Role of diagrams in problem solving: An evaluation of eye-tracking parameters as a measure of visual attention

Ana Susac,[1,2,*] Andreja Bubic,[3] Maja Planinic,[2] Marko Movre,[2] and Marijan Palmovic[4]

[1]*Department of Applied Physics, Faculty of Electrical Engineering and Computing,*
University of Zagreb, Unska 3, 10000 Zagreb, Croatia
[2]*Department of Physics, Faculty of Science, University of Zagreb, Bijenicka 32, 10000 Zagreb, Croatia*
[3]*Chair for Psychology, Faculty of Humanities and Social Sciences,*
University of Split, Sinjska 2, 21000 Split, Croatia
[4]*Laboratory for Psycholinguistic Research, Department of Speech and Language Pathology,*
University of Zagreb, Borongajska cesta 83h, 10000 Zagreb, Croatia

Typical physics textbook problems often include supportive diagrams that visualize the physical situation although the potential benefits of providing such diagrams is not yet fully established. We used eye tracking to explore the role of supportive diagrams in problem solving. Including a supportive diagram with the text of the problem improved students' percentage of correct answer in one of the six problems used in the study. Eye-tracking data showed that students typically spent less time on the text of the problem if they were presented with a diagram, but the total viewing time did not change. When a diagram was presented students split their attention between the diagram and the text without speeding up problem solving. Cognitive load theory and dual coding theory suggest that giving information in two formats (verbal and visual) might reduce extraneous cognitive load and leave more cognitive resources available for further steps in problem solving. However, this does not necessarily lead to a higher percentage of correct answers to the problem, because supportive diagrams influence only one step of the complex process of problem solving. In addition to the role of diagrams, we evaluated different eye-tracking measures as measures of visual attention during physics problem solving. It seems that the fixation duration is rather constant, and not always sensitive to the manipulation in the task. On the other hand, dwell time and the number of fixations show more variability across problems and participants, so they seem to be appropriate measures of the visual attention. Since dwell time and fixation number are dependent measures, and they show a similar pattern of responses, in most cases it seems sufficient to report only one of them.

Introduction.—Teaching and learning physics problem solving is an important physics education research (PER) area. Various problem-solving strategies have been proposed, and all of them include some form of drawing diagrams. For example, the PER group at the University of Minnesota suggests sketching a picture in the first step of understanding and visualizing the problem [1]. In the next step of describing the physics of the problem, students are asked to draw a diagram or a graph that helps in understanding the problem (e.g., a free-body diagram). Nevertheless, reports on the role of diagrams in problem solving are disparate. It has been shown that students who draw diagrams are more successful in problem solving [2]. A reasonable inference would be to prompt students to draw diagrams or to include diagrams in problems as a scaffolding. Some reports suggested that students who were given a diagram describing a physical situation were less likely to draw additional, expertlike diagrams (e.g., to draw vectors) and were less successful in problem solving than those who were prompted to draw the diagrams [3–5]. On the other hand, prompting novice students to draw force diagrams may result in lower success in problem solving [6]. More research is needed to determine the role of diagrams in problem solving. A recent large-scale study on usefulness of supportive diagrams has shown their small positive effect on students' scores [7].

In this study, we use eye tracking to explore the role of supportive diagrams in problem solving. Measurement of eye movements provides detailed information on the underlying cognitive processes during problem solving than conventional assessment methods which provide only

*Corresponding author.
ana.susac@fer.hr

participants' score. Eye tracking also gives insight into the role of visual attention during problem solving in mathematics and science [8,9]. However, there is a limited number of PER studies that have used eye tracking [10–22]. In addition to information on spatial and temporal characteristics of visual attention, eye-tracking data may provide information on cognitive load. The cognitive load theory was developed to facilitate learning and problem solving [23]. Cognitive load theory specifies three types of cognitive load: intrinsic, extraneous, and germane [24,25]. Intrinsic cognitive load is related to the specific material or skill being learned whereas extraneous cognitive load is associated with the way the material or skill is presented to a learner. Germane cognitive load is produced by the construction of schemas, i.e., permanent stores of knowledge, and is desirable because it supports learning of new materials or skills. According to the dual-coding theory, both visual and verbal representations are used to represent information [26].

The cognitive load theory [23] and the dual-coding theory [26] suggest that the supportive diagrams could be useful in the initial phase of problem solving because presenting the physical situation in two formats (verbal and visual) might reduce the extraneous cognitive load and leave more cognitive resources for further steps in problem solving. Previous non-PER studies showed that various eye-tracking measures can be used as correlates of work load during different tasks [27,28]. However, tasks in PER studies are complex and different from the typical visual search or memory tasks, and the advantages and disadvantages of different eye-tracking measures in PER tasks should be explored. A recent PER study used eye tracking to investigate the effect of intrinsic, extraneous, and germane cognitive load on different eye-tracking measures during multimedia lessons [29]. The results provided insight into the complex interrelations among different eye-tracking measures. In this study, our goal is to further evaluate application of eye-tracking measures during physics problem solving.

We aim to answer the following research questions:

(i) Do diagrams help students in problem solving?
(ii) What can eye-tracking reveal about the role of diagrams in problem solving?
(iii) Which eye-tracking parameters are appropriate measures of visual attention during physics problem solving?

Methods.—Participants: Sixty undergraduate students from the Department of Physics, University of Zagreb participated in this study. All participants were senior years prospective physics teachers. Each participant gave an informed written consent before taking part in the experiment.

Materials: Six multiple-choice questions on energy were used in this study (see the Supplemental Material [30,31]). The questions were always presented in the same order. In one experimental group, the first, third, and fifth questions were presented with diagrams, while the remaining questions were presented without diagrams. In another experimental group, the second, fourth, and sixth questions were presented with diagrams.

Procedure: Eye-movement data were recorded as in our previous study [14].

Data analysis: The recorded eye-movements data were analyzed using BeGaze software. Saccade duration and number of regressions are often used eye-tracking measures, but might be difficult to interpret for our task because they depend on the spatial layout of the problem and the related diagram. Thus, we evaluate and compare the following eye-tracking measures: dwell time (viewing time), number of fixations, and average fixation duration. We defined three areas of interest (AOIs) for each question and calculated the chosen eye-tracking measures for each AOI. AOIs were rectangles that included the text of the problem (*question*), multiple choice answers (*multiple choice*), and the related diagram (*diagram*). In condition "no diagram," AOI *diagram* was an empty rectangle at the same place as the diagram in "diagram" condition.

Students' responses were graded as either correct or incorrect. The χ^2 tests, analyses of variance (ANOVAs) and Bonferroni corrected Student's *t*-tests were conducted with a threshold for significance of $p = 0.05$.

Results.—Analysis of students' responses: Students' percentages of correct answers on all questions presented with and without diagrams are shown in Fig. 1. The χ^2 tests showed statistically significant differences in percentages of correct answers between the two conditions (diagram, no diagram) only for question Q3 [$\chi^2(1) = 6.24$, $p = 0.01$]. Overall, the difference between the percentages of correct answers on questions with diagram and with no diagram was not statistically significant [$\chi^2(1) = 1.63$, $p > 0.05$].

Analysis of eye-tracking data for question Q3: As we found a positive effect of the supportive diagram on students' percentages of correct answers for question Q3, we further explored eye-tracking data for that question. We compared three eye-tracking measures (dwell time, number of fixations, and average fixation duration) for AOIs *question*, *multiple choice,* and *diagram* (Fig. 2). Dwell time was longer for the AOI question if a diagram was not presented [$t(58) = 4.04$, $p = 0.0004$], but there was no difference for the AOI multiple choice [$t(58) = 0.43$, $p > 0.05$]. Correspondingly, the fixation number was larger for the AOI *question* if a diagram was not presented [$t(58) = 3.85$, $p = 0.0006$] and there was no difference for the AOI multiple choice [$t(58) = 0.17$, $p > 0.05$]. Average fixation duration was not significantly different for both AOIs, question and multiple choice [$t(58) = 2.13$, $p > 0.05$; $t(58) = 1.62$, $p > 0.05$, respectively]. As expected, most participants did not look at the AOI diagram if the diagram was not presented. Thus, we did not statistically compare eye-tracking measures for AOI diagram.

When dwell times and fixation numbers were summed across three AOIs, and fixation duration averaged across
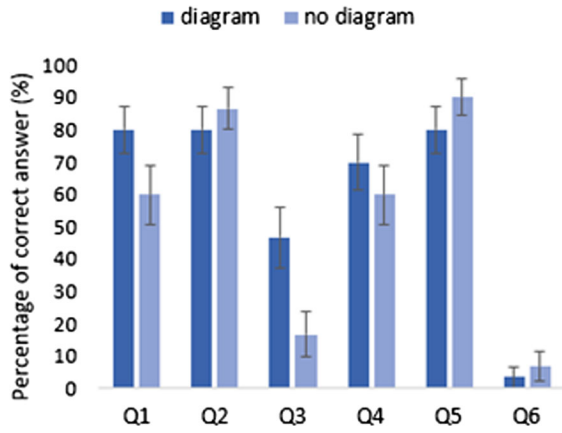
FIG. 1. Comparison of students' percentages of correct answers on all questions (Q1–Q6) between the two conditions (diagram, no diagram). The error bars represent 1 SEM.

three AOIs, they did not differ for the two conditions [$t(58) = 1.27$, $p > 0.05$; $t(58) = 1.09$, $p > 0.05$; $t(58) = 0.77$, $p > 0.05$, respectively].

Corresponding analysis was conducted for correct and incorrect problem solvers (see the Supplemental Material [30]). The results showed that correct problem solvers had shorter dwell time and a smaller number of fixations compared to incorrect problem solvers, whereas no statistically significant difference was found for the average fixation duration.

Analysis of eye-tracking data for all six questions: Since the diagram decreased the dwell time and fixation number for question Q3, we evaluated its effect on the three eye-tracking measures in AOI question for all questions (Fig. 3). Three two-way mixed-design ANOVAs were conducted on dwell time, number of fixations, and average fixation duration with factors diagram and question. The results obtained showed a statistically significant main effect of both factors on all three eye-tracking measures, whereas the interaction effect was significant for dwell time and fixation number (Table I). Dwell time and fixation number decreased when a diagram was presented in
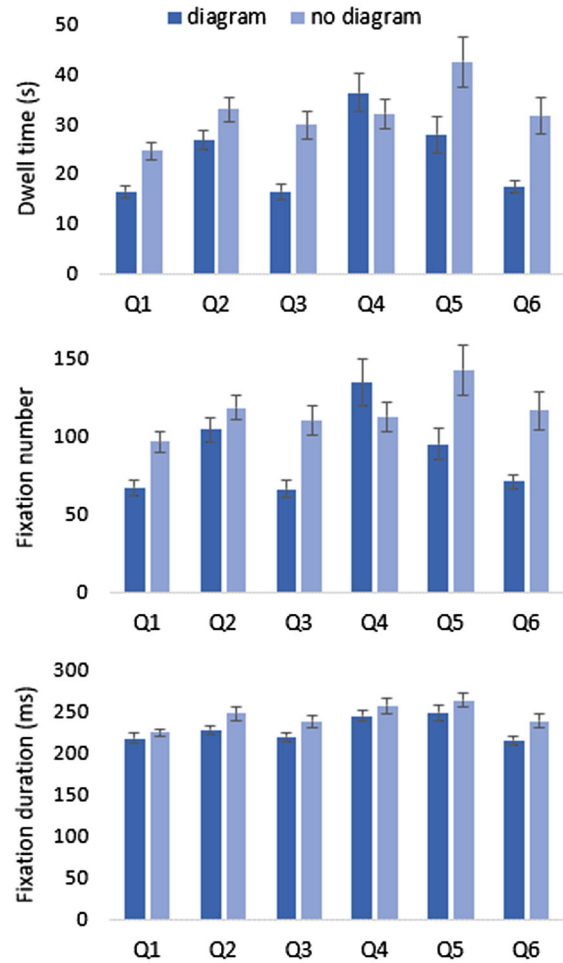

FIG. 3. Dwell time, number of fixations, and average fixation duration for the AOI *question*, calculated for the two conditions (diagram, no diagram) and all questions (Q1–Q6). The error bars represent 1 SEM.

questions Q1 [$t(58) = 3.88$, $p = 0.002$; $t(58) = 3.64$, $p = 0.004$, respectively], Q3 [$t(58) = 4.04$, $p = 0.001$; $t(58) = 3.85$, $p = 0.002$, respectively] and Q6 [$t(58) = 3.70$, $p = 0.003$; $t(58) = 3.56$, $p = 0.005$,
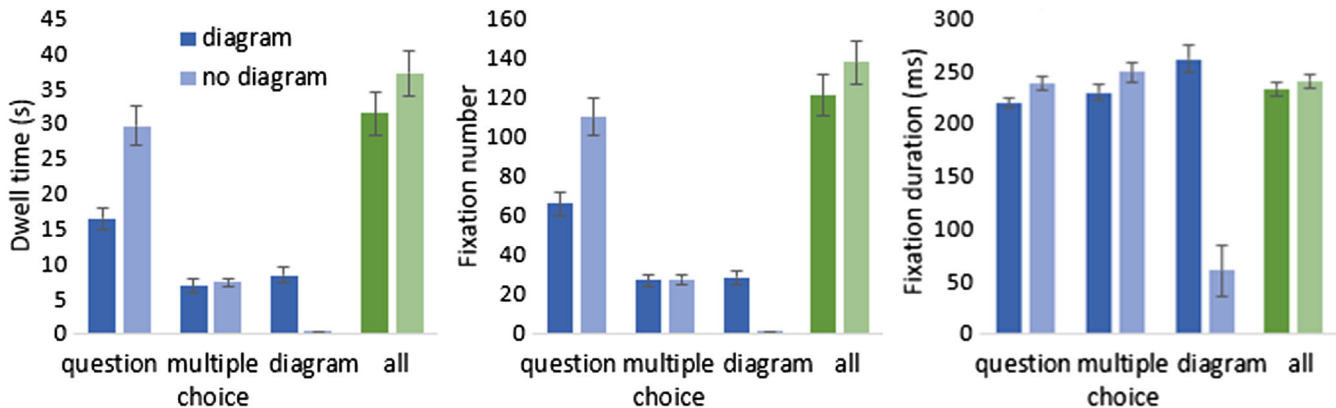

FIG. 2. Dwell time, number of fixations, and average fixation duration for three AOIs (*question*, *multiple choice*, and *diagram*) and all three summed AOIs (shown in green) on question Q3. The error bars represent 1 SEM.
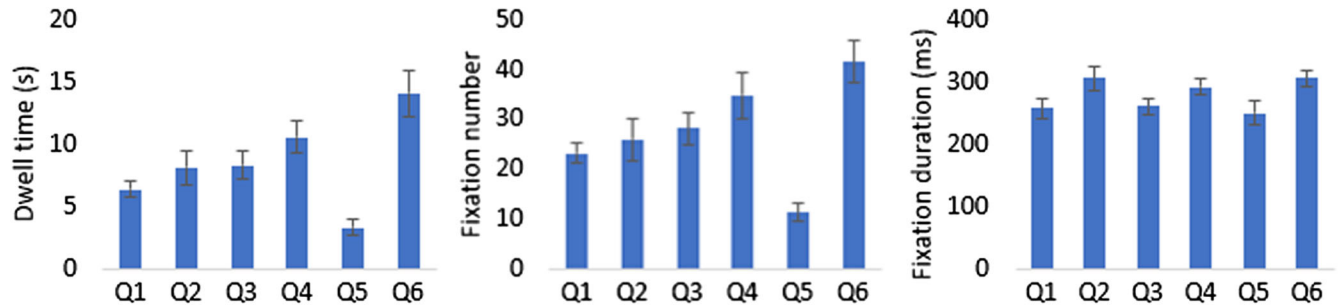
FIG. 4. Dwell time, number of fixations, and average fixation duration for the AOI *diagram*, for all questions (Q1–Q6). The error bars represent 1 SEM.

respectively]. The presence of a diagram did not statistically significantly change dwell time and fixation number in questions Q2 [$t(58) = 2.05$, $p > 0.05$; $t(58) = 1.28$, $p > 0.05$, respectively], Q4 [$t(58) = -0.88$, $p > 0.05$; $t(58) = -1.23$, $p > 0.05$, respectively] and Q5 [$t(58) = 2.31$, $p > 0.05$; $t(58) = 2.54$, $p > 0.05$, respectively].

Furthermore, eye tracking provided data on the time that a participant spent attending to the diagram. Figure 4 shows the three eye-tracking measures for all six questions. One-way ANOVAs revealed that the dwell time, number of fixations, and average fixation duration were significantly different across the questions [$F(5, 145) = 9.49$, $p < 0.0001$, $\eta_p^2 = 0.247$; $F(5, 145) = 9.91$, $p < 0.0001$, $\eta_p^2 = 0.255$; $F(5, 145) = 2.88$, $p = 0.016$, $\eta_p^2 = 0.090$]. Figure 4 reveals that participants spent the shortest time attending the diagram in problem Q5.

Regarding our second research question on the evaluation of eye-tracking measures, visual inspection of

Figs. 2, 3, and 4 reveals similar patterns for dwell time and fixation number whereas fixation duration shows smaller variation. Roughly, dwell time is proportional to fixation number multiplied by average fixation duration. Thus, for rather constant average fixation duration, dwell time and fixation number have similar patterns. This conclusion was confirmed by inspecting individual participants' data (see the Supplemental Material [30]). There is more variability in the total dwell time and total fixation number compared to average fixation duration. To quantify the variability in different eye-tracking measures, we calculated the coefficient of variation (CV) as a ratio of standard deviation (SD) and mean (Table II). CV of the total dwell time and total fixation number is about 3 times larger than CV of average fixation duration.

Discussion.—Supportive diagrams had a positive effect on students' percentages of correct answers in one of the six problems used in the study. This result is in agreement with the previous studies that reported no effect

TABLE I. Results of two-way ANOVAs conducted on dwell time, fixation number, and fixation duration for the AOI *question* with between-subjects factor diagram (diagram vs no diagram) and within-subjects factor question (Q1–Q6).

	Diagram			Question			Interaction		
	F (df)	p	η_p^2	F (df)	p	η_p^2	F (df)	p	η_p^2
Dwell time	20.14 (1, 58)	$<10^{-4}$	0.258	9.25 (5, 290)	$<10^{-4}$	0.137	3.28 (5, 290)	0.007	0.530
Fixation number	14.53 (1, 58)	$<10^{-4}$	0.200	7.03 (5, 290)	$<10^{-4}$	0.108	4.23 (5, 290)	0.001	0.068
Fixation duration	6.89 (1, 58)	0.01	0.106	9.94 (5, 290)	$<10^{-4}$	0.146	0.45 (5, 290)	>0.05	0.008

TABLE II. Mean, standard deviation (SD) and coefficient of variation (CV) of total dwell time, total fixation number, and average fixation duration for each problem.

	Dwell time (s)		Fixation number		Fixation duration (ms)	
Question	Mean ± SD	CV	Mean ± SD	CV	Mean ± SD	CV
Q1	33 ± 11	0.33	120 ± 41	0.34	240 ± 33	0.14
Q2	40 ± 16	0.40	142 ± 54	0.38	250 ± 38	0.15
Q3	35 ± 18	0.51	130 ± 60	0.46	237 ± 36	0.15
Q4	45 ± 26	0.58	158 ± 96	0.61	255 ± 43	0.17
Q5	43 ± 29	0.67	144 ± 85	0.59	260 ± 47	0.18
Q6	40 ± 19	0.48	140 ± 62	0.44	250 ± 45	0.18

or a small positive effect of supportive diagrams on problem solving [7]. The reason why the diagram in question Q3 had a positive effect on students' percentages of correct answers could be that it emphasized the initial equal elevations of the bricks (see the Supplemental Material [30]). Although this diagram does not provide any additional information that is not present in the text, it could have influenced student reasoning similarly as visual cues in Madsen *et al.* [16]. One interpretation may be that a diagram might be beneficial if it highlights some information for problem solving. Overall, our results suggest that, while diagrams can help students in problem solving, the effect is small and rarely statistically significant.

Analysis of eye-tracking data showed that students who were presented with a supportive diagram usually spent less time reading the text of the problem, but the total viewing time was the same for the two conditions (diagram, no diagram). When a diagram was presented students split their attention between the diagram and the text without speeding up problem solving, which was in agreement with our previous findings [12]. Cognitive load theory combined with the dual coding theory implies that information in both formats (verbal and visual) reduces extraneous cognitive load and helps students to better understand the physical situation described in the text. However, better understanding of the physical situation does not necessarily lead to more efficient problem solving. Even if diagrams were useful in the initial phase of problem visualization, they would not guarantee more efficient implementation of the subsequent phases in problem solving. Eye-tracking data do not seem to agree with the hypothesis that presenting information in both formats significantly improves the overall efficiency in problem solving. Further studies are needed to assess the effect of presenting information in both formats on different steps in problem solving.

One goal of this study was to determine which eye-tracking parameters are appropriate measures of visual attention during physics problem solving. Our results show that fixation duration is rather constant across questions and participants, suggesting that it might not be sensitive enough to reveal all differences in visual attention caused by some task manipulation. Although fixation duration is used as a measure of the visual attention, it is possible that it is not so suitable for physics problem solving. On the other hand, dwell time and fixation number show larger variance and sensitivity to task manipulation and therefore seem more appropriate eye-tracking measures for exploring physics problems solving. Similar results are reported in a previous study of website complexity [28] and PER study on kinematic graphs [20]. Patterns of dwell time and fixation number show analogous effects of task manipulation because these two measures are dependent, therefore it might be adequate to report only one of these two measures. In this study, we compared the most used eye-tracking parameters in the PER studies. In future studies, other eye-tracking measures should also be evaluated.

We found that dwell time was sensitive to extraneous load that corroborates the results from Zu *et al.* [29]. However, Zu *et al.* reported significant sensitivity of the mean fixation duration to extraneous load that was not found in our study. A possible explanation for this discrepancy could include differences in the design of the study and corresponding data analysis. Further research with different study designs is needed to explore the relationship between eye-tracking measures and cognitive load.

---

[1] P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving, Am. J. Phys. **60**, 627 (1992).

[2] A. Mason and C. Singh, Helping students learn effective problem solving strategies by reflecting with peers, Am. J. Phys. **78**, 748 (2010).

[3] A. Maries and C. Singh, Should students be provided diagrams or asked to draw them while solving introductory physics problems?, AIP Conf. Proc. **1413**, 263 (2012).

[4] A. Maries and C. Singh, To use or not to use diagrams: The effect of drawing a diagram in solving introductory physics problems, AIP Conf. Proc. **1513**, 282 (2013).

[5] A. Maries and C. Singh, A good diagram is valuable despite the choice of a mathematical approach to problem solving, AIP Conf. Proc. **1513**, 31 (2013).

[6] A. F. Heckler, Some consequences of prompting novice physics students to construct force diagrams, Int. J. Sci. Educ. **32**, 1829 (2010).

[7] Z. Chen, N. Demirci, Y. J. Choi, and D. E. Pritchard, To draw or not to draw? Examining the necessity of problem diagrams using massive open online course experiments, Phys. Rev. Phys. Educ. Res. **13**, 010110 (2017).

[8] M. J. Tsai, H. T. Hou, M. L. Lai, W. Y. Liu, and F. Y. Yang, Visual attention for solving multiple-choice science problem: An eye-tracking analysis, Comput. Educ. **58**, 375 (2012).

[9] A. Susac, A. Bubic, J. Kaponja, M. Planinic, and M. Palmovic, Eye movements reveal students' strategies in simple equation solving, Int. J. Sci. Math. Educ. **12**, 555 (2014).

[10] T. van Gog, F. Paas, and J. J. G. Van Merriënboer, Uncovering expertise-related differences in trouble-shooting performance: Combining eye movement and concurrent verbal protocol data, Appl. Cogn. Psychol. **19**, 205 (2005).

[11] M. Kozhevnikov, M. A. Motes, and M. Hegarty, Spatial visualization in physics problem solving, Cogn. Sci. **31**, 549 (2007).

[12] A. Susac, A. Bubic, P. Martinjak, M. Planinic, and M. Palmovic, Graphical representations of data improve student understanding of measurement and uncertainty: An eye-tracking study, Phys. Rev. Phys. Educ. Res. **13,** 020125 (2017).

[13] P. Klein, J. Viiri, S. Mozaffari, A. Dengel, and J. Kuhn, Instruction-based clinical eye-tracking study on the visual interpretation of divergence: How do students look at vector field plots?, Phys. Rev. Phys. Educ. Res. **14,** 010116 (2018).

[14] A. Susac, A. Bubic, E. Kazotti, M. Planinic, and M. Palmovic, Student understanding of graph slope and area under a graph: A comparison of physics and nonphysics students, Phys. Rev. Phys. Educ. Res. **14,** 020109 (2018).

[15] 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).

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

[17] A. Madsen, A. Rouinfar, A. M. Larson, L. C. Loschky, and N. S. Rebello, Can short duration visual cues influence students' reasoning and eye movements in physics problems?, Phys. Rev. ST Phys. Educ. Res. **9,** 020104 (2013).

[18] A. Rouinfar, E. Agra, A. M. Larson, N. S. Rebello, and L. C. Loschky, Linking attentional processes and conceptual problem solving: Visual cues facilitate the automaticity of extracting relevant information from diagrams, Front. Psychol. **5,** 1094 (2014).

[19] M. Kekule, Students' approaches when dealing with kinematics graphs explored by eye-tracking research method, Proceedings of the Frontiers in Mathematics and Science Education Research Conference (Science Education research Group at Eastern Mediterranean University, Famagusta, 2014), pp. 108–117, ISSN 2301-251X.

[20] M. Kekule, Students' different approaches to solving problems from kinematics in respect of good and poor performance, Proceedings of the International Conference on Contemporary Issues in Education, Dubai (2015), pp. 126–134.

[21] S. C. Chen, H. C. She, M. H. Chuang, J. Y. Wu, J. L. Tsai, and T. P. Jung, Eye movements predict students' computer-based assessment performance of physics concepts in different presentation modalities, Comput. Educ. **74,** 61 (2014).

[22] J. Han, L. Chen, Z. Fu, J. Fritchman, and L. Bao, Eye-tracking of visual attention in web-based assessment using the Force Concept Inventory, Eur. J. Phys. **38,** 045702 (2017).

[23] J. Sweller, Cognitive load during problem solving: Effects on learning, Cogn. Sci. **12,** 257 (1988).

[24] J. Sweller, Cognitive load theory, learning difficulty, and instructional design, Learn. Instr. **4,** 295 (1994).

[25] J. Sweller, Element interactivity and intrinsic, extraneous, and germane cognitive load, Educ. Psychol. Rev. **22,** 123 (2010).

[26] J. M. Clark and A. Paivio, Dual coding theory and education, Educ. Psychol. Rev. **3,** 149 (1991).

[27] K. F. Van Orden, W. Limbert, S. Makeig, and T.-P. Jung, Eye activity correlates of workload during a visuospatial memory task, Human Factors **43,** 111 (2001).

[28] Q. Wang, S. Yang, M. Liu, Z. Cao, and Q. Ma, An eye-tracking study of website complexity from cognitive load perspective, Decision Support Systems **62,** 1 (2014).

[29] T. Zu, J. Hutson, L. C. Loschky, and N. S. Rebello, Use of eye-tracking technology to investigate cognitive load theory, Proceedings of the Physics Education Research Conference 2017, Cincinnati, OH (AIP, New York, 2017), p. 472.

[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.15.013101 for the problems used in the study, the comparison of correct and incorrect problem solvers, and participants' individual dwell time, fixation number, and fixation duration.

[31] N. Brkovic, Zbirka zadataka iz fizike (LUK, Zagreb, 2001).