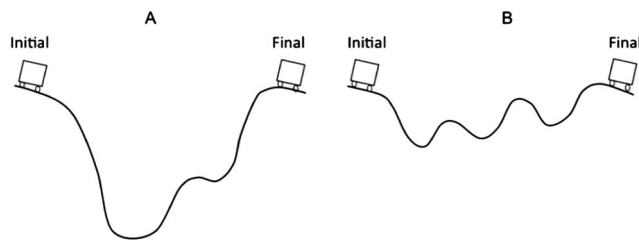
B. Materials

1. Design of study problems

The materials consisted of four sets of related conceptual introductory physics problems in which an accompanying diagram was necessary to answer the problem. Each of these problems had been studied previously in the physics education research literature, and found to have diagram features that students consistently used to produce incorrect answers, which we refer to as "novicelike" areas of the diagram. The problem diagrams also contained areas which one needs to attend to in order to answer the question correctly, called "expertlike" areas. These expert- and novicelike areas were spatially separate and the problem required a solver to make some comparison between the areas to come to the correct answer. In the next section we point out the novicelike and expertlike areas of the diagrams for each of the four problem sets used. These were determined through interviews with novice physics students and confirmed with previous studies from physics education research literature that used the same problems. A more complete description of the features of these problems, details of these interviews, and previous studies of these problems are reported in Madsen *et al.* [15]. Additionally, we found significant differences in the time spent fixating in these diagram areas based on the correctness of participants' answers. Those who answered incorrectly spent more

time fixating in the novicelike areas while those who answered correctly spent more time fixating in the expertlike areas [15]. Three of the six problems discussed in our previous study were on kinematics graphs. We choose to

Assume frictional effects can be ignored. How does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?

Two balls roll along the paths shown below. The position of the balls is shown at equal time intervals of one second each. When does Ball B have the same speed as Ball A?
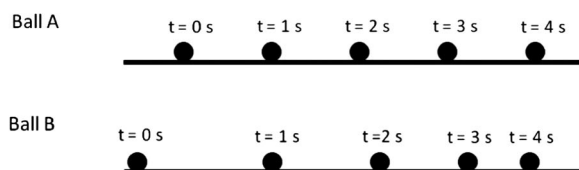
Rank the changes in potential energy during the skier's descent down each slope from greatest to least.
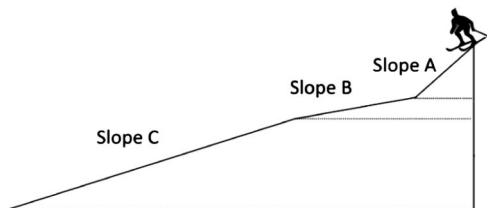
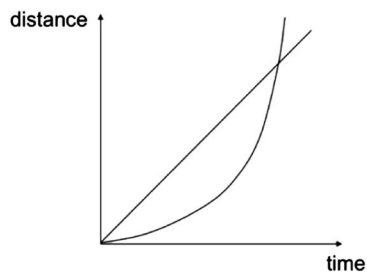The motion of two objects is represented in the graph below. When are the two objects moving with the same speed?

FIG. 1. Four ''initial'' problems taken from Madsen *et al.* (Ref. [15]) and used in the current study. Shown from top to bottom are the roller coaster, ball, Skier, and graph problems.

use one of these kinematic graph questions as well as the three other questions for a total of four problems in the current study. We refer to these problems as the roller coaster, ball, skier, and graph problems (see Fig. 1).

2. Sequence of problems

Each problem set consisted of an ''initial'' problem, four ''similar'' problems, and a ''transfer'' problem. All problems were open- ended and contained a diagram which one had to use in order to answer the problem. The similar problems in each set had the same problem statement as the initial problem and the same surface features. The novicelike area of the diagram for each was manipulated in a way that would change the answer one would give if answering based on a novicelike conception. For example, in the similar problem set shown in Fig. 2, the number and depth of bumps on the roller coaster track and horizontal distance between carts were varied. If a student used the features of the track to determine the speed of the carts, their answer would be different for each of these problems. The transfer problem in each problem set had different surface features but tested the same concept as the initial and similar problems. For example, the roller coaster transfer problem shown in Fig. 3 contained two tracks with different start and end heights. The student needed to reason about the potential and kinetic energy of tracks with the same difference in height.
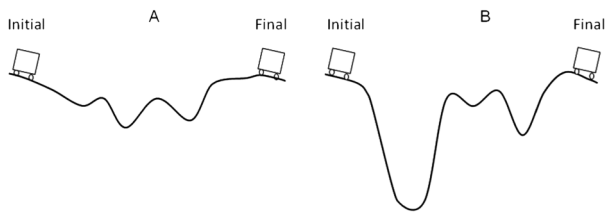
3. Design of and rationale for the visual cues

Participants in the ''cued'' condition were shown dynamic visual cues overlaid on the similar problem diagrams. The cues used in this study were designed to combine integration and selection cues [11] and were also designed to mimic the eye movements of those who answered the same problems correctly in our previous study [15]. There was a large variation in eye movements from one individual to another while viewing the diagrams in these physics problems, so the visual cues did not mimic the eye movements of correct solvers exactly. Instead, video playback of the correct solvers' eye movements was viewed repeatedly and special attention was paid to the eye movements in and around the relevant area of interest. Similarities between participants were observed, and visual cues were modeled after these patterns.
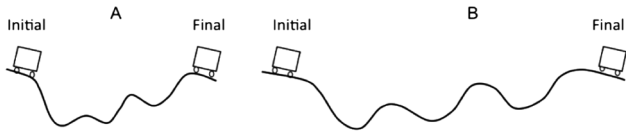
Cues on all four problem sets were intended to prompt selection and integration of expertlike elements in the problem diagrams, and below we describe how the cues could achieve this on each problem set. Importantly, our characterization of the cues as prompting selection and integration is from the target ''expert'' perspective (since the student perspective often corresponds to the ''novice'' perspective). On the ''roller coaster'' problem, the expertlike area of the diagram (as determined in our previous

Assume frictional effects can be ignored. How does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?

Assume frictional effects can be ignored. How does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?

Assume frictional effects can be ignored. How does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?
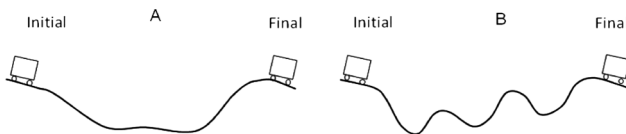

Assume frictional effects can be ignored. How does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?

FIG. 2. Example of four ''similar'' problems used in the roller coaster problem set in the study.

Assume frictional effects can be ignored. How does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?
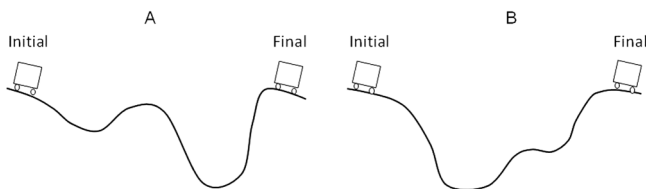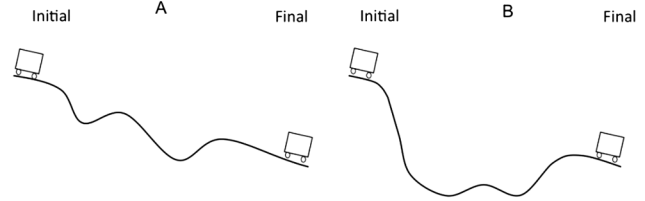

FIG. 3. Example of a transfer problem used in the roller coaster problem set in the study.

students who answered incorrectly cited the shape of the tracks as the reason the final speeds of the carts would differ. The cues in this study helped participants to select the heights of the roller coaster carts and not attend to the shape of the roller coaster tracks, which is the most commonly used feature when giving an incorrect answer. They also aimed to help participants integrate the heights of the roller coaster carts, which were spatially separated, so they could compare the heights of the initial and final carts.

The expertlike areas for the ''ball'' problem were the distances between the balls on track A and track B when both balls have moved the same distance in the same 1 sec interval (see Fig. 10 in the Appendix). So, for this problem the cues aimed to help the students compare the distances between balls during the same time period (integration), for example, by comparing the distance between the balls on track A and track B between 1 and 2 sec. The novicelike area for this problem was the point when the balls are at the same position at the same time. Here, students who answer incorrectly often explain that if the balls on tracks A and B have the same position at the same moment in time, they are moving at the same speed. So, the cues were also intended to help the students select distances between successive balls and not compare the positions of the balls at the same time.

The expertlike areas for the ''skier'' problem were the changes in heights of each slope, as these heights are

Assume frictional effects can be ignored. How does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?
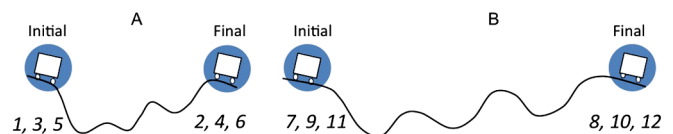
FIG. 4 (color online). Roller coaster similar problem used in the study with visual cues overlaid. The blue dots are the visual cues and the numbers in italics show the sequence of animated cues (the numbers were not seen by study participants).

study [15]) was the relative heights of the roller coaster carts. The relationship between heights of the roller coaster carts is needed to determine the potential energy of each at the beginning and end of the path and then relate this to the amount of kinetic energy and finally the speed [16]. So, in this problem the cues moved between the roller coaster carts to help students compare heights of the roller coaster carts (Fig. 4). The novicelike areas of the diagram were the roller coaster tracks, as

directly related to the change in potential energy for each slope (see Fig. 11 in the Appendix). So, the cues were designed to help the student to compare the changes in heights of each slope (integration). The novicelike areas of the diagram are the slopes, as students who answer incorrectly compare the steepness of the slopes and relate these to the relative changes in potential energy. So, the cues were also designed to help the students select the changes in height and ignore the steepness of each slope.

The expertlike area for the "graph" problem was the point when the slope of the two lines was the same, as this is when the two objects were moving with the same speed (see Fig. 12 in the Appendix). So, the cues aimed to help the students judge the slope of the curved line at several points by tracing out an invisible tangent line at a given time and comparing this to the slope of the straight line at the same time (integration). The novicelike area(s) for this problem was (were) the point(s) where the two lines crossed. At the crossing point(s), the two objects had the same position, but not the same speed. So, the cues were intended to help the solvers to attend to the slopes of the lines and not the points where the line(s) crossed (selection).

C. Apparatus

Participants were presented with physics problems on a computer screen viewed at a distance of 24 in. using a chin and forehead rest to minimize participants' extraneous head movements. The resolution of the computer screen was set to 1024×768 pixels with a refresh rate of 85 Hz. Each physics problem subtended $33.3° \times 25.5°$ of visual angle. Eye movements were recorded with an EyeLink 1000 desktop mounted eye-tracking system (Ref. [17]), which had an accuracy of less than $0.50°$ of visual angle. An eye movement was classified as a *saccade* (i.e., in motion) if the eye's acceleration exceeded $8500°/s^2$ and the velocity exceeded $30°/s$. Otherwise, the eye was considered to be in a *fixation* (i.e., stationary at a specific spatial location). A nine-point calibration and validation procedure was used at the beginning of the experiment. Participants' verbal explanations and gestures were recorded with a Flip video camcorder.

D. Study design and procedure

To ensure that the participants had sufficient prerequisite knowledge of the concepts tested in the study problems, each participant completed a pretest, which consisted of four open-ended questions designed to gauge their understanding of speed and potential energy. Each participant took part in an individual session lasting between 30 and 60 min. In each session they were first given an explanation of what to expect and the eye tracker was calibrated. Next, participants were instructed to spend as much time as needed on each question and answer with
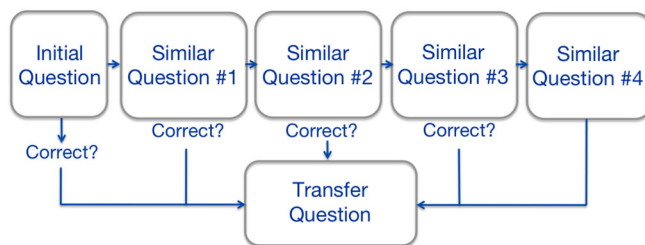

FIG. 5 (color online). Flow chart showing how the initial problem, similar problems, and transfer problems were administered to students in each of four problem sets.

a verbal explanation of their reasoning when ready. Participants in the cued condition were told that colored shapes may appear on some of the problems and when these appeared they should follow them with their eyes. No further information about the purpose of the cues was given to participants.

Each participant was randomly assigned to the cued condition or the noncued condition. Equal numbers of participants were assigned to each condition. The research design is shown in Fig. 5. First, students answered the initial problem to demonstrate their current level of understanding. If they answered incorrectly, they saw a series of similar problems, which contained the same problem statement as the initial problem, tested the same concept, and contained a diagram with similar surface features. When the student answered a similar problem correctly, they were shown the transfer problem. This process continued until a maximum of four similar problems had been viewed by the participant, after which the participant was presented the transfer problem regardless of whether they answered the similar problem correctly or incorrectly. All participants viewed the four sets of problems in the same order.

Whenever a student was ready to provide an answer and explanation for a problem, they indicated this by pressing any key on the keyboard, at which point the problem displayed on the computer would become slightly smaller in size to indicate to the student that they had successfully pressed a key. The participants then explained their answer and reasoning to the experimenter and were able to point to areas on the computer screen if necessary. The experimenter used a predefined rubric to determine if the given answer and explanation were correct or incorrect. If the answer and/or reasoning were vague, the experimenter would ask for clarification. Once the experimenter had sufficient information to determine the correctness of the answer, the experiment would proceed.

Participants in the cued condition saw moving colored shapes overlaid on the similar problems. Moving colored shapes were used because color and motion have been found to be the most predictive of attentional selection because of their high perceptual salience [18]. The cues

used for the roller coaster problem are shown in Fig. 2 and those for the ball, skier, and graph problems are shown in Figs. 10–12 in the Appendix. Each colored shape appeared 4 sec after the problem was presented to give the participant time to read the problem statement (although the problem statement for each similar problem was the same). The cues then appeared for 500 msec at 12 positions in the diagram for a total cueing time of 6 sec. This 6 sec time period was chosen based on the successful cueing work of Thomas and Lleras [8], in which visual cues were shown for 4 sec. After the cues ended, participants could spend as much time as they wanted on the problem.

III. CONNECTION BETWEEN THEORETICAL BACKGROUND AND CURRENT STUDY

We apply representational change theory to understand how visual cues can help learners solve physics problems. Ohlsson [10] conceptualizes insight as "initial failure followed by eventual success." He explains that insight occurs when the problem solver is competent to solve the problem before them, reaches an impasse in the problem solving process, and then successfully breaks this impasse. Representational change theory is valid for problem solving processes in which this impasse-insight sequence occurs. We claim that this sequence is likely to occur in our study since we chose problems that, by their very nature, lent themselves to impasse and insight. We used introductory conceptual physics problems requiring students to activate specific conceptual resources [19]. Since these questions are not given in any particular context, such as the end of a chapter or during lecture, the students must first recognize the appropriate concept. If they cannot recognize an appropriate concept, they may reach an impasse, which could be resolved when they see the visual cues and focus on relevant information. Further, since these questions are conceptual, once a student recognizes the appropriate concept to use, the solution does not require going through a long series of mental steps or calculations before getting to an answer. Instead, the student applies the appropriate conceptual resource and can quickly recognize an answer.

Importantly, the problem diagrams also contain visual information consistent with an incorrect novicelike answer. In our previous study, we found that students who answered these problems incorrectly attended to this novicelike visual information and activated conceptual resources which led to the wrong answer [15]. So to answer a problem correctly, the student must not only realize the appropriate concept, but must also suppress the use of these novicelike concepts that lead to incorrect answers. Students may also reach an impasse when they repeatedly see very similar problems overlaid with visual cues but recognize that they are not answering these problems correctly. During this process, the visual cues draw students' attention to areas they had previously ignored. The combination of being presented with similar problems several times with their attention being redirected to an area they previously found irrelevant could cause students to second-guess their previous answer. As they reconsider the diagram areas highlighted by visual cues, they may resolve their impasse with an insight, activate appropriate conceptual resources, and answer the problem correctly.

In order to resolve an impasse, we hypothesize that visual cues can serve to help the student rerepresent a problem in their mind. In line with representational change theory, the purpose of visual cues is to help the student replace an existing unproductive representation with a productive one, or add to their existing representation until it is adequate to solve the problem. In the current study we explore visual cues that we believe help rerepresentation occur through elaboration and reencoding, but not constraint removal.

Elaboration is useful for a learner who has gathered insufficient information to form a productive mental representation of the problem, and has thus reached an impasse. Integration cues can facilitate the addition of critical new information to the representation by helping the learner attend to information in a particular order and/or help the learner make comparisons between different elements of the diagram. A learner attending to the information provided by these cues is prompted to activate previously dormant relevant information from long-term memory and eventually encode a new representation for the problems. The problems that we used contained expertlike elements in the diagram that were spatially separate and needed to be compared and integrated in order to answer correctly. Integration cues can add information to the learner's current mental representation by helping them make these necessary connections. To determine the most useful way to design the integration cues for these problems, we used the eye movements of correct solvers on the same problems from our previous study [15], looked for patterns in the way correct solvers viewed the expertlike elements, and modeled our visual cues on these patterns.

Reencoding, unlike elaboration, involves not just adding new information, but instead backtracking through previous layers of the problem solving process, eliminating unproductive layers in their mental representation of the problem and creating new productive layers. The reencoding process is especially important for the problems used in our study, as the diagrams for these problems each contain an area consistent with the most common incorrect answer. This feature of the problems makes it more likely that the students will activate unproductive naïve concepts when reasoning to an answer. In order to help them reencode the problem representation in a scientifically accurate way, selection cues could be used. Rather than provide new

information, these cues prompt the learner to ignore irrelevant information and attend to relevant information for solving the problem. If the learner attends to the previously ignored relevant information, this in turn may activate previously dormant relevant prior knowledge from long-term memory, and the learner may eventually encode a new and more correct representation for the problem. In a similar study of expert chess players solving problems with two possible solutions, researchers found that when the players had found the first solution they reported looking for a better one, though the eye-movement record indicated they continued to look at features of the problem related to the solution they had already found [20]. Although they tried to seek out the better solution, their attention was fixated on their first idea. Participants in our previous study who answered these problems incorrectly spent more time looking at areas consistent with a novice than those who answered correctly [15]. If they are similar to the chess players, they may have tried to consider other solutions, but kept their attention fixated on areas consistent with their first idea. Selection cues can improve problem solving by helping solvers ignore the novicelike areas of the diagram, by instead attending to the expertlike areas, and use information in those areas to create a new mental representation of the problem.

IV. ANALYSIS AND RESULTS

Participants were only included in our analysis if they correctly answered pretest questions demonstrating knowledge of the concepts tested in the study problems. The pretests were scored as correct or incorrect by one of the researchers. When a participant's answer was unclear, two researchers discussed the answer and agreed on a conclusion. There were cases where a participant did not demonstrate adequate understanding of one of the concepts tested, so their data for that concept were not included in this analysis. Further, we only included participants with usable eye-movement data files. There were four participants whose eye-tracking data files became corrupted and could not be used.

A. Improvements to problem solving performance with visual cues

We first investigated the problem solving performance of participants by comparing how often those in the cued and noncued conditions who had answered the initial problem incorrectly answered one of the similar problems correctly. It is necessary to look only at the subgroup of students who answered the initial problem incorrectly, because we wished to test the effect of the cues only on those who answered the initial problem incorrectly. We compared the aggregate number of participants in the cued and noncued conditions who gave an incorrect answer on the initial problem and then gave a correct answer and explanation on one of the four similar
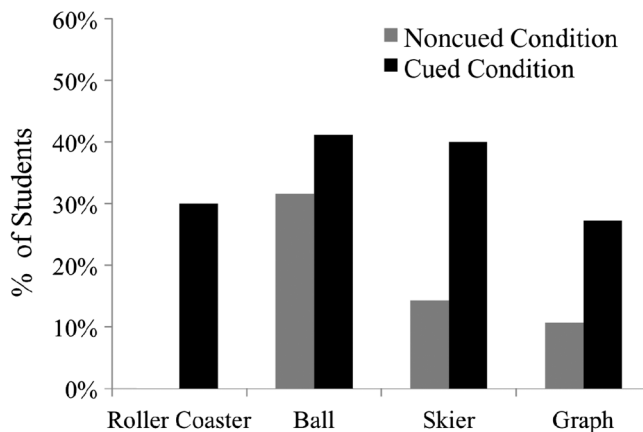

FIG. 6. Comparison of percentage of participants in cued and noncued conditions who gave the correct answer and reasoning on a similar problem. For example, 30% = 6/20 students in the cued condition answering correctly on a similar roller coaster problem.

problems. The percentages of students in each condition who answered a similar problem correctly are displayed in Fig. 6. Fisher's exact test [21] was employed to test the significance of the difference between the cued and noncued conditions in the proportion of students who correctly answered a similar problem. Fisher's exact test is used with categorical data, which is encountered when participants are classified in two different ways, and small sample sizes. In our case, the two different ways of classification are as follows: (1) whether a participant belongs to the cued or noncued condition or (2) whether a participant did or did not correctly answer a similar problem (after answering the initial problem incorrectly). Fisher's exact test examines whether students in one condition are more likely to change to answer a similar problem correctly than students in the other condition on the same problem set. Results of Fisher's exact test are shown in Table I. The total number of students included for each problem set is different and does not result in 63 total participants because we included only those who answered the initial problem incorrectly, had satisfactorily answered the pretest questions, and had usable eye-movement data files.

We found that a statistically greater number of participants in the cued condition answered a roller coaster similar problem correctly ($p = 0.012$). This means that a mere 6 sec of visual cueing for which the participants did not know the purpose resulted in significantly more students going from the wrong understanding of the roller coaster problem to the scientifically correct understanding on a very similar problem. It is promising to find a difference using such a short intervention. We did not find significant differences on the ball, skier, or graph problems. Inferences on why this was the case will be reviewed in Sec. VI.

TABLE I. Summary of results of Fisher's exact test comparing those who did and those who did not answer a similar problem correctly for the cued and noncued conditions. An asterisk indicates a significant difference at the $\alpha = 0.05$ level

Problem set	Condition	Answered similar problem correctly	Did not answer similar problem correctly	p
Roller coaster	Cued	6	14	0.012*
	Noncued	0	19	
	Six participants were not included because of insufficient pretest scores, four participants' eye-movement files were unusable, and 14 answered the initial problem correctly			
Ball	Cued	7	10	0.228
	Noncued	6	13	
	One participant's eye-movement file was unusable, and 26 answered the initial problem correctly			
Skier	Cued	4	6	0.142
	Noncued	2	12	
	Six participants were not included because of insufficient pretest scores, one participant's eye-movement file was unusable, and 32 answered the initial problem correctly			
Graph	Cued	6	16	0.098
	Noncued	3	25	
	Three participants' eye-movement files were unusable, two participants did not complete this problem, and eight answered the initial problem correctly			

B. Changes in eye movements on similar problems

We next investigated how the visual cues influenced participants' eye movements while viewing the similar problems. Prior to the experiment, participants in the cued group were told that they might see colored shapes appear on the screen and when they saw the shapes they should follow them with their eyes. Participants were not informed when they would see the shapes. The eye-movement record revealed that there were individual differences in how closely participants actually followed the moving colored shapes with their eyes. We investigated the possibility that participants who did not follow the shapes closely did not benefit as much from the visual cue as those who watched each segment of the cue. We employed a scan path analysis that looks at both spatial and temporal aspects of eye movements. A scan path is the collection of fixations and saccades one makes over time. We specifically used ScanMatch [22], which is an algorithm that compares two scan paths at a time and computes a number which represents their similarity in space and time. It is based on the Needleman-Wunsch algorithm used to compare DNA sequences. The ScanMatch algorithm overlays a labeled grid onto the image of interest and recodes the ordered locations and fixation durations into a sequence of letters. Longer fixations result in repeated letters in the sequence (Fig. 7). The letter sequences of two sets of eye movements are then compared to each other to calculate a similarity score. Letters near each other in the grid receive a higher score than those with greater spatial separation. The similarity score is normalized so that a score near one represents two sequences of eye movements that are identical spatially and temporally.

We isolated participants' eye movements while the cues were being shown for the first similar problem (since all cued participants saw this problem). We completed this analysis for participants in the cued condition and compared their eye movements to the scan path of the visual cues using the ScanMatch algorithm. We then compared the ScanMatch scores of those who had changed to a correct answer on a similar problem to those who had not. This comparison revealed how the performance of participants in the cued group on the similar problems
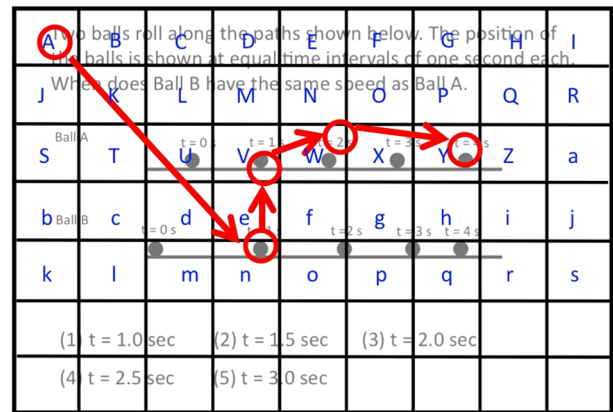
Normal Sequence: A n V W Y
With Temporal binning : A A A n n V V V W Y Y

FIG. 7 (color online). Example of ScanMatch algorithm converting a scan path into the letter sequence which was used to calculate the similarity score. Red circles represent fixations, red arrows represent saccades.

TABLE II. ScanMatch scores for cued participants who did and participants who did not answer a similar problem correctly. An asterisk indicates a significant difference at the $\alpha = 0.05$ level

Problem	ScanMatch score ($\pm$ standard error)	
	Changed to correct answer on similar problem	Did not change to correct answer on similar problem
Roller coaster*	0.588 ± 0.031 ($n = 6$)	0.379 ± 0.042 ($n = 14$)
Ball	0.552 ± 0.046 ($n = 7$)	0.557 ± 0.044 ($n = 10$)
Skier	0.700 ± 0.037 ($n = 4$)	0.588 ± 0.058 ($n = 6$)
Graph	0.652 ± 0.034 ($n = 6$)	0.595 ± 0.028 ($n = 16$)

varied based on how well they followed the cues with their eyes. To do the comparison, we employed the Kruskal-Wallis one-way analysis of variance in SPSS. This test is the nonparametric method to compare two or more independent groups and is the equivalent of the one-way ANOVA. This test was appropriate for our analysis since we had small group sizes that did not form a normal distribution. Average ScanMatch scores and standard errors are reported in Table II. We found a significant difference in ScanMatch scores between those who had answered a similar problem correctly and those who had

not for the roller coaster problem only [$H(3) = 9.939$, $p = 0.019$]. We did not find statistically significant differences on the ball, skier, or graph problems. This means that on the roller coaster problem the participants who answered a similar problem correctly followed the visual cues more closely on similar problem 1. This suggests that on this problem there is a connection between how well one follows the visual cues with their eyes and if they change from an incorrect to correct answer and verbal explanation of their reasoning. We do not suggest a causal mechanism, but will explore this finding more in Sec. V. For the other three problems, such a relationship between following the cues and performance on the problems was not found.

C. Problem solving performance on transfer problems

Showing that visual cues have the potential to help students give the correct answer and reason about a problem is an encouraging result, but we will only have evidence that some kind of learning has occurred if students can subsequently answer a related question with no cues. To investigate this possibility, we analyzed the correctness of those in the cued and noncued conditions on the transfer problem for each problem set. We once again used Fisher's exact test to test for a difference in the number of students who had answered the transfer problem correctly in the cued and noncued conditions. Results are shown in Table III. Figure 8 displays the percentage of students in the cued and noncued conditions who answered each transfer problem correctly. We found that a statistically greater number of participants in the cued condition answered the

TABLE III. Summary of results of Fisher's exact test comparing those who did and did not answer the transfer problem correctly for the cued and noncued conditions. An asterisk indicates a significant difference at the $\alpha = 0.05$ level

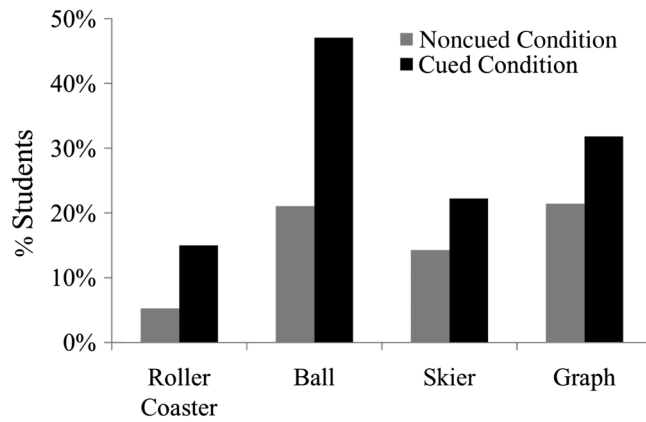| Problem set	Condition	No. of participants providing correct answer	No. of participants providing incorrect answer	p
Roller coaster	Cued	3	17	0.263
	Noncued	1	18	
	Six participants were not included because of insufficient pretest scores, four participants' eye-movement files were unusable, and 14 answered the initial problem correctly			
Ball	Cued	8	9	*0.039
	Noncued	3	16	
	One participant's eye-movement file was unusable and 26 answered the initial problem correctly			
Skier	Cued	2	7	0.370
	Noncued	2	12	
	Six participants were not included because of insufficient pretest scores, one participant's eye-movement file was unusable, and 32 answered the initial problem correctly			
Graph	Cued	7	15	0.181
	Noncued	6	22	
	Three participants' eye-movement files were unusable, two participants did not complete this problem and eight answered the initial problem correctly			

FIG. 8.   Comparison of percentage of participants in cued and noncued conditions who gave the correct answer and reasoning on the transfer problem for each problem set in the study. For example, 15% = 3/20 students in the cued condition answering correctly on the roller coaster transfer problem.
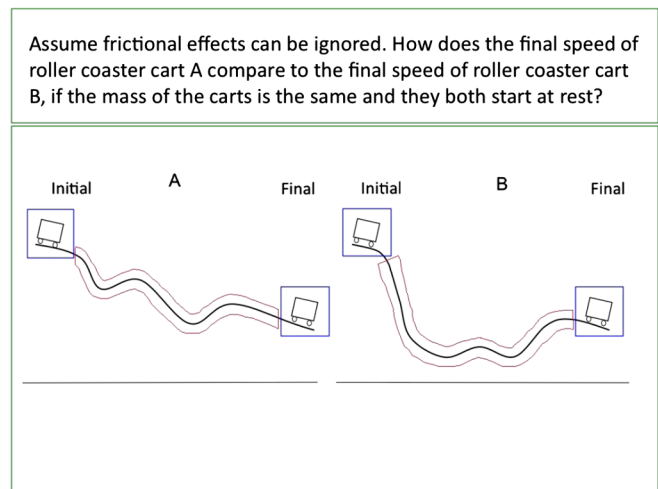

FIG. 9 (color online).   Expertlike and novicelike definitions of areas of interest (AOI) for the roller coaster transfer problem. The expertlike AOIs are around the roller coaster carts (shown in blue) while the novicelike AOIs are around the tracks (shown in red). The green boxes outline the problem statement AOI and the diagram AOI.

ball transfer problem correctly and gave the correct reasoning ($p = 0.039$). We also note that the raw percentage correct on the transfer problem was higher for those in the cued condition than the noncued condition for all four problem sets.

D. Changes in eye movements on transfer problems

The purpose of the visual cues was to redirect visual attention to relevant areas and help students integrate different important elements in a physics diagram. It may be that the brief visual cues did not help students answer the transfer problem correctly, although it may have influenced their visual attention. To test this idea, we completed an areas of interest (AOI) analysis on the eye movements on the transfer problem. To do this we defined two types of areas in each diagram, the expertlike and novicelike areas. The definitions for the novicelike areas of interest came from the individual think-aloud interviews and eye-tracking analysis reported in our previous study [15]. The expertlike areas were defined by expert raters and are also described in [15]. For example, for the roller coaster problem, we defined the expertlike AOI around the roller coaster carts, as expert raters determined that the relative heights of the carts are required to judge the final speeds of the carts on each track. We defined the novicelike AOI around the roller coaster tracks as we found through individual interviews and literature investigating a similar problem that those who answered incorrectly did so using features of the track. We mimicked the AOI definitions used in the previous study in the current analysis for each problem set. The AOIs for the roller coaster problem are pictured in Fig. 9. The eye tracker used in this study had an average error of 0.5 deg of visual angle, so the AOIs were defined to be 0.5 deg of visual angle from the edge of the desired region or element in the diagram.

After defining the novicelike and expertlike AOIs in the problem diagrams, we determined the amount of time each participant spent fixating in these areas and divided by the total time they spent fixating on the diagram to normalize for differences in viewing speeds. We then compared the percentage of time participants in the cued and noncued conditions spent in the novicelike and expertlike AOIs. If visual cues had positively influenced the eye movements of those in the cued condition, we would expect to see larger percentages of time in the expertlike AOIs than those in the noncued condition. Further, we would expect to see smaller percentages of time in the novicelike areas than those in the noncued condition. If cues had no influence on the eye movements of those in the cued condition, we would expect no differences in the percentage of time in either the expertlike or novicelike AOIs based on condition.

We compared the percentage of time spent in the novicelike and expertlike AOIs using a one-way ANOVA with *percentage of time in AOI* as the dependent variable and *cued or noncued condition* as the independent variable for each transfer problem. We remind the reader that we included only the eye movements of those participants who had answered the initial problem incorrectly and thus had been shown the similar problem(s). The results are displayed in Table IV. For the roller coaster problem, we found that participants in the cued condition spent a statistically higher percentage of fixation time in the expertlike AOI and a smaller percentage in the novicelike AOI. We also found for the ball and skier problems participants in the cued condition spent a smaller percentage of fixation time in the novicelike

TABLE IV. Mean percentage fixation time spent ($\pm$ standard error) on the transfer problems for expertlike and novicelike AOIs for participants in the cued and noncued conditions.

	Problem set	Cued	Noncued	ANOVA results	p
Expertlike AOI	Roller coaster*	18.5 ± 2.2 ($n = 21$)	9.7 ± 1.7 ($n = 19$)	$F(1, 38) = 9.573$	0.004
	Ball	28.4 ± 3.3 ($n = 17$)	21.1 ± 3.9 ($n = 19$)	$F(1, 34) = 2.022$	0.164
	Skier	0.5 ± 0.3 ($n = 9$)	1.0 ± 0.6 ($n = 14$)	$F(1, 21) = 0.451$	0.509
	Graph	6.3 ± 1.0 ($n = 21$)	6.7 ± 1.4 ($n = 28$)	$F(1, 48) = 0.039$	0.844
Novicelike AOI	Roller coaster*	18.0 ± 2.1 ($n = 21$)	29.5 ± 3.1 ($n = 19$)	$F(1, 38) = 9.835$	0.003
	Ball*	4.3 ± 1.4 ($n = 17$)	10.9 ± 2.4 ($n = 19$)	$F(1, 34) = 5.372$	0.027
	Skier*	18.2 ± 2.3 ($n = 9$)	49.0 ± 3.5 ($n = 14$)	$F(1, 21) = 42.105$	<0.001
	Graph	8.0 ± 1.2 ($n = 21$)	11.6 ± 1.4 ($n = 28$)	$F(1, 48) = 3.427$	0.070

AOIs. This indicates that the cues helped participants in the cued condition allocate more visual attention to the expertlike area (on the roller coaster problem) which contained information needed to answer correctly and allocate less visual attention to the novicelike areas (on the roller coaster, ball, and skier problems) which contained visual information consistent with scientifically incorrect conceptions. So, seeing the visual cues influenced how participants viewed the transfer problems but not how they answered them.

E. Participants' impressions of the visual cues

Since we did not tell the students the purpose of the visual cues, one might wonder what they thought of the colored shapes flashing on the screen. We present some examples of students' ideas about the cues below. All three excerpts came from transcriptions of audio-recorded conversations that occurred directly after the student had completed all study problems. Students' names have been changed below to numbers S1–S3 to protect their anonymity.

Interviewer: When you saw the colored shapes, what did you think of them?
S1: Why are these moving everywhere?
Interviewer: Did you have any idea what they were?
S1: No, I was just following them like I was told to.

Interviewer: Did you have any thoughts about what those dots were for?
S2: I mean to see if they had the same mass, it was like comparing the carts it seemed like. It would go back and forth, in the cart one at least.
S2: Did you think they were trying to get you to do anything?

S2: Oh, didn't think about it at the time, I guess it was trying to tell you they were at the same height. Didn't even think about that … or the blue dots the other time.

Interviewer: What did you think of those little dots that were moving?
S3: I actually felt like they … After I saw the dots then I thought more about where my eyes were looking when I was looking the graph and it seemed like my eyes afterward did … I felt like I was looking at the same places that I had just been looking with the colored stuff …

We see from these examples that there was great variability in students' ideas about the cues. The cues were interpreted differently by individual students, which may have contributed to the differences in the effects of the cues. We will revisit students' views of the cues in Sec. VI.

V. CONCLUSIONS

In this study we find some evidence that short duration, dynamic visual selection and integration cues can improve students' problem solving performance on introductory conceptual physics problems as participants were able to correctly answer and reason about problems they were previously unable to. Of the four problem sets used, we found significantly more students changed to a correct answer after seeing the visual cues on only the roller coaster problem. Through the lens of representational change theory, this suggests that the cues may have helped the students overcome an impasse and mentally rerepresent that problem so that productive concepts or pieces of knowledge could be retrieved from long-term memory and applied. We did not find this difference on the other three problem sets.

We also investigated how the dynamic visual cues influenced participants' eye movements while viewing the cues. We looked for relationships between how well each participant followed the visual cues with their eyes and if they changed from an incorrect answer on the initial problem to a correct answer and reasoning on a similar problem. To investigate this relationship we calculated similarity scores between their eye-movement scan path and the path of the visual cues using the ScanMatch algorithm and compared these similarity scores between those who had and those who had not correctly answered a similar problem. We found that for the roller coaster problem those who successfully answered and reasoned about a similar problem had statistically significantly higher similarity scores (ScanMatch scores). This means that the participants who correctly answered and reasoned about the roller coaster problem were following the cue more closely with their eyes than students who did not answer or reason about the problem correctly. This suggests a link between how well participants attended to the visual cues and how helpful the cue was at implicitly influencing their reasoning about the problem. We did not find such correspondence between how well eye movements followed the cues and problem solving improvements on the other three problems. This follows from the fact that the cues were not effective at helping students answer a similar problem correctly; thus, there is no reason to expect to find the above-mentioned correspondence between eye movements to the cues and improvements on those problems.

From an educational standpoint, it is not sufficient that visual cues help students answer problems containing those cues. Instead, the educational value of visual cueing would be more apparent if after seeing visual cues repeated on several similar problems, students could then successfully answer and reason about related but different problems without cues, which we have called transfer problems in the current study. We compared the correctness of answers and reasoning on the transfer problems associated with each of the four problem sets between those who had seen visual cues and those who had not. We found a significant difference in transfer problem correctness between conditions for the ball problem only, with a greater number of participants in the cued condition answering this problem correctly. Thus, we find some evidence that repeatedly showing novices visual cues on related problems may help them form a productive mental representation on similar future problems viewed without cues. Nevertheless, no such relationship was found for the other three problem sets.

We also investigated how seeing the dynamic visual cues on similar problems may have influenced participants' visual attention on the transfer problems. We compared the percentage of fixation time spent in novicelike and expertlike areas of interest between those in the cued and noncued conditions on the transfer problems associated with all four problem sets. We found that on the roller coaster problem participants in the cued condition spent a significantly greater percentage of time looking in the expertlike AOI and a significantly smaller percentage of time in the novicelike AOI than those in the noncued condition. We also found for the ball and skier problems those in the cued condition spent a significantly smaller percentage of fixation time looking in the novicelike AOI. This result suggests that seeing the cues on this problem had an influence on participants' visual attention on subsequent uncued problems and helped them to pay more attention to the expertlike elements (in one of four problem sets used) and less attention to the novicelike elements (in three of the four problem sets used). This result is promising, as we know from previous work that participants who answer a problem correctly spend more time looking at the expertlike areas of the problem and those who answer incorrectly spend more time looking at the novicelike areas. Helping participants look at relevant areas and ignore distracting areas when no visual cues are present could be a first step to helping them reason correctly about the problem.

This work adds to the building body of research in physics education on the importance of visual attention in physics problem solving [23–26]. As educators and researchers, we often overlook the way our students view visual representations in physics. This study provides some evidence that where a student looks in a visual representation can influence their reasoning about it, especially when the representation contains relevant and irrelevant elements. This research suggests that computer-assisted instruction may benefit from the use of visual cues to guide learners' attention to relevant areas of figures, especially for those in which novices often attend to the wrong information.

VI. LIMITATIONS AND FUTURE WORK

We find some evidence that visual cues overlaid on static problems can help a student answer similar problems correctly, a transfer problem correctly, and can even influence visual attention on the transfer problem so that participants spend more time looking at relevant areas and ignoring irrelevant areas. But we find this to be true for only some of the problem sets used in this study. This finding also raises the question, why did we not see the same positive results on the other problem sets? Possible answers to this question motivate our future work.

First, we speculate on why the cues on the roller coaster problem were effective at helping students answer the similar problems correctly, but not the cues on the ball, skier, or graph problems. Upon examining the cues, we noticed that the roller coaster cues were especially simple.

The simple back and forth motion highlighting the roller coaster carts was repeated several times. On the other hand, the visual cues used in the ball, skier, and graph problems moved in a more complex pattern. For example, in the ball problem, the cue moved between balls in track A at a given time period, then moved between balls on track B at the same time period, and were then repeated with a different set of balls. This pattern was shown only once for 6 sec and it is likely that the pattern was simply too complicated to draw significant meaning from it in such a short time. Further, we designed the cues to mimic the patterns that correct solvers used when viewing the problem diagrams. When helping someone who does not know how to solve a problem, showing them what an expert does may not be helpful, as an expert's problem solving process is likely streamlined and condensed and based on deep knowledge of why this section of the diagram is important. There are also many types of visual cues that could be applied to a given problem, and we tried only one type, namely, flashing colored shapes that helped the participants select and integrate important diagram elements. In future research, it would be better to use very simple cues that are easily encoded and understood by students. To ensure that the cues are ''student friendly'' one should first test various versions of visual cues on a given problem with introductory physics students in individual interviews. During such interviews, it would help to observe the problem solving process of an introductory student (as opposed to an expert) and offer different types of cues starting with the most implicit moving to the most explicit and illustrative. For example, on the roller coaster problem, one could start by highlighting the carts and dimming the tracks, and then we could try highlighting the carts in a temporal order (as we did in this study). If one found that this was unhelpful for a student, one could add even more information to the problem by overlaying lines under each cart representing the vertical height of each cart. Researchers could also vary how long the cues were shown. This process should be continued with different cues, cueing times, and students until we find the ideal cues.

Additionally, there may only be certain types of problems that lend themselves to improvement through visual cueing. In the successful cueing study of Thomas and Lleras, the problem they used did not require any specialized content knowledge. In our work, students must possess certain physics content knowledge to be able to provide a correct answer and explanation. It may be that implicit cueing is less effective with problems that require content knowledge. Further, we have explored only four problems in this study. A multitude of problems could potentially be tested in future studies. It could also be that the order in which the problems are presented influences the usefulness of the cue. The roller coaster problem was presented first each time and was the only problem the cues were found to influence. In future studies, the order of

cue problems should be randomized to balance out any order effects.

In this study, similar to the work of Thomas and Lleras [8], we did not tell students the purpose of the cues. We hoped that the cues would implicitly influence students to rerepresent a problem and overcome an impasse. While we found evidence that the cues were helpful on the roller coaster problem, they were not helpful on the other three problem tests. We asked a subset of participants what they thought of the visual cues and found great variability in students' ideas about the cues. It may be that students' impressions of the cues influenced how useful the cues were at helping the students. In a future similar study, one could carefully probe students' impressions of the cues at the end of the session and conduct a detailed analysis of the relationship between their views of the cues and their problem solving success. It may be even more beneficial to tell students that the cues are helpful and draw their attention to relevant diagram elements. We are currently conducting a study investigating the benefit of these types of visual cues when students know their purpose.

Further, in this study we found, on one of the four problems tested, a difference in the similarity scores (ScanMatch) between those who did answer a similar problem correctly and those who did not. This means that there was a difference in how well the participants followed the cues with their eyes and this difference was related to their success on this similar problem. We predict that participants will follow the cues more closely and purposefully if they know they are helpful (as opposed to just being random flashing shapes), and the cues are more likely to facilitate correct problem solving since the students are actually looking at them. Additionally, in future studies researchers could include a training session where students gain experience following practice cues with their eyes. This will familiarize students with the task before they see the actual cues, and we predict their ability to follow the cues closely will increase.

We also found differences between the cued and noncued groups on only one of the four transfer problems tested. It appears that the three transfer problems that showed no difference were too difficult for this level of student, as very few students in either group answered these problems correctly. It is also possible that, although we viewed the transfer problems as closely related to the similar problems, the students did not view them this way, and thus were unable to apply what they gained from the cues to the transfer problems. It also may be that the students did not gain anything from the cues that could be applied to solve the transfer problems, thus producing the lack of differences between the cued and noncued groups. Based on the current study, we cannot determine the precise reason for finding differences between the cued and noncued groups on only one of four transfer problems.

In future studies, it would be profitable to first test the transfer problems with students in individual or group interviews to gain insight into how the students view the transfer problems and the connections they see between the similar and transfer problems. Once one is confident that the level of the transfer problems is appropriate and that students clearly interpret the problems as being similar to the initial problems, stronger conclusions about the effectiveness of the cues could be drawn.

Further, the cued group showed superior performance on the roller coaster similar problem, but not on the roller coaster transfer, as one would expect. Instead, it was the ball transfer problem on which the cued group outperformed the noncued group, though there were no differences in performance between groups' similar problems for the ball problem set. There is no clear reason for why this occurred. Using the improvements suggested in this section, we will need to see if such a result is replicable, and if so, what it means.

Students' level of prior physics knowledge may also influence the effectiveness of the visual cues. In the current study, we used a simple pretest to ensure that the participants had the minimum level of conceptual knowledge to answer the questions. In future studies, researchers could control for prior knowledge more closely with a thorough pretest and students' current and previous semester physics grades to use as covariates in the analyses. Further, the time between the relevant material being covered in the student's physics course and the time of the visual cueing session may be important in determining how useful the cues are for the students. In the current study, students participated in the study throughout the course of a semester, though this was after they had covered the relevant material in their physics course. In the future, visual cueing sessions could be condensed to a given week in the semester so that all students have approximately the same time between exposure to the material in class and participation in the study.

In conclusion, there is much work to be done to understand the factors that lead to effective visual cues in physics problem solving. This study suggests that cueing can potentially serve as effective conceptual scaffolding for novice physics students, but much further work will be necessary to develop both a sound theory of cueing that is applicable to a wide array of physics problems and methods of cueing that are more reliably effective in facilitating learning of physics concepts.

APPENDIX

Problems used in study with visual cues overlaid (Figs. 10–12). The colored shapes in each problem are the visual cues and the numbers in italics show sequence of animated cues (the numbers were not seen by study participants). Each colored shape was seen by participants for 0.5 sec.

Two balls roll along the paths shown below. The position of the balls is shown at equal time intervals of one second each. When does Ball B have the same speed as Ball A?
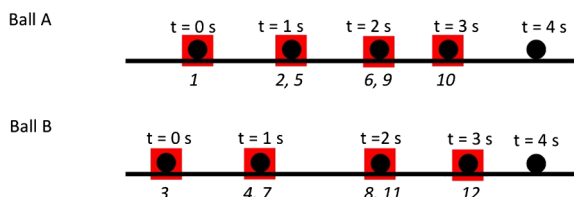

FIG. 10 (color online).   Ball problem overlaid with visual cues.

Rank the changes in potential energy during the skier's descent down each slope from greatest to least.
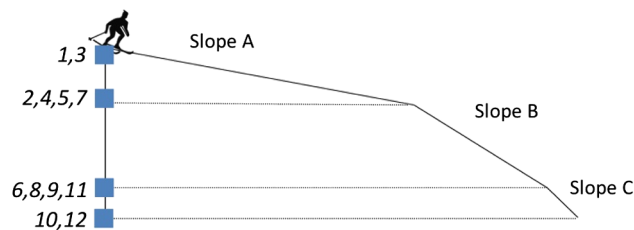

FIG. 11 (color online).   Skier problem overlaid with visual cues.

The motion of two objects is represented in the graph below. When are the two objects moving with the same speed?
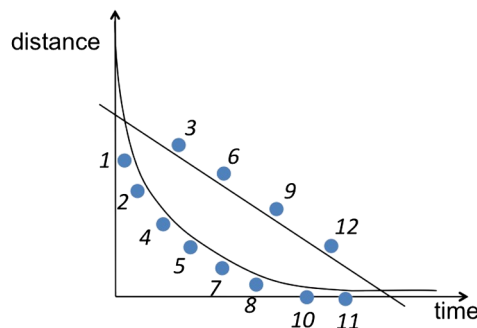

FIG. 12 (color online).   Graph problem overlaid with visual cues.

[1] R. Desimone and J. Duncan, Neural mechanisms of selective visual attention, Annu. Rev. Neurosci. **18**, 193 (1995).

[2] A. M. Treisman and G. Gelade, A feature-integration theory of attention, Cogn. Psychol. **12**, 97 (1980).

[3] D. E. Irwin and R. D. Gordon, Eye movements, attention and trans-saccadic memory, Vis. Cognit. **5**, 127 (1998).

[4] D. J. Simons and C. F. Chabris, Gorillas in our midst: Sustained inattentional blindness for dynamic events, Perception **28**, 1059 (1999).

[5] E. M. Richard and R. E. Mayer, *Multimedia Learning* (Cambridge University Press, Cambridge, England, 2001).

[6] E. R. Grant and M. J. Spivey, Eye movements and problem solving: Guiding attention guides thought, Psychol. Sci. **14**, 462 (2003).

[7] K. Duncker and L. S. Lees, On problem-solving, Psychol. Monogr. **58**, I (1945).

[8] L. E. Thomas and A. Lleras, Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition, Psychonomic Bull. Rev. **14**, 663 (2007).

[9] L. E. Thomas and A. Lleras, Swinging into thought: Directed movement guides insight in problem solving, Psychonomic Bull. Rev. **16**, 719 (2009).

[10] S. Ohlsson, in *Advances in the Psychology of Thinking*, edited by M. T. Keane and K. J. Gilhooley (Harvester-Wheatsheaf, London, 1992), p. 1.

[11] B. B. de Koning, H. K. Tabbers, R. M. J. P. Rikers, and F. Paas, Towards a framework for attention cueing in instructional animations: Guidelines for research and design, Educ. Psychol. **21**, 113 (2009).

[12] R. K. Lowe, Search strategies and inference in the exploration of scientific diagrams, Educ. Psychol. **9**, 27 (1989).

[13] P. N. Johnson-Laird, *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Harvard University Press, Cambridge, MA, 1983), Vol. 6.

[14] C. M. Sorensen, A. D. Churukian, S. Maleki, and D. A. Zollman, The New Studio format for instruction of introductory physics , Am. J. Phys. **74**, 1077 (2006).

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

[16] We recognize that the ''similar'' roller coaster problems state that the coaster starts from rest, but between the cart's initial and final positions a section of track extends slightly above the initial height of the cart. In this case the cart could not reach the intended final position. We learned of this mistake after the study was complete. We did ask students to explain out loud their answer and reasoning for each of these problems. For the roller coaster similar problems, we did not have any participants explain that the height of the final incline slightly exceeded the initial height, so we do not believe that the students noticed and responded to this error. Students who answered incorrectly always cited the shape of the middle section of the tracks (where there were no cases of a hill exceeding the initial height), the horizontal distance between carts, or the number of ''bumps'' as the reason the final speeds were different.

[17] http://www.sr-research.com

[18] R. Carmi and L. Itti, Visual causes versus correlates of attentional selection in dynamic scenes, Vision Res. **46**, 4333 (2006).

[19] D. Hammer, Student resources for learning introductory physics, Am. J. Phys. **68**, S52 (2000).

[20] M. Bilalić, P. McLeod, and F. Gobet, Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect, Cognition **108**, 652 (2008).

[21] R. A. Fisher, On the interpretation of χ^2 from contingency tables, and the calculation of P, J. R. Stat. Soc. **85**, 87 (1922).

[22] F. Cristino, S. Mathôt, J. Theeuwes, and I. D. Gilchrist, ScanMatch: A novel method for comparing fixation sequences, Behav. Res. Meth. Instrum. Comput. **42**, 692 (2010).

[23] D. Rosengrant, C. Thomson, and T. Mzoughi, Comparing experts and novices in solving electrical circuit problems with the help of eyetracking, AIP Conf. Proc. **1179**, 249 (2009).

[24] A. Feil and J. P. Mestre, Change blindness as a means of studying expertise in physics, J. Learn. Sci. **19**, 480 (2010).

[25] A. D. Smith, J. P. Mestre, and B. H. Ross, Eye-gaze patterns as students study worked-out examples in mechanics, Phys. Rev. ST Phys. Educ. Res. **6**, 020118 (2010).

[26] R. H. Tai, J. F. Loehr, and F. J. Brigham, An exploration of the use of eyegaze tracking to study problemsolving on standardized science assessments, Int. J. Res. Meth. Educ. **29**, 185 (2006).