

# How do multimedia hints affect students' eye movements on conceptual physics problems?

Xian Wu<sup>1</sup>, John Hutson<sup>2</sup>, Lester C. Loschky<sup>2</sup> and N. Sanjay Rebello<sup>3</sup>

<sup>1</sup>*Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506*

<sup>2</sup>*Department of Psychological Sciences, Kansas State University, Manhattan, KS 66506*

<sup>3</sup>*Department of Physics & Astronomy, Purdue University, 525 Northwestern Ave., West Lafayette, IN, 47907*

**Abstract.** We investigated the effect of hint modality on students' eye movements on conceptual physics problems with diagrams. We recruited 57 students enrolled in a physics class for future elementary teachers. The participants were randomly assigned to conditions with no hints, visual hints, text hints, audio hints, and all possible hint modality combinations. We found that different hint modalities affect students' eye movements differently and the difference of students' eye movements relates to their problem-solving performance. The results of this study are different from the predictions based on Cognitive Theory of Multimedia Learning (CTML). Our results suggest that the cognitive process in physics problem solving may not be fully explained by CTML and therefore more research might be necessary in this area.

**PACS:** 01.40.Fk

## I. INTRODUCTION

Previous studies [1] have shown that learners' eye movements can reflect cognitive processes in diagram-based problem solving in physics. Correct solvers attend to the part of a diagram that is closely related to correct answer or "expert area" more closely than incorrect solvers. Correct solvers also look at the part of a diagram that is closely related to incorrect answer or "novice area" less frequently than the incorrect solvers. To investigate whether drawing a solver's visual attention to the expert areas in a problem diagram could improve problem-solving performance, a follow-up study [2] found that visual cues on the expert areas could improve performance on physics problems with diagrams. Moreover, it also found that the cued participants spent less time on the expert areas of the transfer problems than the uncued participants. All of this evidence showed that problem-solving performance on physics problems containing a diagram is related to solvers' visual attention, which could be directly manipulated by visual cues.

## II. MOTIVATION

The long-term goal of this study is to build a computer-assisted instruction in physics problem solving. To reach this goal, the modality of information should be carefully studied. In addition to visual modality, information can also be conveyed to problem solvers in auditory and textual modalities. In a multimedia-learning environment where a learner receives such information, it would be important to

understand how he/she attends to and responds to such information. To answer this question, we used eye-tracking technology to probe how solvers split their visual attention on multimedia hints, and how this process is affected by hint modalities. In conjunction with problem solving performance data, solvers' eye movements can tell us which parts of problem diagrams they attended to, and how their attention affected their performance. Our research questions are: (i) How does hint modality affect correctness of solutions on problems and subsequent transfer problems? (ii) How does hint modality affect solvers' visual attention on the diagram of the problem in which the hint is provided?

## III. THEORETICAL BACKGROUND

Mayer's Cognitive Theory of Multimedia Learning (CTML) [3] suggests that the auditory channel is more effective in conveying information than the text channel (modality principle). Additionally, presenting information via more channels is not necessarily better (redundancy principle). The modality and redundancy principles are based on one of the main theoretical assumptions of Mayer's CTML, which is the dual-channel assumption. It states that learners have two separate systems to process pictorial and verbal information. If the learner needs to process multiple pieces of information in different channels, learning is hindered. Conversely, when the learner uses multiple channels (visual and auditory) to process one piece of information, the learning process is enhanced. There are few eye-tracking studies on effects of modalities. Schmidt-Weigand and his colleagues [4] found participants spending

more time studying the diagram with narration than those with on-screen text. They also found that their participants read text more intensively in a self-paced condition compared to a computer-paced condition. These studies provide further support for CTML.

#### IV. METHOD

All of the participants (N=57) were recruited from a conceptual physics class for non-science students, mostly sophomores and juniors, at a U.S. Midwestern university. A vast majority (>90%) of the students were pre-elementary education majors, and a vast majority (>90%) were female. Very few of the students had any prior physics class in high school and none of them had a college physics class prior to this class. Students received extra credit equal to 5% of the course grade for participating in a 60-minute long interview containing four problem sets. The problem sets were based on our previous work [5], and named – Ball, Graph, Skier, and Roller Coaster. Each set consisted of one initial problem, six training problems, a near transfer problem, and a far transfer problem presented on a computer screen. The initial problems, training problems, and the near transfer problems were designed based on the same physics concept and representation but a different context. The far transfer problems used the same physics concept, but different representation and context from the other problems. See Figs. 1, 2, 4, and 5 respectively for examples of problems.

To study students’ visual attention on problems, we defined expert and novice areas of interest (AOIs). Expert AOIs are portions of the diagram relevant to the correct solution. Novice AOIs are the portions of the diagram consistent with the most common incorrect solution documented in literature and our previous study [6]. The eye tracker we used has an average error of 0.5 degrees of visual angle so both expert areas and novice areas were defined to be 0.5 degrees of visual angle from the edge of the desired regions.

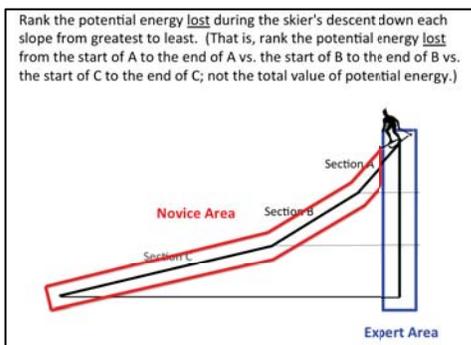


FIG 1. The initial problem from the skier problem set showing the Expert and Novice Areas.

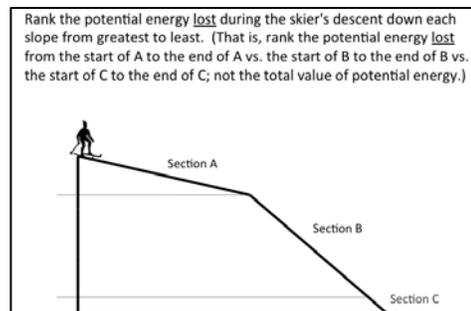


FIG 2. Example of a training problem.

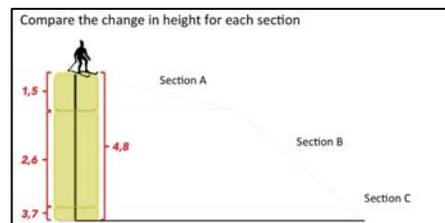


FIG 3. Visual and text hints on skier problem.

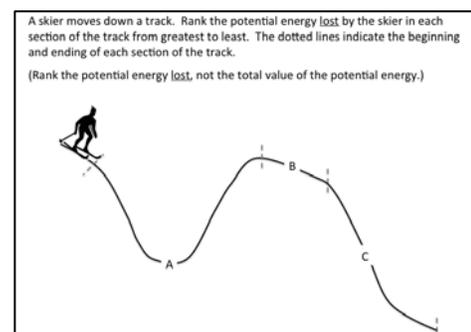


FIG 4. The near transfer problem from skier problem set.

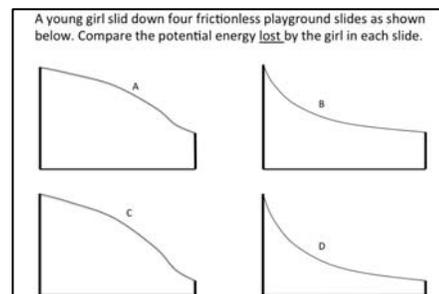


FIG 5. The far transfer problem from skier problem set.

Our goal was to test the effects of multimedia hints modalities of on students’ eye movements during problem solving. Specifically, we wanted to test the effects and interactions between text hints and audio hints with visual hints present. A 2(text hint / no text hint) × 2(audio hint / no audio hint) full factorial design was used. Participants were randomly assigned to one of the four conditions. Participants in each condition received hints with different modalities such as visual only, visual + text, visual + audio, and visual + text + audio when they solved the training

problems. Participants were not provided with any hints on the initial, near transfer, or far transfer problems in any condition. We focus on problem-solving performance on the training problems of the Skier problem set.

## V. RESULTS

### A. Problem-solving performance

We were interested in comparing participants' problem solving performance with different hint modalities. A one-way ANOVA showed no significant difference between conditions on initial problem performance,  $F(3, 53) = .60, p = .616$ . A two-way ANOVA was conducted to test the main effects of text and audio hints, and their interaction, on training problem performance. The results indicated that there was a significant main effect of text hints,  $F(1, 371) = 11.40, p = .001$ . There was also a main effect of audio hints,  $F(1, 371) = 13.21, p < .001$ . There was no significant interaction between text and audio hints,  $F(1, 371) = 3.13, p = .078$ .

A post-hoc comparison showed that the solvers without audio hints outperformed the solvers with audio hints ( $47.5\% \pm 3.7\%$  vs.  $31.4\% \pm 3.3\%$ , respectively). Each number in the parentheses includes the mean and standard deviation. The solvers with text hints significantly outperformed solvers without text hints ( $45.4\% \pm 3.5\%$  vs.  $31.8\% \pm 3.6\%$ , respectively). Further, the solvers with visual + text hints outperformed the solvers with visual + audio hints ( $61.5\% \pm 5.4\%$  vs.  $26.4\% \pm 5.2\%$ , respectively). This result is inconsistent with Mayer's modality principle, which predicts that audio information should facilitate learning better than textual information. The visual + text + audio condition significantly underperformed the visual + text condition ( $34.4\% \pm 4.3\%$  vs.  $61.5\% \pm 5.4\%$ , respectively), consistent with the idea of Mayer's redundancy principle, which is *less is better*.

A one-way ANOVA found no significant difference between conditions on either the near transfer problem performance,  $F(3, 53) = .80, p = .498$ , or the far transfer problem performance,  $F(3, 53) = 1.07, p = .372$ . Another one-way ANOVA showed significant differences in participants' performance on the initial, near transfer, and far transfer problems,  $F(2, 168) = 3.90, p = .022$ . Post-hoc tests using the Bonferroni correction revealed a significant improvement from initial to near transfer problem performance ( $1.8\%$  vs.  $15.8\%$ , respectively) and from initial to far transfer problem performance ( $1.8\%$  vs.  $15.8\%$ , respectively). However, there was no significant difference between near and far transfer problem performance.

### B. Dwell time in areas of interest

The duration of each hint is about eight seconds. It means every participant has about eight seconds within which to allocate his or her visual attention. However, it is clear that the defined expert and novice AOIs have different sizes and this difference should be taken into account to make a fair comparison. To do this, we divided the percentage of dwell time in an AOI by the percentage of area of that AOI, to generate a new variable called PTPA (percent time over percent area), also known as the domain relevant ratio [6].

We conducted a 2 (expert area / novice area)  $\times$  2 (correct answer / incorrect answer)  $\times$  4 (visual hint / visual + text hint / visual + audio hint / visual + text + audio hint) ANOVA to compare the PTPA on expert areas and novice areas with different hint modalities. There was a main effect on hint modality,  $F(3, 640) = 3.74, p = .110$ ; a main effect on different areas,  $F(1, 640) = 71.30, p < .001$  and a two-way interaction between hint modality and areas,  $F(3, 640) = 4.91, p = .002$ . We found no other significant main effects or interactions. Since the visual hints are highlighting the expert AOIs, it would be natural for correct solvers to pay more visual attention on expert AOIs. We were interested in how hint modalities manipulated solvers' visual attention on both expert and novice AOIs.

The statistical analysis with Bonferroni correction showed that solvers in the visual + text condition attended to the expert AOIs significantly more often than the visual hints only ( $p = .014$ ) condition and visual + audio hints ( $p = .039$ ) condition. There was no significant difference in the PTPA on the expert areas between visual + text hints and visual + text + audio hints ( $p = 1.000$ ). Solvers in the visual + text + audio hint condition attended to the expert areas more often than visual hints only condition ( $p = .050$ ). None of the other comparisons were statistically significant, and these results were not affected by whether or not the participants solved the problems correctly ( $p = .956$ ).

As for the PTPA data on novice AOIs, the statistical analysis with Bonferroni correction showed that correct solvers spent less time on the novice AOIs than the incorrect solvers ( $p = .047$ ). The solvers in the visual + text condition spent less time on the novice AOIs than those in the visual only ( $p = .018$ ) and visual + audio ( $p = .028$ ) conditions. No other comparisons were statistically significant, and these results were not affected by whether or not the participants solved the problems correctly ( $p = .797$ ).

## VI. DISCUSSION

We found that different hint modalities facilitated our participants to solve training problems differently. Unlike the prediction based on Mayer's CTML, visual + text hints facilitated solving training problems significantly better than other hint modalities. The performance of visual + audio group is significantly worse than visual only group. Our results are inconsistent with the *dual-channel assumption* from CTML. However, this study is confined to a particular set of problems and a particular set of students. Therefore it is not possible to generalize this result to physics problem solving. Further research is needed to examine whether this assumption holds in physics problem solving. The auditory information might be competing against the information from other modalities, while the text information is complimentary with the graphical information. Therefore, processing auditory information may create a heavier cognitive load and hinder problem solving.

The eye movement data generally supported our explanation of the performance data. The participants who received visual + text hints attended to the expert areas longer than participants in any of the other three conditions (significantly longer than visual hints only and visual + audio hints, only slightly longer than visual + text + audio hints). They attended to the novice areas shorter than participants in all the other conditions (significantly shorter

than visual hints only and visual + audio hints, only slightly shorter than visual + text + audio hints). The results suggest that adding text hints with visual hints did not split the solvers' visual attention. On the contrary, adding text hints made the solvers spend more time on the expert areas highlighted by visual hints.

## VII. LIMITATIONS

One limitation of our study is that we accumulated the eye movement data across all six training problems during which the hints remained the same. In the visual + text hints condition, the solvers may have revisited the expert areas more frequently after they have read the text hints on successive training problems. A detailed analysis of eye movements on successive training problems is necessary. Another limitation is that we failed to find a significant difference between the visual + text condition and visual + text + audio condition on the PTPA on either expert areas or novice areas even though the performance difference between these two conditions was significant. This departs from prior research [2] and must be investigated further.

## ACKNOWLEDGMENTS

This work is supported in part by the U.S. National Science Foundation grants No. 1138697 and 1348857.

- [1] A. M. Madsen, Ph. D. thesis, Kansas State University, 2012.
- [2] A. Rouinfar, E. Agra, A. M. Larson, L.C. Loschky, and N.S. Rebello, *Front. Psychol.* **5**, 1094, (2014).
- [3] R.E. Mayer, *Multimedia Learning*, 41-160 (Cambridge Univ. Press, Cambridge; New York, 2001).
- [4] F. Schmidt-Weigand, A. Kohnert, and U. Glowalla, *Learn. Instr.* **20**(2), 100-110 (2010).
- [5] A. Rouinfar, E. Agra, J. Murray, A. Larson, L. Loschky, and N. Rebello, *Can Visual Cues and Correctness Feedback Influence Students' Reasoning?*, presented at the Physics Education Research Conference 2013, Portland, OR, 2013.
- [6] S. Fletcher-Watson, J.M. Findlay, S. R. Leekam, and V. Benson, *Perception* **37** (2008).