Annual Report for Period: 10/2011 - 09/2012
Submitted on: 07/10/2012
Principal Investigator: Rebello, N. Sanjay .
Award ID: 1138697
Organization: Kansas State University
Submitted By:
Rebello, N. Sanjay - Principal Investigator
Title:
FIRE: Exploring Visual Cueing to Facilitate Problem Solving in Physics

Project Participants

Senior Personnel

Name: Rebello, N. Sanjay
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Loschky, Lester
Worked for more than 160 Hours: Yes
Contribution to Project:

Post-doc

Graduate Student

Name: Larson, Adam
Worked for more than 160 Hours: Yes
Contribution to Project:
Assist the project team with (i) developing the protocols for the eye-tracking experiments (ii) programming the eye-tracker and accessories for the experiment (iii) converting the eye-tracking data into a form that can be analyzed by other personnel on the project

Name: Madsen, Adrian
Worked for more than 160 Hours: Yes
Contribution to Project:
Assist the project team with (i) developing and designing the interview protocols; (ii) analyzing qualitative and quantitative data from eye-tracking interviews with students, and (iii) creating publishable papers and presenting talks to disseminate the research The doctoral dissertation of this student will focus on the various aspects of the research and analysis of eye-tracking interviews.

Name: Rouinfar, Amy
Worked for more than 160 Hours: Yes
Contribution to Project:
Assist the project team with (i) developing and designing the interview protocols; (ii) analyzing qualitative and quantitative data from eye-tracking interviews with students, and (iii) creating publishable papers and presenting talks to disseminate the research The doctoral dissertation of this student will focus on the various aspects of the research and analysis of eye-tracking interviews.

Undergraduate Student

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

Other Collaborators or Contacts

David Irwin (University of Illinois)
Jose P. Mestre (University of Illinois)
Brian Ross (University of Illinois)
Eric Weibe (North Carolina State University)

Activities and Findings

Research and Education Activities: (See PDF version submitted by PI at the end of the report)
Please see attached file.

Findings: (See PDF version submitted by PI at the end of the report)
Please see attached file.

Training and Development:
P.I. Rebello: Learned the methods of research in cognitive psychology, particularly visual cognition research, by participating in the group meetings with the Loschky group, attending the Cognitive Science Brownbag series and presenting there.

P.I. Loschky: Learned about the issues pertaining to physics education research, particularly problems solving. Also, realized the potential applications of his research on visual cognitin to problem solving in STEM education.

Graduate Student Larson: Learned about the issues pertaining to physics education research, particularly problems solving. Also, realized the potential applications of his research on visual cognition to problem solving in STEM education.

Graduate Student Madsen: Learned the methods of research in cognitive psychology, particularly visual cognition research, by participating in the group meetings with the Loschky group, attending the Cognitive Science Brownbag series and presenting there.
Work done on this project contributes in substantial part to Ms. Madsen's Ph.D. dissertation.

Graduate Student Rouinfar: Learned the methods of research in cognitive psychology, particularly visual cognition research, by participating in the group meetings with the Loschky group, attending the Cognitive Science Brownbag series and presenting there.
Work done on this project contributes in substantial part to Ms. Rouinfar's Ph.D. dissertation.

Outreach Activities:
Ms. Madsen shared the methods and some of the prelimiarly findings of ths study in a local high school physics classroom.

Journal Publications

Madsen, AM; Larson, AM; Loschky, LC; Rebello, NS, "Differences in visual attention between those who correctly and incorrectly answer physics problems", PHYSICAL REVIEW SPECIAL TOPICS-PHYSICS EDUCATION RESEARCH, p. , vol. 8, (2012). Published, 10.1103/PhysRevSTPER.8.01012

Books or Other One-time Publications

Web/Internet Site

URL(s):
http://web.phys.ksu.edu/fire

Description:
This is the main website for this grant award. It contains:
-- Project Proposal
-- Summary of Studies
-- Publications
-- Talks and Posters
-- Personnel

Other Specific Products

Contributions

Contributions within Discipline:

Contributions to Other Disciplines:

Contributions to Human Resource Development:

Contributions to Resources for Research and Education:

Contributions Beyond Science and Engineering:

Conference Proceedings

Special Requirements

Special reporting requirements: None
Change in Objectives or Scope: None
Animal, Human Subjects, Biohazards: None

Categories for which nothing is reported:

Organizational Partners
Any Book
Any Product
Contributions: To Any within Discipline
Contributions: To Any Other Disciplines
Contributions: To Any Human Resource Development
Contributions: To Any Resources for Research and Education
Contributions: To Any Beyond Science and Engineering
Any Conference

Exploring Visual Cueing to Facilitate Problem Solving in Physics

PI: N. Sanjay Rebello
Co-PI: Lester C. Loschky
Kansas State University
Project Award Number: 1138697

Annual Report for Year I: 2011-2012

PROJECT FINDINGS

During the first year of this grant we made the following findings:

- **STUDY 1** Findings: See details below.
- **STUDY 2** (PILOT) Findings: See details below.
- Met with the **Advisory Board** via Skype: See attached…
  - Executive summary provided to committee prior to meeting.
  - Report on themes of the conversation at the meeting, prepared after the meeting.
- Also attached:
  - **Journal Paper**: *Physical Review Special Topics – Physics Education Research*
  - **Conference Papers**:
    - *National Association for Research in Science Teaching* (*NARST 2012*)
    - *American Educational Research Association* (*AERA 2012*)
    - *Eye Tracking Research & Applications* (*ETRA 2012*)
  - Conference **Posters** & **Talks**:
    - Talk at *Eye Tracking Research & Applications* (*ETRA 2012*)
    - Poster at *Eye Tracking Research & Applications* (*ETRA 2012*)

STUDY 1: Differences in visual attention between those who correctly and incorrectly answer physics problems

FINDINGS

STUDY 1A: INTERVIEWS TO DETERMINE NOVICE LIKE AREAS OF PROBLEMS

The six problems included in this analysis showed consistent incorrect reasoning patterns. These answer patterns align well with previous findings in the literature. This gave us confidence that the definitions of novice-like areas of interest, for each physics problem, do indeed represent the most common novice-like answers of the larger population of introductory physics students.

STUDY 1B: DETERMINING DIFFERENCES IN VISUAL SELECTIVE ATTENTION BASED ON CORRECTNESS OF PROBLEM SOLUTION

Results from Area of Interest (AOI) Analysis

Mixed factorial 2 x 6 ANOVAs with proportion of time in each AOI type as the dependent variable and problem number and correctness of answer as independent variables were conducted for all three AOI types.

Full trial period: We found a significant main effect for correctness of answer as well as for problem number for all three AOI types. We were looking for a main effect of correctness, as this would indicate there are differences in percentage of time spent in an AOI between those who answered correctly and those who answered incorrectly. The main effect of correctness addresses our research questions and will be further analyzed below. The main effect of problem number indicates there is at least one difference in proportion of time in each AOI type between different problems. We were not interested in how the proportion of time spent fixating varies between problems, as this is not relevant to our research questions, so the effect of problem number will not be further analyzed. We found a significant interaction between problem number and correctness of answer in the perceptually salient AOI. This means the relationship between correctness and time spent in the perceptually salient area is different across problems. This interaction is not relevant to our research question and will not be further investigated

First Two Seconds After Leaving Problem Statement: We found that on five out of six problems used in Study 1B, those who answered the problem correctly spent a higher percentage of total viewing time fixating on thematically relevant areas in the problem diagram. Those who answered correctly likely had the domain knowledge needed to solve each problem, and therefore spent more time viewing the relevant areas in each diagram. This result is consistent with previous findings where those with high levels of domain knowledge in a discipline, such as identifying fish locomotion, art, and chess, spend more time looking at areas of diagrams and pictures relevant to a task. Our finding is evidence for top-down processes playing a key role in guiding visual attention when solving physics problems correctly.

We also found that on five out of six problems, those who answered the problem incorrectly spent a higher percentage of total viewing time looking at areas of the diagram consistent with a novice-like response. Furthermore, on the one problem that did not quite reach statistical significance ($p = .058$) the effect was in the same direction as the other five problems. These novice-like AOIs were determined through individual interviews described in Study 1A, and were consistent with the physics education literature describing common student misconceptions. Importantly, the finding that incorrect solvers spent more time fixating on novice-like areas is evidence for their visual attention being guided by top-down processes. However, instead of attention being guided by scientifically correct domain knowledge, incorrect problem solvers' attention was guided by novice-like misconceptions. Thus, when solving physics problems, top-down processing plays a key role in guiding visual selective attention either to thematically relevant areas, or novice-like areas, depending upon the scientific correctness of a student's physics knowledge.

Concerning the effects of bottom-up processes in guiding attention during physics problem solving, we found that those who answered incorrectly spent more time in perceptually salient areas during the full problem period on only two of the six problems. Nevertheless, for five of the six problems the effect was in the predicted direction, such that incorrect problem solvers spent a higher percentage of total time fixating on the perceptually salient AOIs than the correct problem solvers. However, four of those effects were not statistically significant. A likely explanation for this result is that in these two problems, the perceptually salient AOI partially or completely overlapped with the novice-like AOI, which was not the case for the other four problems. We have already shown that those who answered the problem incorrectly spent significantly more time fixating on the novice-like AOIs on two of the six problems than those who answered the problem correctly. So the significant result for these problems for the perceptually salient AOI is likely due to this AOI overlapping with the novice-like AOI. This result also seems to indicate that attending to the perceptually salient area is not necessarily a good predictor of correctness. These results appear to be consistent with a study of change blindness that found that problem solvers seldom notice changes in color, even though color is most perceptually salient. Thus, when considering the full time period of problem solving, perceptual salience appears to have played a minimal role in guiding the attention of incorrect physics problem solvers. Nevertheless, previous vision research has suggested that the effects of bottom-up perceptual salience on eye movements are limited to the first two seconds of viewing a stimulus. Thus, this seeming null result could be argued to have resulted from diluting the effect of saliency by including eye-movement data from the entire duration of the trial, rather than only the first two seconds.

To reanalyze the data including only the first two seconds of viewing a diagram, we completed a mixed factorial 2 x 6 ANOVA with proportion of time in each AOI type as the dependent variable and problem number and correctness of answer as independent variables for all three AOI types for the first two seconds of viewing the diagram. We were looking for a main effect of correctness, as this would indicate there are differences in percentage of time spent in an AOI between those who answered correctly and those who answered incorrectly. For the first two seconds after leaving the problem statement, we found no main effect for correctness of answer for any of the AOI types. So, there are no significant differences in proportion of time spent fixating in the AOI types between those who answered correctly and those who answered incorrectly for any of the problems and no further analysis was conducted.

We did find a main effect for problem number for the novice-like and perceptually salient AOIs. This means for each of these AOIs, there is at least one difference in proportion of time between the different problems when considering the data for all participants. We were not interested in how the proportion of time spent fixating varies between problems, as this is not relevant to our research questions. We also found a significant interaction between problem number and correctness of answer in the thematically relevant AOI. This means the relationship between correctness and time spent in the thematically relevant area is different across problems. This interaction also does not address our research questions, and is not analyzed further.

The reanalysis of the data for the first two seconds of viewing the diagram found no statistically significant differences between correct and incorrect solvers on any of the problems for the perceptually salient AOI. Indeed, there were no statistically significant differences between correct and incorrect solvers in time spent in the thematically relevant or novice-like AOIs. In sum, we found no support for the hypothesis that perceptual salience influences visual selective attention more for incorrect problem solvers during the first two seconds of diagram viewing. This result is consistent with previous studies that have shown that top-down influences on visual attention tend to dominate bottom-up influences when a viewer is given a specific goal or task. Nevertheless, such null results for the effects of bottom-up saliency on visual attention are consistent with our own results, which considered both the full problem solving time period, and only the first two seconds, and found little if any effects.

However, before completely rejecting the hypothesis that bottom-up saliency affects attentional selection during physics problem solving, we must consider two observations that provide partial support for it. First, it may be that the early effect of perceptual salience on eye movements was present; however, the data lacked sufficient statistical power to detect it. Some support for this explanation is shown by comparing the mean difference for the correct versus incorrect problem solvers for the perceptually salient AOIs for the first two seconds of viewing the diagram. Specifically, the percentage of time spent looking in the perceptually salient AOI is higher for incorrect solvers than correct problem solvers on five of the six problems, though not statistically significantly so. Thus, it is possible that a larger study with more observations might show this effect to be statistically significant. Secondly, the perceptual salience model proposed by Itti and Koch predicted that early in scene viewing eye movements are more influenced by bottom-up perceptual information than top-down knowledge. Therefore, the saliency model would predict that early in viewing a physics problem, correct and incorrect problem solvers would not have had sufficient amount time to apply their (correct or incorrect) top-down knowledge to guide their attention to thematically relevant or novice-like areas of the diagram. If so, during the first two seconds of viewing the diagram, there should be no difference between correct and incorrect problems solvers' percentage of total fixation time in either the thematically relevant or novice-like AOIs. The data supports this hypothesis, which shows that there is no significant difference in viewing time for thematically relevant AOIs between correct and incorrect problem solvers. In sum, the data showed essentially no influence by top-down domain knowledge during the first two seconds of diagram viewing, though such effects were statistically significant later in time, when considering the full problem solving time period. Thus, based on the above two observations, we must withhold complete rejection of the hypothesis that bottom-up salience affects the visual selective attention of incorrect physics problem solvers. Even so, such an interpretation of the data should be made cautiously since it is based on null effects. Future studies will be required in order to explicitly test the effects of bottom-up and top-down information on early and late visual selective attention processes in eye movements.

Results from Scan Path Analysis

We did not find significant differences in ScanMatch scores be-tween those in the C-C comparisons and those in the I-I compari-sons on five of the six problems analyzed in this study. This evi-dence is consistent with the hypothesis that the attention of incorrect solvers is primarily directed by top-down naïve theories and not the relative perceptual salience of the elements. This finding aligns well with our previous findings [Madsen et al. 2011] that showed no significant difference in the percentage of fixation time in the perceptually salient areas of the diagram during the full problem period, or the first two seconds of viewing the diagram, when the effects of perceptual salience should be most pronounced. It also aligns well with the findings showing significant differences in the percentage of time incorrect solvers spent in the novice-like areas of the diagram and the percentage of time correct solvers spent in the thematically-relevant areas of the diagram.

We found significant differences between the I-I and C-I compari-sons on three of the six problems. These differences were expected as we have previously seen that correct solvers and incorrect solvers spend different amounts of time looking at thematically-relevant and novice-like elements in the problem, so their scan paths scores are likely to be different. It is curious that we did not find that the I-I comparison and the C-C comparison had higher ScanMatch scores than the C-I comparison on all of the problems. The problems used in the study included a text problem statement, diagram, and multiple-choice answers. The hypotheses set forward in this study assumed a similar reading pattern of the problem statement and answer choices for all participants. The hypotheses were formed assuming only differences in how the participants looked at the diagram. Differences in reading the problem statement and answer choices may have overwhelmed small differences in diagram viewing, resulting in no difference in the ScanMatch scores of the C-C and I-I comparisons compared to the C-I comparison.

These findings may have implications for educational interventions aimed at helping novices learn to answer such conceptual questions correctly. Researchers in physics education have devoted much attention to addressing these consistent wrong answer patterns by changing the way students think about how the world works. If it were true that this problem had an underlying perceptual component, these interventions would need to instead help students learn how to ignore salient elements and focus instead on thematically-relevant elements. The results of this study suggest that wrong answers have roots in the incorrect ways students think about how the world works, not how a problem diagram looks. So it seems that the educational interventions used to improve student understanding are on the right track.

STUDY 2 (PILOT)**: Using Dynamic Cues to Influence Reasoning** (NOTE: This is the pilot for a more detailed STUDY 2 to be conducted in the second year of the grant.)

FINDINGS

In this study we find some evidence that viewing a physics problem overlaid with short duration visual cues can indeed help students correctly answer and reason about problems they were previously unable to. Of the four problem sets used, we found on one of these problem sets significantly more students changed to a correct answer after seeing cues. It is not enough though, to provide visual cues to help students answer a given set of problems. In looking at transfer problem performance, we found nearly significant differences on the ball and graph transfer problems with the cue group outperforming the no cue group. Thus, we find some evidence that repeatedly showing novices visual cues on related problems may help them to properly apply the factual knowledge on similar future problems viewed without cues.

We also found some differences in eye movements of those who changed to a correct answer on a similar problem and those who did not. Those who changed to a correct answer on a similar problem for the rollercoaster problem followed the visual cues more closely than those who did not change to a correct answer. Thus, there may be a relationship between how well a participant follows the visual cues with their eyes, and how helpful these cues are. This suggests that following the cues closely is related to changing to a correct answer. Further, we looked for evidence that seeing cues changes the way in which one views future problems with no cues and found no evidence for this on the ball transfer problem.

While we did find some results that point to the usefulness of cueing, we also saw no difference between cue and no cue groups in the number of similar problems answered correctly on three of the four problem sets tested. There are many reasons that the cueing may have failed. First, we have previously discussed how the cue pattern may have been too complex on these three problems. In the future, we can use simple repetitive cues as well as increase the time the cues are seen. We can also change the type of cue we are using, for example instead of moving colored shapes modeling correct solvers' eye movements, we could use lines or bars to illustrate comparisons in the diagram that correct solvers make. Additionally, there may be only certain types of problems that lend themselves to improvement through visual cueing. We have only explored four problems in this study. There are a plethora of problems that can be categorized in a variety of ways and 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 will be randomized to balance out any order effects.

We also found differences between the cue and no cue groups on two of the four transfer problems tested. As mentioned earlier it may be that the two 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 the researchers viewed the transfer problems as closely related to the similar problems, though the students did not view them this way, and thus were unable to apply what they gained from the cues to the transfer problems. In other words, the transfer problems, though deemed to be near transfer problems by the researchers, were perceived to be far transfer problems by the participants in our study. A problem may be perceived as near or far transfer depending upon whether the problem solver perceives the two problems to be different in surface feature or deep structure. So, it seems that although the 'similar' and 'transfer' problems were deemed to differ only in surface feature by the researchers, the participants in our study appear to have perceived them as being different in deep structure as well.

Previous studies with visual cues in several domains have also found mixed results on the effectiveness of cueing. There is much work to be done to understand the factors that lead to helpful cues. This study offers some hope that cueing can potentially serve as effective conceptual scaffolding for novice physics students, but much work is necessary to perfect this method.

Visual Cueing in Physics Problem Solving
Advisory Board Meeting
March 16, 2012

SUMMARY REPORT TO ADVISORY COMMITTEE

The overarching goal is to explore and possibly exploit the link between cognition and eye movements in the context of physics problem solving. Several studies have shown consistent incorrect answer patterns for some conceptual physics problems (see Docktor & Mestre, 2011 for a review). Until recently, several cognitive top-down explanations were provided, including misconceptions or misapplication of conceptual resources. Alternatively, Heckler (2011) has suggested a bottom-up perceptual basis for incorrect answers. Perceptually salient and plausibly relevant features in a problem capture students' attention through bottom-up processes, leading to incorrect answers, while less salient, albeit thematically relevant features are not considered by the student.

Our first study (Madsen et. al, 2011) found that incorrect solvers spend statistically significantly larger percentage of their dwell time in the 'novice-like' areas of a problem than the correct solvers, but found no such difference in 'perceptually salient' areas. A subsequent analysis (Madsen, et. al., 2012) showed that on most problems the scan paths (Cristino et al. 2010) of incorrect solvers were no more similar to each other than the scan paths of correct solvers. These results appear to be consistent with the top-down basis for incorrect answer patterns.

Our second study explored the possibility of using visual cues to facilitate problem solving in physics. Prior research (Thomas & Lleras, 2006; Grant & Spivey, 2003) shows that visually cueing attention in a way that embodies the problem solution can improve problem solving performance. We found some evidence that viewing a physics problem overlaid with short duration visual cues can indeed help students correctly answer and reason about problems they were previously unable to. We also found some evidence that repeatedly showing novices visual cues on related problems may help improve their performance on transfer problems viewed without cues. Finally, we also found that participants who changed their answer to a correct answer after cueing had a significantly higher percentage of saccades that closely followed the visual cues than those that did not follow the cues. This result suggests that following the cues closely is related to changing to a correct answer. However, we found no evidence that seeing cues changes participants' eye movements on the transfer problem. Given the small number of participants in this study, we are continuing to collect data on this study.

Our most recent (current) study revisits the issue of salience that we had begun to explore in the first study. Many of the problems in the first study had the novice-like areas overlapping with the perceptually salient areas. Thus, it was difficult to distinguish whether top-down or bottom-up processes were driving the answer patterns. In the current study we have controlled for perceptual salience by creating three versions of each problem – one in which the novice-like region has the highest salience, another in which the thematically relevant region has the highest salience, and yet another in which the two areas have equal salience. This study will enable us to test whether it is the top-down or bottom-up processes that drive students' eye movements and answer patterns to these problems.

Based on the studies so far, we plan to investigate the following research questions going forward:

- What factors of a cue (e.g. duration, cue strength should we experiment with?
    - Should we increase cue duration?
        - Could this depend on what experts do?
        - If so, would this be arbitrary or non-generalizable to different contexts and problems?
    - Should we increase cue strength? If so how do we do this? Some possibilities are:
        - Increasing cueing strength based on bottom-up cueing (e.g. multiple cues presented on the same figure)
        - Increasing the strength of the cue by perceptually grouping the conceptual information needed to answer the problem.
    - Should we use both exogenous and endogenous cues?
        - We assume, the impact of the perceptual salience in the first two seconds is affecting cognition after the first two seconds. However, this is just the first step in the processes regarding thinking about the problem. So we:
            - First use an exogenous cue to attract attention. This may get people thinking about the appropriate construct.
            - Next, provide a stronger cue that could be either perceptually salient or it could represent the conceptual information needed to solve the problem.
        - In this way, the concept (top-down) is primed by salient information (bottom-up)

- Do we need to tell people that the cue will help them?
    - In normal instruction, teachers give students explicit instruction. But the use of cues seems an inherently implicit instructional strategy.
        - How does this relate to the distinction between explicit declarative knowledge (e.g., of a principle or concept from a lecture or textbook) versus implicit procedural knowledge (e.g., used when solving a problem)?

- How many times should we repeat the cueing in similar problems without making the students aggravated?

We hope to gain insights and advice from the Advisory Committee on these questions.

FIRE Visual Cueing Project Advisory Board Meeting
Friday March 16, 2012

Themes of Conversation

Present:
Advisory Board: David Irwin, Jose Mestre, Brian Ross, Eric Wiebe
KSU Visual Cueing Project Group: Adrian Madsen, Adam Larson, Amy Rouinfar, Lester Loschky, N. Sanjay Rebello

Below is a bulleted list of the main themes that seemed to emerge from the conversation during the Advisory Board Meeting.

Explicit Versus Implicit Cues

- General agreement that explicit cues are preferable.
    - Tell students what you are going to do with the cues--what you will illustrate--and then give the cues.
    - If students see cues "cold" may not know what to do with them.
- Thomas and Lleras used implicit cueing to show interesting effect in a problem where there was a single new thing the participant needed to think about. In education, we want the strongest effect and its likely explicitly telling students that cues are helpful will lead to this.
- The goal is transfer, telling students that cues are helpful gives a better chance of seeing transfer.

Duration of Cues

- Time scale of problem solving in physics is very slow, so repeat the cues every few seconds rather than increase the duration, i.e., present for 4 seconds, wait for 2 second, present again for 4 seconds, etc. Maybe increase the intensity of the cues as you repeat them.
- Repeating them with increasing emphasis might help avoid adaptation.

Cue Design

- Don't design cues based on what experts do. There may be cues that are more pedagogically salient/effective for novices.
- Instead, do a task analysis to determine the steps for a novice to successfully solve the problem.
    - This is in fact what we did when we developed the pilot study on cueing. However, we need to do it more formally now, with an eye to eventually generating a generalizable theoretical framework within which we do it.
- Relate the steps from the task analysis to the three types of cues: selection, organization and integration.
    - Go beyond selection, use cues to compare heights on skier problem, don't just have participants attend to heights.

  - There are several cognitive steps needed to solve many of the example problems and the number of steps varies in each problem. For example on distance time graph, we can highlight slope but students need to know how to relate slope of d vs t graph to velocity. It also takes more steps to pick the slopes as the right answer than it does to pick the crossing point (wrong answer)
- Cues could explicitly show how to transfer from one problem to another.
- Ultimately want a general model of cueing
  - Want a general model of how to design cues that can be applied to any problem, by anyone. Need general principles to create this model.
  - Start with cues that work for sets of problems (and cues that don't) and a general theory can emerge from this.

Purpose of Cues

- Need to make the purpose of the cues more explicit. Are we trying to improve transfer? Make minimalistic cues?
- Cues should scaffold novices understanding and help them pull pieces together and make sense of material.
- Cues should relay the purpose and procedure needed to extract necessary information from the diagram.
- Cues can help student's abstract information from problems into a general model that will be useful for transfer. The visual support should help with the abstraction.
  - May need a visual intermediary to help with abstraction of concepts. Visually link to an intermediate abstraction.

Cueing in Reading vs. Physics

- Poor readers make dysfunctional eye movements compared to good readers. At one point it was thought that training poor readers to make better eye movements would improve their reading but this failed. Poor readers are poor because of top-down reasons, not because they make poor eye movements.
- In physics, novices don't even know what to look at, but looking in itself is not adequate. Need to think about top down organization and integration.

Differences in visual attention between those who correctly and incorrectly answer physics problems

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

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

(Received 28 December 2011; published 11 May 2012)

This study investigated how visual attention differed between those who correctly versus incorrectly answered introductory physics problems. We recorded eye movements of 24 individuals on six different conceptual physics problems where the necessary information to solve the problem was contained in a diagram. The problems also contained areas consistent with a novicelike response and areas of high perceptual salience. Participants ranged from those who had only taken one high school physics course to those who had completed a Physics Ph.D. We found that participants who answered correctly spent a higher percentage of time looking at the relevant areas of the diagram, and those who answered incorrectly spent a higher percentage of time looking in areas of the diagram consistent with a novicelike answer. Thus, when solving physics problems, top-down processing plays a key role in guiding visual selective attention either to thematically relevant areas or novicelike areas depending on the accuracy of a student's physics knowledge. This result has implications for the use of visual cues to redirect individuals' attention to relevant portions of the diagrams and may potentially influence the way they reason about these problems.

DOI: 10.1103/PhysRevSTPER.8.010122 PACS numbers: 01.40.Fk

I. INTRODUCTION

Often diagrams in physics problems contain information that is both relevant to the solution of the problem and information that is irrelevant. Students commonly use this irrelevant information as they reason their way to an incorrect answer, when, in fact, they should simply ignore it. The use of irrelevant information in student answers has been observed in many studies, such as those by McDermott looking at common student difficulties in understanding motion [1,2].

A convenient way of measuring what learners pay attention to is to measure their eye movements. In the current study, we measure students' saccades (i.e., when eyes are in motion) and fixations (i.e., when eyes are stationary at a specific spatial location) to measure where they attend in physics problems. Attention is guided by two sources of information, one internal and the other external, referred to as top-down and bottom-up information, respectively. Bottom-up information is based on the physical characteristics of stimuli on which a person fixates, and the visual processes that work on such information tend to be very fast and involve primitive brain areas early in the visual stream [3,4]. The influence of bottom-up information on attention is generally explained in terms of the relative *perceptual salience* of elements of the visual stimuli [5–7]. Perceptually salient regions of an image tend to be those with relatively greater contrast in terms of color, orientation, intensity, or motion compared to the other image elements. Perceptually salient elements are argued to automatically capture attention through primitive visual mechanisms [8,9]. Itti, Koch, and Niebur [5,6] have created a computational algorithm to produce a salience map of a scene or diagram, using contrasts of light intensity, orientation (e.g., of lines), and color. Such salience maps have been found to predict significantly greater than chance where people will fixate as they view images [10,11] though top-down factors (described below) have been shown to have even larger effects on where people fixate [12–14]. Models of the effects of saliency on eye movements argue that one's attention first selects the location with highest salience, this location is then fixated, and after the information there has been sufficiently processed, one's attention moves to the next most salient spatial location. Carmi and Itti [7] studied the effects of saliency as a function of viewing time and found that their perceptual salience model best predicted the first six or seven fixations when viewing a scene (see also Parkhurst, Law, and Niebur [10]). For the average viewer, this is equivalent to about the first 2 sec of viewing. This suggests that bottom-up processes are more dominant in the first 2 sec of viewing, with top-down processes exerting a greater influence on eye movements thereafter.

In the domain of physics, it has been proposed by Heckler [15] that the consistent wrong answer pattern by novices on introductory physics problems is in part a result of their attention being directed to the most perceptually salient and plausibly relevant features in a problem. He explains that the most salient features capture attention through perceptual processes and less salient features have little opportunity to be considered, leading to an incorrect answer. Student answer patterns are cited as evidence for these perceptually driven responses; however, no eye-movement data supporting this hypothesis are provided.

However, some researchers [13] have found that perceptual saliency, as assessed by Itti's model, did poorly in accounting for the paths that viewers' eyes took when given a search task. For instance, in the study by Hegarty, Canham, and Fabrikant [16], university students viewed weather maps and were tasked to determine wind direction. The researchers found no evidence to indicate that over the full trial period participants looked at the perceptually salient areas of the weather maps based on Itti's algorithm. However, the researchers did not limit their analysis to only the first 2 sec of viewing, when the effect of saliency driven bottom-up processes should be most pronounced.

Top-down information and the processes that act upon it are based on the viewer's prior knowledge, task goals, and expectancies. Top-down effects on attention tend to be mediated by higher brain areas and occur later in the time course of vision [17,18]. Most importantly for the current study, it has been observed that experts in a domain attend to task-relevant portions of a diagram more than novices in that domain. Thus, the expertise of these individuals helps to guide their visual attention in the diagram. Jarodzka *et al.* [19] studied the visual attention of both novices and experts who viewed videos of unfamiliar fish swimming and classified the type of locomotion. The authors found that experts spent significantly more time fixating on relevant areas of the video than biology students, who had the necessary background knowledge for differentiating types of locomotion but little practice in this classification task. The authors also found that novices spent more time than experts fixating on areas irrelevant for determining locomotion. Similar studies have measured eye movements of experts when viewing art [20] and playing chess [21], and have shown that the increased domain knowledge in these fields affects where people fixate while performing domain-relevant visual tasks. Thus, important differences in the eye movements of experts, who possess the necessary domain knowledge, versus novices, who do not possess such knowledge, can be seen by tracking their eye movements while they are carrying out domain-relevant tasks [22–24].

Visual attention allocation in the discipline of physics may work somewhat differently than the previously discussed disciplines as our everyday interactions with the physical world may help us develop ideas about how it works without any formal instruction. Thus, novice reasoning may be influenced by top-down knowledge, which may be based on either correct or incorrect representations of the physical world. Physics education research has cataloged a pattern of consistently incorrect answers to many common physics questions. These patterns, called misconceptions [25,26], may be a result of stable mental entities built up through years of interaction with the physical world and through schooling. These consistently incorrect answer patterns have also been explained in terms of a misapplication of small chunks of information, referred to as resources [27], which students develop through their experience with the world. In a physics class, they may bring together groups of resources to answer questions, and may apply inappropriate resources to a given situation. Conversely, these consistently incorrect answer patterns may be the result of students categorizing scientific ideas into inappropriate ontological categories [28]. However, while the precise cognitive processes that lead to these consistently incorrect answer patterns are still being debated, all the proposed explanations rely in some way on ''domain knowledge'' about how the world works. Thus, for the purposes of this paper, we will refer to the cognitive underpinnings of these consistently incorrect answer patterns in physics as *novicelike misconceptions*. Therefore, a key question addressed in the current study is whether both experts' scientifically correct domain knowledge and beginners' novicelike misconceptions exert top-down influences on visual attention when viewing physics problems. If novicelike conceptions do influence eye movements when answering physics problems, then participants who provide incorrect answers should spend more time fixating on irrelevant areas of a diagram than the relevant or perceptually salient areas of the diagram.

The interaction between perceptual salience and level of domain knowledge is also important to consider. A study by Lowe [29] found that the written responses of meteorology students who studied animated weather maps and recorded generalizations about them primarily contained information extracted from perceptually salient areas of the weather maps. However, a more recent study by Hegarty, Canham, and Fabrikant [16] showed an interesting interaction between bottom-up salience and top-down knowledge in guiding attention while looking at weather maps. The authors investigated this interaction by recording participants' eye movements as they viewed static weather maps in which the relative salience of task-relevant and task-irrelevant information had been manipulated. Before instruction, participants spent more time fixating on task-irrelevant areas when they were the most perceptually salient elements on the map. However, after instruction, there was no difference in the time spent fixating on task-irrelevant information regardless of which elements had been made most perceptually salient. Thus, while both of these studies show that novice learners are strongly influenced by areas of a diagram that are perceptually salient, the study by Hegarty, Canham, and Fabrikant shows that the effect of perceptual salience on attention decreases as domain knowledge increases.

Previous research has shown that there is competition for attention between bottom-up and top-down processes as people view visual stimuli. The key question addressed in the current study is how these processes interact when answering physics problems. We use eye-movement data to infer the extent to which bottom-up and top-down processes influence people's attention as they answer introductory conceptual physics questions containing diagrams.

We hypothesize that those with adequate domain knowledge to correctly answer a problem will spend more time fixating on thematically relevant areas of a diagram that provide the solution to the problem than on irrelevant areas of the diagram. Conversely, we predict that those who answer incorrectly will spend more time fixating elsewhere in the diagram. More specifically, based on previous research in physics education concerning novicelike misconceptions, which consistently lead to incorrect answers, we hypothesize that those answering the problem incorrectly will spend more time fixating on areas of the diagram consistent with a novicelike misconception. These participants will initially attend to perceptually salient areas of the diagram, but will quickly disengage their attention from these areas and instead attend to novicelike areas. Such effects would suggest a strong role for top-down factors in guiding attention while solving physics problems involving diagrams.

Alternatively, it has been shown that perceptual salience has a larger influence on novice learners' eye movements than those with more domain knowledge. Based on this finding, we could predict that the fixated locations of those who answer incorrectly are more likely to be influenced by perceptual salience than those who have adequate domain knowledge. Such effects would suggest a strong role for bottom-up factors in guiding attention during physics problem solving with diagrams. Thus, a key question is whether the attention of people who answer physics problems incorrectly is more influenced by the top-down factor of novicelike misconceptions or by the bottom-up factor of the perceptually salient areas of the diagram.

Specifically, we examine the following three-part research question:

How does the correctness or incorrectness of one's answer to a physics problem involving a diagram relate to the time spent *looking at* those areas of the diagram that are

(a) thematically relevant to the problem's solution?
(b) consistent with novicelike misconceptions? Or
(c) perceptually salient?

II. STUDY 1: INTERVIEWS TO DETERMINE NOVICELIKE AREAS OF INTEREST

A. Study 1: Methodology

In order to define areas of a physics problem diagram that contain visual information related to a novicelike misconception, we conducted individual interviews with students enrolled in an introductory psychology course. We specifically looked at the interview segments where participants provided incorrect answers to the physics problems and observed the areas of the diagram that students identified and discussed while giving their verbal explanation. This information was used to define "novicelike" areas of interest (AOI), or specific areas of the diagram in which a participant who answered incorrectly would use to come to their answer. These areas of interest will be used in the analysis for study 2.

1. Participants

The participants were 13 students (eight females) enrolled in an introductory psychology course. All of the students had taken at least one physics course in high school, though some had taken an introductory physics course at the university level as well. They were given course credit for participation.

2. Materials

The materials consisted of 10 multiple-choice conceptual physics problems covering various topics in introductory physics including energy, kinematics, and graphing of motion (see the Appendix for a list of problems). Each problem contained a diagram that had a thematically relevant visual component that students needed to attend to in order to correctly answer the question. For example, in problem 4 (see the Appendix), to compare the speeds of ball A and ball B, one must attend to the distances between the balls at each time interval and ignore the point where the balls are aligned spatially. So, the distance between balls at 2 and 3 sec is the relevant area to attend to. These problems were chosen based on prior experience of the researchers which indicated that these problems could be answered using common naive conceptions documented in physics education literature [1,2,30].

3. Procedure

Each participant took part in an individual session which was between 20 and 40 min long. At the beginning of the session, participants were given a short explanation of the goal of the interview and the purpose of the research. Further, they were instructed to think aloud and explain their reasoning process as they answered each question. They were told they might be asked additional clarifying questions during their explanations. Participants were given one problem at a time, each printed on an $8\ 1/2 \times 11$ sheet of paper. They were allowed to write or draw on the problems as they deemed necessary. If a participant's answer was not clear, the interviewer asked questions to clarify the meaning of the explanation. Participants' verbal explanations, gestures, and writing on the paper were recorded with a Flip video camera.

B. Study 1: Analysis

The purpose of these interviews was to determine which portion of each diagram was attended to by incorrect problem solvers. Therefore, only the interview segments where the participant gave a final incorrect answer were included in the analysis. A phenomenological approach was used to code the interviews [31]. Table I contains the answers and reasoning provided by participants who answered the problems incorrectly. Four of the 10 problems used in the interviews showed no consistent answering patterns among incorrect solvers after a first pass analysis. These problems are not included here, as there were no identifiable novicelike areas to be utilized in study 2.

C. Study 1: Results and conclusion

The six problems included in this analysis (see the Appendix) showed consistent incorrect reasoning patterns. These answer patterns align well with previous findings in the literature. Student difficulties with distance versus time graphs were studied extensively by McDermott, Rosenquist, and van Zee [2] and Beichner [32]. McDermott, Rosenquist, and van Zee interviewed students at all levels of introductory college physics as well as high school physics and physical science students. They found when students responded to a problem very similar to problem 2 used in our study, they often selected the point where the graph crossed the x axis because "the position was going from positive to negative," instead of correctly choosing the point on the graph where the slope was zero. In a similar study, Trowbridge and McDermott [33] found that a common student misconception is the idea that when two objects have reached the same spatial position they have the same speed. In their study, Trowbridge and McDermott used a problem very similar to problem 4 in our study, and found that a substantial number of students chose the instant when the balls passed each other as the time when they were moving at the same speed. In problem 4 in our study, this instant of the balls passing is at 1 sec, which is the most common incorrect answer we observed. Conflating position and speed is also observed in problems 3 and 7 in our study. In problem 7, we observed students incorrectly choosing the point where the graphs of two objects crossed as the point when the objects were moving at the same speed. This crossing point is the place where the objects have the same position, but not the same speed. In problem 3, we observed students choosing the points where the graph crosses the x axis as the place where the object's speed is zero. These crossing points are the places where the object has a zero position

TABLE I. Number of students providing each answer and reasoning on conceptual physics questions with a diagram.

Question no. and description	Answer	Reasoning	No. of responses
Q1. Roller coaster	Final speed B > final speed A	Compares drops and climbs on tracks A and B	2
		Height of initial drop on track A > height of initial drop on track B	2
	Final speed A > final speed B	Compares drops and climbs on tracks A and B	5
Q2. Distance time graph 1	Point C	Distance changes from positive to negative	5
Q3. Distance time graph 2	Point A	Distance is zero	2
		Distance and time are zero	2
	Points A and C	Distance and time are zero	1
	Point C	Distance goes from negative to positive	1
Q4. Balls on tracks	1 second	Balls at the same position at same time	5
	1.5 seconds	The balls are the same and have same acceleration	1
		Comparing distances between balls on track B	1
Q7. Distance time graph 3	Points A and E	At point A objects have traveled zero distance at $t = 0$ seconds, at point E objects are at same position at same time	2
	Point E	Objects traveled same distance in same time	3
		That is the point where the lines cross	2
Q10. Skier on slope	$B > C = A$	Steepness of slope influences speed	1
	$B > C > A$		1
	$A > B > C$	Steepness of slopes influences speed, kinetic energy and potential energy	2
		Steepness of slope directly related to change in potential energy	1
	$B > C > A$	Relates slope, height and width of segment to change in potential energy	1

relative to the origin, but not a zero speed. So the incorrect answers we observed on problems 3 and 7 align well with this documented student difficulty. Viennot [1,34] also investigated student difficulties with force and motion. She surveyed about 2000 university and high school students in France, Belgium, and Britain and found that students often attempted to account for differences present in a diagram that may or may not be related to the problem solution. This is consistent with our findings in problems 1 and 10. In problem 1, tracks A and B are different, though one only needs to notice that the initial and final heights are the same, so the final speeds will be the same. Students who answered incorrectly in our study discussed the differences between the tracks to explain their answers. On problem 10, one needs to notice that the heights of each slope are the same. Those who answered incorrectly in our study primarily reasoned using the fact that the slopes were changing.

In sum, there was strong agreement between our interview findings and documented student difficulties in the literature. This gave us confidence that the definitions of novicelike areas of interest, for each physics problem, do indeed represent the most common novicelike answers of the larger population of introductory physics students.

III. STUDY 2: DETERMINING DIFFERENCES IN VISUAL SELECTIVE ATTENTION BASED ON CORRECTNESS OF PROBLEM SOLUTION

A. Study 2: Methodology

1. Participants

There were 24 participants in the study (three females, two were graduate students and one was a psychology student) with two different levels of experience in physics. Ten participants were first-year through fifth-year Ph.D. students in physics who had either taught an introductory physics course or had been a teaching assistant for an introductory physics lab. One participant was a postdoctoral candidate in physics who had received his Ph.D. within the last two years and had teaching experience. Thirteen participants were enrolled in an introductory psychology course and had taken at least one physics course in high school, though some had also taken an introductory physics course at the university level. The Ph.D. students and postdoctoral candidate participated as volunteers and the psychology students received course credit for their participation. Because we were looking to compare those who answered the physics problems correctly to those who answered incorrectly, we selected participants with a broad range of experience. We expected that the Ph.D. students would answer correctly, while the psychology students might answer incorrectly, though we knew that this might not always be the case since there is a wide distribution of expertise among introductory physics students and physics graduate students [35].

2. Materials

The materials consisted of the six multiple-choice introductory physics problems analyzed in study 1 (see the Appendix).

3. 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 [36], 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.

4. Procedure

Each participant took part in an individual session lasting 20–40 min. At the beginning of the session, participants were given a short explanation of what to expect in the study. After calibrating the eye-tracking system, if the validation's mean error was $\leq 0.50°$ of visual angle, the experiment began, otherwise the calibration and validation was repeated until successful. Next, the participant was instructed to silently answer 10 multiple-choice questions while their eye movements were recorded. Participants indicated their answer to each question using number keys on the keyboard. Between questions, a calibration drift correction procedure was done to ensure proper calibration throughout the experiment. This procedure required the participant to fixate on a small white dot in the middle of a gray screen and press a key. Pressing the key caused the screen to advance to the next problem when the participant's fixation was within a predefined area around the white dot. Finally, each participant was asked to provide a cued verbal retrospective report [37] for which they were shown a replay of their eye movements on each problem and asked to explain their thought processes (either after watching the replay of their eye movements or concurrently while watching them). This method has been found to produce more in-depth explanations than without viewing one's eye movements. If a participant's explanation was unclear, they were asked follow-up questions. Participants were given unlimited time to answer the questions and provide retrospective verbal reports. Verbal explanations and gestures were recorded with a Flip video camcorder.

B. Study 2: Analysis

To analyze participants' eye fixations, we defined areas of interest (AOIs) for specified areas of each diagram. These AOIs were used to determine the total fixation time (i.e., the total amount of time the participant spent fixating on a given AOI). There were three different types of AOIs identified for each physics problem analyzed in study 1. These types were thematically relevant AOIs, perceptually salient AOIs, and novicelike AOIs. The definition for the thematically relevant AOI came from three independent raters, one physics professor and two Ph.D. students in physics, who indicated, on each of the problems, the area which contained visual information necessary to answer the problem. The definition for the perceptually salient AOI in each problem was determined using an implementation of the Itti, Koch, and Niebur saliency map algorithm in MATLAB [38]. This MATLAB toolbox produced a heat map representation of relative saliency over the entire diagram for each problem (see Fig. 1). The area on the diagram with the highest rating of saliency was used to define the perceptually salient AOI. If there were several portions of the diagram with the highest level of perceptual salience, according to the salience map, then all of these areas were used when defining the perceptually salient AOI.

The novicelike AOI was defined based on the interviews described above in study 1. Figure 2 shows the thematically relevant, novicelike and perceptually salient areas of the problem whose heat map is shown in Fig. 1.

The areas of the diagram referred to by the majority of the interviewees from study 1 who answered the problem

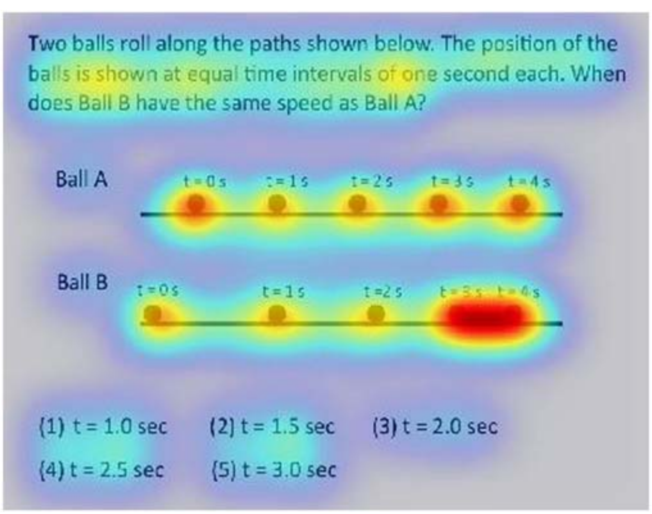


FIG. 1 (color online). Heat map of perceptual salience created using the Itti, Koch, and Niebur salience algorithm. Red indicates area of highest perceptual salience.

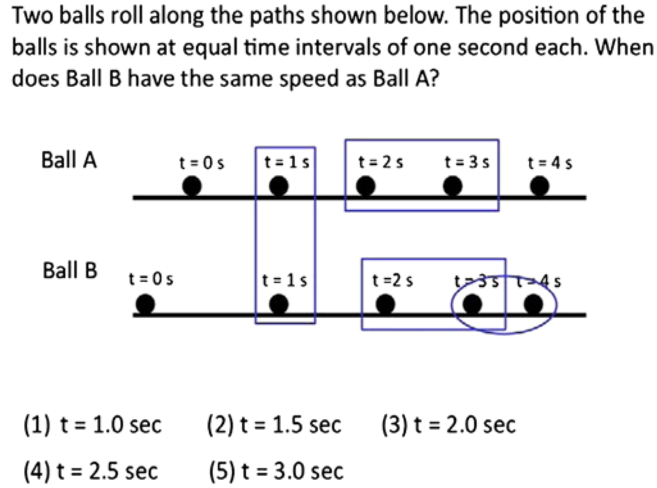


FIG. 2 (color online). Thematically relevant AOI is the distance between balls at 2–3 sec. Novicelike AOI is when the balls are at the same position, at 1 sec. Perceptually salient AOI is oval around ball B at 3 and 4 sec.

incorrectly were defined as the novicelike AOI for each of the problems. These areas are listed in Table II.

These thematically relevant, perceptually salient, and novicelike AOIs were applied to the problems analyzed in study 1. Additionally, an AOI containing the entire diagram was applied to each of the problems. The total amount of time each participant spent fixating on each AOI was determined (total fixation time), as well as the total time spent looking at the entire diagram. To account for differences in total viewing time on each problem, the percentage of time spent in each respective AOI was determined by dividing the total viewing time, for each participant, in a specified AOI by the total time spent viewing the entire diagram [39]. The percentage of time spent in each type of interest area was compared between students who answered the problem correctly and those who answered incorrectly for the entire problem set. There were a few instances where the eye-movement data file were corrupted for a participant on a single problem. In this case, the participant's data were not included in the analysis.

We were also interested in determining if perceptual salience played a greater role in influencing eye

TABLE II. Novicelike AOIs defined based on the most common incorrect student responses in study 1.

Problem	Novicelike AOI
1	Roller coaster tracks
2	Point where graph crosses x-axis
3	Origin of graph
4	Point where balls A and B line up spatially
7	Point where graphs of two objects cross
10	Slopes A, B and C

movements in the first two seconds of viewing the problem diagram. To do this, we determined the first time the participant's eye left the problem statement to look elsewhere. Applying the same AOIs described previously, we selected 2 sec of fixation data immediately following the transition from reading the problem statement to looking elsewhere in the problem. It should be noted that not all participants read the problem statement, viewed the diagram, and then the answer choices. Some participants looked from the problem statement to the diagram very briefly and then continued reading and some went from the problem statement to the answer choices. Thus, the first 2 sec of fixation data represents many different patterns of viewing. We then converted the fixation time from the first two seconds to a percentage and compared the percentage of time spent in each type of interest area between students who answered the problems correctly versus those who answered incorrectly.

C. Study 2: Results and discussion

Mixed factorial 2×6 ANOVAs (analysis of variance) with *proportion of time in each AOI type* as the dependent variable and *problem number* and *correctness of answer* as independent variables were conducted for all three AOI types. Results for the full trial period are reported in Table III. Results for the first 2 sec of viewing the diagram are reported in Table V.

1. Full trial period

For the full trial period, we found a significant main effect for correctness of answer as well as for problem number for all three AOI types. We were looking for a main effect of correctness, as this would indicate there are differences in percentage of time spent in an AOI between those who answered correctly and those who answered incorrectly. The main effect of correctness addresses our research questions and will be further analyzed below. The

TABLE III. Results of mixed factorial ANOVA for all three AOI types for full problem period. * indicates an interaction of the effects of problem number and correctness of answer.

	Thematically relevant AOI		Novicelike AOI		Perceptually salient AOI	
Effect	F	p	F	p	F	p
Problem no.	$F(5, 128) = 8.9$	<0.001	$F(5, 128) = 14.1$	<0.001	$F(5, 128) = 18.5$	<0.001
Correctness of answer	$F(1, 128) = 48.8$	<0.001	$F(1, 28) = 34.0$	<0.001	$F(1, 128 = 26.3$	<0.001
Problem no. * correctness of answer	$F(5, 128) = 0.88$	0.500	$F(5, 128) = 0.58$	0.716	$F(5, 128) = 4.6$	0.001

TABLE IV. Mean percentage time spent (± standard error) and results of one-way ANOVA during entire problem period for thematically relevant, novicelike, and perceptually salient AOIs for participants who answered the question correctly or incorrectly.

AOI Type	Problem	Answered correctly	Answered incorrectly	F	p
Thematically relevant	1	46.6 (±5.5) ($n = 11$)	33.2 (±5.7) ($n = 11$)	$F(1, 20) = 2.9$	0.107
	2[a]	24.4 (±2.9) ($n = 13$)	11.6 (±3.3) ($n = 10$)	$F(1, 21) = 8.6$	0.008
	3[a]	28.5 (±4.1) ($n = 18$)	8.9 (±2.3) ($n = 6$)	$F\ (1, 22) = 7.1$	0.014
	4[a]	49.8 (±3.9) ($n = 14$)	25.5 (±4.1) ($n = 9$)	$F(1, 21) = 17.5$	<0.001
	7[a]	36.7 (±5.5) ($n = 15$)	10.3 (±2.1) ($n = 9$)	$F(1, 22) = 13.1$	0.002
	10[a]	29.0 (±5.0) ($n = 11$)	15.1 (±2.7) ($n = 13$)	$F(1, 22) = 6.6$	0.018
Novicelike	1[a]	22.3 (±4.5) ($n = 11$)	43.5 (±7.3) ($n = 11$)	$F(1, 20) = 6.0$	0.020
	2[a]	12.7 (±3.3) ($n = 13$)	27.2 (±4.8) ($n = 10$)	$F(1, 21) = 6.6$	0.018
	3[a]	19.8 (±3.7) ($n = 18$)	39.4 (±5.4) ($n = 6$)	$F(1, 22) = 7.5$	0.012
	4	18.1 (±2.5) ($n = 14$)	26.8 (±3.9) ($n = 9$)	$F(1, 21) = 4.0$	0.058
	7[a]	12.6 (±2.6) ($n = 15$)	25.0 (±6.0) ($n = 9$)	$F(1, 22) = 4.7$	0.041
	10[a]	41.2 (±6.6) ($n = 11$)	62.2 (±5.1) ($n = 13$)	$F(1, 22) = 6.5$	0.018
Perceptually salient	1	6.6 (±1.9) ($n = 11$)	13.0 (±2.5) ($n = 11$)	$F(1, 20) = 4.1$	0.056
	2	19.3 (±4.1) ($n = 13$)	28.2 (±4.9) ($n = 10$)	$F(1, 21) = 1.9$	0.179
	3[a]	9.5 (±2.2) ($n = 18$)	30.5 (±4.6) ($n = 6$)	$F(1, 22) = 20.1$	0.001
	4	11.9 (±1.7) ($n = 14$)	9.0 (±2.2) ($n = 9$)	$F(1, 22) = 1.1$	0.316
	7[a]	19.1 (±3.0) ($n = 15$)	39.5 (±5.6) ($n = 9$)	$F(1, 22) = 12.3$	0.002
	10	4.2 (±1.1) ($n = 11$)	6.3 (±1.6) ($n = 13$)	$F(1, 22) = 1.1$	0.305

[a]Indicates significant difference, $p < 0.05$.

TABLE V. Results of mixed factorial ANOVA for all three AOI types for first 2 sec of viewing the diagram. * indicates an interaction of the effects of problem number and correctness of answer.

Effect	Thematically relevant AOI F	p	Novicelike AOI F	p	Perceptually salient AOI F	p
Problem no.	$F(5, 128) = 2.10$	0.069	$F(5, 128) = 6.72$	<0.001	$F(5, 128) = 10.7$	<0.001
Correctness of answer	$F(1, 128) = 0.495$	0.483	$F(1, 28) = 2.03$	0.156	$F(1, 128 = 2.47$	0.119
Problem no. * correctness of answer	$F(5, 128) = 2.30$	0.048	$F(5, 128) = 0.036$	0.999	$F(5, 128) = 0.671$	0.646

main effect of problem number indicates there is at least one difference in proportion of time in each AOI type between different problems. We were not interested in how the proportion of time spent fixating varies between problems, as this is not relevant to our research questions, so the effect of problem number will not be further analyzed. We found a significant interaction between problem number and correctness of answer in the perceptually salient AOI. This means the relationship between correctness and time spent in the perceptually salient area is different across problems. This interaction is not relevant to our research question and will not be further investigated.

The main effect of correctness was further analyzed for each of the six different problems using a one-way ANOVA with *percentage of time* for all three AOI types as the dependent variable and *correctness of answer* as the independent variable. Results of one-way ANOVAs for each type of AOI for the full trial period are reported in Table IV. Mean percentage of fixation time and standard error for the correct and incorrect responders for each question are also shown in Table IV. The footnote indicates a significant difference at the $\alpha = 0.05$ level.

We found that on five out of six problems used in study 2, those who answered the problem correctly spent a higher percentage of total viewing time fixating on thematically relevant areas in the problem diagram (Table IV). Those who answered correctly likely had the domain knowledge needed to solve each problem, and therefore spent more time viewing the relevant areas in each diagram. This result is consistent with previous findings where those with high levels of domain knowledge in a discipline, such as identifying fish locomotion [19], art [20], and chess [21], spend more time looking at areas of diagrams and pictures relevant to a task. Our finding is evidence for top-down processes playing a key role in guiding visual attention when solving physics problems correctly.

We also found that on five out of six problems, those who answered the problem incorrectly spent a higher percentage of total viewing time looking at areas of the

TABLE VI. Mean percentage fixation time spent (± standard error) during the first 2 sec after leaving the problem statement for thematically relevant, novicelike, and perceptually salient AOIs for participants who answered the question correctly or incorrectly

AOI Type	Problem	Answered correctly	Answered incorrectly
Thematically relevant	1	13.5 (±6.8) ($n = 11$)	31.1 (±6.3) ($n = 11$)
	2	10.9 (±2.9) ($n = 13$)	8.6 (±3.4) ($n = 10$)
	3	9.7 (±3.1) ($n = 18$)	9.7 (±5.0) ($n = 6$)
	4	26.5 (±5.0) ($n = 14$)	11.9 (±6.5) ($n = 9$)
	7	17.6 (±6.5) ($n = 15$)	17.6 (±2.4) ($n = 9$)
	10	13.0 (±4.2) ($n = 11$)	9.7 (±4.1) ($n = 13$)
Novicelike	1	2.6 (±1.4) ($n = 11$)	9.4 (±2.7) ($n = 11$)
	2	9.4 (±4.3) ($n = 13$)	13.0 (±6.2) ($n = 10$)
	3	12.1 (±3.2) ($n = 18$)	15.2 (±9.0) ($n = 6$)
	4	17.6 (±4.2) ($n = 14$)	22.3 (±6.1) ($n = 9$)
	7	17.4 (±4.7) ($n = 15$)	20.8 (±7.6) ($n = 9$)
	10	30.7 (±7.0) ($n = 11$)	34.6 (±5.2) ($n = 13$)
Perceptually salient	1	0.7 (±0.7) ($n = 11$)	2.5 (±1.8) ($n = 11$)
	2	10.8 (±3.2) ($n = 13$)	21.8 (±8.1) ($n = 10$)
	3	8.3 (±2.7) ($n = 18$)	9.0 (±4.1) ($n = 6$)
	4	2.5 (±2.5) ($n = 14$)	2.3 (±2.3) ($n = 9$)
	7	23.2 (±4.4) ($n = 15$)	32.5 (±8.0) ($n = 9$)
	10	10.9 (±4.9) ($n = 11$)	11.6 (±3.4) ($n = 13$)

diagram consistent with a novicelike response (Table IV). Furthermore, on the one problem that did not quite reach statistical significance ($p = 0.058$) the effect was in the same direction as the other five problems. These novicelike AOIs were determined through individual interviews described in study 1, and were consistent with the physics education literature describing common student misconceptions. Importantly, the finding that incorrect solvers spent more time fixating on novicelike areas is evidence for their visual attention being guided by top-down processes. However, instead of attention being guided by scientifically correct domain knowledge, incorrect problem solvers' attention was guided by novicelike misconceptions. Thus, when solving physics problems, top-down processing plays a key role in guiding visual selective attention either to thematically relevant areas, or novicelike areas, depending upon the scientific correctness of a student's physics knowledge.

Concerning the effects of bottom-up processes in guiding attention during physics problem solving, we found that those who answered incorrectly spent more time in perceptually salient areas during the full problem period on only two of the six problems, namely, problems 3 and 7. Nevertheless, for five of the six problems the effect was in the predicted direction, such that incorrect problem solvers spent a higher percentage of total time fixating on the perceptually salient AOIs than the correct problem solvers. However, four of those effects were not statistically significant. A likely explanation for this result is that in problems 3 and 7, the perceptually salient AOI partially or completely overlapped with the novicelike AOI (Figs. 3 and 4), which was not the case for the other four problems. We have already shown that those who answered the problem incorrectly spent significantly more time fixating on the novicelike AOIs on problems 3 and 7 than those who answered the problem correctly. So the significant result for problems 3 and 7 for the perceptually salient AOI is likely due to this AOI overlapping with the novicelike AOI. This result also seems to indicate that attending to the perceptually salient area is not necessarily a good predictor of correctness. These results appear to be consistent with a study of change blindness that found that problem solvers seldom notice changes in color, even though color is most perceptually salient [40]. Thus, when considering the full time period of problem solving, perceptual salience appears to have played a minimal role in guiding the attention of incorrect physics problem solvers. Nevertheless, previous vision research has suggested that the effects of bottom-up perceptual salience on eye movements are limited to the first 2 sec of viewing a stimulus [7]. Thus, this seemingly null result could be argued to have resulted from diluting the effect of saliency by including eye-movement data from the entire duration of the trial, rather than only the first 2 sec. We therefore reanalyzed the data including only the first 2 sec that participants spent viewing the diagram.

2. First 2 sec after leaving problem statement

To reanalyze the data including only the first 2 sec of viewing a diagram, we completed a mixed factorial 2×6 ANOVA with *proportion of time in each AOI type* as the dependent variable and *problem number* and *correctness of answer* as independent variables for all three AOI types for the first 2 sec of viewing the diagram. These results are reported in Table V. We were looking for a main effect of correctness, as this would indicate there are differences in percentage of time spent in an AOI between those who

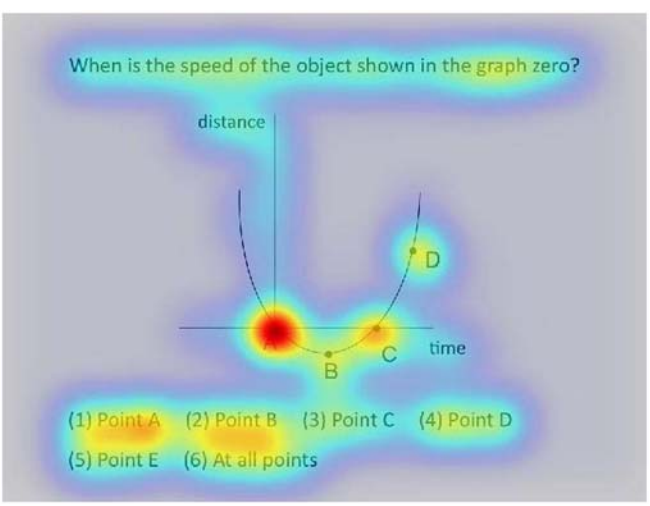


FIG. 3 (color online). Itti, Niebur, and Koch saliency map for problem 3. The perceptually salient AOI overlapped the novicelike AOI, which was at the origin of the graph.

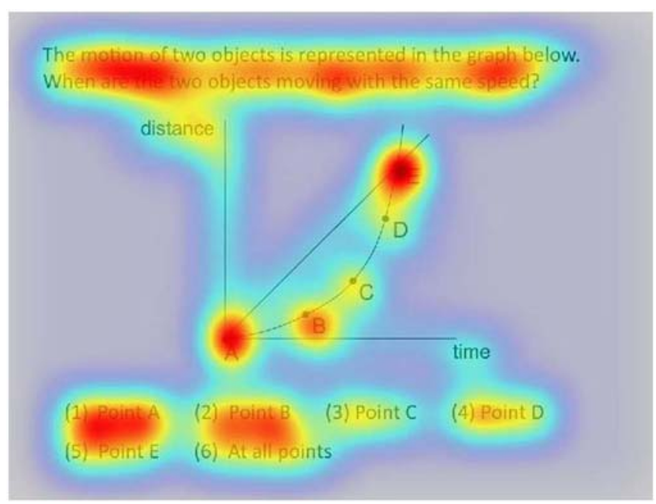


FIG. 4 (color online). Itti, Niebur, and Koch saliency map for problem 7. The perceptually salient AOI partially overlapped with the novicelike AOI, which was at the point where the two lines cross.

answered correctly and those who answered incorrectly. For the first 2 sec after leaving the problem statement, we found no main effect for correctness of answer for any of the AOI types. So, there are no significant differences in proportion of time spent fixating in the AOI types between those who answered correctly and those who answered incorrectly for any of the problems and no further analysis was conducted.

We did find a main effect for problem number for the novicelike and perceptually salient AOIs. This means for each of these AOIs, there is at least one difference in proportion of time between the different problems when considering the data for all participants. We were not interested in how the proportion of time spent fixating varies between problems, as this is not relevant to our research questions. We also found a significant interaction between problem number and correctness of answer in the thematically relevant AOI. This means the relationship between correctness and time spent in the thematically relevant area is different across problems. This interaction also does not address our research questions, and is not analyzed further.

The mean percentage of fixation time spent looking in thematically relevant, novicelike, and perceptually salient AOIs for participants who answered the question correctly and incorrectly for the first 2 sec of viewing the diagram is displayed in Table VI. As mentioned above, there are no significant differences between the percentage of fixation time for correct and incorrect solvers shown in this table.

The reanalysis of the data for the first 2 sec of viewing the diagram found no statistically significant differences between correct and incorrect solvers on any of the problems for the perceptually salient AOI. Indeed, there were no statistically significant differences between correct and incorrect solvers in time spent in the thematically relevant or novicelike AOIs. In sum, we found no support for the hypothesis that perceptual salience influences visual selective attention more for incorrect problem solvers during the first 2 sec of diagram viewing. This result is consistent with previous studies [e.g., [12,33]] that have shown that top-down influences on visual attention tend to dominate bottom-up influences when a viewer is given a specific goal or task. Nevertheless, such null results for the effects of bottom-up saliency on visual attention are consistent with our own results, which considered both the full problem solving time period and only the first 2 sec, and found little if any effects.

However, before completely rejecting the hypothesis that bottom-up saliency affects attentional selection during physics problem solving, we must consider two observations that provide partial support for it. First, it may be that the early effect of perceptual salience on eye movements was present; however, the data lacked sufficient statistical power to detect it. Some support for this explanation is shown by comparing the mean difference for the correct versus incorrect problem solvers for the perceptually salient AOIs for the first 2 sec of viewing the diagram (Table VI). Specifically, the percentage of time spent looking in the perceptually salient AOI is higher for incorrect solvers than correct problem solvers on five of the six problems, though not statistically significantly so. Thus, it is possible that a larger study with more observations might show this effect to be statistically significant. Secondly, the perceptual salience model proposed by Itti and Koch [6] predicted that early in scene viewing eye movements are more influenced by bottom-up perceptual information than top-down knowledge. Therefore, the saliency model would predict that early in viewing a physics problem, correct and incorrect problem solvers would not have had sufficient amount of time to apply their (correct or incorrect) top-down knowledge to guide their attention to thematically relevant or novicelike areas of the diagram. If so, during the first 2 sec of viewing the diagram, there should be no difference between correct and incorrect problem solvers' percentage of total fixation time in either the thematically relevant or novicelike AOIs. The data support this hypothesis, which shows that there is no significant difference in viewing time for thematically relevant AOIs between correct and incorrect problem solvers. In sum, the data showed essentially no influence by top-down domain knowledge during the first 2 sec of diagram viewing, though such effects were statistically significant later in time, when considering the full problem solving time period. Thus, based on the above two observations, we must withhold complete rejection of the hypothesis that bottom-up salience affects the visual selective attention of incorrect physics problem solvers. Even so, such an interpretation of the data should be made cautiously since it is based on null effects. Future studies will be required in order to explicitly test the effects of bottom-up and top-down information on early and late visual selective attention processes in eye movements.

IV. FUTURE DIRECTIONS AND LIMITATIONS

Overall, these findings motivate the use of visual cues to redirect individuals' attention to relevant portions of the diagrams and potentially influence the way they reason about these questions. The problems used in study 2 all contained AOIs consistent with novicelike misconceptions. Those who answered incorrectly spent more time looking at these novicelike AOIs. One way to help incorrect problem solvers pay attention to the relevant areas of a problem diagram is to overlay dynamic visual cues on it. These cues should have very high perceptual salience, perhaps using color or motion cues, in order to reliably attract the problem solver's attention. Visual cues have been found to facilitate comprehension in several contexts, such as insight problems [41] and educational animations [42]. Grant and Spivey [41] studied an insight problem where one must determine how to use lasers to kill an inoperable tumor without harming the healthy tissue surrounding the tumor. To solve this problem, one must use several weak lasers at different spatial positions surrounding the tumor, so as not to damage the healthy tissue, but at the point at which the lasers converged, it would have a high enough intensity to kill the tumor cells. They found more participants correctly solved the problem when the task-relevant information in the diagram, namely, the healthy tissue, was made more perceptually salient by increasing and decreasing its width. Many studies using visual cues to focus viewers' attention on relevant information have been conducted using animations. In one of these studies, de Koning *et. al.* [43] used a spotlight cueing technique to focus learners' attention on the valves of the heart in an animation of the cardiovascular system. He found those who viewed the animation with the cues had higher comprehension and transfer scores on post-test questions about heart valves and the cardiovascular system. These examples and many others suggest that visual cues overlaid on physics problems such as those in the current study may help students to ignore the novicelike AOIs of diagrams and instead pay attention to the thematically relevant AOIs in order to reason in a scientifically correct manner about the problem.

This study describes only a limited number of introductory physics problems. To increase the generalizability of our conclusions, the study should be repeated with more problems from other areas of introductory physics and with students having a wider range of prior knowledge of physics. Additionally, the study could be improved by using a larger number of participants, which would increase the statistical power of the study and enable us to more thoroughly test the perceptual saliency hypothesis. Furthermore, the conclusions we have drawn about the influence of perceptual salience on visual attention must remain tentative as we only used one computational model of visual salience (albeit the most famous one) and in some of the problems the perceptually salient AOI overlapped the novicelike AOI. In future work several different models of saliency will be used and only problems where the perceptually salient AOI does not overlap any other AOI will be used. Thus, future research should include similar studies using the suggestions discussed above as well as studying the influence that visual cues overlaid on such problems have on students' visual attention and the correctness of their problem answers.

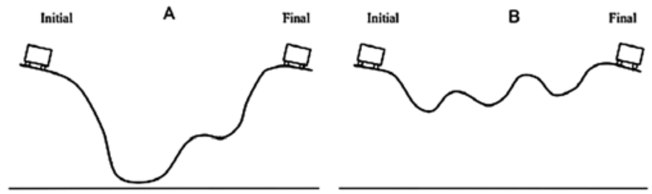
ACKNOWLEDGMENTS

We gratefully acknowledge the contributions of Dr. Elizabeth Gire for her participation in discussions of the design of the questions used in this study.

APPENDIX

Various problems used in studies 1 and 2 are shown in Figs. 5–10.

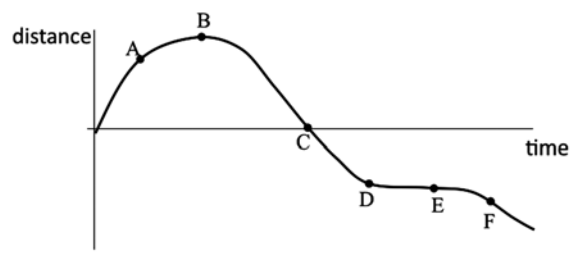
If frictional effects can be ignored, how does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?



(1) The cart A is moving faster at the final position
(2) The cart B is moving faster at the final position
(3) Carts A and B have the same speed at the final position
(4) There is not enough information to decide

FIG. 5. Problem 1 used in studies 1 and 2.

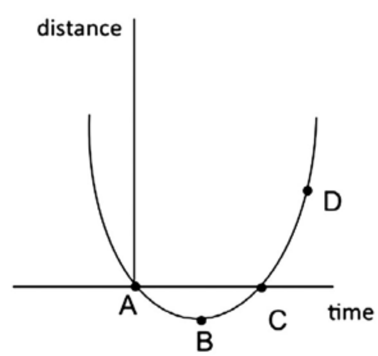
At which point on the graph is the object turning around (moving away then coming back)?



(1) A (2) B (3) C (4) D (5) E (6) F

FIG. 6. Problem 2 used in studies 1 and 2.

When is the speed of the object shown in the graph zero?



(1) Point A (2) Point B (3) Point C (4) Point D
(5) Point E (6) At all points

FIG. 7. Problem 3 used in studies 1 and 2.

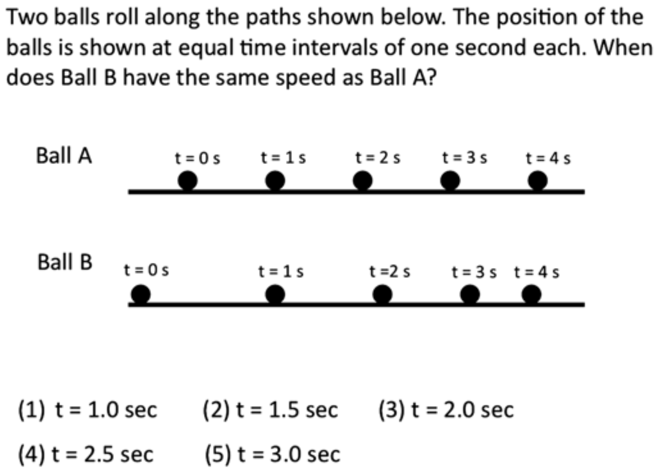


FIG. 8. Problem 4 used in studies 1 and 2.

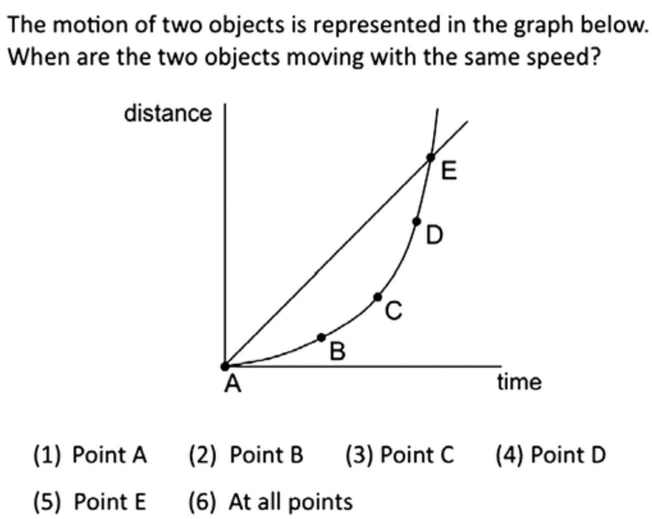


FIG. 9. Problem 7 used in studies 1 and 2.

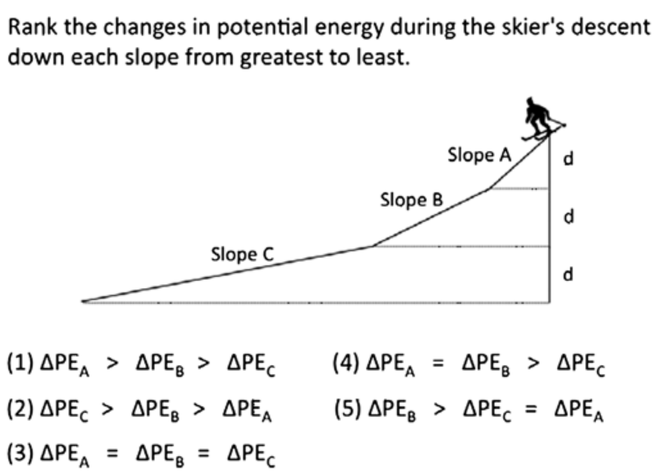


FIG. 10. Problem 10 used in studies 1 and 2.

[1] L. C. McDermott, Research on conceptual understanding in mechanics, Phys. Today **37**, No. 7, 24 (1984).

[2] L. C. McDermott, M. L. Rosenquist, and E. H. van Zee, Student difficulties in connecting graphs and physics: Examples from kinematics, Am. J. Phys. **55**, 503 (1987).

[3] M. Corbetta *et al.*, A common network of functional areas for attention and eye movements, Neuron **21**, 761 (1998).

[4] A. C. Nobre *et al.*, Covert visual spatial orienting and saccades: Overlapping, NeuroImage **11**, 210 (2000).

[5] L. Itti, C. Koch, and E. Niebur, A model of saliency-based visual attention for rapid scene analysis, IEEE Trans. Pattern Anal. Mach. Intell. **20**, 1254 (1998).

[6] L. Itti and C. Koch, A saliency-based search mechanism for overt and covert shifts of visual attention, Vision Res. **40**, 1489 (2000).

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

[8] S. E. Boehnke and D. P. Munoz, On the importance of the transient visual response in the superior colliculus, Curr. Opin. Neurobiol. **18**, 544 (2008).

[9] A. A. Kustov and D. L. Robinson, Shared neural control of attentional shifts and eye movements, Nature (London) **384**, 74 (1996).

[10] D. Parkhurst, K. Law, and E. Niebur, Modeling the role of salience in the allocation of overt visual attention, Vision Res. **42**, 107 (2002).

[11] D. J. Parkhurst and E. Niebur, Scene content selected by active vision, Spatial Vis. **16**, 125 (2003).

[12] W. Einhäuser, U. Rutishauser, and C. Koch, Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli, J. Vis. **8**, 2 (2008).

[13] T. Foulsham and G. Underwood, How does the purpose of inspection influence the potency of visual salience in scene perception?, Perception **36**, 1123 (2007).

[14] J. M. Henderson *et al.*, in *Eye Movements: A Window on Mind and Brain*, edited by R. V. Gompel *et al.* (Elsevier, Amsterdam, 2007), p. 537.

[15] A. F. Heckler, in *Psychology of Learning and Motivation: Cognition In Education*, edited by J. P. Mestre and B. H. Ross (Academic, Oxford, England, 2011).

[16] M. Hegarty, M. S. Canham, and S. I. Fabrikant, Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task, J. Exp. Psychol. Learn. Mem. Cogn. **36**, 37 (2010).

[17] D. L. Sheinberg and G. J. Zelinsky, in *Perception and Cognition: Advances in Eye Movement Research*, edited by G. d'Ydewalle and J. V. Rensbergen (North-Holland, Amsterdam, 1993), p. 333.

[18] S. Van der Stigchel, M. Meeter, and J. Theeuwes, Eye movement trajectories and what they tell us, Neurosci. Biobehav. Rev. **30**, 666 (2006).

[19] H. Jarodzka, K. Scheiter, P. Gerjets, and T. van Gog, In the eyes of the beholder: How experts and novices interpret dynamic stimuli, Learn. Instr. **20**, 146 (2010).

[20] J. R. Antes and A. F. Kristjanson, Discriminating artists from nonartists by their eye-fixation patterns, Percept. Mot. Skills **73**, 893 (1991).

[21] N. Charness, E. M. Reingold, M. Pomplun, and D. M. Stampe, The perceptual aspect of skilled performance in

chess: Evidence from eye movements, Mem. Cogn. **29**, 1146 (2001).

[22] 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).

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

[24] R. H. Tai, J. F. Loehr, and F. J. Brigham, An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments, Int. J. Res. Meth. Educ. **29**, 185 (2006).

[25] J. L. Docktor and J. P. Mestre, *A Synthesis of Discipline-Based Education Research in Physics*, White paper commissioned by the National Research Council, 2011, http://www7.nationalacademies.org/bose/DBER_Docktor_October_Paper.pdf.

[26] E. Etkina, J. Mestre, and A. O'Donnell, in *The Cognitive Revolution in Educational Psychology*, edited by J. M. Royer (Information Age Publishing, Greenwich, CT, 2005), p. 119.

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

[28] M. T. H. Chi, in *Cognitive Models of Science: Minnesota Studies in the Philosophy of Science*, edited by R. Giere (University of Minnesota Press, Minneapolis, MN, 1992), p. 129.

[29] R. K. Lowe, Extracting information from an animation during complex visual learning, Eur. J. Psychol. Educ. **14**, 225 (1999).

[30] L. C. McDermott and E. F. Redish, Resource letter: PER-1: Physics education research, Am. J. Phys. **67**, 755 (1999).

[31] F. Marton, Phenomenography-research approach to investigating different understandings of reality, J. Thought **21**, 28 (1986).

[32] R. J. Beichner, Testing student interpretation of kinematics graphs, Am. J. Phys. **62**, 750 (1994).

[33] D. E. Trowbridge and L. C. McDermott, Investigation of student understanding of the concept of velocity in one dimension, Am. J. Phys. **48**, 1020 (1980).

[34] L. Viennot, Spontaneous reasoning in elementary dynamics, Eur. J. Sci. Educ. **1**, 205 (1979).

[35] A. Mason and C. Singh, Assessing expertise in introductory physics using categorization task, Phys. Rev. ST Phys. Educ. Res. **7**, 020110 (2011).

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

[37] T. van Gog, F. Paas, J. van Merriënboer, J. G. Jeroen, and P. Witte, Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting, J. Exp. Psychol. Appl. **11**, 237 (2005).

[38] J. Harel, C. Koch, and P. Perona, Graph-based visual saliency, in *Advances in Neural Information Processing Systems 19*, edited by B. Schölkopf, J. Platt, and T. Hoffman (MIT Press, Cambridge, MA, 2007). pp. 545–552.

[39] J. Cohen *et al.*, *Applied Multiple Regression–Correlation Analysis for the Behavioral Sciences* (Lawrence Erlbaum Associates, Mahwah, NJ, 2003).

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

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

[42] 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. Rev. **21**, 113 (2009).

[43] B. B. de Koning, H. K. Tabbers, R. M. J. P. Rikers, and F. Paas, Attention cueing as a means to enhance learning from an animation, Appl. Cogn. Psychol. **21**, 731 (2007).

Guiding Attention on Physics Problems Using Visual Cues Modeled After Experts' Eye Movements

Adrian Madsen, Adam Larson, Amy Rouinfar, Allison Coy, Lester Loschky, and N. Sanjay Rebello

Kansas State University

Please address all correspondence to:
Adrian Madsen
Kansas State University
Department of Physics
403 Cardwell Hall
Manhattan, KS 66502
(785) 532-1612
adrianc@phys.ksu.edu

Paper presented at the
Annual Meeting of the National Association for Research in Science Teaching
Indianapolis, IN, March 25-28, 2012

Abstract

To maximize learning one must ensure most of the learner's cognitive resources are spent on relevant tasks and avoid instructional environments that facilitate focusing on the irrelevant. To help novice learners focus on elements relevant for learning it may be helpful to give them insight into the way experts allocate their visual attention, for example, by using visual cues. To design appropriate cues, we must first understand how those who solve problems correctly allocate their visual attention by recording their eye movements. In Study 1, we record eye movements of introductory and graduate physics students while answering conceptual physics problems containing a diagram to determine differences in visual attention. We use the eye movements of those who answer these questions correctly to design visual cues for Study 2. In Study 2, we overlay these dynamic visual cues onto the same physics problems and find evidence of increased conceptual understanding in novices who view the problems overlaid with cues.

Guiding Attention on Physics Problems Using Visual Cues Modeled After Experts' Eye Movements

Introduction

To maximize learning one must ensure most of the learner's cognitive resources are spent on relevant tasks and avoid instructional environments that facilitate focusing on the irrelevant. Mayer's cognitive theory of multimedia learning (2001) 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 one's working memory. Often learners are faced with learning environments that impose a high cognitive load and max out the limited capacity of working memory. To help alleviate this problem, visual cues can be used. de Koning, Tabers, Rikers and Paas (2009) has devised a framework to describe three specific functions of cueing which include guiding learners' attention to essential information, emphasizing organization, and making the relations between elements more salient to foster their integration. There exists a large body of research which utilizes visual cues in a variety of contexts to increase learning in animations and static problems (de Koning, Tabers, Rikers & Paas, 2007, Grant & Spivey, 2003, Kriz & Hegarty, 2007, Mautone & Mayer, 2001, Ozcelik, Arslan-Ari & Cagiltay, 2007).

The first function of cueing is especially interesting to us as it works to ensure that cognitive resources are spent appropriately by helping the learner to focus primarily on relevant elements. This frees up working memory and allows for real learning to occur. In physics classes, students are often faced with diagrams or animations of the real world which contain elements relevant to the task at hand as well as elements which are present in the real world, but not useful for answering the given question. To help students select the relevant and ignore the irrelevant, we study the use of dynamic visual cues overlaid on static physics problems containing a diagram. But before we can design cues to guide learners' attention, we must first understand where learners who answer questions incorrectly and correctly look when viewing physics problems containing diagrams. We need this information to determine what information in the diagrams is relevant for the correct answer and what should be ignored.

Research in many disciplines has used eye-tracking technology to investigate the differences in the visual attention of experts and novices. It has been observed that experts in a domain attend to task-relevant portions of a diagram more than novices in that domain. Jarodzka, Scheiter, Gerjets, & van Gog (2009) studied the visual attention of both novices and experts when viewing videos of unfamiliar fish swimming and classifying the type of locomotion. The authors found that experts spent significantly more time fixating on relevant areas of the video than biology students, who had the necessary background knowledge for differentiating types of locomotion but little practice in this classification task. The authors also found that novices spent more time than experts fixating on areas irrelevant for determining locomotion. Similar studies have measured eye movements of experts when viewing art (Antes & Kristjanson, 1991) and playing chess (Charness, Reingold, Pomplun, & Stampe, 2001), and have shown that the increased domain knowledge in these fields affects where people fixate while performing domain-relevant visual tasks. Based on these studies and other previous research, we conclude that novice and expert learners, who differ in their domain specific knowledge, allocate visual attention differently. Experts' knowledge drives them to attend to thematically relevant areas, or those portions of the diagram relevant to the task at hand, while novices lack of knowledge leaves them to view irrelevant areas of a diagram.

Here we report on two different studies. Study 1 investigates the differences in eye movements of those who answer physics problems correctly and incorrectly, where the critical information needed to answer the problem is contained in a diagram. Study 2 uses the results of Study 1 to design visual cues based on eye movements of those who answered the questions correctly. These visual cues are overlaid on the physics problems and shown to introductory physics students. We compare the eye movements and reasoning of students who saw visual cues and those who did not.

Study 1: Comparison of Eye Movements of Correct and Incorrect Solvers

Method

There were 24 participants in the study (3 females) with two different levels of experience in physics. Ten participants were first-year through fifth-year PhD students in physics who had taught an introductory physics course or been a teaching assistant for an introductory physics lab. One participant was a postdoctoral candidate in physics who had received his PhD within the last two years and had teaching experience. Thirteen participants were enrolled in an introductory psychology course and had taken at least one physics course in high school, though some had also taken an introductory physics course at the university level. The PhD students and post-doc participated as volunteers and the psychology students received course credit for their participation. As we were looking to compare those who answered the physics problems correctly to those who answered incorrectly, we selected participants with a broad range of experience. We expected that the PhD students would answer correctly, while the psychology students might answer incorrectly, though we knew that this might not always be the case since there is a wide distribution of expertise among introductory physics students and physics graduate students (Mason & Singh, 2011). The participants viewed 10 multiple-choice introductory physics problems. Participants were presented with physics problems on a computer screen viewed at a distance of 24 inches using a chin and forehead rest to minimize participants' extraneous head movements. The resolution of the computer screen was set to 1024 x 768 pixels with a refresh rate of 85 Hz. Each physics problem subtended 33.3° x 25.5° of visual angle. Eye movements were recorded with an EyeLink 1000 desktop mounted eye-tracking system (http://www.sr-research.com), 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 8,500°/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.

Each participant took part in an individual session lasting 20-40 minutes. At the beginning of the session, participants were given a short explanation of what to expect in the study. After calibrating the eye tracking system, if the validation's mean error was ≤ 0.50° of visual angle, the experiment began—otherwise the calibration and validation was repeated until successful. Next, the participant was instructed to silently answer 10 multiple-choice questions while their eye movements were recorded. Between questions, a calibration drift correction procedure was done to ensure proper calibration throughout the experiment. Participants indicated their answer to each question using number keys on the keyboard. Finally, each participant was asked to provide a cued verbal retrospective report (van Gog, Paas, van Merriënboer & Witte, 2005) for which they were shown a replay of their eye movements on each problem and asked to explain their thought processes (either after watching the replay of their eye movements or concurrently while watching them). This method has been found to produce more in-depth explanations than without viewing one's eye movements. If a participant's explanation was unclear, they were asked follow up questions. Participants were given unlimited time to answer the questions and provide retrospective verbal reports.

Analysis and Results

To analyze participants' eye movements, we created areas of interest (AOIs) which specified areas of the diagram that were used to determine the fixation time i.e., the total amount of time the participant spent looking at a given region (see Figure 1). There were two types of AOIs defined for a subset of the problems, these types being novice-like AOIs and thematically relevant AOIs. The thematically relevant AOIs are those which one would need to attend to in order to correctly answer the physics problem. The thematically relevant AOIs were defined by three independent raters, one physics professor and two physics graduate students. The AOI definitions were compared and any discrepancies were resolved through discussion. The definition for the novice-like AOIs comes from a previous interview study (Madsen, Larson, Loschky & Rebello, 2012) where 13 students (eight females) enrolled in an introductory psychology course participated in individual think aloud interviews. All of the participants had taken at least one physics course in high school, though some had taken an introductory physics course at the university level as well. The students were shown a set of 10 conceptual physics problems with a diagram and asked to explain the reasoning, which led them to their answer. The answers and reasoning of those who answered incorrectly were coded and analyzed to produce the definitions for the novice-like AOI. On four of the problems, there was not a consistent area those who answered incorrectly cited for their answer, so these four problems were not included in this analysis.

The fixation time in the AOI for each participant on each problem was determined from eye tracking data. To account for any differences in the total viewing time on each problem, the fixation time in each AOI was divided by the total viewing time for the diagram on a given problem to get the percentage of time an individual spent in a particular AOI. For each problem, the percentage of time spent in each type of interest area was compared between participants who answered the problem correctly and those who did not using a one-way ANOVA with *percentage of time* for all three AOI types as the dependent variable and *correctness of answer* as the independent variable. Mean percentage of fixation duration and standard error for the correct and incorrect responders for each question are shown in Table 1. We found that on five out of six problems analyzed, participants who correctly answered the question spent significantly more time looking at the thematically relevant areas of the diagrams than those who answered the same question incorrectly. There were also five problems where there was a significant difference between the correct and incorrect responders in the percentage of fixation time in the novice-like portions of the diagrams. Thus, we have identified several problems which are good candidates for visual cueing, as these problem diagrams contain areas which correct solvers spend more time looking at as well as areas that incorrect solvers spend more time looking at. These findings are also consistent with previous findings (Jarodzka et. al, 2009; Antes & Kristjanson, 1991; Charness et. al, 2001) that experts spend more time looking at relevant visual elements while novices spend more time looking at irrelevant visual elements. It should be noted that in our study we compared the visual attention of those who answered the problems correctly to those who answered incorrectly, where previous studies compared participants' eye movements based on expertise.

Study 2: Using Dynamic Visual Cues to Influence Reasoning

Method

Participants in the study were 55 individuals concurrently enrolled in an introductory algebra-based physics course. To ensure sufficient prerequisite knowledge, each completed a pre-test, which consisted of four open-ended questions gauging their understanding of speed and potential energy. The pre-tests 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.

Participants took part in individual sessions lasting between 30 and 60 minutes. 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 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.

Participants were presented with physics problems on a computer screen viewed at a distance of 24 inches using a chin and forehead rest to minimize participants' extraneous head movements. The resolution of the computer screen was set to 1024 x 768 pixels with a refresh rate of 85 Hz. Each physics problem subtended 33.3° x 25.5° of visual angle. Eye movements were recorded with an EyeLink 1000 desktop mounted eye-tracking system which had an accuracy of less than 0.50° of visual angle. 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.

The materials consisted of four sets of conceptual physics problems covering energy and kinematics, which were found to have significant differences in the way correct and incorrect solvers answered them in Study 1 (See Figure 1). It should be noted that these four specific problems were chosen from the six problems analyzed in Study 1 because they tested distinct concepts in physics. There were three problems in Study 1 that tested a similar physics concept, so only one of these problems was included in Study 2.

Within each problem set, there was an "initial" problem, four "similar" problems and a "transfer" problem. All problems were open-ended and contained a diagram with the necessary information needed to answer the problem. 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 and tested the same concept and contained a diagram with similar surface features. When the student answered a similar problem correctly, they saw the transfer problem. The surface features of the transfer problems were different than the initial and similar problems, though the concept tested was the same. All participants viewed the four sets of problems in the same order.

Whenever a student was ready to answer a problem, they indicated this by pressing any key on a keyboard, at which point the problem displayed on the computer would become slightly smaller in size (this was so that the student knew they had successfully pressed a key). The participant 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 pre-defined 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 group saw colored shapes overlaid on the similar problems appear four seconds after the problem was initially seen. Colored shapes were used because color is known to capture one's attention because of its high visual salience. Each colored shape appeared for 500 ms at 12 different positions in the diagram for a total cueing time of six seconds. The visual cues were designed to mimic the eye movements of those who answered the same problems correctly in Study 1. There is a large variation in eye movements from one individual to another while viewing the diagrams in these physics problems, so the visual cues could 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 thematically relevant area of interest. Similarities between participants were observed, and visual cues modeled after these patterns. Further, the cues could have remained static and simply drawn participants' attention to the relevant areas of the problem, but we hoped by modeling the way in which correct solvers viewed the thematically relevant areas and compared elements within these areas, the cues would give the participants more insight into how to correctly answer the problems.

Analysis and Findings

In order to determine if these visual cues had a positive influence on participants' answers and reasoning, we compared the number of students who answered the "initial" problem incorrectly and then changed to a correct answer and reasoning on a "similar" problem. We used a Mann-Whitney U test to compare the number of participants in the cue and no cue groups who changed to a correct answer on any of the four similar problems seen in a problem set. We found a significant difference on the roller coaster problem ($p = .002$) where six students in the cued group ($N = 18$) changed to the correct answer while zero students in the no cue group ($N = 14$)

made this change. There were no significant differences found between the cue and no cue groups in the number of changes to a correct answer on the "ball," "skier," or "graph" similar problems. It is interesting that only one of the four problems sets studied seemed to show a difference between the cue and no cue groups. There are many possible reasons for this difference. One is the design of the cues themselves. The rollercoaster problems used repetitive simple cues while the cue patterns on the other three problems were more complex. It is possible that in the short six second cueing period, the participants couldn't draw meaning from the more complex cue patterns, and thus they were ineffective. This hypothesis will be tested in future studies.

To determine if visual cueing is useful for learning beyond the problem being cued, we compared the correctness of the answer and reasoning on transfer problems between the cue group and no cue group for each problem set. Figure 3 shows the percentage of participants who answered the transfer problem correctly after answering the initial problem incorrectly. We compared the cue and no cue groups' performances on the transfer problems using the Mann-Whitney U test. We found that there is a nearly significant difference for the ball transfer problem ($p = .06$) and the graph transfer problem ($p = .054$). There was no difference found for the roller coaster and skier transfer problems. These results suggest the visual cues in the ball and graph problem sets positively influenced performance on some of the related transfer problems.

It is curious that there is no difference between the cue and no cue groups on the transfer problem from the rollercoaster problem set, though we did find a difference in the number of students who answered similar problems correctly after answering the initial problem incorrectly. It seems that this transfer problem may have been difficult for this level of student as only one out of 14 students in the no cue group answered correctly while two out of 18 students in the cue group answered correctly. A similar explanation could account for no difference between the cue and no cue groups on the skier problem. It is also possible that the students did not view the concept tested in the transfer problem to be similar to that tested in the similar problems, even though the researchers did see these similarities. In other words, from the point of view of our students, the transfer task can be interpreted as primarily a far transfer task, while the similar problems constituted near transfer and thus were not as challenging for the students. These interpretations from students' perspectives may be expected given that the similar problems shared the same surface features with each other, rather than the transfer task which had different surface features, and only shared the deep structure with the training problems.

Discussion and Conclusion

In this study we find some evidence that viewing a physics problem overlaid with short duration visual cues can indeed help students correctly answer and reason about problems they were previously unable to. Of the four problem sets used, we found on one of these problem sets significantly more students changed to a correct answer after seeing cues. It is not enough though, to provide visual cues to help students answer a given set of problems. In looking at transfer problem performance, we found nearly significant differences on the ball and graph transfer problems with the cue group outperforming the no cue group. Thus, we find some evidence that repeatedly showing novices visual cues on related problems may help them to properly apply the factual knowledge on similar future problems viewed without cues.

While we did find some results that point to the usefulness of cueing, we also saw no difference between cue and no cue groups in the number of similar problems answered correctly on three of the four problem sets tested. There are many reasons that the cueing may have failed.

First, we have previously discussed how the cue pattern may have been too complex on these three problems. In the future, we can use simple repetitive cues as well as increase the time the cues are seen. We can also change the type of cue we are using, for example instead of moving colored shapes modeling correct solvers' eye movements, we could use lines or bars to illustrate comparisons in the diagram that correct solvers make. Additionally, there may be only certain types of problems that lend themselves to improvement through visual cueing. We have only explored four problems in this study. There are a plethora of problems that can be categorized in a variety of ways and 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 cued problems will be randomized to balance out any order effects.

We also found differences between the cue and no cue groups on two of the four transfer problems tested. As mentioned earlier it may be that the two 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 the researchers viewed the transfer problems as closely related to the similar problems, though the students did not view them this way, and thus were not able to apply what they gained from the cues to the transfer problems.

Previous studies with visual cues in several domains have also found mixed results on the effectiveness of cueing. There is much work to be done to understand the factors that lead to helpful cues. This study offers some hope that cueing can potentially serve as effective conceptual scaffolding for novice physics students, but much work is necessary to perfect this method.

References

Antes, J. R., & Kristjanson, A. F. (1991). Discriminating artists from non-artists by their eye-fixation patterns. *Perceptual and Motor Skills*, 73, 893-894.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition*, 29, 1146-1152.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*(2), 113-140.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology, 21*(6), 731.

Grant, E. R. & Spivey, M.J. (2003). Eye movements and problem solving. *Psychological Science, 14*(5), 462.

Jarodzka, H. Scheiter, K., Gerjets, P., & van Gog, T. (2010) In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, 20(2), 146–154.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies, 65*(11), 911-930.

Lowe, R.K., (2003). Animation and learning: Selective processing of information in dynamic graphics, *Learning and Instruction.* 13, 157–176.

Madsen, A., Larson, A., Loschky, L., & Rebello, N.S. (2012) Differences in visual attention between those who correctly and incorrectly answer physics problems. *Physical Review Special Topics Physics Education Research.* Submitted.

Mason, A. & Singh, S. (2011) Assessing expertise in introductory physics using categorization tasks. *Physical Review Special Topics Physics Education Research.* 7(2), 0201101-02011017.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*(2), 377-389.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*(1), 110-117.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J. (1999). *Instructional design*. Melbourne: ACER Press.

Van Gog, T., Paas, F., van Merriënboer, J.J.G., Witte, P. (2005) Uncovering the Problem-Solving Process: Cued Retrospective Reporting Versus Concurrent and Retrospective Reporting. *Journal of Experimental Psychology, Applied*, 11(4), 237-244.

Table 1

Mean percentage time spent (± std err) during entire problem period for thematically-relevant and novice-like AOIs for participants who answered the question correctly/incorrectly (indicates a significant difference, p < .05).*

Thematically Relevant AOI		
Problem #	Answered Correctly	Answered Incorrectly
1	46.6 (± 5.5) (n = 11)	33.2 (± 5.7) (n = 11)
2*	24.4 (± 2.9) (n = 13)	11.6 (± 3.3) (n = 10)
3*	28.5 (± 4.1) (n = 18)	8.9 (± 2.3) (n = 6)
4*	49.8 (± 3.9) (n = 14)	25.5 (± 4.1) (n = 9)
7*	36.7 (± 5.5) (n = 15)	10.3 (± 2.1) (n = 9)
10*	29.0 (± 5.0) (n = 11)	15.1 (± 2.7) (n = 13)
Novice-Like AOI		
	Answered Correctly	Answered Incorrectly
1*	22.3 (± 4.5) (n = 11)	43.5 (± 7.3) (n = 11)
2*	12.7 (± 3.3) (n = 13)	27.2 (± 4.8) (n = 10)
3*	19.8 (± 3.7) (n = 18)	39.4 (± 5.4) (n = 6)
4 (p=.058)	18.1 (± 2.5) (n=14)	26.8 (±3.9) (n=9)
7*	12.6 (± 2.6) (n = 15)	25 (± 6.0) (n = 9)
10*	41.2 (± 6.6) (n = 11)	62.2 (± 5.1) (n = 13)

Table 2

Number of students in cued and no cue group who answered a similar problem correctly after answering the initial problem incorrectly (indicates a significant difference, p<.05).*

	Rollercoaster Problem*		Ball Problem		Skier Problem		Graph Problem	
Changed to Correct Answer	Cued (N = 18)	No Cue (N=14)	Cued (N=10)	No Cue (N=14)	Cued (N=11)	No Cue (N=7	Cued (N=17)	No Cue (N=22)
	6	0	6	4	2	2	4	1

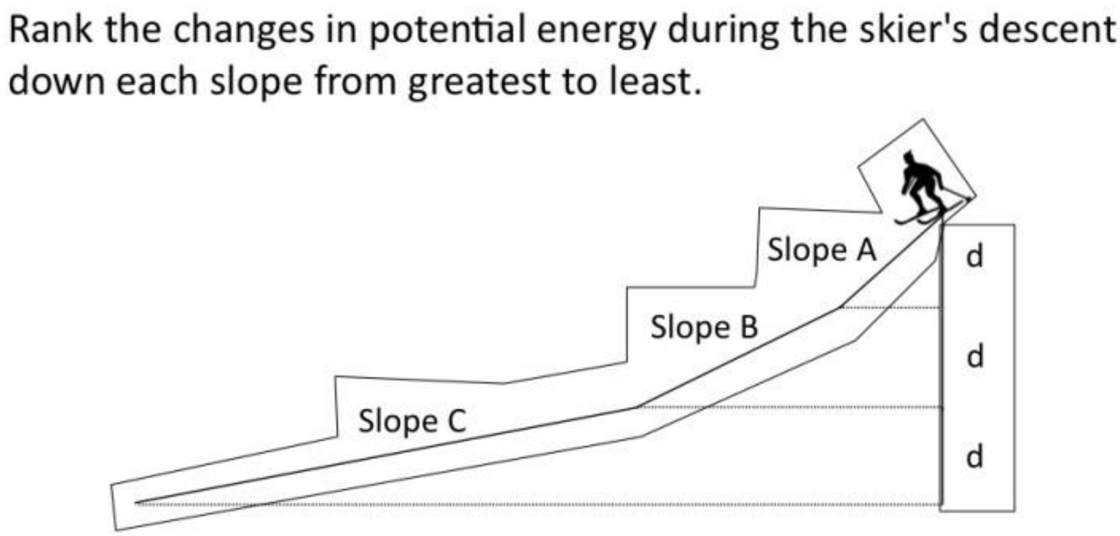


Figure 1. Problem 10 used in Study 1. Novice-like AOI along slope. Thematically relevant AOI along height of hill.

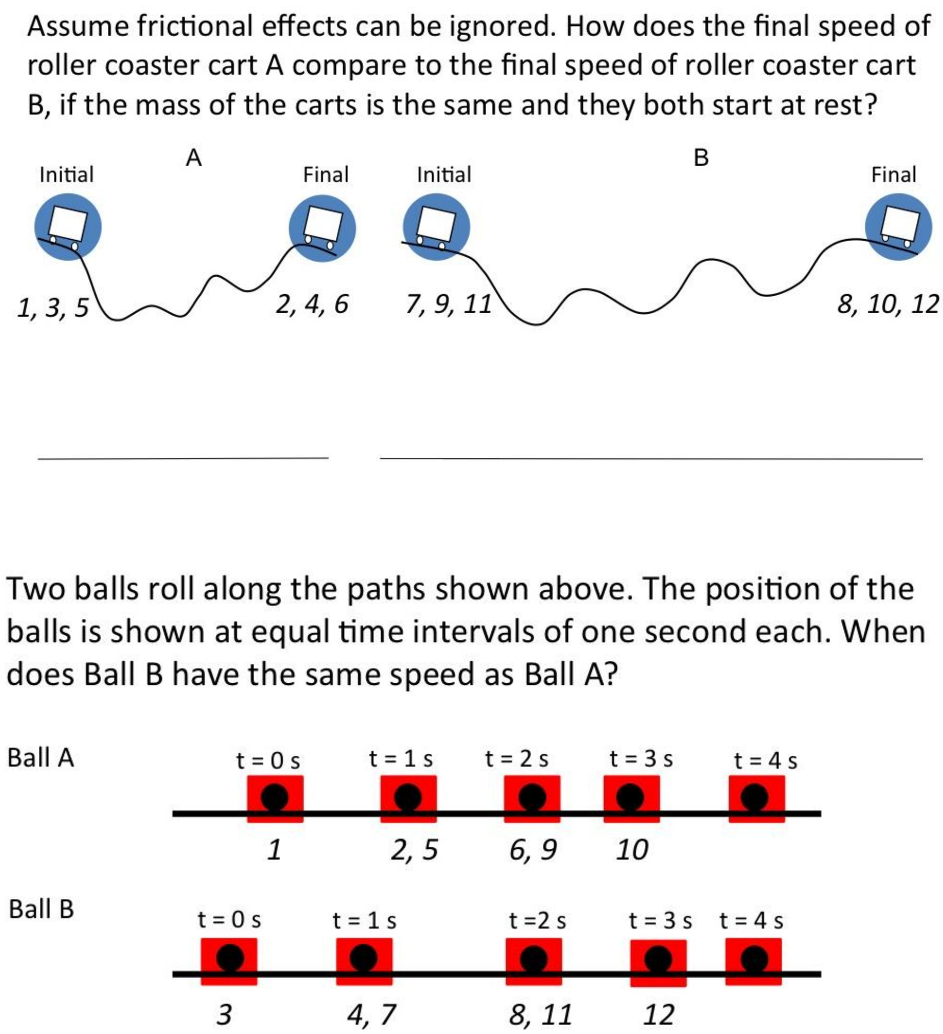


*Figure 2***.** Problems 1 and 2 used in Study 2. Colored shapes are visual cues overlaid on the problem diagram. Numbers in italics show the sequence of animated cues (the numbers were not seen by study participants).

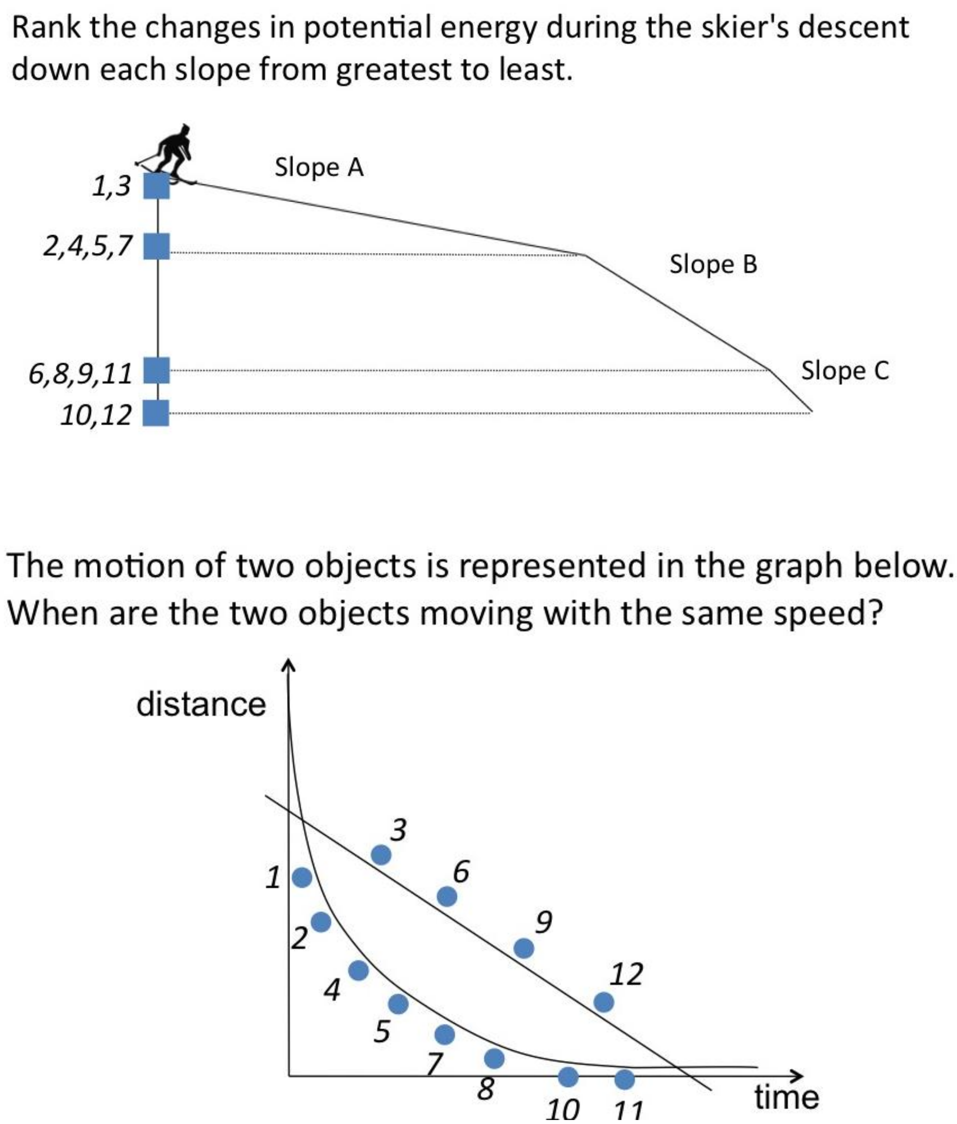


*Figure 3***.** Problems 3 and 4 used in Study 2. Colored shapes are visual cues overlaid on the problem diagram. Numbers in italics show the sequence of animated cues (the numbers were not seen by study participants).

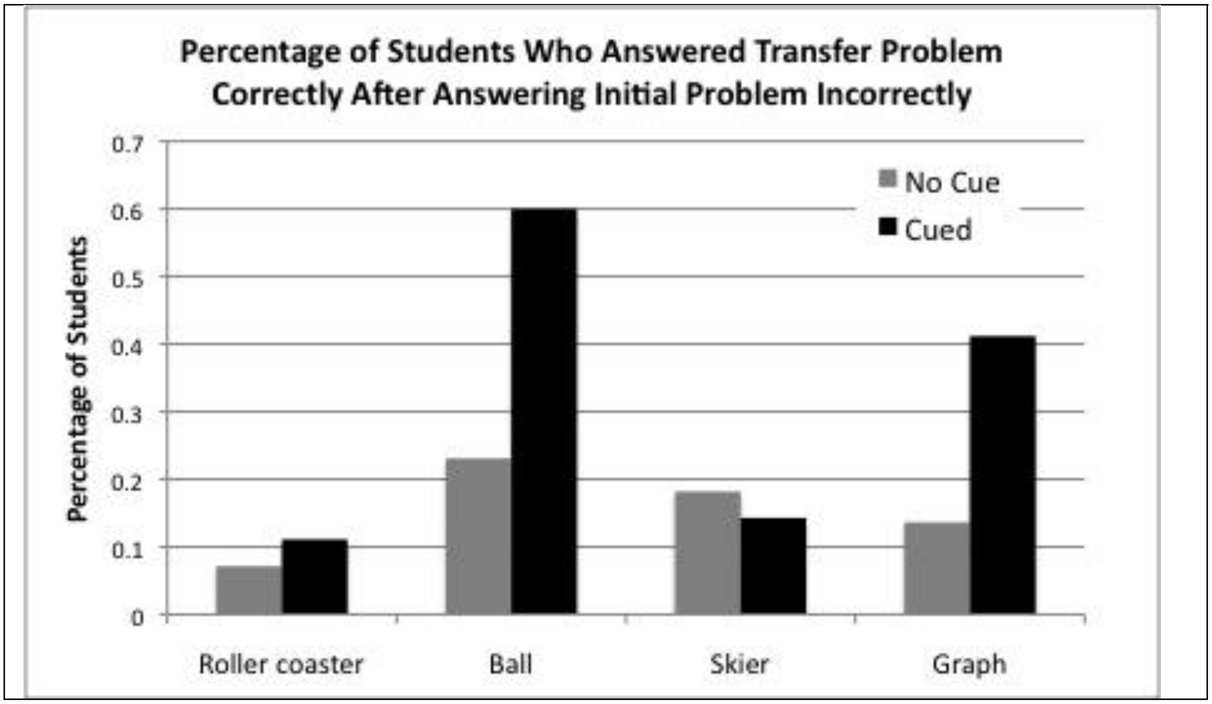


Figure 4. Percentage of students in “cued” and “no cue” conditions who answered initial problem incorrectly, but answered transfer problem correctly.

Influence of Visual Cues on Eye Movements and Reasoning in Physics Problems

Adrian Madsen
Adam Larson
Amy Rouinfar
Allison Coy
Lester Loschky
N. Sanjay Rebello

Kansas State University

Please address all correspondence to:
Adrian Madsen
Kansas State University
Department of Physics
403 Cardwell Hall
Manhattan, KS 66502
(785) 532-1612
adrianc@phys.ksu.edu

Paper presented at the
Annual Meeting of the American Education Research Association
Vancouver, BC, April 13-17, 2012

Influence of Visual Cues on Eye Movements and Reasoning in Physics Problems

Adrian Madsen
Adam Larson
Amy Rouinfar
Allison Coy
Lester Loschky
N. Sanjay Rebello

Kansas State University

Visual cues overlaid on diagrams and animations can reduce cognitive load by drawing attention to relevant areas. Additionally, cues can increase speed and accuracy by causing learners to view a diagram in a pattern related to a problem's solution. We investigate the effects of visual cueing on students' eye movements and reasoning on introductory physics problems with a diagram. Students in the treatment group were shown an initial problem, and if they answered that incorrectly, they were shown a series of problems each with moving shapes cueing the correct solution. Students in the control group were also provided a series of problems, but without any visual cues. Students in both groups were asked to verbally explain their reasoning after each question, and were provided a transfer problem without cues at the end. We report on students' eye movements while answering the questions and verbal reasoning for their answers.

Introduction

To maximize learning one must ensure most of the learner's cognitive resources are spent on relevant tasks and avoid instructional environments that facilitate focusing on the irrelevant. Mayer's cognitive theory of multimedia learning (2001) 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 one's working memory. Often learners are faced with learning environments that impose a high cognitive load and max out the limited capacity of working memory. To help alleviate this problem, visual cues can be used. de Koning, Tabers, Rikers and Paas (2009) have devised a framework to describe three specific functions of cueing which include guiding learners' attention to essential information, emphasizing organization, and making the relations between elements more salient to foster their integration. There exists a large body of research which utilizes visual cues in a variety of contexts to increase learning in animations and static problems (de Koning, Tabers, Rikers & Paas, 2007, Grant & Spivey, 2003, Kriz & Hegarty, 2007, Mautone & Mayer, 2001, Ozcelik, Arslan-Ari & Cagiltay, 2007).

The first function of cueing is especially interesting to us as it works to ensure that cognitive resources are spent appropriately by helping the learner focus primarily on relevant elements. This frees up working memory and allows for real learning to occur. In physics classes, students are often faced with diagrams or animations of the real world which contain elements relevant to the task at hand as well as elements which are present in the real world, but not useful for answering the given question. To help students select the relevant and ignore the irrelevant, we study the use of dynamic visual cues overlaid on static physics problems containing a diagram.

This study builds on previous research that investigated where learners who answer questions incorrectly and correctly look when viewing physics problems containing diagrams (Madsen, Larson, Loschky & Rebello, 2012). In this study, Madsen et al. recorded eye movements of 24 individuals on six different conceptual physics problems where the necessary information to solve the problem was contained in a diagram. The problems also contained areas consistent with a novice-like response. Participants ranged from those who had only taken one high school physics course to those who had completed a Physics PhD. They found that participants who answered correctly spent a higher percentage of time looking at the relevant areas of the diagram, and those who answered incorrectly spent a higher percentage of time looking in areas of the diagram consistent with a novice-like answer. Thus, there is a significant difference in the way correct and incorrect solvers of these physics problems view the diagram. This lays the foundation for the current study, as it confirms the need for visual cues to redirect the attention of incorrect solvers from irrelevant areas of the diagram to relevant areas of the diagram.

In the current study, we use a subset of the physics problems used in Madsen et al. (2012) as well as the eye movements of those who responded correctly to design the visual cues. We aim to answer the following research questions:

1. Do dynamic visual cues patterned after experts' eye movements scaffold students' understanding of physics concepts?
2. Does students' ability to apply a given concept to a new problem improve after seeing visual cues on similar problems?
3. Do students' eye movements change on current and subsequent problems a result of seeing dynamic visual cues?

Methodology

Participants in the study were 55 individuals concurrently enrolled in an introductory algebra-based physics course. To ensure sufficient prerequisite knowledge, each completed a pre-test, which consisted of four open-ended questions gauging their understanding of speed and potential energy. The pre-tests 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.

Participants took part in individual sessions lasting between 30 and 60 minutes. 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 a verbal explanation of their reasoning when ready. Participants in the cue condition were told that colored shapes may appear on some of the problems and when these appeared, they should follow them with their eyes.

Participants were presented with physics problems on a computer screen viewed at a distance of 24 inches using a chin and forehead rest to minimize their extraneous head movements. The resolution of the computer screen was set to 1024 x 768 pixels with a refresh rate of 85 Hz. Each physics problem subtended 33.3° x 25.5° of visual angle. Eye movements were recorded with an EyeLink 1000 desktop mounted eye-tracking system which had an accuracy of less than 0.50° of visual angle. 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.

The materials consisted of four sets of conceptual physics problems covering energy and kinematics, which were found to have significant differences in the way correct and incorrect solvers answered them (Madsen et al. 2012). It should be noted that these four specific problems were chosen from the six problems analyzed because they tested distinct concepts in physics. We refer to these problems as the roller coaster (Figures 1 and 2), ball (Figure 3), skier (Figure 4) and graph (Figure 5) problems

The research design is shown in Figure 6. Within each problem set, there was an "initial" problem, four "similar" problems, and a "transfer" problem. All problems were open-ended and contained a diagram with the necessary information to answer the problem. 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 and tested the same concept and contained a diagram with similar surface features. When the student answered a similar problem correctly, they saw 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 he/she answered the similar problem correctly or incorrectly. The surface features of the transfer problems were different than the initial and similar problems, though the concept tested was the same. All participants viewed the four sets of problems in the same order.

Whenever a student was ready to answer a problem, they indicated this by pressing any key on a keyboard, at which point the problem displayed on the computer would become slightly smaller in size (this was so that the student knew 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 pre-defined 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 cue group saw colored shapes overlaid on the similar problems appear four seconds after the problem was initially seen. Colored shapes were used because color is known to capture one's attention because of its high visual salience. The cues used for the roller coaster, ball, skier, and graph problems are shown in Figures 2, 3, 4 and 5 respectively. Each colored shape appeared for 500 ms at 12 different positions in the diagram for a total cueing time of six seconds. The visual cues were designed to mimic the eye movements of those who answered the same problems correctly in Madsen et al. (2012). 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 could 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 thematically relevant area of interest. Similarities between participants were observed, and visual cues were modeled after these patterns. Further, the cues could have remained static and simply drawn participants' attention to the relevant areas of the problem, but we hoped by modeling the way in which correct solvers viewed the thematically relevant areas and compared elements within these areas, the cues would give the participants more insight into how to correctly answer the problems.

Results

Changes to Correct Answer on Similar Problems

Table 1 shows the number of students who answered the initial problem incorrectly and then changed to a correct answer and reasoning on a similar problem. Using a Mann-Whitney U test to compare the number of participants in the cue and no cue groups who changed to a correct answer, we found a significant difference on the roller coaster problem ($p=.002$) where six students in the cue group ($N=18$) changed to the correct answer while zero students in the no cue group ($N=14$) made this change. There were no significant differences between groups on the ball, skier, or graph problems. The particular aspects of each problem and associated cues will be further analyzed to determine where the differences in effectiveness originate.

Transfer Problem Correctness

To determine if visual cueing is useful for learning beyond the problem being cue, participants answered a transfer problem for each problem set. Figure 7 shows the percentage of participants who answered the transfer problem correctly after answering the initial problem incorrectly. We compared the cue and no cue groups performances on the transfer problems using the Mann-Whitney U test. We found that there is a nearly significant difference for the ball transfer problem ($p=.06$) and the graph transfer problem ($p=.054$). There was no difference found for the roller coaster and skier transfer problems. These results suggest the visual cues in the ball and graph problem sets positively influenced performance on the related transfer problems.

Eye Movements on Roller Coaster Problem

To further investigate the positive effect of the visual cues on the problems, we looked at the eye movements of the students. First, we investigated how well students in the cue group followed the visual cues with their eyes. To do this, we created four interest areas where the cues began and ended (around the roller coaster carts). We then counted the number of saccades each participant made between these interest areas and the total number of saccades made within the diagram during the four seconds that the cues appeared. On the roller coaster problem, 52.6% of the saccades of those in the cue group followed the cues. In the no cue group only 0.96% of the saccades were in a pattern similar to the cues (though the no cue group did not see any cues, and we did not expect their eyes to move in the pattern of the cues spontaneously).

Next we looked for a correlation between following the cues closely with the eyes and changing to the correct answer on a similar problem. We counted saccades between the same areas of interest described above for those in the cue group only. Using a one-way ANOVA, we found a significant difference in the percentage of saccades that correctly followed the cues between those who changed to a correct answer on a similar problem and those who did not ($F(1,14)=10.8$, $p=.005$). Students who answered a similar problem correctly made 85.5% of their saccades in a manner that followed the cues. Students who did not answer any of the similar problems correctly made only 46.4% of their saccades in a manner that followed the cues. This suggests a relationship between closely following the visual cues and coming to the correct answer on the roller coaster similar problems.

Eye Movements on Ball Problem

On the ball problem, we found a nearly significant difference in transfer problem performance between groups. In the cue group, 60% of students answered the transfer problem correctly while in the no cue group, 23.1% answered correctly. This suggests that seeing the visual cues positively influenced performance on the transfer problem. To further investigate this finding, we looked at the eye movements on this problem.

The visual cues used in the similar ball problems had the students compare the distances between balls at each time interval. If the visual cues influenced how students look at the transfer problem, we anticipate that the cue group would show a greater number of saccades comparing the distance between balls. In the case of the transfer problem, these would be vertical saccades within interest areas 1 and 2 (Figure 8). We found that for those in the cue group, 23.0% of their saccades were within interest areas 1 and 2. For those in the no cue group, 24.1% of their saccades were within interest areas 1 and 2. Thus, no differences were found between the way those in the cue and no cue groups look at the distances between the balls on the ball transfer problem. This suggests that this difference in performance on the transfer problem is not reflected in the participants' eye movements. We further looked at the percentage of saccades within interest areas 1 and 2 of those in each group who answered correctly versus incorrectly. We anticipated that those who answered correctly would display more saccades within the interest areas. These results are displayed in Table 2. Once again, we find no differences between the cue and no cue groups in this analysis. This suggests that the visual cues are not changing the way participants view the ball transfer problem.

Conclusion

In this study we find some evidence that viewing a physics problem overlaid with short duration visual cues can indeed help students correctly answer and reason about problems they were previously unable to. Of the four problem sets used, we found on one of these problem sets significantly more students changed to a correct answer after seeing cues. It is not enough though, to provide visual cues to help students answer a given set of problems. In looking at transfer problem performance, we found nearly significant differences on the ball and graph transfer problems with the cue group outperforming the no cue group. Thus, we find some evidence that repeatedly showing novices visual cues on related problems may help them to properly apply the factual knowledge on similar future problems viewed without cues.

We also found some differences in eye movements of those who changed to a correct answer on a similar problem and those who did not. Those who changed to a correct answer on a similar problem for the rollercoaster problem followed the visual cues more closely than those who did not change to a correct answer. Thus, there may be a relationship between how well a participant follows the visual cues with their eyes, and how helpful these cues are. This suggests that following the cues closely is related to changing to a correct answer. Further, we looked for evidence that seeing cues changes the way in which one views future problems with no cues and found no evidence for this on the ball transfer problem.

While we did find some results that point to the usefulness of cueing, we also saw no difference between cue and no cue groups in the number of similar problems answered correctly on three of the four problem sets tested. There are many reasons that the cueing may have failed. First, we have previously discussed how the cue pattern may have been too complex on these three problems. In the future, we can use simple repetitive cues as well as increase the time the cues are seen. We can also change the type of cue we are using, for example instead of moving colored shapes modeling correct solvers' eye movements, we could use lines or bars to illustrate comparisons in the diagram that correct solvers make. Additionally, there may be only certain types of problems that lend themselves to improvement through visual cueing. We have only explored four problems in this study. There are a plethora of problems that can be categorized in a variety of ways and 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 will be randomized to balance out any order effects.

We also found differences between the cue and no cue groups on two of the four transfer problems tested. As mentioned earlier it may be that the two 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 the researchers viewed the transfer problems as closely related to the similar problems, though the students did not view them this way, and thus were unable to apply what they gained from the cues to the transfer problems. In other words, the transfer problems, though deemed to be near transfer problems by the researchers, were perceived to be far transfer problems by the participants in our study. A problem may be perceived as near or far transfer depending upon whether the problem solver perceives the two problems to be different in surface feature or deep structure. So, it seems that although the 'similar' and 'transfer' problems were deemed to differ only in surface feature by the researchers, the participants in our study appear to have perceived them as being different in deep structure as well.

Previous studies with visual cues in several domains have also found mixed results on the effectiveness of cueing. There is much work to be done to understand the factors that lead to helpful cues. This study offers some hope that cueing can potentially serve as effective conceptual scaffolding for novice physics students, but much work is necessary to perfect this method.

REFERENCES

Carmichael, A., Larson, A., Gire, E., Loschky, L. and Rebello, N.S. (2010). How Does Visual Attention Differ Between Experts and Novices on Physics Problems? In *AIP Conference Proceedings* vol. 1289, 93-96.

de Koning, B. B., Tabbers, H. K., Rikers, Remy M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*(2), 113-140.

de Koning, B. B. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology, 21*(6), 731.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*(2), 111-122.

Grant, E. R. & Spivey, M. J., (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*(5), 462.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies, 65*(11), 911-930.

Madsen, A., Larson, A., Loschky, L., & Rebello, N.S. (2012) Differences in visual attention between those who correctly and incorrectly answer physics problems. *Physical Review Special Topics: Physics Education Research.* Accepted.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*(2), 377-389.

Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology, 99*(3), 640-652.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*(1), 110-117.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J. (1999). *Instructional design*. Melbourne: ACER Press.

Thomas, L. E. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin Review, 14*(4), 663.

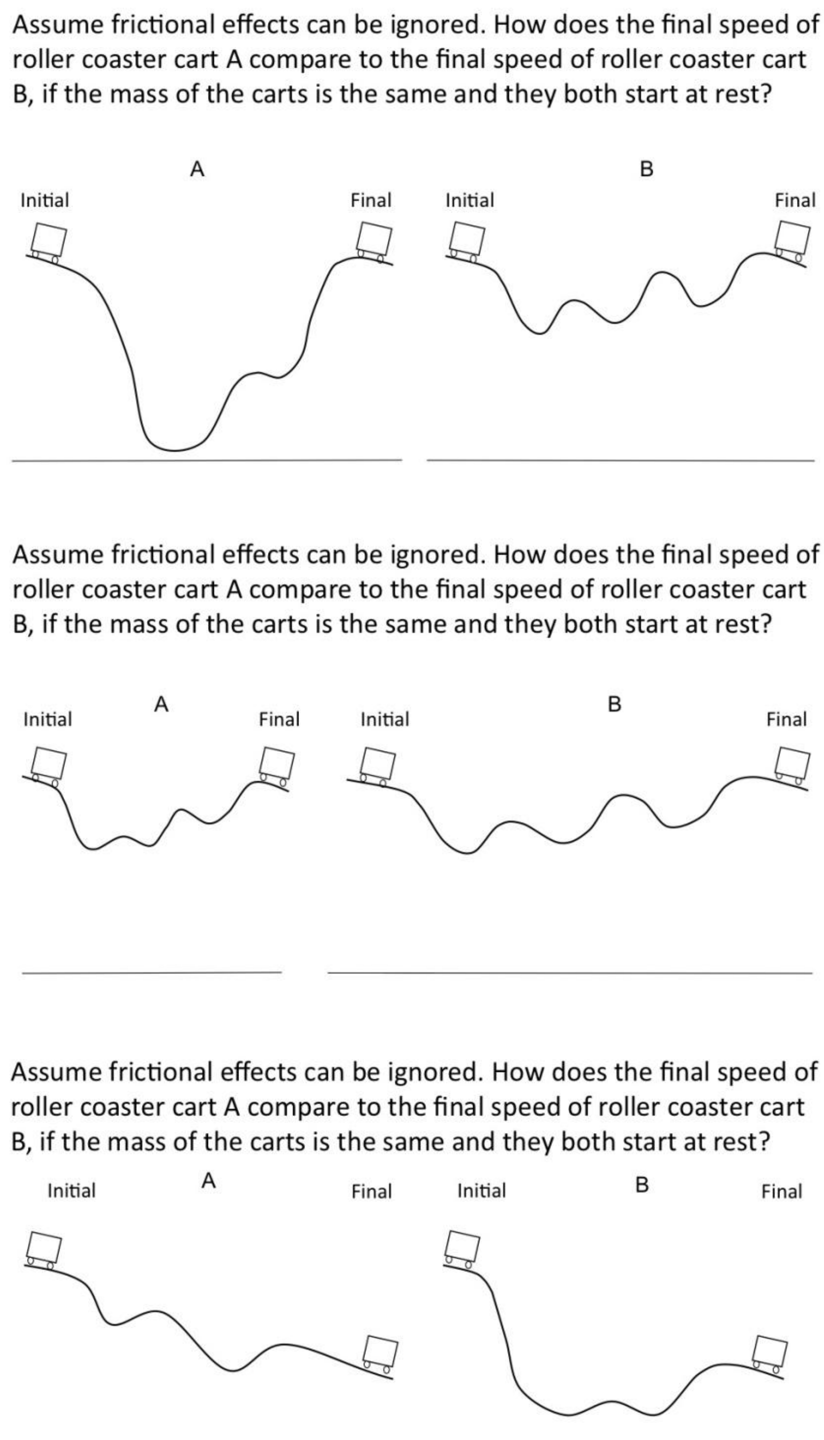


Figure 1. Example of initial problem (top), similar problem (middle), and transfer problem (bottom) used in study.

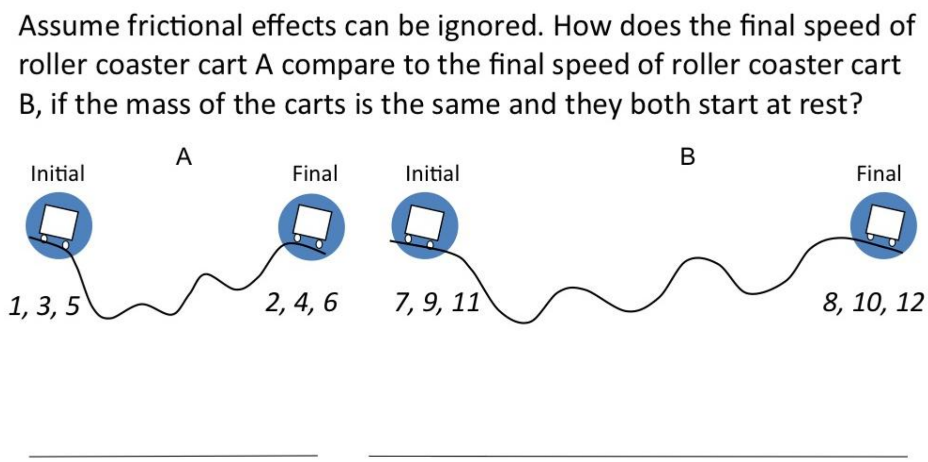


Figure 2. Problem 1 used in study. Blue circles are the visual cues overlaid on the diagram. Numbers in italics show sequence of animated cues (the numbers were not seen by study participants).

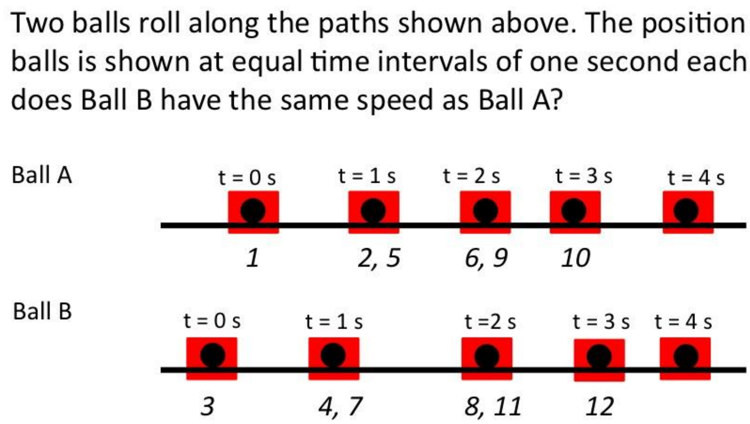


Figure 3. Problem 2 used in study. Red squares are the visual cues overlaid on the diagram. Numbers in italics show sequence of animated cues (the numbers were not seen by study participants).

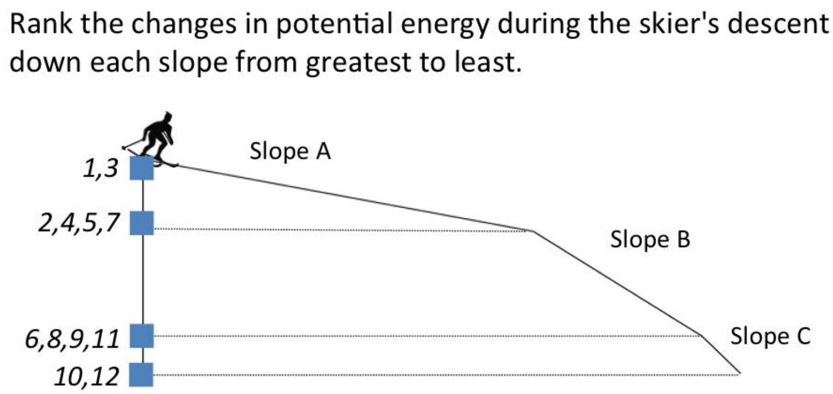


Figure 4. Problem 3 used in study. Blue squares are the visual cues overlaid on the diagram. Numbers in italics show sequence of animated cues (the numbers were not seen by study participants).

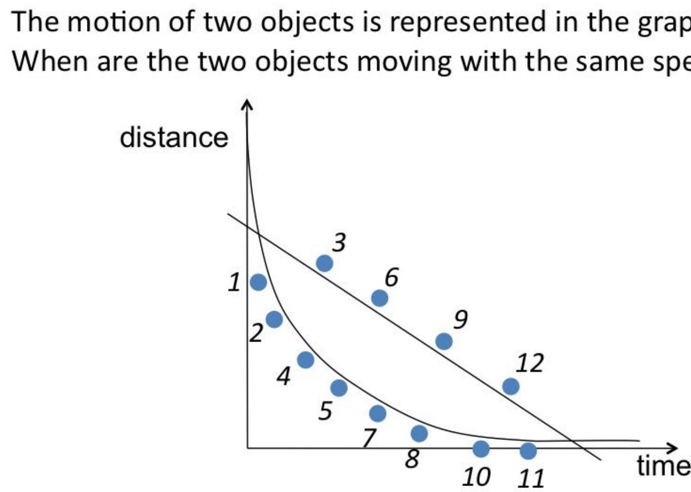


Figure 5. Problem 4 used in study. Blue dots are the visual cues overlaid on the diagram. Numbers in italics show sequence of animated cues (the numbers were not seen by study participants).

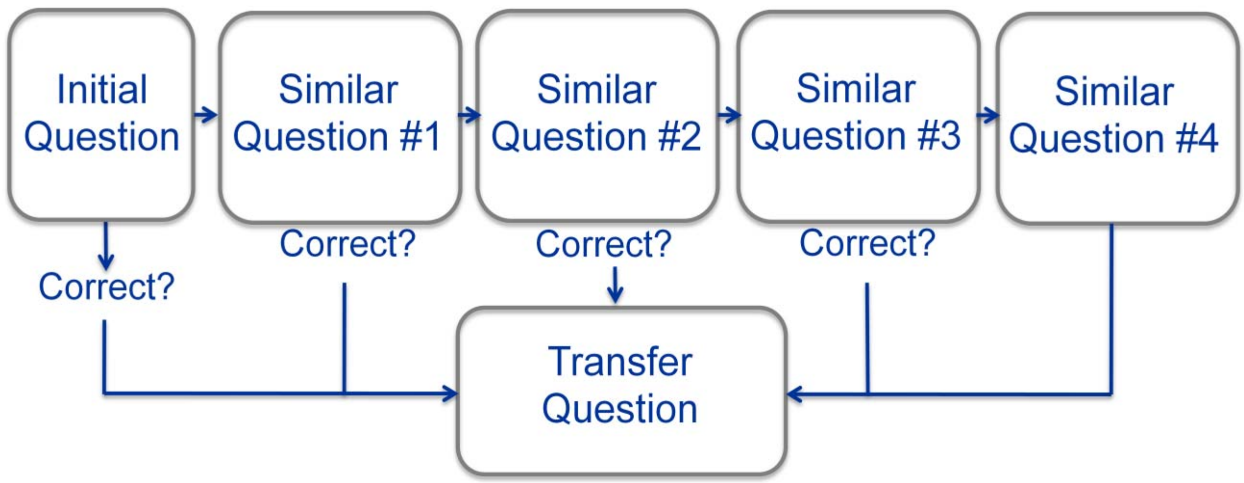


Figure 6. Flow chart showing how the initial problem, similar problems, and transfer problems were administered to students in each problem set.

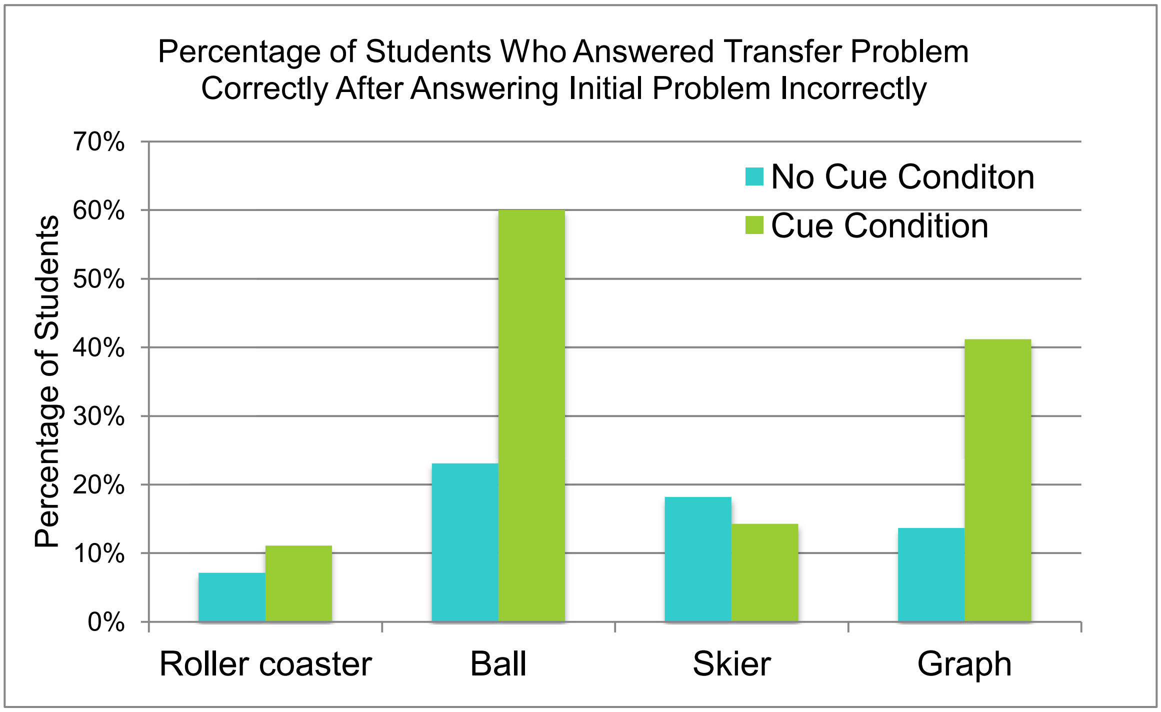


Figure 7. Percentage of students in "cue" and "no cue" conditions who answered initial problem incorrectly, but answered transfer problem correctly.

Ball A begins riding downward in an elevator at the same time Ball B is dropped from the roof of an adjacent building. The position of the balls is shown at equal time intervals of one second. When does Ball B have the same speed as Ball A?

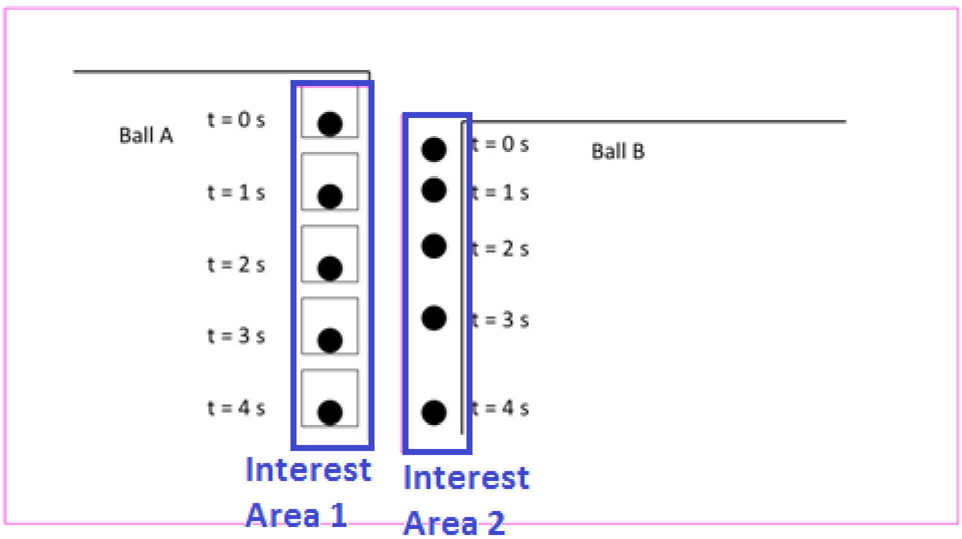


Figure 8. In the ball transfer problem the leftmost blue rectangle is interest area 1 and rightmost blue rectangle is interest area 2. These interest areas were used when analyzing eye movements.

Table 1. Number of students in cue and no cue group who answered a similar problem correctly after answering the initial problem incorrectly (* indicates a significant difference, p<.05).

	Rollercoaster Problem*		Ball Problem		Skier Problem		Graph Problem	
	Cue (N=18)	No Cue (N=14)	Cue (N=10)	No Cue (N=14)	Cue (N=11)	No Cue (N=7)	Cue (N=17)	No Cue (N=22)
Number Changing to Correct Answer	6	0	6	4	2	2	4	1

Table 2. Comparison of participants who answered correctly versus incorrectly in the cue and no cue groups on the ball transfer problem (Figure 8). We compared the percentage of saccades made vertically within interest area 1 or interest area 2.

		% Saccades Within Interest Areas 1 and 2
Cue	Correct Answer	18.2%
	Incorrect Answer	26.1%
No Cue	Correct Answer	17.4%
	Incorrect Answer	28.6%

Using ScanMatch Scores to Understand Differences in Eye Movements Between Correct and Incorrect Solvers on Physics Problems

Adrian Madsen
Kansas State University
adrianc@ksu.edu

Adam Larson
Kansas State University
adlarson@ksu.edu

Lester Loschky
Kansas State University
loschky@ksu.edu

N. Sanjay Rebello
Kansas State University
srebello@ksu.edu

Abstract

Using a ScanMatch algorithm we investigate scan path differences between subjects who answer physics problems correctly and incorrectly. This algorithm bins a saccade sequence spatially and temporally, recodes this information to create a sequence of letters representing fixation location, duration and order, and compares two sequences to generate a similarity score. We recorded eye movements of 24 individuals on six physics problems containing diagrams with areas consistent with a novice-like response and areas of high perceptual salience. We calculated average ScanMatch similarity scores comparing correct solvers to one another (C-C), incorrect solvers to one another (I-I), and correct solvers to incorrect solvers (C-I). We found statistically significant differences between the C-C and I-I comparisons on only one of the problems. This seems to imply that top down processes relying on incorrect domain knowledge, rather than bottom up processes driven by perceptual salience, determine the eye movements of incorrect solvers.

CR Categories: J.2 [Computer Applications]: Physical Science and Engineering – Physics

Keywords: eye movements, attention, scan path, ScanMatch, problem solving, physics

1 Introduction

Researchers have found consistent patterns of wrong answers to many simple conceptual physics questions [Trowbridge and McDermott 1980; McDermott et. al 1987]. Several cognitive top-down explanations have been provided, including misconceptions formed through interactions with the natural world or misapplication of conceptual resources [Docktor and Mestre 2011]. However, recent claims by Heckler [2011] have suggested a perceptual basis for students' incorrect answers, which are based on attention being directed to the most perceptually salient and plausibly relevant features in a problem. The most salient features capture attention through perceptual processes and less salient features have little opportunity to be considered. Heckler shows some evidence for perceptually-driven responses; however, no eye movement data supporting this hypothesis is provided. Further, he does not provide a specific definition of salience. Therefore, incorrect answers may be governed either by top-down processes relying on incorrectly learned or applied information, or by bottom-up perceptual processes resulting in certain elements capturing attention and leading to activation of reasoning resources based on these elements.

An eye-movement study was used to test these competing hypotheses. Introductory and graduate physics students answered conceptual physics problems regarding a diagram [Madsen et al., 2011]. Three areas of interest (AOIs) were defined for each diagram. First, thematically-relevant AOIs that contained information necessary to correctly answer the question were determined by experts in physics. Second, novice-like AOIs were defined based on coded interview data from novices [Madsen et al. 2011], and third, perceptually salient AOIs were defined as the area(s) on the diagram with the highest saliency rating according to the salience maps produced by a computational algorithm [Itti 2000]. For each problem, the percentage of time spent in each type of interest area was compared between students who answered the problem correctly and those who answered the problem incorrectly.

If top-down cognitive processes utilizing naïve theories or misapplied information were directing attention in physics problems, then those who answer the problems incorrectly should spend more time looking at the novice-like AOIs than those who answer correctly. If perceptual salience captures attention and leads students to an incorrect answer, then more time should be spent looking at perceptually salient AOIs. We found that in five of six problems, those who answered incorrectly spent significantly more time looking at the novice-like AOI than those who answered correctly. No differences were found between correct and incorrect solvers in the perceptually salient AOIs. However, it is important to note that Carmi and Itti [2006] studied the effects of perceptual saliency as a function of viewing time. They found that their model of perceptual salience performed best on the first six to seven fixations when viewing a scene. For the average viewer, this is equivalent to about the first two seconds of viewing. In light of this finding, we also compared the amount of time spent in the perceptually salient AOI during the first two seconds of viewing the diagram for those who answered correctly versus incorrectly. No significant differences were found between those who answered correctly versus incorrectly, although the data were in the predicted direction (i.e., the raw percentage of time spent in the perceptually salient AOI was higher for those who answered incorrectly on five of the six problems analyzed). Thus, it may be either that the small number of fixations observed in the first two seconds of diagram viewing lacked the statistical power to find an effect, or there may simply be no effect between those who answer correctly versus incorrectly on the viewing time of perceptually salient elements of the diagram.

In this paper, we will expand on our previous work [Madsen et al. 2011] to further investigate the role of perceptual salience in guiding the attention of those who incorrectly answer conceptual physics questions containing a diagram. A scan path analysis was performed using an algorithm called ScanMatch [Cristino et al. 2010], which is based on the Needleman-Wunsch algorithm used to compare DNA sequences. ScanMatch bins a saccade sequence both spatially and temporally and then recodes this information to create a sequence of letters which represents the location, duration, and order of the fixations. The letter sequences of two sets of eye movements are then compared to each other to calculate a similarity score. A similarity score near one represents two sequences of eye movements that are very similar spatially and temporally. The ScanMatch analysis requires no decisions to be made about the data a priori, for example, one does not have to define AOIs based on an experimenter's definition or rating. Therefore, it is possible that differences exist in sets of eye movement data that are not detected by looking at fixation durations in AOIs.

We will compare the average ScanMatch scores produced by comparing the correct solvers to one another (C-C comparison), the incorrect solvers to one another (I-I comparison), and the correct solvers to the incorrect solvers (C-I comparison).

We hypothesize that if the incorrect solvers are being primarily led by the perceptual salience of the elements in the diagram, then it is likely that they will attend to the same elements in a similar order. For example, attention would be first guided to the most perceptually salient region, followed by the next most salient region, and so on [Itti 2000]. Thus, the I-I comparison would have higher ScanMatch scores than the C-C comparison, who might attend to perceptually salient areas early on in diagram viewing; however, the variable onset of top-down processes on eye movements would result in greater temporal and spatial variability of gaze towards thematically-relevant elements in the diagram, resulting in lower ScanMatch scores. The I-I and C-C groups would also have higher ScanMatch scores than the C-I group, since the correct solvers and incorrect solvers are known to spend different amounts of times looking at thematically-relevant and novice-like elements [Madsen et al. 2011; Carmichael et al. 2010].

Conversely, if top-down processes are directing the attention of incorrect solvers, namely some form of naïve theory, the ScanMatch score of the I-I comparison should be similar to that of the C-C comparison. The domain knowledge possessed by those in both comparison groups, whether correct or incorrect knowledge, guides their attention to look at certain elements of the problem, but not in a particular order. Once again, the I-I comparison and the C-C comparison should have higher ScanMatch scores than the C-I comparison.

In summary:

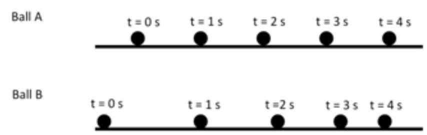
Hypothesis 1: If perceptual salience is primarily influencing the attention of incorrect solvers, the I-I comparison will have higher ScanMatch scores than the C-C comparison.

Hypothesis 2: If top-down processes utilizing naïve theories are primarily influencing the attention of incorrect solvers, the I-I comparison and the C-C comparison will have similar ScanMatch scores, and these will both be higher than the C-I comparison.

2 Methodology

There were 24 participants (three females), with two different levels of experience in physics. Ten participants were PhD students in physics and one was a postdoctoral researcher in physics; all had taught an introductory physics course. Thirteen participants were introductory psychology students who had taken at least one physics course in high school, though some had also taken a physics course at the university. The PhD students and post-doc voluntarily participated while the psychology students received course credit. Since we sought to compare those who answered the physics problems correctly versus incorrectly, we selected participants with a broad range of experience. We expected the PhD students to answer correctly, while the psychology students might answer incorrectly, though we know that this may not always be the case as it has been shown there is a wide distribution of expertise among introductory physics students and physics graduate students [Mason and Singh, 2011]. The materials consisted of 10 multiple-choice conceptual physics problems covering various topics in introductory physics. For an example, see Figure 1. Each problem contained a diagram with a thematically-relevant visual component that students needed to attend to in order to answer correctly. These problems also contained areas consistent with naïve conceptions documented in physics education literature [McDermott and Redish 1999].

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.



(1) t = 1.0 sec (2) t = 1.5 sec (3) t = 2.0 sec
(4) t = 2.5 sec (5) t = 3.0 sec

Figure 1. *Problem 4 used in this study. Thematically-relevant area is distance between balls between 2 and 3 seconds. Novice-like area is area where balls are lined up spatially at 1 second.*

The physics problems were presented to participants on a computer screen. Participants used a chin and forehead rest that was 24 inches from the screen. The screen had a resolution of 1024 by 768 pixels and a refresh rate of 85 Hz. Eye movements were recorded with an EyeLink 1000 desktop mounted eye-tracking system, which had an accuracy of less than 0.50° of visual angle. The images subtended 33.3° x 25.5° of visual angle. An eye movement was classified as a saccade if acceleration exceeded 8,500°/s^2 and velocity exceeded 30°/s. Participants' verbal explanations and gestures were recorded with a Flip video camera. Each participant took part in an individual session, lasting 20-40 minutes. At the beginning of the session, participants were given a short explanation of the experiment. The eye tracking system was calibrated to the individual using a nine-point calibration and validation procedure, with a threshold agreement of 0.50° visual angle required to begin the experiment. Next, the participant was instructed to silently answer 10 multiple-choice questions, with their head on a headrest, while their eye movements were recorded. Between questions, drift correction was carried out using the central fixation point to ensure proper calibration. Participants indicated their answer to each question using number keys on the keyboard. Finally, each participant was asked to provide a verbal cued retrospective report [Van Gog 2005] for which they were shown a replay of their eye movements on each problem and they were asked to explain their thought processes. This method has been found to produce more depth of explanation than a retrospective report without viewing one's eye movements. If a participant's explanation was unclear, follow-up questions were asked of him/her. Participants were not given any time limits.

3 Analysis and Results

We used the ScanMatch toolbox for MatLab [Cristino et. al 2010] to compare the scan paths of our participants based on the correctness of their answers given for each problem. The ScanMatch algorithm compares the sequence and durations of fixations in a pair-wise fashion and produces a numerical score representing the similarity of the scan paths both spatially and temporally. A score of one indicates that the scan paths being compared are identical while a score of zero represents no relationship between the scan paths. We calculated ScanMatch scores for three different comparisons of participants' scan paths. The correct-correct comparison (C-C) contained scores comparing each participant who answered a question correctly to one another. The incorrect-incorrect comparison (I-I) contained scores comparing each participant who answered a question incorrectly to one another. Finally, the correct-incorrect comparison (C-I) contained scores comparing those who answered correctly to those who answered incorrectly. We then completed a one-way ANOVA[1] comparing the ScanMatch scores of the C-C comparison, I-I comparison, and C-I comparison for each problem. When we obtained a significant result, we used post-hoc contrasts to determine which comparisons contained a significant difference. We then referenced the mean score values for each comparison to determine the direction of this difference. When homogeneity of variance was violated, we used the Games-Howell test for the post-hoc contrasts, otherwise we used Tukey's HSD test for the contrasts. In the previous study [Madsen et al. 2011] for which this analysis is a follow-up, the eye movements of only six of the 10 problems participants viewed were analyzed. This is because we found that four of the problems did not contain a consistent novice-like area of interest. On those four problems, participants who answered incorrectly reasoned from a wide variety of areas in the problem diagram. Without a precise definition for the novice-like area of interest, these problems could not be included in the original analysis. This scan path analysis is a follow-up on the previous analysis, so we analyze only those six problems included in the original study.

We found statistically significant main effects on three of the six problems tested (Table 1). On problem 1, the ANOVA showed a statistically significant main effect of comparison, $F(2,220)=7.324$, $p=.001$. The contrasts revealed that the I-I comparison had significantly higher ScanMatch scores than the C-I comparison ($p<.001$). Problem 2 also showed significant main effect of comparison, $F(2,250)=6.308$, $p=.002$. The contrasts showed that the I-I comparison ($p<.001$) had a higher ScanMatch score than the C-I comparison. Further, the I-I comparison had a significantly higher score than the C-C comparison ($p=.005$). A significant main effect was also found for problem 10, $F(2,273)=3.583$, $p=.029$. On this problem, the I-I comparison had a significantly higher ScanMatch score than the C-I comparison ($p=.05$). There were no differences found between comparisons on problems 3, 4 and 7.

[1] When using the one-way ANOVA, we recognize there may be issues with the homogeneity of variance because of the unequal sample sizes between correct and incorrect responders. For this reason, we used corrected post-hoc contrasts (Games-Howell test) when this assumption was violated. Further, we employed a non-parametric procedure [Feusner and Lukoff, 2008] to confirm the ANOVA results and found general agreement.

Problem	Comparison	Mean	SD (+/-)
1*	C-C (n=47)	.396	.068
(n=11 correct	I-I (n=55)	.414	.056
n=11 incorrect)	C-I (n=121)	.370	.080
2*	C-C (n=90)	.330	.151
(n=14 correct	I-I (n=36)	.413	.047
n=10 incorrect)	C-I (n=127)	.371	.119
3	C-C (n=137)	.351	.093
(n=17 correct	I-I (n=21)	.400	.108
n=7 incorrect)	C-I (n=119)	.364	.100
4	C-C (n=90)	.379	.088
(n=14 correct	I-I (n=35)	.398	.055
n=9 incorrect)	C-I (n=126)	.362	.088
7	C-C (n=105)	.312	.125
(n=15 correct	I-I (n=36)	.311	.119
n=9 incorrect)	C-I (n=135)	.298	.112
10*	C-C (n=55)	.333	.086
(n=11 correct	I-I (n=78)	.368	.091
n=13 incorrect)	C-I (n=143)	.340	.078
*this indicates a significant difference at the p=.05 level			

Table 1. *Mean ScanMatch score for C-C, I-I, and C-I comparison for each problem used in the study.*

Figure 2 shows a box and whiskers plot comparing the ScanMatch scores of each group averaged over the problems in Table 1.

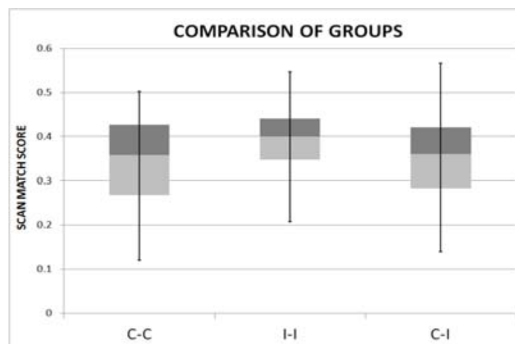


Figure 2. Box and whiskers plot showing the median, max, min and 1st and 3rd quartile of the ScanMatch scores for each group.

4 Conclusion

We did not find significant differences in ScanMatch scores between those in the C-C comparisons and those in the I-I comparisons on five of the six problems analyzed in this study. This evidence is consistent with the hypothesis that the attention of incorrect solvers is primarily directed by top-down naïve theories and not the relative perceptual salience of the elements. This finding aligns well with our previous findings [Madsen et al. 2011] that showed no significant difference in the percentage of fixation time in the perceptually salient areas of the diagram during the full problem period, or the first two seconds of viewing the diagram, when the effects of perceptual salience should be most pronounced. It also aligns well with the findings showing significant differences in the percentage of time incorrect solvers spent in the novice-like areas of the diagram and the percentage of time correct solvers spent in the thematically-relevant areas of the diagram.

We found significant differences between the I-I and C-I comparisons on three of the six problems. These differences were expected as we have previously seen that correct solvers and incorrect solvers spend different amounts of time looking at thematically-relevant and novice-like elements in the problem, so their scan paths scores are likely to be different. It is curious that we did not find that the I-I comparison and the C-C comparison had higher ScanMatch scores than the C-I comparison on all of the problems. The problems used in the study included a text problem statement, diagram, and multiple-choice answers. The hypotheses set forward in this study assumed a similar reading pattern of the problem statement and answer choices for all participants. The hypotheses were formed assuming only differences in how the participants looked at the diagram. Differences in reading the problem statement and answer choices may have overwhelmed small differences in diagram viewing, resulting in no difference in the ScanMatch scores of the C-C and I-I comparisons compared to the C-I comparison.

These findings may have implications for educational interventions aimed at helping novices learn to answer such conceptual questions correctly. Researchers in physics education have devoted much attention to addressing these consistent wrong answer patterns by changing the way students think about how the world works. If it were true that this problem had an underlying perceptual component, these interventions would need to instead help students learn how to ignore salient elements and focus instead on thematically-relevant elements. The results of this study suggest that wrong answers have roots in the incorrect ways students think about how the world works, not how a problem diagram looks. So it seems that the educational interventions used to improve student understanding are on the right track.

5 Limitations and Future Work

The manner in which participants read the problem statement and answer choices may be interfering with our goal of looking for differences in scan paths while viewing the diagram specifically. To address this issue, this work will be repeated with the text and diagram on two separate slides, which can be toggled between by pressing a button on the keyboard. In this new setup, the scan paths of the participants' first view the diagram can be compared to one another to look for influences of visual salience or naïve theories. Additionally, further studies will not use multiple-choice problems, as we have seen some participants rely on a strategy of eliminating distracter answer choices instead of reasoning through the problem on their own. Instead, participants will indicate when they are ready to answer and will give a verbal explanation of their answer and reasoning. Further, the physics topics covered in these problems are limited. It would be useful to expand the number of topics covered by using a larger variety of problems. This will allow us to determine if the conclusions drawn from this work are context-dependent or generalizable to a wider range of physics problems.

More importantly, follow-up studies will explore the hypothesis that cueing students' while they look at physics problems will improve their accuracy in solving them. Because our previous work [Madsen et al. 2011] has shown that those who answer such questions correctly look at the thematically-relevant AOIs more than the novice-like AOIs, we can test the hypothesis that cueing students to look at the thematically-relevant areas will improve the accuracy of their answers, and that doing this repeatedly will improve their accuracy on conceptually similar transfer problems.

References

CARMI, R. AND ITTI, L. 2006. Visual causes versus correlates of attentional selection in dynamic scenes. In *Vision Research,* vol. 46, 4333.

CARMICHAEL, A., LARSON, A., GIRE, E., LOSCHKY, L. AND REBELLO, N.S. 2010. How Does Visual Attention Differ Between Experts and Novices on Physics Problems? In *AIP Conference Proceedings* vol. 1289, 93-96.

CRISTINO, F., MATHOT, S., THEEUWES, J. AND GILCHRIST, I.D. 2010. ScanMatch: A novel method for comparing fixation sequences. In *Behavior Research Methods* vol. 42, 692.

DOCKTOR, J.L. AND MESTRE, J.P. 2011. A Synthesis of Discipline-Based Education Research in Physics. White paper commissioned by the *National Research Council.*

FEUSNER, M. & LUKOFF, B., 2008. Testing for statistically significantly differences between groups of scan patterns. In *Proceedings of the 2008 Symposium on Eye Tracking Research.* 43-46.

HECKLER, A.F. 2011, The Ubiquitous Patterns of Incorrect Answers to Science Questions: The Role of Automatic, Bottom-Up Processes. In *Psychology of Learning and Motivation: Cognition In Education*, J. P. Mestre and B. H. Ross Eds. Academic Press, Oxford, UK, 227-268.

ITTI, L. AND KOCH, C. 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. In *Vision Research* vol. 40, 1489-1506.

MCDERMOTT, L.C., ROSENQUIST, M.L. AND VAN ZEE, E.H. 1987. Student difficulties in connecting graphs and physics: Examples from kinematics. In *American Journal of Physics* vol. 55, 503.

MCDERMOTT, L.C. AND REDISH, E.F. 1999. Resource letter: PER-1: Physics education research. In *American Journal of Physics* vol. 67, 755.

MADSEN, A., LARSON, A., LOSCHKY, L., AND REBELLO, N. S., 2011. Differences in visual attention between those who correctly and incorrectly answer physics problems. In *Physical Review Special Topics Physics Education Research,* submitted.

MASON, A. AND SINGH, C. 2011. Assessing expertise in introductory physics using categorization task. In *Physical Review Special Topics Physics Education Research* vol 7, 020110.

TROWBRIDGE, D. AND MCDERMOTT, L. 1980. Investigation of student understanding of the concept of velocity in one dimension. In *American Journal of Physics* vol. 48, 1020.

VAN GOG, T. 2005. Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. In *Applied Cognitive Psychology* vol. 19, 205.

Guiding Attention on Physics Problems Using Visual Cues Modeled After Experts' Eye Movements

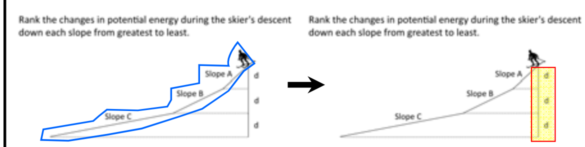
Adrian Madsen[1], **N. Sanjay Rebello**[1] Adam Larson[2]
Amy Rouinfar[1], Lester Loschky[2]
[1]KSU Physics Department
[2]KSU Psychology Department

Supported in part by NSF Grants REC-1138697 & DGE-0841414

2

Research Goals

- Understand differences in visual attention between successful and unsuccessful physics problem solvers.
- Use the attentional patterns of successful solvers to design visual cues for unsuccessful solvers.



Relevant Literature 3

Visual Attention

Influencing Factors : Top-down and bottom-up processes

- Top-down processes
  - require mental effort and rely ones' goals as well as knowledge
  - dominate in learners with higher domain knowledge
- Bottom-up processes
  - faster, more primitive mechanism that drives attention based on the features of the stimuli
  - those with lower domain knowledge may respond more to perceptual salience

Relevant Literature 4

Attentional Cueing

- Framework for Attentional Cueing (deKoning, et. al. 2009)
  - Visual cueing can facilitate selection, organization and integration of information contained in visualizations.
- Grounded in:
  - Cognitive Theory of Multimedia Learning (Mayer, 1997, 2001)
    - Selection, Organization and Integration (SOI) involved in meaningful learning with multiple modalities.
  - Cognitive Load Theory (Sweller, 1988, 1989)
    - Working memory is limited.
    - Cueing can reduce extraneous cognitive load, direct attention to relevant areas, free up mental resources (Britton, 1982)

Overview of Studies 5

Overview

Study 1: Determine visual attention differences between correct and incorrect solvers

↓

Correct solvers' eye movements

↓

Study 2: Explore the use of attentional cueing to improve problem solving

Study 1: Research Questions 6

Study 1: Research Questions

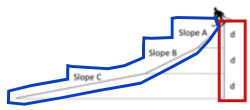
How does the correctness or incorrectness of ones' answer to a physics problem involving a diagram relate to the time spent looking at those areas of the diagram that are:

- thematically relevant to the problem's solution?
- consistent with novice-like misconceptions?

Study 1: Determine visual attention differences between correct and incorrect solvers

- 10 Ph.D. physics students and 14 intro psychology students who have taken at least high school physics
- Six introductory physics problems with diagram

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



(1) $\Delta PE_A > \Delta PE_B > \Delta PE_C$
(2) $\Delta PE_C > \Delta PE_B > \Delta PE_A$
(3) $\Delta PE_A = \Delta PE_B = \Delta PE_C$
(4) $\Delta PE_A = \Delta PE_B > \Delta PE_C$
(5) $\Delta PE_B > \Delta PE_C = \Delta PE_A$

Novice-like areas from interviews (confirmed by previous research)

Thematically relevant areas defined by experts

[1]Itti, 2000; [2]Harel, 2011

Study 1: Analysis

- Determined fixation time (total time spent) in each area of interest (AOI) for each participant.
- % of total viewing time in diagram calculated for each AOI.
- Compared % fixation time for correct vs. incorrect solvers for novice-like and thematically relevant AOI's.

[1]Itti, 2000; [2]Harel, 2011; [3]Cristino et al., 2010

Study 1: Results

Differences in % of Fixation Time

Thematically Relevant AOI	
Prob 1	Correct = Incorrect
Prob 2	Correct > Incorrect
Prob 3	Correct > Incorrect
Prob 4	Correct > Incorrect
Prob 7	Correct > Incorrect
Prob 10	Correct > Incorrect

Novice-Like AOI	
Prob 1	Incorrect > Correct
Prob 2	Incorrect > Correct
Prob 3	Incorrect > Correct
Prob 4	Incorrect = Correct (p=.058)
Prob 7	Incorrect > Correct
Prob 10	Incorrect > Correct

One-way ANOVA α =.05

Study 1: Conclusions

% of fixation time of …

- ✓ correct solvers > incorrect solvers in thematically relevant areas for five of six problems.
- ✓ Incorrect solvers > correct solvers in novice-like areas for five of six problems.

- Incorrect solvers are attending to irrelevant elements in the problems' diagrams.
- Can visual cues can help direct attention to thematically relevant areas?
- Have record of correct solvers' eye movements to use when designing visual cues.

Study 2: Explore the Use of Attentional Cueing to Improve Problem Solving

Hypothesis

Visual cues based on expert eye movements can improve problem solving in physics

Study 2: Research Questions

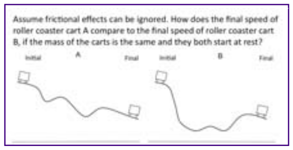
- Can a six-second visual cue modeled after expert eye movements really help students answer physics questions?

If 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?

(1) The cart A is moving faster at the final position
(2) The cart B is moving faster at the final position
(3) Carts A and B have the same speed at the final position
(4) There is not enough information to decide

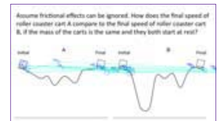
Study 2: Research Questions

- Can a 6-second visual cue modeled after expert eye movements really help students answer physics questions?
- Does students' ability to answer transfer problems improve after seeing visual cues?

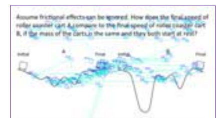


Study 2: Research Questions

- Can a 6-second visual cue modeled after expert eye movements really help students answer physics questions?
- Does students' ability to answer transfer problems improve after seeing visual cues?
- Do cues influence students' eye movements on current and subsequent problems?



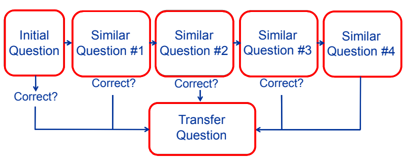
Eye movements during cues



Eye movements after cues have ended

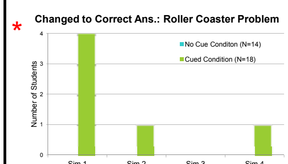
Study 2: Method

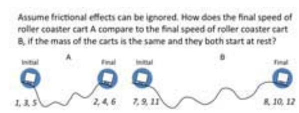
- 1st and 2nd semester algebra-based physics students
- Cued group (N=22) and No Cue group (N=23)
- Online pre-test
- Dynamic cues modeled after experts' eye movements.



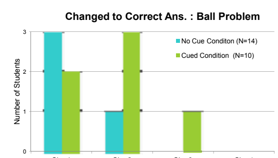
Study 2: Results – Answers on Similar Problems

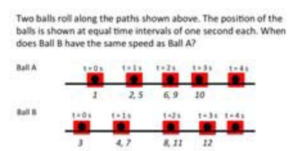
*





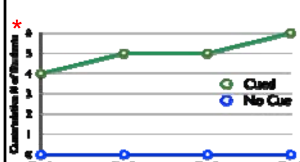
*Mann-Whitney U Test, p=.002

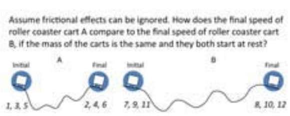




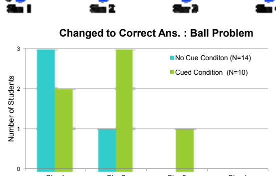
Study 2: Results – Answers on Similar Problems

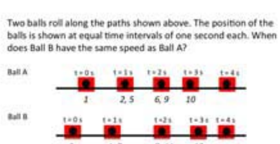
*



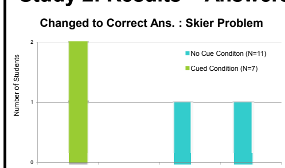


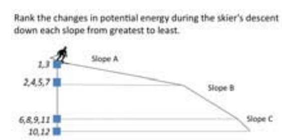
*Mann-Whitney U Test, p=.002

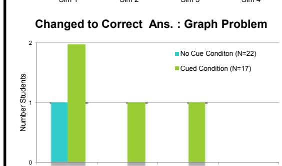


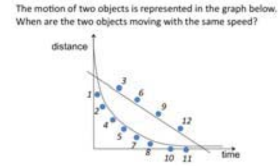


Study 2: Results – Answers on Similar Problems (cont'd)



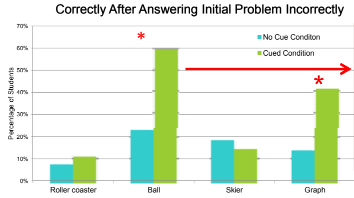






Study 2: Results – Answers on Transfer Problem

% of Students Who Answered Transfer Problem Correctly After Answering Initial Problem Incorrectly

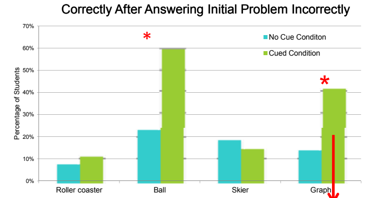


When does Ball B have the same speed as Ball A?

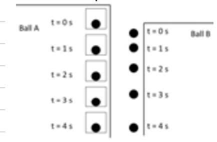
Ball transfer problem.

Study 2: Results – Answers on Transfer Problem

% of Students Who Answered Transfer Problem Correctly After Answering Initial Problem Incorrectly



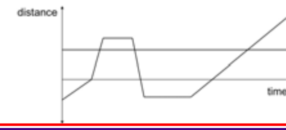
When does Ball B have the same speed as Ball A?



Ball transfer problem.

When are the two objects moving with the same speed?

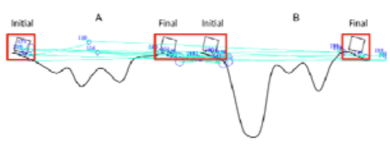
Graph transfer problem.



Study 2: Results – Eye Movements During Cues

% of Saccades in Pattern Similar to Cues: During Cues

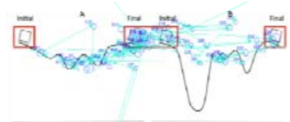
Cued Correct	Cued Incorrect
86%	46%



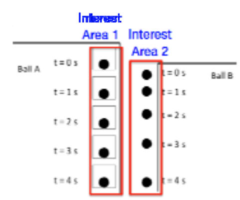
Eye movements during cues.

Study 2: Results – Eye Movements After Cues

After cues ceased: no difference between cued and no cued in % of saccades similar to cues



No difference between cued and no cued in % of saccades on transfer prob.



Study 2: Conclusions

- In one case, short duration visual cues helped students answer conceptual physics problems they were previously unable to.
- Visual cues can influence transfer problem performance.
  - Those who saw visual cues answered ball and graph transfer problems more correctly.
- Following cues closely with eyes is related to getting correct answer in one problem.
- Seeing visual cues does not seem to influence eye movements after cues cease (roller coaster) or on transfer problem (ball).

Future Studies

- Increase number of similar problems: Six (instead of four) per set.
- Increase cue duration: 50 seconds (Boucheix, 2010)
- Assure participant looking at first step of cue before cues begin to move : Gaze contingent cues
- Have an additional condition: cued and told that cue will be useful
- One no cue group + Three cued groups:
  - Selection cued,
  - Organization cued, and
  - Integration cued

[1]Boucheix, 2010

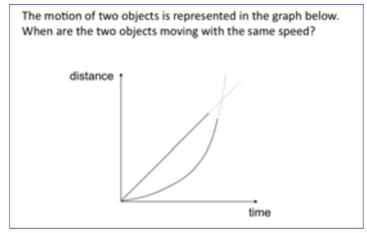
Thank you.
srebello@phys.ksu.edu

Paper is on NARST CD

Future Studies (cont'd)

One no-cue group + Three cued groups, 15-20 participants per group

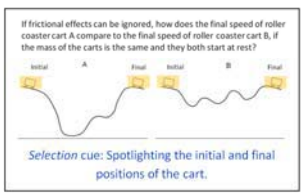
- Selection



Future Studies (cont'd)

One no-cue group + Three cued groups, 15-20 participants per group

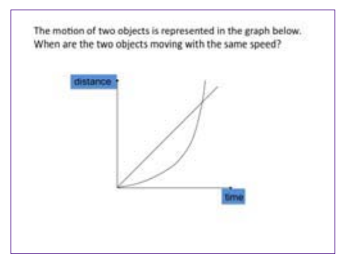
- Selection



Future Studies (cont'd)

One no-cue group + Three cued groups, 15-20 participants per group

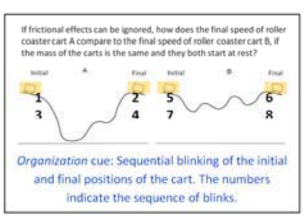
- Selection
- Organization



Future Studies (cont'd)

One no-cue group + Three cued groups, 15-20 participants per group

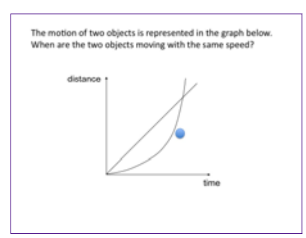
- Selection
- Organization



Future Studies (cont'd)

One no-cue group + Three cued groups, 15-20 participants per group

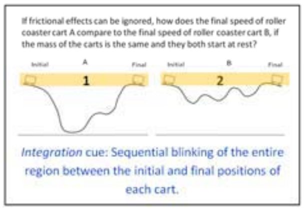
- Selection
- Organization
- Integration



Future Studies (cont'd)

One no-cue group + Three cued groups, 15-20 participants per group

- Selection
- Organization
- Integration



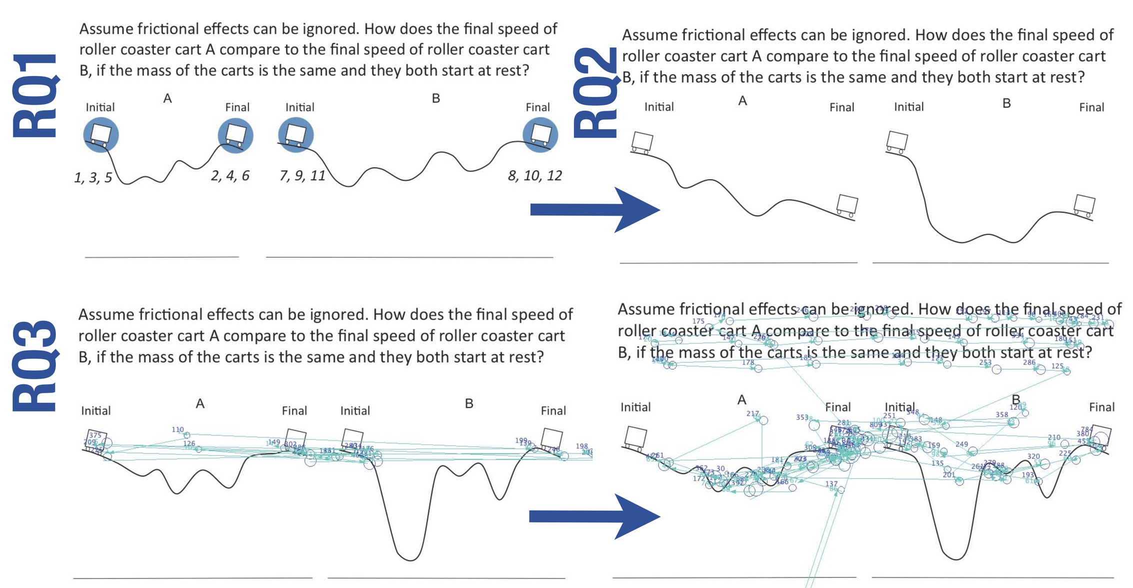
Influence of Visual Cues on Eye Movements and Reasoning in Physics Problems

Adrian Madsen, Adam Larson, Amy Rouinfar, Allison Coy, Lester Loschky & N. Sanjay Rebello, *Kansas State University*

Can a 6-second visual cue modeled after experts' eye movements overlaid on a physics question really help students come to the right answer?

RESEARCH QUESTIONS

1. Do visual cues modeled after experts' eye movements help students answer physics questions?
2. Does students' ability to answer transfer problems improve after seeing visual cues?
3. Do cues influence students' eye movements on current and subsequent problems?



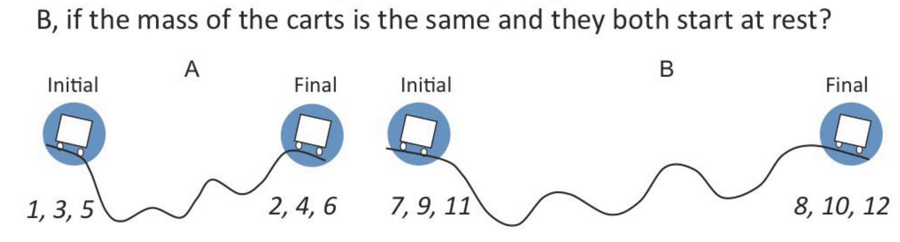
METHOD

Online pre-test (4 open-ended questions on kinematics & energy) → Instructions & calibration of eye tracker → Initial Question → Similar Question #1 → Similar Question #2 → Similar Question #3 → Similar Question #4

Correct? → Transfer Question

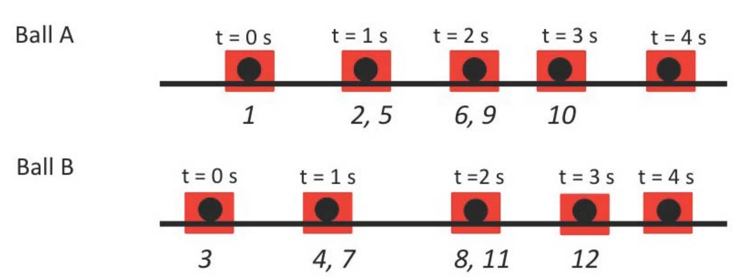
Students in cued group saw visual cues on "similar" problems.

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?



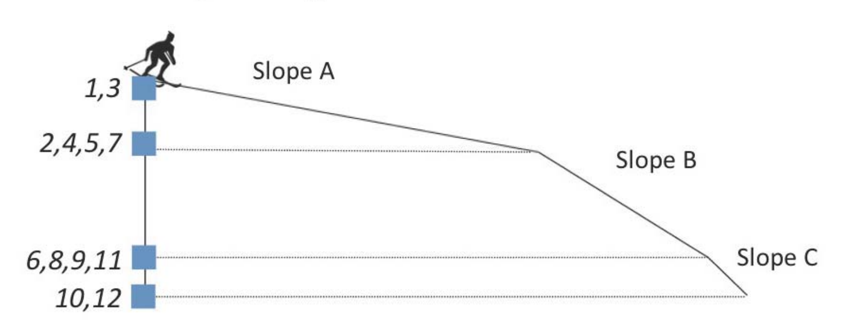
Problem 1: "Roller coaster" problem

Two balls roll along the paths shown above. 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?



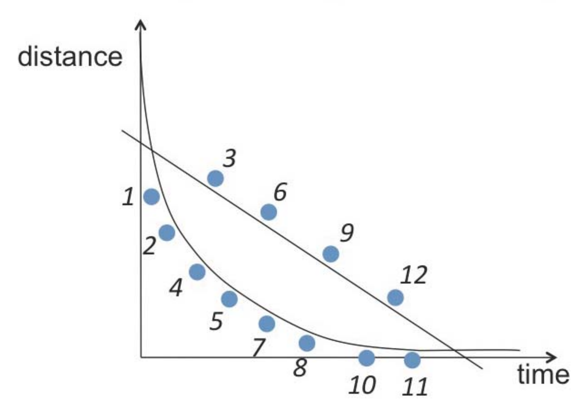
Problem 2: "Ball" problem

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



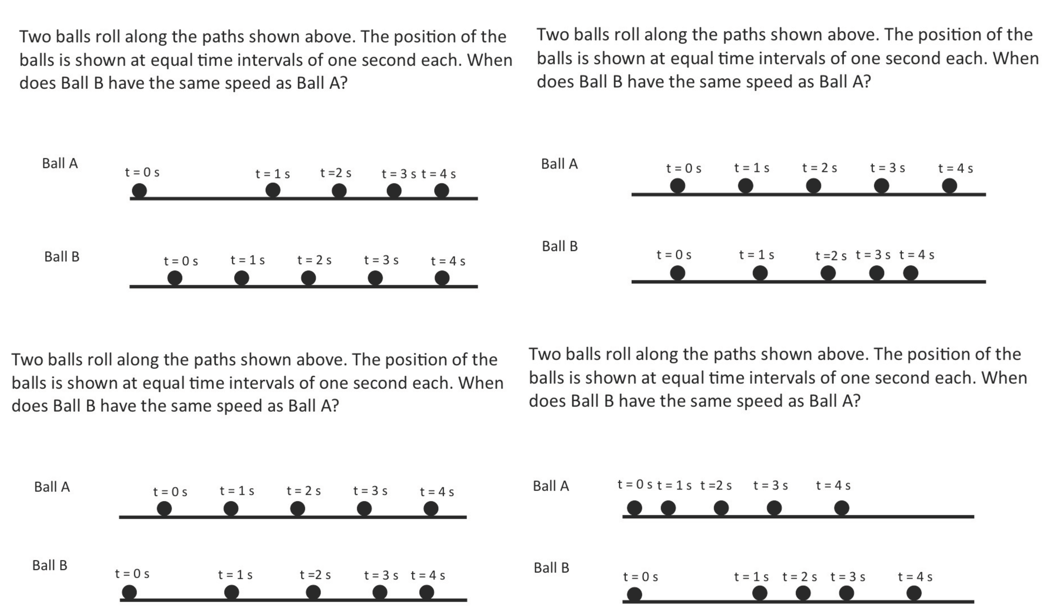
Problem 3: "Skier" problem

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

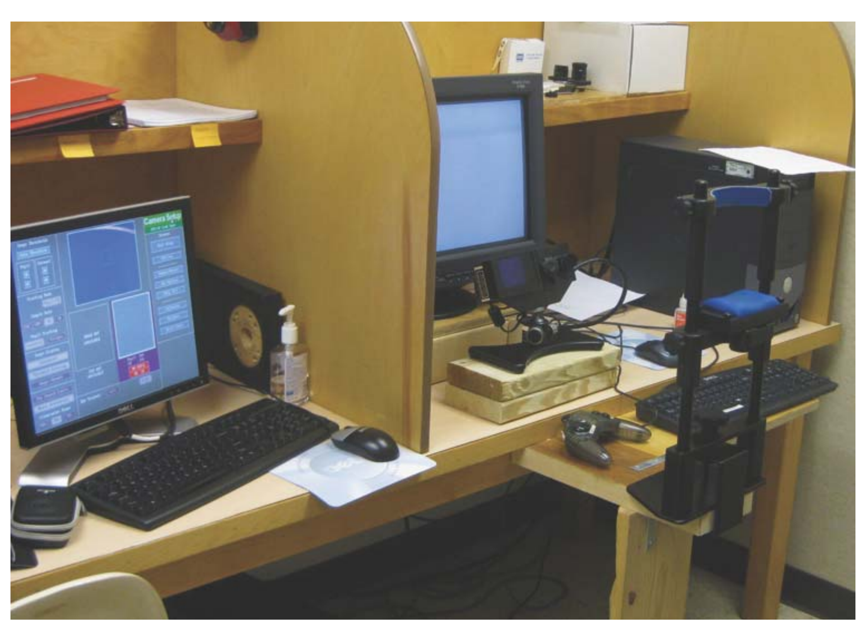


Problem 4: "Graph" problem

Colored shapes appeared for 500 ms each, for a total cueing time of 6 seconds. Numbers in italics represent order in which cues appeared.



Four "similar" problems for ball problem set. Similar problems have the same problem statement, but different diagrams.

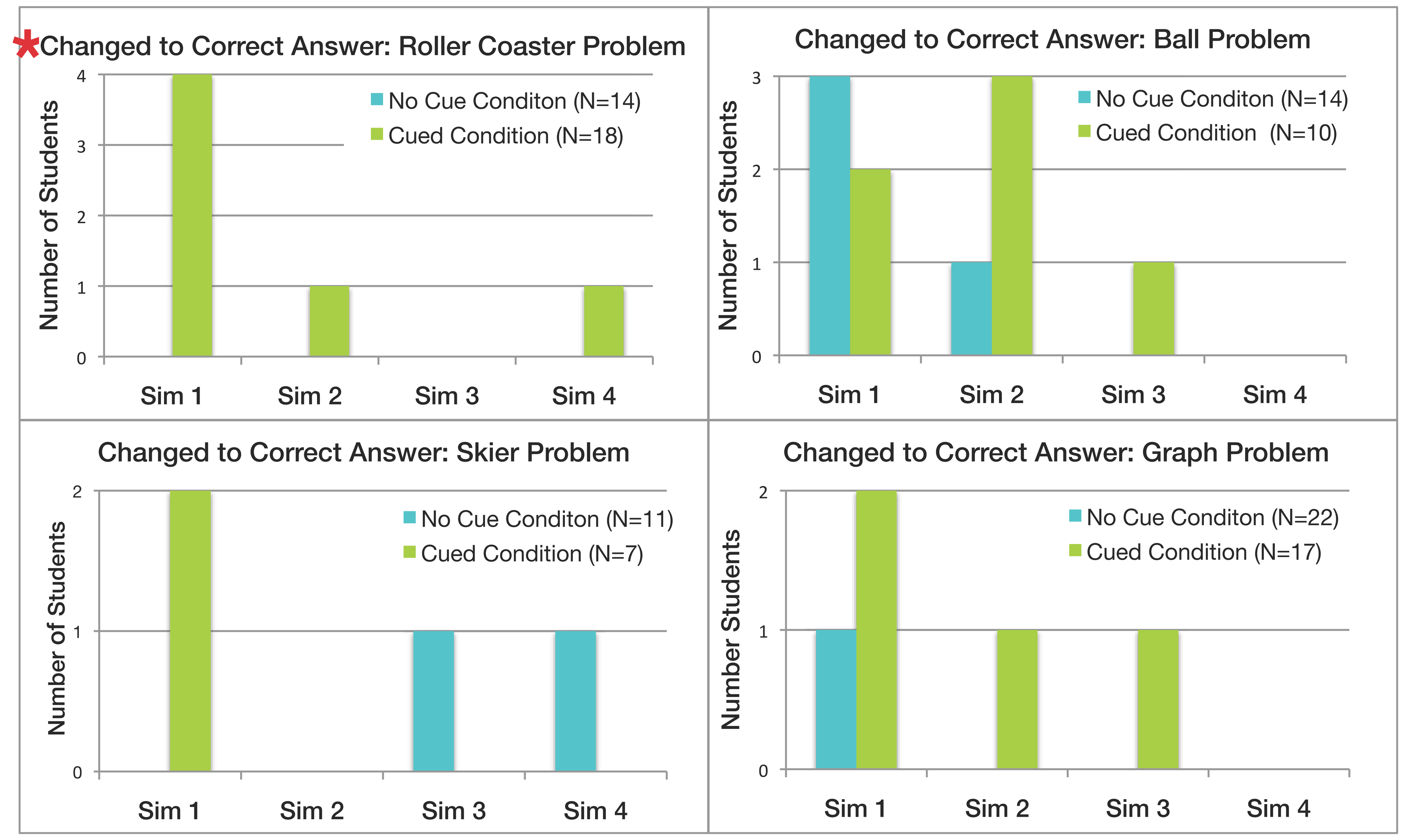


Eye tracker used for study.

RESULTS

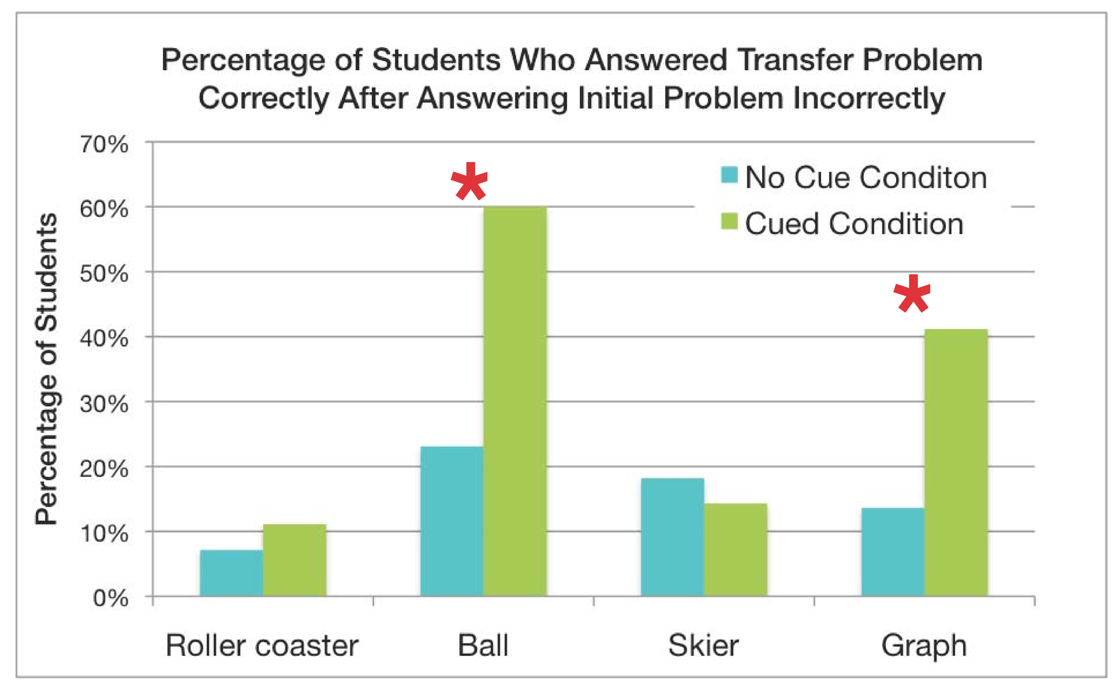
Correctness of Similar Problems

- **Students who answered initial problem incorrectly saw up to four similar problems.**
- **Graphs to the right show number of students who answered one of similar problems correctly.**
- **Significant difference between number of students in each group who answered roller coaster "similar" problems correctly. (Mann-Whitney U test p=.002)**



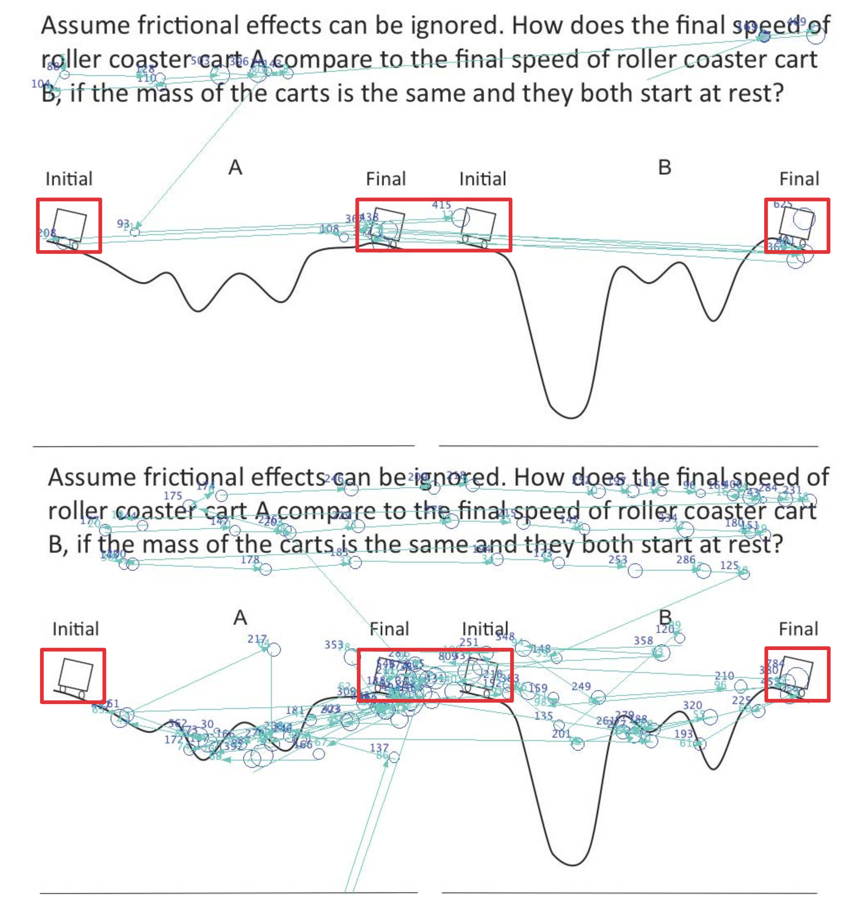
Correctness of Transfer Problems

- **After giving correct answer on similar problem, students saw transfer problem without cues.**
- **Graph below shows % of students who answered transfer correctly after answering initial incorrectly.**
- **Nearly significant difference on ball transfer problem (p=.06) and graph transfer problem (p=.054).**



Eye Movements: Roller Coaster Problem

- **Determined number of horizontal saccades between roller coaster carts while viewing cues and after cues ended.**
- **Significant difference in % of saccades following cues for those in cued group who answered correctly versus incorrectly (F(1,14)=10.8, p=.005).**
- **No difference in horizontal saccades between carts after cues ended between groups.**



	Cued	No Cue
Percentage of Saccades in Pattern Similar to Cues: *During Cues*	52.6%	0.96%* (Did not see cues)
	Cued: Correct	Cued: Incorrect
	85.5%	46.4%*

	Cued	No Cue
Percentage of Saccades in Pattern Similar to Cues: *After Viewing Cues*	2.1%	1.0%

CONCLUSIONS

- **In some cases, short duration visual cues can help students answer conceptual physics questions that they were previously unable to answer (roller coaster problems).**
- **Visual cues can influence transfer problem performance. Those who saw visual cues answered ball and graph transfer problems more correctly.**
- **Following cues closely with eyes is related to getting correct answer on roller coaster problems.**
- **Seeing visual cues doesn't seem to influence eye movements after cues cease on roller coaster problems.**

*This work is supported by KSU NSF GK-12 Program under NSF DGE-0841414.

Wednesday, July 27, 11

Using Scan Match Scores to Understand Differences in Eye Movements Between Correct and Incorrect Solvers in Physics

Adrian Madsen, Adam Larson, Lester Loschky & N. Sanjay Rebello, *Kansas State University*

OBJECTIVE

Understand how top-down and bottom-up processes influence incorrect problem solvers in physics.

PROBLEM: Consistent Wrong Answer Patterns in Physics

Two explanations types for consistent wrong answer patterns in physics: cognitive and perceptual.

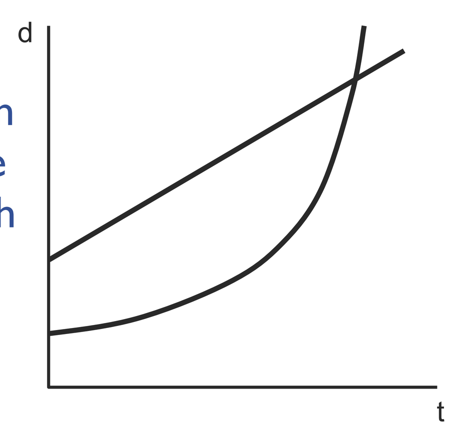
Cognitive
- Misconceptions based on naïve theories[1]
- Misapplication of resources[2]
- Miscategorized ontology[3]

Top-down Processes

Perceptual
- Attention initially caught by perceptually salient, plausible & relevant elements.
- Student answers based on perceptually salient elements.[4]

Bottom-up Processes

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



METHOD

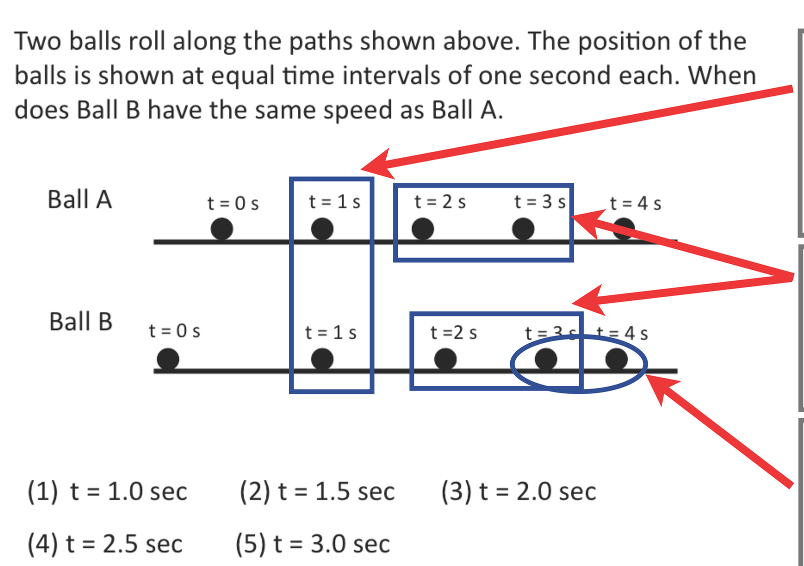
Participants: 10 PhD students in physics with teaching experience and 14 introductory psychology students who have taken a physics course.

Eye Tracker: Eye Link 1000 desktop mounted eye tracker.

Instructions and calibration of eye tracker → Answer 10 multiple-choice questions while eye movements recorded → Explain reasoning for answers while watching playback of eye movements

BACKGROUND: Areas of Interest Analysis of Eye Movements

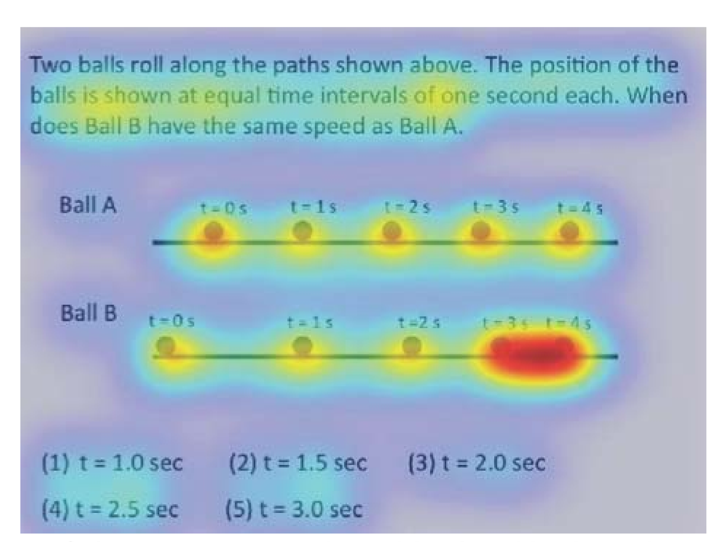
Determined % fixation time for novice-like, thematically relevant and perceptually salience areas of interest (AOI) for full problem duration. Compared for correct and incorrect solvers.



Novice-like areas from interviews (confirmed by previous research)

Thematically relevant areas defined by experts

Areas of high perceptual salience[1-2]



Saliency map produced by Itti, Koch and Niebur algorithm.[5]

Results of AOI Analysis (full problem duration)

	% of Fixation Time
Thematically Relevant AOI	Correct > Incorrect (5 of 6 problems)
Novice-Like AOI	Incorrect > Correct (5 of 6 problems)
Perceptually Salient AOI	No differences found

Further Analysis of Perceptually Salient AOI

- Compared % fixation time for first 2 seconds of viewing diagram.
  - Effects of perceptual salience most pronounced < 2 sec.
- No statistically significant differences, but raw % of fixation time > for incorrect solvers on 5 of 6 problems.
- No effect or lacking statistical power?

Evidence for top-down processes primarily influencing attention of correct and incorrect solvers.

REFERENCES

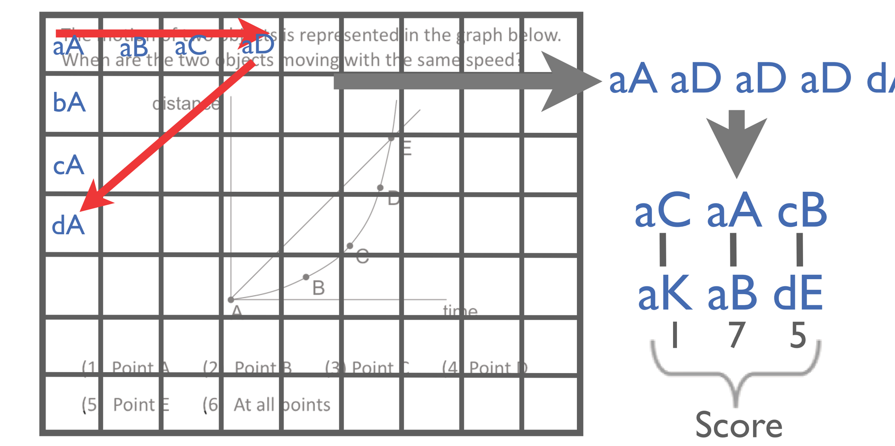
1. L. C. McDermott, M. L. Rosenquist, and E. H. van Zee, American Journal of Physics 55 (1987).
2. D. Hammer, American Journal of Physics 68 (2000).
3. M. T. H. Chi, in Cognitive Models of Science: Minnesota Studies in the Philosophy of Science, edited by R. Giere (University of Minnesota Press, Minneapolis, MN, 1992), pp. 129
4. A. F. Heckler, in Psychology of Learning and Motivation: Cognition In Education, edited by J. P. Mestre and B. H. Ross (Academic Press, Oxford, UK, 2011)
5. L. Itti, C. Koch, and E. Niebur, Pattern Analysis and Machine Intelligence, IEEE Transactions 20 (1998).
6. F. Cristino, S. Mathot, J. Theeuwes, and I.D. Gilchrist, Behavior Research Methods 42 (2010).

EIDROP An NSF GK-12 Program Kansas State University

ANALYSIS: ScanMatch

- ScanMatch: scan path analysis based on Needleman-Wunsch algorithm used for DNA sequencing.[6]
- Algorithm converts scan path to letter sequence and compares pairs of sequences, seeking optimal alignment by maximizing similarity score.
- Higher score indicates scan paths with strong similarity temporally and spatially.
- Calculated ScanMatch score comparing:
  - Correct solvers to one another (CC)
  - Incorrect solvers to one another (II)
  - Correct to incorrect solvers (CI)

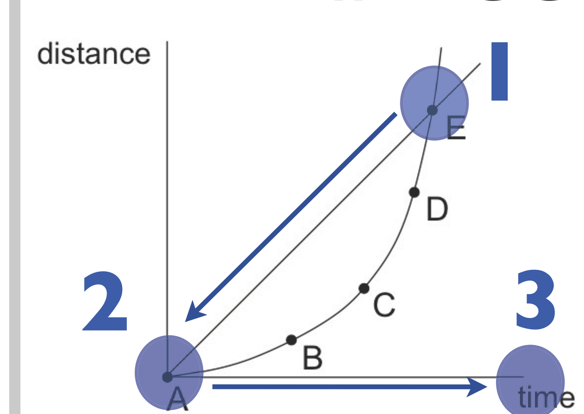
- Letter sequence binned temporally.
- Elements in each sequence compared & scored based on distance apart in grid.
- Gaps (& gap penalties) included to maximize score.
- Score normalized to maximum of one.



HYPOTHESES

Hypothesis 1: If bottom-up processes based on perceptual salience of primarily influence attention of incorrect solvers:

II > CC > CI

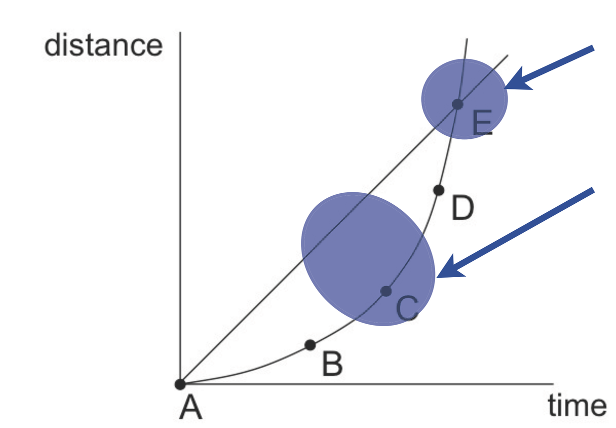


Example scan path if perceptual salience is primary influence.

- Salience models predict order & location of fixations = higher ScanMatch scores
- Correct solvers attend to similar regions of diagram in varying order = lower ScanMatch scores

Hypothesis 2: If top-down processes utilizing naïve theories primarily influence attention of incorrect solvers:

II = CC, II & CC > CI



Incorrect solvers fixate most in novice-like AOI.

Correct solvers fixate most in thematically relevant AOI.

Both attend to other elements in varying order.

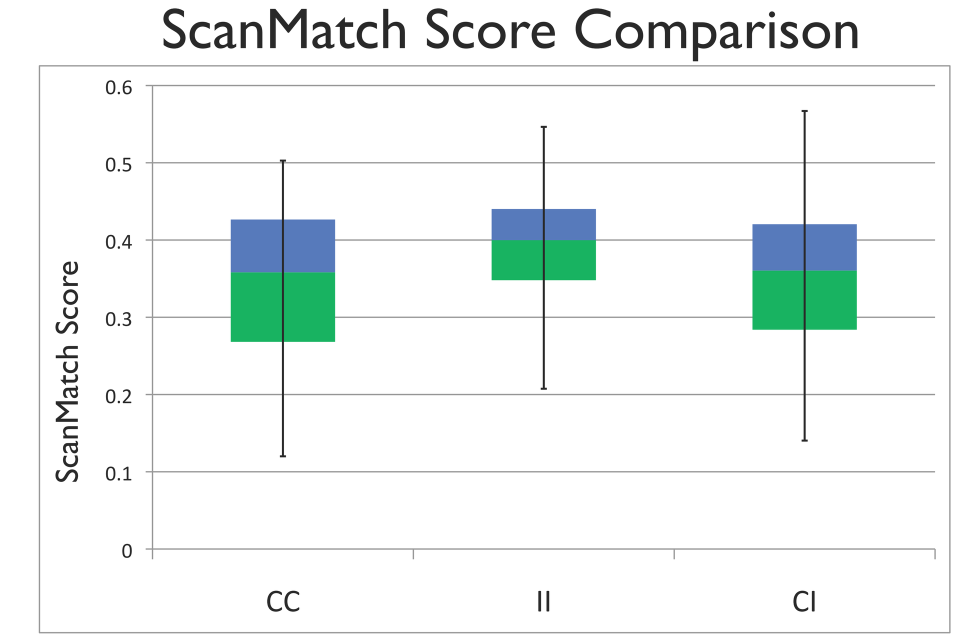
RESULTS

Completed one-way ANOVA comparing ScanMatch scores of CC, II & CI comparisons for each problem. If significant result obtained, used post-hoc contrasts to determine statistically significant comparisons.

Summary of Significant Differences

Problem	Comparison
1	II > CI (p<.001)
2	II > CI (p<.001)
	II > CC (p=.005)
3	no significant differences
4	
7	
10	II > CI (p=.05)

ScanMatch Score Comparison



Box and whiskers plot with median, max, min and 1st and 3rd quartile of the ScanMatch scores for each comparison

CONCLUSION

- **Found evidence for top-down naive theories primarily influencing attention of incorrect solvers.**
  - No differences between CC and II comparisons on 5 of 6 problems.
  - Consistent with previous finding, incorrect solvers greater % fixation time in novice-like AOI
- Did not find CC & II > CI as expected. Differences in way participants read elements of problem may lead to noise in the data.

Using Scan Match Scores to Understand Differences in Eye Movements Between Correct and Incorrect Solvers

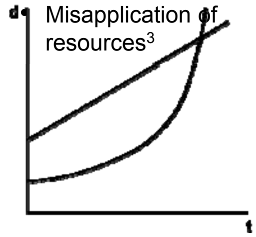
Adrian Madsen
adrianc@ksu.edu

Adam Larson, Lester Loschky and N. Sanjay Rebello
Kansas State University

Funded in part by the NSF GK-12 Program

Consistent Wrong Answer Pattern in Physics

- Cognitive Explanations
- Misconceptions[1,2]
- Misapplication of resources[3]



Perceptual Explanations

•Attention is initially caught by perceptually salient and plausible elements.

•Student answers are based on perceptually salient elements.[4]

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

[1]Trowbridge and McDermott 1980; [2]McDermott et. al 1987; [3]Hammer, 2000; [4]Heckler, 2011

Consistent Wrong Answer Pattern in Physics

Cognitive Explanations

•Misconceptions[1,2]

•Misapplication of resources[3]

Perceptual Explanations

•Attention is initially caught by perceptually salient and plausible elements.

•Student answers are based on perceptually salient elements.[4]

Top-down processes

Bottom-up processes

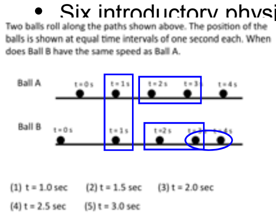
[1]Trowbridge and McDermott 1980; [2]McDermott et. al 1987; [3]Hammer, 2000; [4]Heckler, 2011

Consistent Wrong Answer Pattern in Physics

How do top-down and bottom-up processes influence incorrect problem solvers in physics?

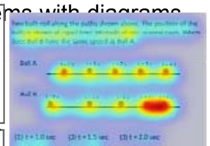
Study Design

- 10 Ph.D. physics students and 14 intro psychology students who have taken at least high school physics
- Six introductory physics problems with diagrams



Novice-like areas from interviews (confirmed by previous research)

Thematically relevant areas defined by experts



Areas of high perceptual salience[1-2]

[1]Itti, 2000; [2]Harel, 2011

Area of Interest (AOI) Analysis

- Determined percentage of fixation time novice-like, thematically relevant and perceptually salience AOIs for full problem duration.
  - Compared for correct and incorrect solvers.

	% of Fixation Time
Thematically Relevant AOI	Correct > Incorrect (5 of 6 problems)
Novice-Like AOI	Incorrect > Correct (5 of 6 problems)
Perceptually Salient AOI	No differences found

Area of Interest (AOI) Analysis

7

- Determined percentage of fixation time novice-like, thematically relevant and perceptually salience AOIs for full problem duration.
  - Compared for correct and incorrect solvers.

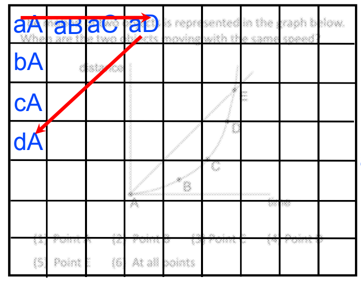
Evidence for dominance of top-down processes

	% of Fixation Time
Thematically Relevant AOI	Correct > Incorrect (5 of 6 problems)
Novice-Like AOI	Incorrect > Correct (5 of 6 problems)
Perceptually Salient AOI	No differences found

ScanMatch Analysis

8

- Scan path analysis based on Needleman-Wunsch algorithm.



Normal Sequence:
aA aD dA

Temporal binning (50 ms bins):
aA aA aD aD aD dA

[1]Cristino, Mathot, Theeuwes, Gilchrist, 2010

ScanMatch Analysis

9

- Algorithm seeks optimal alignment of letter sequences by maximizing similarity score.

aC aA cB
cD aK aB dE

5

aC aA cB

2

Similarity Matrix: predefined matrix of scores based on distance

[1]Cristino, Mathot, Theeuwes, Gilchrist, 2010

Similarity Score Comparisons

10

Correct solvers to all other correct solvers	CC	Incorrect solvers to all other incorrect solvers	II

- If incorrect solvers are more heavily influenced by bottom-up processes, II > CC[1-2]
- If top-down processes are dominant for incorrect solvers, CC = II

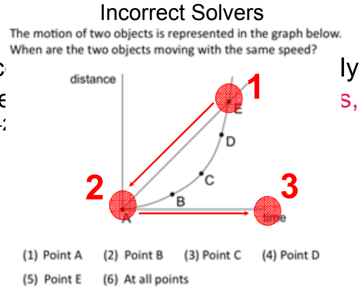
[1]Hegarty, Canham, Fabrikant, 2010; [2] Itti, 2000

Similarity Score Comparisons

11

- If inc... ly influe... s, II > CC[1-2]

Incorrect Solvers



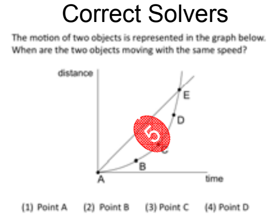
[1]Hegarty, Canham, Fabrikant, 2010; [2] Itti, 2000

Similarity Score Comparisons

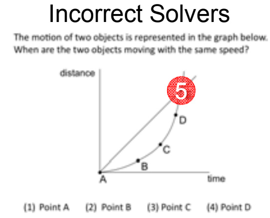
12

- If top-down processes are dominant for incorrect solvers, CC = II

Correct Solvers



Incorrect Solvers



[1]Hegarty, Canham, Fabrikant, 2010; [2] Itti, 2000

13

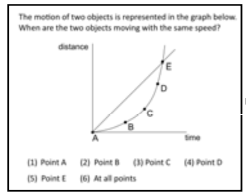
Evidence for Dominance of Top-Down Processes

- Found no significant differences between CC and II on 5 of 6 problems.
- Consistent with previous findings that incorrect solvers spend more % of time in novice-like AOI.

14

Future Work

Current Study



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

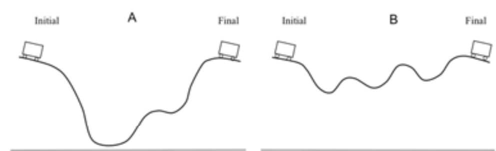
15

Thank you.

adrianc@ksu.edu

16

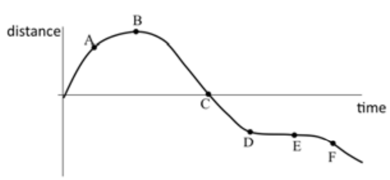
If frictional effects can be ignored, how does the final speed of roller coaster cart A compare to the final speed of roller coaster cart B, if the mass of the carts is the same and they both start at rest?



(1) The cart A is moving faster at the final position
(2) The cart B is moving faster at the final position
(3) Carts A and B have the same speed at the final position
(4) There is not enough information to decide

17

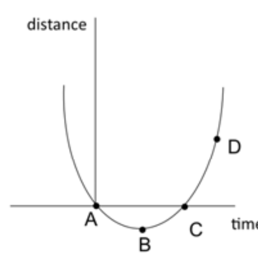
At which point on the graph is the object turning around (moving away then coming back)?



(1) A (2) B (3) C (4) D (5) E (6) F

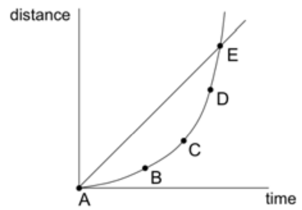
18

When is the speed of the object shown in the graph zero?



(1) Point A (2) Point B (3) Point C (4) Point D
(5) Point E (6) At all points

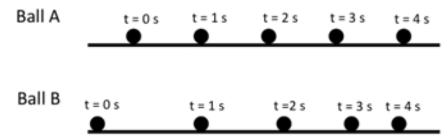
The motion of two objects is represented in the graph below. When are the two objects moving with the same speed?



(1) Point A (2) Point B (3) Point C (4) Point D
(5) Point E (6) At all points

20

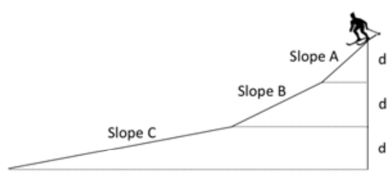
Two balls roll along the paths shown above. The position of the balls is shown at equal time intervals of one second each. When does Ball B have the same speed as Ball A.



(1) t = 1.0 sec (2) t = 1.5 sec (3) t = 2.0 sec
(4) t = 2.5 sec (5) t = 3.0 sec

21

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



(1) $\Delta PE_A > \Delta PE_B > \Delta PE_C$
(2) $\Delta PE_C > \Delta PE_B > \Delta PE_A$
(3) $\Delta PE_A = \Delta PE_B = \Delta PE_C$
(4) $\Delta PE_A = \Delta PE_B > \Delta PE_C$
(5) $\Delta PE_B > \Delta PE_C = \Delta PE_A$

Exploring Visual Cueing to Facilitate Problem Solving in Physics

PI: N. Sanjay Rebello
Co-PI: Lester C. Loschky
Kansas State University
Project Award Number: 1138697

Annual Report for Year I: 2011-2012

PROJECT ACTIVITIES

During the first year of this grant we completed the following project activities:

- Completed **STUDY 1**
  - For METHODOLOGY see below.
  - For Results of Study 1 see FINDINGS.
- Completed **STUDY 2** (**PILOT**)
  - For METHODOLOGY see below.
  - For Results of Study 2 see FINDINGS.
- Met with the **Advisory Board** via Skype: See details in FINDINGS.
- **Published** the findings of **STUDY 1**: See reprint of journal paper in FINDINGS.
- **Presented** the findings of **STUDY 1&2** at Conferences: See details in FINDINGS.

STUDY 1: Differences in visual attention between those who correctly and incorrectly answer physics problems

METHODOLOGY

In order to define areas of a physics problem diagram that contain visual information related to a novice-like misconception, we conducted individual interviews (STUDY 1A) with students enrolled in an introductory psychology course. We specifically looked at the interview segments where participants provided incorrect answers to the physics problems and observed the areas of the diagram that students identified and discussed while giving their verbal explanation. This information was used to define "novice-like" areas of interest (AOI), or specific areas of the diagram in which a participant who answered incorrectly would use to come to their answer. These areas of interest were used in the analysis for STUDY 1B in which we tracked students eyes as they solved the problems.

STUDY 1A: INTERVIEWS TO DETERMINE NOVICE LIKE AREAS OF PROBLEMS

Participants: The participants were 13 students (eight females) enrolled in an introductory psychology course. All of the students had taken at least one physics course in high school, though some had taken an introductory physics course at the university level as well. They were given course credit for participation.

Materials: The materials consisted of 10 multiple-choice conceptual physics problems covering various topics in introductory physics including energy, kinematics, and graphing of motion. Each problem contained a diagram that had a thematically relevant visual component that students needed to attend to in order to correctly answer the question. For example, in Problem 4 (see Appendix), to compare the speeds of ball A and ball B, one must attend to the distances between the balls at each time interval and ignore the point where the balls are aligned spatially. So, the distance between balls at two seconds and three seconds is the relevant area to attend to. These problems were chosen based on prior experience of the researchers which indicated that these problems could be answered using common naïve conceptions documented in physics education literature.

Procedure: Each participant took part in an individual session which was between 20 and 40 minutes long. At the beginning of the session, participants were given a short explanation of the goal of the interview and the purpose of the research. Further, they were instructed to think aloud and explain their reasoning process as they answered each question. They were told they might be asked additional clarifying questions during their explanations. Participants were given one problem at a time, each printed on an 8 1/2 x 11 sheet of paper. They were allowed to write or draw on the problems as they deemed necessary. If a participant's answer was not clear, the interviewer asked questions to clarify the meaning of the explanation. Participants' verbal explanations, gestures, and writing on the paper were recorded with a Flip video camera.

Analysis: The purpose of these interviews was to determine which portion of each diagram was attended to by incorrect problem solvers. Therefore, only the interview segments where the participant gave a final incorrect answer were included in the analysis. A phenomenological approach was used to code the interviews. Four of the 10 problems used in the interviews showed no consistent answering patterns among incorrect solvers after a first pass analysis. These problems are not included here, as there were no identifiable novice-like areas to be utilized in Study 1B.

STUDY 1B: DETERMINING DIFFERENCES IN VISUAL SELECTIVE ATTENTION BASED ON CORRECTNESS OF PROBLEM SOLUTION

Participants: There were 24 participants in the study (three females, two were graduate students and one was a psychology student) with two different levels of experience in physics. Ten participants were first-year through fifth-year PhD students in physics who had either taught an introductory physics course or been a teaching assistant for an introductory physics lab. One participant was a postdoctoral candidate in physics who had received his PhD within the last two years and had teaching experience. Thirteen participants were enrolled in an introductory psychology course and had taken at least one physics course in high school, though some had also taken an introductory physics course at the university level. The PhD students and post-doc participated as volunteers and the psychology students received course credit for their participation. Because we were looking to compare those who answered the physics problems correctly to those who answered incorrectly, we selected participants with a broad range of experience. We expected that the PhD students would answer correctly, while the psychology students might answer incorrectly, though we knew that this might not always be the case since there is a wide distribution of expertise among introductory physics students and physics graduate students.

Materials: The materials consisted of the six multiple-choice introductory physics problems analyzed in Study 1A.

Apparatus: Participants were presented with physics problems on a computer screen viewed at a distance of 24 inches using a chin and forehead rest to minimize participants' extraneous head movements. The resolution of the computer screen was set to 1024 x 768 pixels with a refresh rate of 85 Hz. Each physics problem subtended 33.3 x 25.5 of visual angle. Eye movements were recorded with an EyeLink 1000 desktop mounted eye-tracking system (http://www.sr-research.com), which had an accuracy of less than 0.50of visual angle. An eye movement was classified as a saccade (i.e., in motion) if the eye's acceleration exceeded 8,500°/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.

Procedure: Each participant took part in an individual session lasting 20-40 minutes. At the beginning of the session, participants were given a short explanation of what to expect in the study. After calibrating the eye tracking system, if the validation's mean error was ≤ 0.50 of visual angle, the experiment began, otherwise the calibration and validation was repeated until successful. Next, the participant was instructed to silently answer 10 multiple-choice questions while their eye movements were recorded. Participants indicated their answer to each question using number keys on the keyboard. Between questions, a calibration drift correction procedure was done to ensure proper calibration throughout the experiment. This procedure required the participant to fixate on a small white dot in the middle of a gray screen and press a key. Pressing the key caused the screen to advance to the next problem when the participant's fixation was within a pre-defined area around the white dot. Finally, each participant was asked to provide a cued verbal retrospective report for which they were shown a replay of their eye movements on each problem and asked to explain their thought processes (either after watching the replay of their eye movements or concurrently while watching them). This method has been found to produce more in-depth explanations than without viewing one's eye movements. If a participant's explanation was unclear, they were asked follow-up questions. Participants were given unlimited time to answer the questions and provide retrospective verbal reports. Verbal explanations and gestures were recorded with a Flip video camcorder.

Analysis: Two kinds of analyses were conducted:

AOI Analysis: To analyze participants' eye fixations, we defined areas of interest (AOIs) for specified areas of each diagram. These AOIs were used to determine the total fixation time, (i.e., the total amount of time the participant spent fixating on a given AOI). There were three different types of AOIs identified for each physics problem analyzed in Study 1A. These types were thematically relevant AOIs, perceptually salient AOIs, and novice-like AOIs. The definition for the thematically relevant AOI came from three independent raters, one physics professor, and two PhD students in physics, who indicated, on each of the problems, the area which contained visual information necessary to answer the problem. The definition for the perceptually salient AOI in each problem was determined using an implementation of the Itti, Koch and Niebur saliency map algorithm in MATLAB. This MATLAB toolbox produced a heat map representation of relative saliency over the entire diagram for each problem. The area on the diagram with the highest rating of saliency was used to define the perceptually salient AOI. If there were several portions of the diagram with the highest level of perceptual salience, according to the salience map, then all of these areas were used when defining the perceptually salient AOI.

Scan Path Analysis: We expanded on our previous analysis to further investigate the role of perceptual salience in guiding the attention of those who incorrectly answer conceptual physics questions containing a diagram. A scan path analysis was per-formed using an algorithm called ScanMatch (Cristino et al. 2010) which is based on the Needleman-Wunsch algorithm used to compare DNA sequences. ScanMatch bins a saccade sequence both spatially and temporally and then recodes this information to create a sequence of letters which represents the location, duration, and order of the fixations. The letter sequences of two sets of eye movements are then compared to each other to calculate a similarity score. A similarity score near one represents two sequences of eye movements that are very similar spatially and temporally. The ScanMatch analysis requires no decisions to be made about the data a priori, for example, one does not have to define AOIs based on an experimenter's definition or rating. Therefore, it is possible that differences exist in sets of eye movement data that are not detected by looking at fixation durations in AOIs.

We compared compare the average ScanMatch scores produced by comparing the correct solvers to one another (C-C comparison), the incorrect solvers to one another (I-I comparison), and the correct solvers to the incorrect solvers (C-I comparison).

We hypothesize that if the incorrect solvers are being primarily led by the perceptual salience of the elements in the diagram, then it is likely that they will attend to the same elements in a similar order. For example, attention would be first guided to the most perceptually salient region, followed by the next most salient region, and so on. Thus, the I-I comparison would have higher ScanMatch scores than the C-C comparison, who might attend to perceptually salient areas early on in diagram viewing; however, the variable onset of top-down processes on eye move-ments would result in greater temporal and spatial variability of gaze towards thematically-relevant elements in the diagram, re-sulting in lower ScanMatch scores. The I-I and C-C groups would also have higher ScanMatch scores than the C-I group, since the correct solvers and incorrect solvers are known to spend different amounts of times looking at thematically-relevant and novice-like elements.

Conversely, if top-down processes are directing the attention of incorrect solvers, namely some form of naïve theory, the ScanMatch score of the I-I comparison should be similar to that of the C-C comparison. The domain knowledge possessed by those in both comparison groups, whether correct or incorrect knowledge, guides their attention to look at certain elements of the problem, but not in a particular order. Once again, the I-I comparison and the C-C comparison should have higher ScanMatch scores than the C-I comparison.

STUDY 2 (PILOT)**: Using Dynamic Cues to Influence Reasoning** (NOTE: This is the pilot for a more detailed STUDY 2 to be conducted in the second year of the grant)

METHODOLOGY

Participants in the study were 55 individuals concurrently enrolled in an introductory algebra-based physics course. To ensure sufficient prerequisite knowledge, each completed a pre-test, which consisted of four open-ended questions gauging their understanding of speed and potential energy. The pre-tests 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.

Participants took part in individual sessions lasting between 30 and 60 minutes. 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 a verbal explanation of their reasoning when ready. Participants in the cue condition were told that colored shapes may appear on some of the problems and when these appeared, they should follow them with their eyes.

Participants were presented with physics problems on a computer screen viewed at a distance of 24 inches using a chin and forehead rest to minimize their extraneous head movements. The resolution of the computer screen was set to 1024 x 768 pixels with a refresh rate of 85 Hz. Each physics problem subtended 33.3° x 25.5°of visual angle. Eye movements were recorded with an EyeLink 1000 desktop mounted eye-tracking system which had an accuracy of less than 0.50° of visual angle. 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.

The materials consisted of four sets of conceptual physics problems covering energy and kinematics, which were found to have significant differences in the way correct and incorrect solvers answered them (Madsen et al. 2012). It should be noted that these four specific problems were chosen from the six problems analyzed because they tested distinct concepts in physics. We refer to these problems as the roller coaster, ball, skier and graph problems

Within each problem set, there was an "initial" problem, four "similar" problems, and a "transfer" problem. All problems were open-ended and contained a diagram with the necessary information to answer the problem. 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 and tested the same concept and contained a diagram with similar surface features. When the student answered a similar problem correctly, they saw 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 he/she answered the similar problem correctly or incorrectly. The surface features of the transfer problems were different than the initial and similar problems, though the concept tested was the same. All participants viewed the four sets of problems in the same order.

Whenever a student was ready to answer a problem, they indicated this by pressing any key on a keyboard, at which point the problem displayed on the computer would become slightly smaller in size (this was so that the student knew 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 pre-defined 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 cue group saw colored shapes overlaid on the similar problems appear four seconds after the problem was initially seen. Colored shapes were used because color is known to capture one's attention because of its high visual salience. Each colored shape appeared for 500 ms at 12 different positions in the diagram for a total cueing time of six seconds. The visual cues were designed to mimic the eye movements of those who answered the same problems correctly in Madsen et al. (2012). 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 could 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 thematically relevant area of interest. Similarities between participants were observed, and visual cues were modeled after these patterns. Further, the cues could have remained static and simply drawn participants' attention to the relevant areas of the problem, but we hoped by modeling the way in which correct solvers viewed the thematically relevant areas and compared elements within these areas, the cues would give the participants more insight into how to correctly answer the problems.