Editors' Suggestion

Effect of representation format on conceptual question performance and eye-tracking measures

Ana Susac,[1,*] Maja Planinic,[2] Andreja Bubic,[3] Katarina Jelicic,[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]Department of 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

(Received 9 October 2022; accepted 31 May 2023; published 18 August 2023)

Previous studies have shown the important role of different representations in the teaching and learning of physics. In this study, we used eye tracking to investigate the effect of different representations on the process of answering conceptual questions. We compared students' scores and eye-tracking measures on isomorphic questions which contained graphical, pictorial, and verbal representations. On average, in two-thirds of cases, students were consistent in their answers (correct or incorrect) across all three representations. There was no statistically significant difference in students' scores for different representations. However, eye-tracking measures suggest that it was easiest for students to extract information from verbal representations and most difficult from pictorial representations for the conceptual questions used in this study. These results could be useful to teachers and researchers when creating conceptual questions and, more generally, when teaching with multiple representations.

DOI: 10.1103/PhysRevPhysEducRes.19.020114

I. INTRODUCTION

Physics education research (PER) results have shown the very important role of representations in understanding physical concepts [1]. It has been suggested that students should learn to use multiple representations in problem solving and that it would help them approach problems like physicists [2]. On the other hand, understanding and using different representations proved to be challenging for students. For example, in an early study on the effect of representational format on students' problem-solving performance, Meltzer found that students had difficulties with the use of vector arrows to differentiate forces acting on an object from forces exerted by that object and they had higher scores on the same question in the verbal format [3]. However, students had similar overall scores on quiz questions using verbal, diagrammatic (pictorial), mathematical, and graphical representations. Analogous results were obtained by Kohl and Finkelstein [4] who compared student performance on problems given in mathematical, pictorial, graphical, and verbal representational formats. Although they found differences in student performance on problems with different representations, they did not find one representation that would always be the easiest for students, regardless of the context of the question. When given the possibility to choose a representational format, students preferred pictures over words, graphs, or mathematical expressions, but this did not make them more successful at solving problems.

To investigate students' consistency in understanding different representations, Nieminen *et al.* developed the Representational Variant of the Force Concept Inventory (R-FCI), using the bar chart, graphical, motion map, vectorial, and verbal representations [5]. They found that students' representational consistency significantly depended on the context of the item, and it increased during the instruction. Furthermore, a number of PER studies have shown that if students are taught physics using multiple representations, they use them in solving problems and this leads to their increased performance, e.g., [6–11]. Previous studies did not provide a conclusive answer to the question of whether it is easier to extract the necessary information for solving problems from some representations and whether any representation has an advantage over others. We believe that this is an important issue in physics education, which typically uses many

*Corresponding author.
ana.susac@fer.hr

different representations. Knowing how students react to different representations could have implications for the choice of representation in some settings or for devoting more teaching time to more difficult representations which are considered important for students to master.

Recently, measurement of eye movements has been increasingly used in PER to investigate student attention during problem solving. For a recent review of eye tracking in physics education research, we refer to the systematic literature review by Hahn and Klein [12]. Eye tracking was used to assess students' conceptual understanding in different physics domains such as kinematics [13–20], forces [21–23], conservation laws [24,25], vector fields [26–28], electrical circuits [29], coordinate systems [30,31], wave optics [32,33], measurement uncertainty [34], and various topics [35–39]. In most of these studies, a single representation was used. In the present study, we employed isomorphic questions with graphical, pictorial, and verbal representations to explore students' visual attention while solving questions with different representations.

II. THEORETICAL BACKGROUND

A. Background literature

Theoretical models of learning from multimedia usually assume that adding more representations leads to better learning. Paivio's dual coding theory [40] proposes that information is processed and represented in two separate but interconnected systems: verbal and nonverbal (imagery). According to this theory, learning is more effective when both systems are engaged, and information is presented in a way that allows learners to create connections between the two systems [40]. The cognitive theory of multimedia learning expands on this idea, stating that learning is more effective when multiple sensory channels are engaged [41]. This theory suggests that combining visual and verbal information can enhance learning by providing multiple representations of the same content, which allows learners to build a more comprehensive mental model of the concept being taught.

Contrary to these theoretical models, some studies have shown that adding pictures to a text can have negative effects on learning [42]. According to the cognitive load theory developed by Sweller and his colleagues [43–45], the human cognitive system has a limited capacity for processing information, and learning is more effective when the cognitive load is managed appropriately. The cognitive load can be divided into three categories: intrinsic load, extraneous load, and germane load. The intrinsic cognitive load is the inherent complexity of the learning materials. The extraneous cognitive load refers to the cognitive effort required to process information that is not directly related to the learning goals. The germane cognitive load refers to the cognitive effort required to process and integrate information in a way that supports learning. So, according to the cognitive load theory, adding pictures to text can increase the extraneous cognitive load, which can have negative effects on learning. It can be argued that learners who are presented with text that includes pictures spend cognitive effort on processing the pictures, which reduces the cognitive resources available for processing the text. Therefore, when using pictures in text, it is important to consider their relevance to the learning goals and their potential to increase cognitive load.

Similarly, cognitive load theory can explain how different representations might affect students' answers to conceptual questions. The complexity of the representation can influence the extraneous cognitive load, i.e., it is possible that a different effort is needed to extract necessary information from different representations. For example, it is possible that a verbal statement that one speed is higher than another is easier to understand than if this information has to be inferred from a position versus time graph. Indeed, as mentioned above, PER studies have shown that students had different scores on the same questions using different representations [3–5].

In these previous studies, different students solved conceptual questions with different representations. It would be useful to investigate whether the same effect of different representations would be observed if the same students answer conceptual questions with different representations. In this way, differences between students would be eliminated. Furthermore, it would be possible to quantify how consistent students are in their answers (correct or incorrect) for different representations. Response consistency measures whether a student's answers are reliable and stable over time, and it is closely related to response confidence because students who are confident in their understanding of a concept are more likely to provide consistent answers to related questions. Moreover, when students are less confident in their responses, they need more time to process the item and decide on their responses [46].

B. Eye-tracking measures

Assuming that students process information that they visually attend [47], eye tracking can provide valuable insights into the cognitive processes involved in answering conceptual physics questions, particularly with regard to the role of representations. The use of eye tracking in exploring the role of representations in answering conceptual questions could bring new understandings that otherwise would not be available. By analyzing eye movements, PER researchers can better understand how students process verbal, graphical, and pictorial information, which can help improve physics education and inform instructional design. Therefore, we will introduce basic eye-tracking terms and measures and explain how they can be interpreted.

A fixation is a period of time during which the eyes remain relatively still and focused on a specific point of interest in the visual field. According to the eye-mind *assumption* [47], the brain extracts information from the visual input during fixation. *Just and Carpenter* [47] assumed that "*the eye remains fixated on a word as long as the word is being processed. So, the time it takes to process a newly fixated word is directly indicated by the gaze duration*." *Analogously, we can extend the interpretation of the eye-mind assumption to other visual elements besides words. For example, in this study, we will assume* the eye remains fixated on a specific part of a graph during processing information available from that position. The information collected from fixations can be used to better understand cognitive processes such as visual attention, perception, and decision-making.

On the other hand, a saccade is a rapid eye movement that serves to shift the line of sight from one point of interest to another. Saccades are fast eye movements that occur between fixations and are typically characterized by a high velocity and a short duration (usually only a few milliseconds). Saccades are essential for a variety of visual tasks, such as reading, where the eyes make a series of saccades to move along a line of text or for exploring a visual scene. Eye tracking can detect and measure both fixations and saccades, which gives researchers insights into how people process visual information and allocate their attention over time.

In PER studies, eye-tracking measures based on fixations are mainly used, although there were examples where the saccadic angle and saccade length were used (e.g., [26,27]). Areas of interest (AOIs) are usually defined first and then corresponding eye-tracking measures are evaluated. Commonly used eye-tracking measures include the dwell time (viewing time), the number of fixations, the average fixation duration, and the number of revisits for each AOI [12,48,49].

Dwell time (total fixation duration or total viewing time) refers to the amount of time that a participant spends looking at a particular area of interest. It is the sum of all fixation durations and saccade durations within an AOI. Dwell time can be used to assess how long a participant looks at specific areas of interest, such as key parts of a graph or important words in a text.

The number of fixations refers to the fixation count for a particular area of interest. The fixation number can provide insight into how frequently a person attends an AOI. The higher number of fixations indicates more visual attention to certain areas. It has been shown that, for typical PER tasks, dwell time and fixation number show a similar pattern [25], i.e., they are highly dependent. If the dwell time for an AOI is higher, then the number of fixations is also higher for that AOI.

Average fixation duration refers to the average duration of single eye fixations. It can provide information about how cognitively demanding is information processing. Higher values of average fixation duration typically imply a higher cognitive effort [12].

Revisits refer to the number of times that participants return to a particular area of interest after having looked away. This eye-tracking measure is similar to transitions (number of fixation shifts from one AOI to another) and they both show attention shifts over time. In PER studies, they are used as indicators of integration processes [12]. For example, the higher number of transitions between question stem and options in multiple-choice questions suggests better integration of information presented in question stem and options.

III. RESEARCH QUESTIONS

The importance of using multiple representations in teaching and learning physics is widely recognized, although the development of optimal ways for their use remains an area of ongoing exploration. It is especially important to understand the integration of information from different representations because the use of multiple representations can have both beneficial and adverse effects depending on the context. Eye tracking can serve as an excellent tool to examine the information processing of individual representations as well as the integration of information from multiple representations. Various eye-tracking measures can provide insights into the duration of information processing (dwell times), cognitive processing demand (average fixation duration), and information integration (revisits). It is noteworthy that previous research has not systematically compared different representations concerning these specific processes of information extraction.

In this study, we aim to answer the following research questions:

(1) Do different representations affect students' answers to conceptual questions? Are students consistent in their answers across representations?
(2) Is it equally difficult for students to extract information from different representations? Does any representation have an advantage?

Since we reduced the variability in student responses by having the same students answer the same questions with different representations, we expected a smaller effect of different representations on student responses than in previous studies. However, our assumption was not that isomorphic items will have similar eye-tracking measures (dwell time, average fixation duration, and the number of revisits). We expected differences in *dwell times, average fixation durations, and numbers of revisits* for different representations as indicators of *different duration of information processing, cognitive load, and information integration, respectively.* The goal of this exploratory study was to determine which representations are advantageous and which are more demanding for students.

IV. METHODS

A. Participants

Participants in this study were 38 high school students (aged 18–19 years) in the last (fourth) year of high school. They attended different general education and science-mathematics types of gymnasiums in Zagreb, Croatia, where physics is taught as a compulsory subject all four years. They studied physics for 2 or 3 hours per week (depending on the type of school) during four years of high school and covered topics in mechanics, thermodynamics, electromagnetism, oscillations and waves, optics, and an introduction to modern physics. All participants gave informed written consent before taking part in the study. Three participants did not have acceptable calibration, so we will report data from 35 students.

B. Materials

We developed or modified from the previous studies [4,5] six sets of isomorphic questions which contained graphical, pictorial, and verbal representations. All eighteen multiple-choice test items used in the study are given in the Supplemental Material [50]. Sets of questions were related to free fall (Q1), Newton's second law (Q2), Newton's third law (Q3), conservation of energy (Q4), oscillation (Q5), and photoelectric effect (Q6). The topics of the items were chosen arbitrarily; the only condition was that they were covered in physics classes.

We prepared three versions of the test, with the same order of the topics (Q1, Q2, …) and a counterbalanced sequence of graphical, pictorial, and verbal representations. Each participant answered all 18 items in one of three sequences shown in Table I.

TABLE I. The order of the items in three test versions.

Test version 1	Test version 2	Test version 3
Q1_G	Q1_P	Q1_V
Q2_P	Q2_V	Q2_G
Q3_V	Q3_G	Q3_P
Q4_G	Q4_P	Q4_V
Q5_P	Q5_V	Q5_G
Q6_V	Q6_G	Q6_P
Q1_P	Q1_V	Q1_G
Q2_V	Q2_G	Q2_P
Q3_G	Q3_P	Q3_V
Q4_P	Q4_V	Q4_G
Q5_V	Q5_G	Q5_P
Q6_G	Q6_P	Q6_V
Q1_V	Q1_G	Q1_P
Q2_G	Q2_P	Q2_V
Q3_P	Q3_V	Q3_G
Q4_V	Q4_G	Q4_P
Q5_G	Q5_P	Q5_V
Q6_P	Q6_V	Q6_G

C. Procedure

Eye-movement data were recorded using the SMI screen-based RED-m system (SensoMotoric Instruments GmbH) with a sample rate of 120 Hz integrated with a 17″ TFT LCD monitor. The eye-tracking system was calibrated for each participant before the data recording using a five-point calibration algorithm. Questions were presented on a monitor at a distance of 50 cm from the participant's eyes. Participants chose the answer by clicking the mouse on a, b, c, or d and thus advanced to the next question. There was no time limit to answer the questions.

D. Data analysis

The recorded eye-movement data were analyzed using BeGaze software and the Identification by Velocity-Threshold (IVT) algorithm with a velocity threshold of $40°/\mathrm{s}$. The minimum fixation duration was set to 100 ms and all fixations below the threshold were rejected.

We defined two rectangular areas of interest (AOIs) for each question that included the question stem (AOI *question*) and multiple-choice options in different representations (AOI *representations*). We evaluated the dwell time (viewing time), the number of fixations, the average fixation duration, and the number of revisits for each AOI. As we previously reported, these eye-tracking measures are dependent, and since dwell time and the number of fixations show a similar pattern of responses [25], we will not report here the numbers of fixations.

Student's t tests and χ^2 tests, one-way analyses of variance (ANOVAs), and Tukey's HSD (honestly significant difference) tests were conducted in the analysis of students' scores and eye-tracking data [51,52]. A threshold of $p = 0.05$ was used for determining the level of effect significance within all conducted tests.

V. RESULTS

A. Analysis of students' scores

The students' mean score and standard deviation were $(54 \pm 21)\%$. Figure 1 shows the distribution of students' scores. A high standard deviation and Fig. 1 indicate that the scores are widely spread around the mean value.

Figure 2 shows students' scores on all items. The χ^2 tests revealed that the differences in students' scores on isomorphic items were not statistically significant (all $p > 0.05$).

To assess the overall effect of representations on students' answers, we calculated the percentages of correct answers for each representation. The mean scores and standard deviations were $(55 \pm 22)\%$ for graphical representation, $(54 \pm 26)\%$ for pictorial representation, and $(52 \pm 23)\%$ for verbal representation. One-way ANOVA showed that

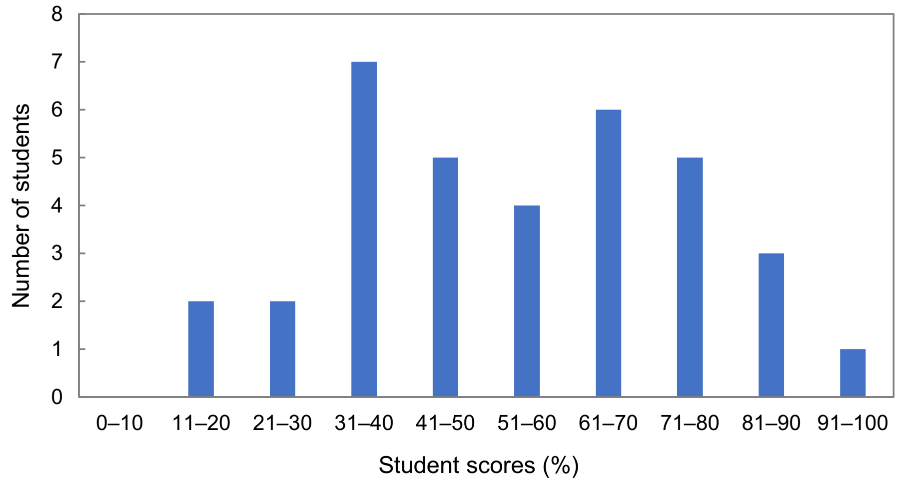

FIG. 1. Distribution of student scores.

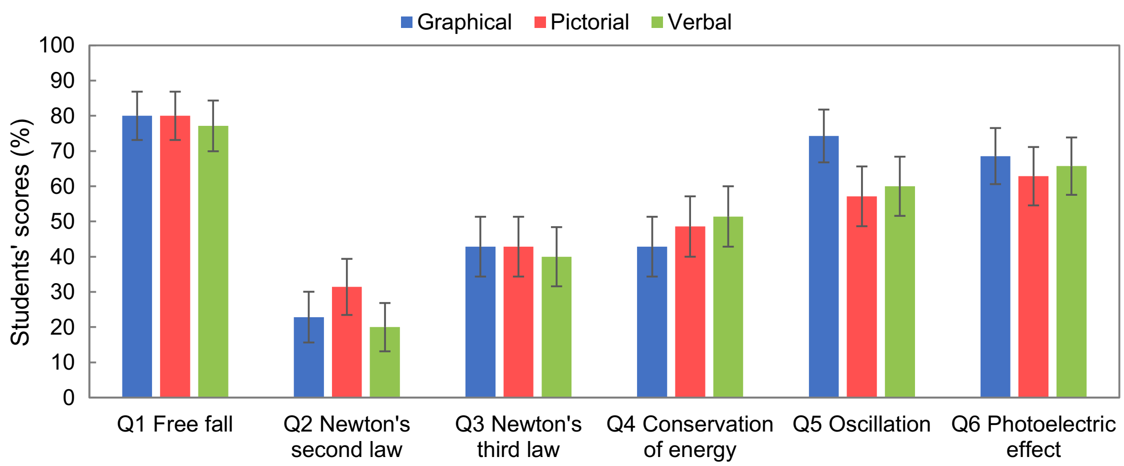

FIG. 2. The mean students' scores on all isomorphic questions with graphical, pictorial, and verbal representations. The error bars represent 1 SEM.

there was no statistically significant effect of representation on students' scores [$F(2,68) = 0.46, p > 0.05, \eta_p^2 = 0.01$].

Although we did not find a significant effect of representations on students' answers, the content and context of the questions influenced the students' scores. Student performance was highest on the free fall question (Q1) and lowest on questions probing understanding of Newton's second and third law (Q2 and Q3).

Furthermore, we wanted to evaluate the consistency of students' responses on isomorphic items, so we calculated the proportion of students who gave the same answer (correct or incorrect) on each set of isomorphic items (Fig. 3). On average, 41% of students gave the same correct answer, whereas 25% gave the same incorrect answer. Overall, about two-thirds of students gave the same answer (correct or incorrect) across all three representations (graphical, pictorial, and verbal).

For example, Fig. 2 shows that about 80% of students answered Q1 correctly for each of all three representations,

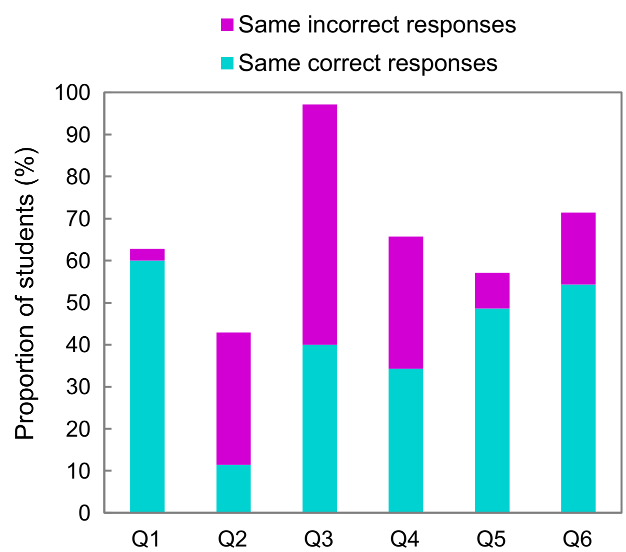

FIG. 3. The proportion of students who gave the same answer (correct or incorrect) on all three isomorphic items for each question (Q1–Q6).

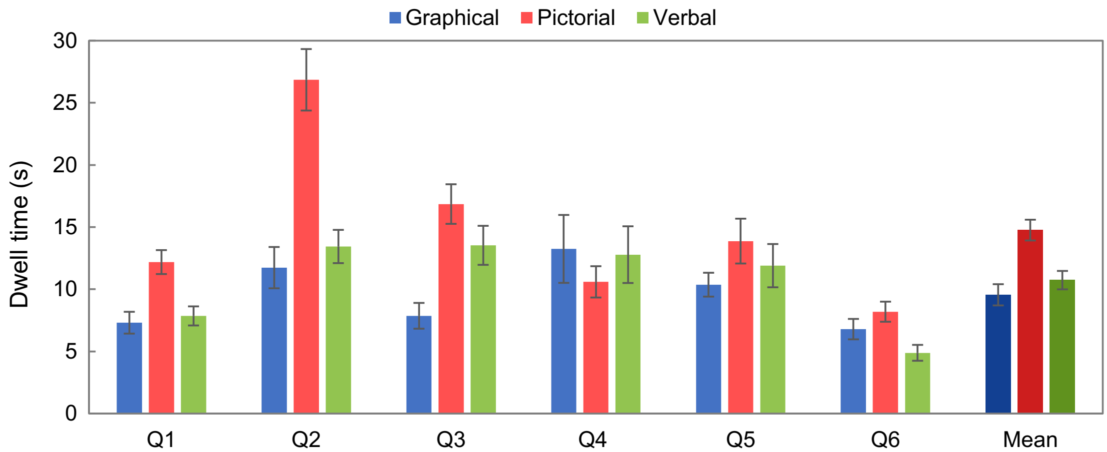

FIG. 4. The dwell time on AOI *representations* for graphical, pictorial, and verbal representations on all questions (Q1–Q6). The last three bars show the mean values of dwell time for each representation. The error bars represent 1 SEM.

whereas Fig. 3 shows that only 60% of students gave the correct answer across all three representations. Furthermore, there is only a very small percentage of those who gave the same incorrect answer across all three representations (about 3%), which indicates that the students did not have a preferred incorrect answer. About 20% of students gave the correct answer for one representation, but they gave at least one incorrect answer for the other two representations. Taken together, these results suggest that probably a portion of the students randomly guessed the correct answer to question Q1 for each of all three representations. On the other hand, on question Q3, students were very consistent in choosing answers across all three representations; 97% of students gave the same answer (correct or incorrect) for all three representations.

Figure 3 indicates that the proportion of students who gave the same incorrect answer was the highest on questions Q2, Q3, and Q4. As many as 57% of students chose the incorrect answer that a truck acts with a larger force on a car than a car on a truck on question Q3 regarding the application of Newton's third law. The preferred incorrect answer on question Q2 was that, after the force is doubled, the speed will also double but will be constant (31% of students chose that distractor in all three isomorphic items). The same proportion of students (31%) chose the same distractor that the kinetic energy of the sled will be equal to a quarter of the total initial energy of the sled when it is at the height $h/4$ above the base of the slope.

B. Analysis of eye-tracking data

The mean dwell time and standard deviation were (12.9 ± 3.8) s for AOI *question* and (11.7 ± 4.1) s for AOI *representations* that contained multiple-choice options. Since our goal was to investigate the effect of different representations on students' answers to conceptual questions, we will report the results on AOI *representations* below.

Figure 4 shows the dwell time on AOI *representations* for all items and mean values for graphical, pictorial, and verbal representations. The mean dwell times and standard deviations were (9.5 ± 5.0) s for graphical representation, (14.8 ± 4.9) s for pictorial representation, and (10.7 ± 4.4) s for verbal representation. One-way ANOVA revealed that representation had a statistically significant effect on dwell time $[F(2, 68) = 27.65,\ p < 0.0001,\ \eta_p^2 = 0.45]$. Pairwise comparisons showed that dwell time was longer for pictorial than graphical and verbal representations (both $p < 0.01$). This finding that dwell time is the longest for pictorial representation is quite consistent across all questions but Q4. By far, the longest dwell was found for pictorial representation on question Q2 (Newton's second law).

Further, we evaluated the average fixation duration for different representations on all questions (Fig. 5). The mean fixation duration and standard deviations were (265 ± 38) ms for graphical representation, (279 ± 37) ms for pictorial representation, and (234 ± 29) ms for verbal representation. Comparison of the average fixation duration showed a statistically significant effect of representation on average fixation duration $[F(2, 68) = 57.29,\ p < 0.0001$, and $\eta_p^2 = 0.63]$. Pairwise comparisons revealed that the average fixation duration was the longest for pictorial and shortest for verbal representation (all $p < 0.01$). Inspection of Fig. 5 suggests that these results are fairly consistent across all questions except Q6, especially the finding that average fixation duration is the shortest for verbal representation.

To explore the integration of information from multiple-choice options and question stem, we evaluated the number of revisits on AOI *representations* (Fig. 6). The mean number of revisits and standard deviations were (3.3 ± 1.8) s for graphical representation, (3.4 ± 1.1) s for pictorial representation, and (2.5 ± 1.4) s for verbal representation. The number of revisits on AOI *representations*

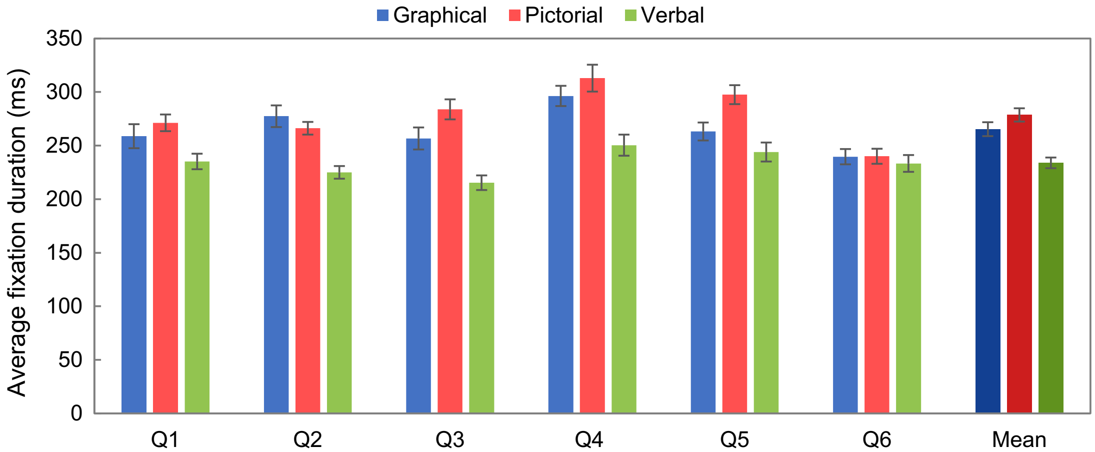

FIG. 5. The average fixation duration on AOI *representations* for graphical, pictorial, and verbal representations on all questions (Q1–Q6). The last three bars show the mean values of fixation duration for each representation. The error bars represent 1 SEM.

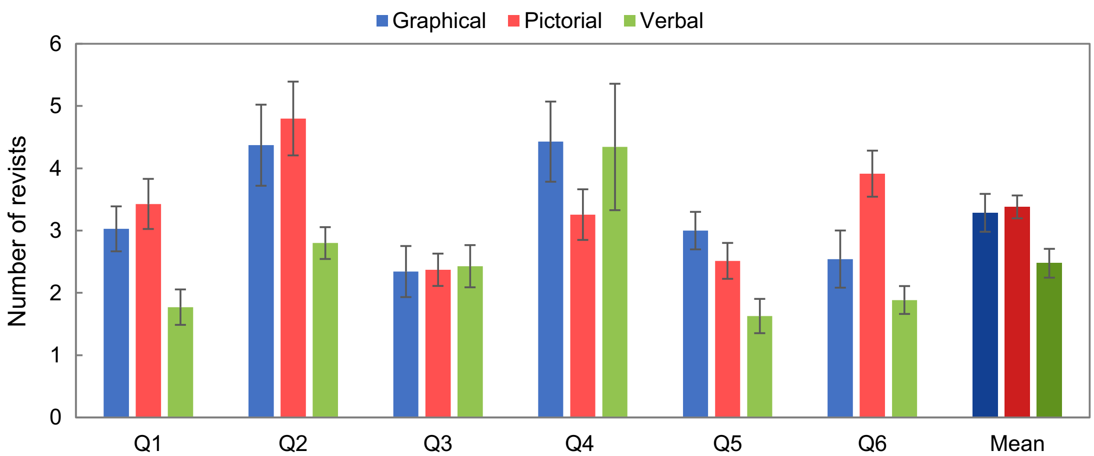

FIG. 6. The number of revisits on AOI *representations* for graphical, pictorial, and verbal representations on all questions (Q1–Q6). The last three bars show the mean values of the number of revisits for each representation. The error bars represent 1 SEM.

was affected by representation [$F(2,68) = 7.63$, $p = 0.001$, and $\eta_p^2 = 0.18$]. Pairwise comparisons showed that the number of revisits was the smallest for verbal representation (both $p < 0.01$). Examination of all items indicates that this result holds for questions Q1, Q2, Q5, and Q6 (Fig. 6).

VI. DISCUSSION

Students' scores were not different for isomorphic items in this study. Although students' average scores on specific questions varied significantly (ranging from 25% on question Q2 to 79% on question Q1), there was no question where students consistently performed better with one representation over the others. The overall result that no one representation helps students to better solve problems corroborates the results of previous studies [3–5]. However, the authors of the previous studies did find statistically significant differences in students' scores for different representations in some items. A possible reason why we did not find such differences is that, in our study, the same group of students solved isomorphic tasks with all representations, whereas Meltzer [3] and Kohl and Finkelstein [4] had different groups of students for different representations. Furthermore, it is possible that the different contexts of the questions also influenced the results because Kohl and Finkelstein concluded that performance on different representations depends on "the specific contextual features of the problems and the representations." Indeed, the most difficult questions were Q2 and Q3, which were related to Newton's second and third laws. This finding is consistent with previous PER studies that showed that students have a lot of misconceptions related to understanding Newton's laws [53]. A previous study showed that question Q2 is the most difficult question on the Force Concept Inventory (FCI) [54].

Further analysis of students' responses revealed that, on average, 66% of students were consistent, i.e., they gave the same answers (correct or incorrect), across all three representations. Interestingly, the students were most consistent on question Q3 about Newton's third law, where they constantly chose either the correct option or the incorrect option related to the well-known and strong preconception that the bigger or heavier object exerts more force [55,56]. Similarly, on question Q2 about Newton's second law, students were more consistent in choosing the incorrect option referring to the common preconception that velocity is proportional to the applied force [53] than the correct option. This finding agrees with the previous report that students are quite consistent on this question across different representations, but they do not give the correct answer [5]. On question Q4 about the conservation of energy, about a third of students were consistent in choosing the incorrect option, most likely because they confused kinetic and potential energy or because they simply followed the numerical cue in question. On question Q6 about the photoelectric effect, some students were consistent in choosing a distractor that the number of emitted electrons will remain the same if the intensity of the laser light increases. On the remaining two questions (Q1 and Q5), students were mostly consistent in the correct answers. Overall, our results indicate that students' consistency across representations depends on the probed concepts and this finding corroborates previous results by Nieminen *et al.* [5].

To explore the effect of representation on the efficacy of information extraction, we analyzed eye-tracking measures for graphical, pictorial, and verbal representations. The dwell time was the longest for pictorial representation and this result was particularly pronounced for questions Q2 and Q1 in which pictorial representation consisted of motion maps that indicated the position of an object at different times. Although students' scores were not statistically significantly different for graphical, pictorial, and verbal representations, i.e., students were not more successful for a certain type of representation, it took them longer to extract relevant information from pictorial representation used in this study than from other representations. This particularly holds for motion maps, for which a previous eye-tracking study has shown that they present greater problems for students than graphical representations [57]. As in the present study, the authors reported that "it was difficult to obtain enough information to arrive at the correct answer at a quick glance" [57].

The remaining two eye-tracking measures, the average fixation duration and the number of revisits, both indicate that verbal representation is the least demanding for students since the average fixation duration was the shortest and the number of revisits was the smallest for verbal representation. Average fixation duration can be considered a measure of cognitive load while the number of revisits is similar to the number of transitions and indicates the integration of information from different sources [12]. For example, experts have shorter fixation durations than nonexperts [58]. More complex tasks usually require a higher cognitive effort to process information, and this can be reflected in the longer average fixation duration. Consequently, the results of this study suggest that it was easiest for students to extract information from verbal representations.

The limitation of this study is a rather small number of isomorphic questions, mostly from mechanics. In future studies, additional isomorphic questions from other areas of physics should be used. Although the number of participants in this study is typical for eye-tracking studies [59], it would be desirable to confirm these findings on different samples, preferably from different educational systems, to be able to draw more general conclusions. Also, if a larger number of students participated, it would be possible to evaluate how a higher level of expertise for a certain representation affects eye-tracking measures. In the following studies, we plan to investigate this question. Furthermore, we are planning to investigate the effect of teaching with multiple representations, which usually happens in the physics classroom and physics textbooks.

In summary, the results of this study show that none of the representations was advantageous in terms of students' scores, but it seems that it was easiest to extract information from verbal representations and most difficult from pictorial representations. This conclusion about verbal representation might be quite general because most likely we decode information from graphical and pictorial representations to some form of verbal (or symbolic) information in order to make inferences. So, verbal representation might be some sort of elementary representation, at least for simple conceptual questions used in this study. Our finding that pictorial representation is the most difficult for the extraction of information might be related to the specific pictorial representations used in this study. From the previous studies, it is known that motion maps and vectors are difficult for students [5,57,60–62], and if questions with different types of pictorial representations were used, the result might be different. It also might be related to the need that students should receive explicit instructions and practice to productively use multiple representations for problem solving [2], and students are more familiar with verbal and graphical representations from textbooks and school instructions than with pictorial ones. However, physics teachers should be aware that it is not equally difficult for students to extract information from different representations. They should take that into account when creating conceptual questions and more generally when teaching with multiple representations.

[1] D. F. Treagust, R. Duit, and H. E. Fischer, *Multiple Representations in Physics Education* (Springer International Publishing, Cham, 2017).
[2] A. Van Heuvelen, Learning to think like a physicist: A review of research-based instructional strategies, Am. J. Phys. **59**, 891 (1991).
[3] D. E. Meltzer, Relation between students' problem-solving performance and representational format, Am. J. Phys. **73**, 463 (2005).
[4] P. B. Kohl and N. D. Finkelstein, Student representational competence and self-assessment when solving physics problems, Phys. Rev. ST Phys. Educ. Res. **1**, 010104 (2005).
[5] P. Nieminen, A. Savinainen, and J. Viiri, Force concept inventory-based multiple-choice test for investigating students' representational consistency, Phys. Rev. ST Phys. Educ. Res. **6**, 020109 (2010).
[6] A. Van Heuvelen and X. Zou, Multiple representations of work–energy processes, Am. J. Phys. **69**, 184 (2001).
[7] P. B. Kohl, D. Rosengrant, and N. D. Finkelstein, Strongly and weakly directed approaches to teaching multiple representation use in physics, Phys. Rev. ST Phys. Educ. Res. **3**, 010108 (2007).
[8] D. Rosengrant, A. Van Heuvelen, and E. Etkina, Do students use and understand free-body diagrams?, Phys. Rev. ST Phys. Educ. Res. **5**, 010108 (2009).
[9] D. McPadden and E. Brewe, Impact of the second semester University Modeling Instruction course on students' representation choices, Phys. Rev. Phys. Educ. Res. **13**, 020129 (2017).
[10] A. Kohnle and G. Passante, Characterizing representational learning: A combined simulation and tutorial on perturbation theory, Phys. Rev. Phys. Educ. Res. **13**, 020131 (2017).
[11] J. Scheid, A. Müller, R. Hettmannsperger, and W. Schnotz, Improving learners' representational coherence ability with experiment-related representational activity tasks, Phys. Rev. Phys. Educ. Res. **15**, 010142 (2019).
[12] L. Hahn and P. Klein, Eye tracking in physics education research: A systematic literature review, Phys. Rev. Phys. Educ. Res. **18**, 013102 (2022).
[13] S. Brückner, J. Schneider, O. Zlatkin-Troitschanskaia, and H. Drachsler, Epistemic network analyses of economics students' graph understanding: An eye-tracking study, Sensors **20**, 6908 (2020).
[14] S. Brückner, O. Zlatkin-Troitschanskaia, S. Küchemann, P. Klein, and J. Kuhn, Changes in students' understanding of and visual attention on digitally represented graphs across two domains in higher education: A postreplication study, Front. Psychol. **11**, 1 (2020).
[15] M. Kekule, Students' approaches when dealing with kinematics graphs explored by eye-tracking research method, Eur. J. Sci. Math. Educ. **2**, 108 (2014).
[16] P. Klein, A. Lichtenberger, S. Kuchemann, S. Becker, M. Kekule, J. Viiri, C. Baadte, A. Vaterlaus, and J. Kuhn, Visual attention while solving the test of understanding graphs in kinematics: An eye-tracking analysis, Eur. J. Phys. **41**, 025701 (2020).
[17] P. Klein, S. Küchemann, S. Brückner, O. Zlatkin-Troitschanskaia, and J. Kuhn, Student understanding of graph slope and area under a curve: A replication study comparing first-year physics and economics students, Phys. Rev. Phys. Educ. Res. **15**, 020116 (2019).
[18] P. Klein, S. Becker, S. Küchemann, and J. Kuhn, Test of understanding graphs in kinematics: Item objectives confirmed by clustering eye movement transitions, Phys. Rev. Phys. Educ. Res. **17**, 013102 (2021).
[19] 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).
[20] C. J. Wu and C. Y. Liu, Eye-movement study of high- And low-prior-knowledge students' scientific argumentations with multiple representations, Phys. Rev. Phys. Educ. Res. **17**, 010125 (2021).
[21] 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 Eye-tracking of visual attention in web-based assessment using the Force Concept Inventory, Eur. J. Phys. **38**, 045702 (2017).
[22] E. Hejnová and M. Kekule, Observing students' problem solving strategies in mechanics by the eye-tracking method, Sci. Educ. **9**, 102 (2018).
[23] M. Kekule and J. Viiri, Students' approaches to solving R-FCI tasks observed by eye-tracking method, Sci. Educ. **9**, 117 (2018).
[24] 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).
[25] A. Susac, A. Bubic, M. Planinic, M. Movre, and M. Palmovic, Role of diagrams in problem solving: An evaluation of eye-tracking parameters as a measure of visual attention, Phys. Rev. Phys. Educ. Res. **15**, 013101 (2019).
[26] 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).
[27] P. Klein, J. Viiri, and J. Kuhn, Visual cues improve students' understanding of divergence and curl: Evidence from eye movements during reading and problem solving, Phys. Rev. Phys. Educ. Res. **15**, 010126 (2019).
[28] S. Küchemann, S. Malone, P. Edelsbrunner, A. Lichtenberger, E. Stern, R. Schumacher, R. Brünken, A. Vaterlaus, and J. Kuhn, Inventory for the assessment of representational competence of vector fields, Phys. Rev. Phys. Educ. Res. **17**, 020126 (2021).
[29] T. van Gog, F. Paas, and J. J. G. van Merriënboer, Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data, Appl. Cogn. Psychol. **19**, 205 (2005).
[30] C. Hoyer and R. Girwidz, Animation and interactivity in computer-based physics experiments to support the documentation of measured vector quantities in diagrams: An eye tracking study, Phys. Rev. Phys. Educ. Res. **16**, 020124 (2020).
[31] S. Küchemann, P. Klein, H. Fouckhardt, S. Gröber, and J. Kuhn, Students' understanding of non-inertial frames of reference, Phys. Rev. Phys. Educ. Res. **16**, 010112 (2020).

[32] A. Susac, M. Planinic, A. Bubic, L. Ivanjek, and M. Palmovic, Student recognition of interference and diffraction patterns: An eye-tracking study, Phys. Rev. Phys. Educ. Res. **16**, 020133 (2020).

[33] A. Susac, M. Planinic, A. Bubic, K. Jelicic, L. Ivanjek, K. Matejak Cvenic, and M. Palmovic, Effect of students' investigative experiments on students' recognition of interference and diffraction patterns: An eye-tracking study, Phys. Rev. Phys. Educ. Res. **17**, 010110 (2021).

[34] 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).

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

[36] 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).

[37] 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).

[38] 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).

[39] J. Škrabánková, S. Popelka, and M. Beitlová, Students' ability to work with graphs in physics studies related to three typical student groups, J. Balt. Sci. Educ. **19**, 298 (2020).

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

[41] R. E. Mayer, Multimedia learning: Are we asking the right questions?, Educ. Psychol. **32**, 1 (1997).

[42] W. Schnotz and M. Bannert, Construction and interference in learning from multiple representation, Learn. Instr. **13**, 141 (2003).

[43] J. Sweller, P. Ayres, and S. Kalyuga, *Cognitive Load Theory* (Springer, New York, 2011).

[44] J. Sweller, J. J. G. van Merriënboer, and F. Paas, Cognitive architecture and instructional design: 20 years later, Educ. Psychol. Rev. **31**, 261 (2019).

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

[46] B.D. Zumbo and A.M. Hubley, *Understanding and Investigating Response Processes in Validation Research* (Springer International Publishing, New York, 2017).

[47] M. A. Just and P. A. Carpenter, A theory of reading: From eye fixations to comprehension, Psychol. Rev. **87**, 329 (1980).

[48] A. Duchowski, *Eye Tracking Methodology. Theory and Practice* (Springer, London, 2007).

[49] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van De Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures* (Oxford University Press, Oxford, 2011).

[50] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.19.020114 for six sets of isomorphic items used in the study, each consisting of graphical, pictorial, and verbal representations.

[51] A. P. Field, *Discovering Statistics Using IBM SPSS Statistics* (Sage, London, 2013).

[52] F. J. Gravetter, L. B. Wallnau, L.-A. B. Forzano, and J. E. Witnauer, *Essentials of Statistics for the Behavioral Sciences*, 10th ed. (Cengage, Boston, 2020).

[53] D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, Phys. Teach. **30**, 141 (1992).

[54] M. Planinic, L. Ivanjek, and A. Susac, Rasch model based analysis of the Force Concept Inventory, Phys. Rev. ST Phys. Educ. Res. **6**, 010103 (2010).

[55] L. Bao, K. Hogg, and D. Zollman, Model analysis of fine structures of student models: An example with Newton's third law, Am. J. Phys. **70**, 766 (2002).

[56] T. I. Smith and M. C. Wittmann, Comparing three methods for teaching Newton's third law, Phys. Rev. ST Phys. Educ. Res. **3**, 020105 (2007).

[57] J. Viiri, M. Kekule, J. Isoniemi, and J. Hautala, Eyetracking the effects of representation on students problem solving approaches, in *Proceedings of the Finnish Mathematics and Science Education Research Association (FMSERA) Annual Symposium*, edited by M. A. Asikainen and P. E. Hirvonen (2017), pp. 88–98.

[58] A. Gegenfurtner, E. Lehtinen, and R. Säljö, Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains, Educ. Psychol. Rev. **23**, 523 (2011).

[59] A. R. Strohmaier, K. J. MacKay, A. Obersteiner, and K. M. Reiss, Eye-tracking methodology in mathematics education research: A systematic literature review, Educ. Stud. Math. **104**, 147 (2020).

[60] A. F. Heckler and T. M. Scaife, Adding and subtracting vectors: The problem with the arrow representation, Phys. Rev. ST Phys. Educ. Res. **11**, 010101 (2015).

[61] P. Barniol and G. Zavala, Test of understanding of vectors: A reliable multiple-choice vector concept test, Phys. Rev. ST Phys. Educ. Res. **10**, 010121 (2014).

[62] A. Susac, M. Planinic, D. Klemencic, and Z. Milin, Sipus, using the Rasch model to analyze the test of understanding of vectors, Phys. Rev. Phys. Educ. Res. **14**, 023101 (2018).