

## Instruction-based clinical eye-tracking study on the visual interpretation of divergence: How do students look at vector field plots?

P. Klein,<sup>1,\*</sup> J. Viiri,<sup>2</sup> S. Mozaffari,<sup>3</sup> A. Dengel,<sup>3,4</sup> and J. Kuhn<sup>1</sup>

<sup>1</sup>*Department of Physics, Physics Education Research Group, University of Kaiserslautern, Erwin-Schrödinger-Strasse 46, 67663 Kaiserslautern, Germany*

<sup>2</sup>*Department of Teacher Education, University of Jyväskylä, 40014 Jyväskylä, Finland*

<sup>3</sup>*Smart Data & Knowledge Services, German Research Center for Artificial Intelligence, Trippstadter Strasse, 122, 67663 Kaiserslautern, Germany*

<sup>4</sup>*Computer Science Department, University of Kaiserslautern, Erwin-Schrödinger-Strasse 46, 67663 Kaiserslautern, Germany*



(Received 13 November 2017; published 28 March 2018)

Relating mathematical concepts to graphical representations is a challenging task for students. In this paper, we introduce two visual strategies to qualitatively interpret the divergence of graphical vector field representations. One strategy is based on the graphical interpretation of partial derivatives, while the other is based on the flux concept. We test the effectiveness of both strategies in an instruction-based eye-tracking study with  $N = 41$  physics majors. We found that students' performance improved when both strategies were introduced (74% correct) instead of only one strategy (64% correct), and students performed best when they were free to choose between the two strategies (88% correct). This finding supports the idea of introducing multiple representations of a physical concept to foster student understanding. Relevant eye-tracking measures demonstrate that both strategies imply different visual processing of the vector field plots, therefore reflecting conceptual differences between the strategies. Advanced analysis methods further reveal significant differences in eye movements between the best and worst performing students. For instance, the best students performed predominantly horizontal and vertical saccades, indicating correct interpretation of partial derivatives. They also focused on smaller regions when they balanced positive and negative flux. This mixed-method research leads to new insights into student visual processing of vector field representations, highlights the advantages and limitations of eye-tracking methodologies in this context, and discusses implications for teaching and for future research. The introduction of saccadic direction analysis expands traditional methods, and shows the potential to discover new insights into student understanding and learning difficulties.

DOI: [10.1103/PhysRevPhysEducRes.14.010116](https://doi.org/10.1103/PhysRevPhysEducRes.14.010116)

### I. INTRODUCTION

A vector field is a structure in which a vector is assigned to every point in (a subset) of space. Vector fields are important in many branches of physics, and students are confronted with them from the very beginning of the university curriculum. Examples include Newton's gravitational field, velocity fields of fluids, and electromagnetic fields.

Divergence is a mathematical concept that applies to vector fields. It is a vector operator that produces a scalar field giving the quantity of a vector field's source at each point in space. Technically speaking, the divergence represents the volume density of the outward flux of a vector field

from an infinitesimal volume around a given point. The divergence operator ( $\vec{\nabla}$ ) occurs, for example, in Maxwell's equations or in continuity equations describing the transport of some quantity. Divergence can be interpreted locally (sources or sinks at specific points) or globally (net flow outwards of an area or volume). This intuition is made precise by the divergence theorem of Gauss.

The concept of divergence can be expressed in many different ways, for example, textual descriptions, symbolic expressions, graphs, pictures, etc. [1]. These different formats are called representations. As vector fields are often expressed in a graphical representation [2], developing a visual understanding of divergence is important for learning. Both interpretations of divergence mentioned above can be applied qualitatively to a vector field representation. We refer to them as the differential strategy (DS) and the integral strategy (IS) and illustrate them in Fig. 1 for a simple but useful case for teaching (see comments below).

\*pklein@physik.uni-kl.de

Published by the American Physical Society under the terms of the *Creative Commons Attribution 4.0 International* license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

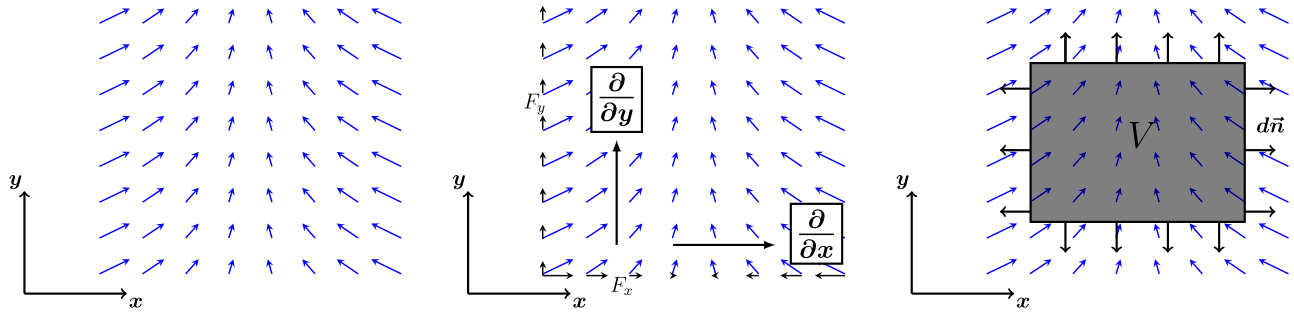


FIG. 1. Graphical representation of a two-dimensional vector field (left). Divergence of this field is constant and negative, as we can prove qualitatively. Using the differential strategy (DS), we conclude that  $F_x$  becomes smaller in the  $x$  direction and  $F_y$  is constant (center). Using the integral strategy (IS), we observe a negative net flux through the left and right boundaries of a test volume, while there is no net flux through the upper and lower boundaries (right). Both arguments hold for every position in the field and for any test volume size.

(DS) Using the differential representation of divergence (e.g., in Cartesian coordinates  $x$  and  $y$ ), that is,

$$\text{div}\vec{F} = \vec{\nabla} \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y}, \quad (1)$$

one can evaluate the divergence of a two-dimensional vector field  $\vec{F}$  by interpreting the partial derivatives. To put it clearly, one must inspect the change of the field components  $F_x$  and  $F_y$  in the  $x$  and  $y$  directions, respectively. We indicate this in Fig. 1 (center) by a vector decomposition and by highlighting the  $x$  and  $y$  directions. If the vector field has constant (zero or nonzero) divergence, this procedure leads to the same result for every position  $(x, y)$  in the field.

(IS) Divergence can also be evaluated using its integral representation, that is,

$$\text{div}\vec{F} = \lim_{V \rightarrow 0} \frac{1}{V} \int_{\partial V} \vec{F} \cdot d\vec{n}. \quad (2)$$

Given a visual field representation, one can estimate divergence with the flux through the boundary  $\partial V$  of a test volume  $V$  in the field (or test areas in the two-dimensional case). Test volumes can take any form, but in many cases they are constructed with respect to the field’s symmetry. For instance, qualitative reasoning is simple if the outer surface normal  $d\vec{n}$  is either parallel or perpendicular to the field vector  $\vec{F}$  (e.g., cuboids or spheres in 3D and rectangles or circles in 2D).

In simple cases that are useful for teaching and assessment, the vector field is two dimensional, only depends on two coordinates,  $x$  and  $y$ , and divergence is constant (either zero or nonzero, independent of position). In other words, the divergence does not vary spatially what can be considered as a restriction for educational and methodological reasons [3]. Students then qualitatively estimate whether divergence

is zero or nonzero (cf. caption of Fig. 1). Both strategies (DS and IS) are highly visual and address the students’ ability to relate mathematical concepts to graphical representations in a qualitative manner. This ability is considered very important for student understanding and problem solving [4–8].

## A. Theoretical background and state of research

### 1. Representational competence

In a more general form, the ability to relate mathematical concepts to various representations contributes to representational competence, that is, being able to use multiple representations for understanding. Research in a variety of domains has demonstrated that visual representations of abstract concepts have the potential to substantially promote learning [7–11]. However, making connections between mathematical concepts and graphical representations is a difficult task that students often do not engage in spontaneously, even though it is critical to their learning [12,13]. Therefore, students need instructional support to make these connections, especially when concepts become more complex and sophisticated. In this line of reasoning, physics education research (PER) of the last decades has demonstrated that using multiple representations in instruction facilitates students’ conceptual understanding [14–18].

In our research, the mathematical concept of divergence is applied in two different ways (DS and IS) to graphical representations of vector field plots. Both strategies emphasize complementary conceptual aspects of the learning material and might have different effects on mental processing; a similar effect has already been explored in a different context [11,19]. Based on research about representational competence, we can assume that a solid visual understanding of both strategies mentioned above constitutes for students conceptual understanding of vector field concepts. This point has also been made by Feynman in a more general way. Feynman has placed great emphasis

on the importance of deriving different formulations for the same physical law, even if they are mathematically equivalent (e.g., Newton's second law, the minimum principle, etc.). He reminded us that different representations of a physical concept can evoke varied mental pictures and can therefore assist in making new discoveries [20]:

*“Psychologically they [representations of physical concepts] are different because they are completely inequivalent when you are trying to guess new laws.”*

Both representations of divergence (DS and IS) are mathematically equivalent (theorem of Gauss), yet they might support learning differently. In this study, we explore whether multiple representational approaches to divergence complement each other to be beneficial for learning or not [21].

## 2. Student understanding of divergence

Previous research has shown that many students lack a coherent picture of the meaning of divergence. For instance, Singh and Maries [22] found that even graduate students had great difficulty in recognizing whether presented vector field diagrams had nonzero divergence and/or curl. Only about half of the students succeeded in this qualitative task without any instruction. However, these students had no difficulty computing the divergence of the same vector fields analytically.

Similar results were found by Bollen *et al.*, who investigated students' understanding of divergence and curl in the context of electromagnetism [23,24]. Based on student interviews and written responses to several questions, they have categorized strategies students used to interpret graphical representations of electromagnetic fields in terms of divergence (or curl) as either *concept based*, *physics based*, *formula based*, *description based*, or *unclear*. Bollen *et al.* found that many students used various strategies inconsistently, suggesting that students lack a structural understanding of the mathematical concepts [23]. They found that students lacked a conceptual understanding of the vector operators and struggled to interpret graphical representations of vector fields. This result was confirmed in a subsequent work [24], strongly indicating that traditional instruction is not sufficient for our students to fully understand the meaning and power of differential operators in electrodynamics. Furthermore, in the framework of conceptual blending, they showed how the lack of conceptual understanding accounts for an improper understanding of Maxwell's equations. Their work underlines the general ideas about the relationship between representational competence and conceptual understanding described in Sec. I A 1 and advocates using specific learning materials based on multiple representations to clarify vector calculus concepts in physics. They conclude [23] (p. 10) as follows:

*“Since we [Bollen et al.] (and many others) believe that these graphical representations are helpful when trying to conceptualize the abstract mathematical structures in vector calculus, we think it would be advisable to put more effort into this kind of exercise in both physics and mathematics instruction.”*

Other work has revealed that students hold common incorrect associations of a nonzero divergence with regions in a diagram where field lines spread apart [25], that students lack a geometrical understanding of partial derivatives in the context of vector calculus [26], and that students tend to think about the common English meaning of the word divergence and decide whether a field looks to be diverging rather than applying a rigorous procedure for determination [26].

There is consensus among these studies that divergence should be taught using more visual representations of vector fields. However, so far, there has been little work to explicitly provide instructions about how visual strategies could be used. As a result, conceptual differences between both representations of divergence (integral and differential) have yet to be explored empirically. While Bollen *et al.* [24] have elaborated the formal differences between both interpretations (p. 7–8), it remains unknown how they are processed cognitively by students and which one promotes better visual understanding of the concept of divergence. Both approaches appear in upper-division physics education and complement each other in terms of local and global understanding of differential operators. To gain a deeper insight into student difficulties on this topic, we recorded their eye movements when employing both visual strategies (DS and IS).

## B. Eye tracking and PER

Eye tracking is a nonintrusive method to obtain information about visual attention and cognitive processing while students solve problems, particularly where visual strategies are involved. A robust body of research has shown that eye gaze may be considered an unbiased indicator of the focus of visual attention [27,28] and that eye-tracking measures can be related to performance measures, particularly in the context of processing visual stimuli [29]. The most often used eye-tracking measures are derived from fixations and saccades (see Fig. 2). While fixations are relatively long periods usually lasting between 100 and 600 ms in which the eye is almost still, saccades are very fast movements of the eye between fixations lasting less than 100 ms. Eye movement data are typically interpreted based on the eye-mind assumption [30], meaning that fixations reflect the attention and cognitive processing at specific locations, and are determined by the perceptual and cognitive analysis of the information at that location. Longer fixations indicate deeper or more difficult processing.

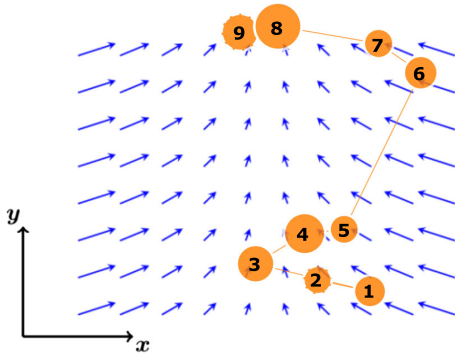


FIG. 2. Excerpt of an individual gaze plot during problem solving. Regular eye movement alternates between fixations (orange circles, centered at the position of fixation) and saccades (lines between fixations). Increasing numbers reflect the chronological order of fixations, and the size of the circles illustrates the duration of the fixation. A short video sequence can be found in the Supplemental Material [31].

Eye-tracking research has received increasing attention within the PER community over the last few years but is still rare compared to, for example, cognitive science. Recent studies have (i) explored students' visual attention in different areas of interest (AOI) to distinguish between experts and novices or between good and poor performances [32–34], (ii) used transition analysis to investigate how students switch back and forth between different AOIs while learning or problem solving to get insight into student expertise or the characteristics of the learning material [35–38], (iii) used classical eye-tracking measures as fixation durations or the number of fixations to predict learning outcomes [39], and (iv) touched on problems and issues peripherally related to PER, such as the troubleshooting of malfunctioning circuits [40], comprehending malfunctioning mechanical devices or how mechanical systems work [41,42], and how spatial ability influences solving kinematics problems with trajectories [43].

Following up previous research, in this work we have used classical eye-tracking measures to explore the differences between both visual strategies and to find correlations to learning outcomes. Moreover, we applied deeper analysis procedures to interpret students' visual behavior when they applied the DS and IS: Concerning the IS, we analyzed whether students used the field's symmetry to balance positive or negative flux through the boundaries of a rectangle. Concerning the DS, we expect that students are looking in horizontal and vertical directions corresponding to the direction of the coordinate axes,  $x$  and  $y$ , respectively; cf. Eq. (1). Therefore, we have analyzed absolute saccadic angles between subsequent fixations. Saccadic directions were previously analyzed in the context of reading [44], general picture or scene viewing [45], and binocular research [46] but—to the best of our knowledge—not in educational studies. More methodological details are given in Sec. II E.

## C. Research questions

Here, we explore the potential of eye tracking when students apply the concept of divergence to visual representations of vector fields to identify student difficulties and mastery of the strategies, as well as differences between the strategies. Eye tracking is complemented with performance measures and findings from student interviews. Specifically, we address the following research questions:

- RQ 1 Which visual strategy promotes better learning according to performance scores?
- RQ 2 What difficulties do students experience while applying the strategies?
- RQ 3 What differences, if any, are there in visual processing between both strategies?
- RQ 4 How well can eye-tracking measures discriminate between correct and incorrect answers or between best performers and worst performers?

In the next section, we give details about the experimental setup to answer the research questions.

## II. METHODS

### A. Participants

Forty-one physics majors from the University of Technology Kaiserslautern (TUK) took part in the experiment (37 male, average 20.9 yr). All participants completed two mechanics courses (calculus-based mechanics and experimental physics) that introduced the concept of divergence. When the experiment was conducted, all students were taking an introductory electromagnetism course. We recruited the students during lecture time without telling them explicitly what the study was about. Participation was voluntary and was compensated with 10 €. Our sample covered 57% of all students in class (total course enrollment was 72). All students had normal or correct-to-normal vision.

As stated by the faculty, the concept of divergence has already been used several times in the electromagnetism lecture. This means that students have seen graphical representations of vector fields before and have also been instructed using typical textbook examples of vector fields with zero and nonzero divergence. However, there was no previous instruction aiming explicitly at assessing the divergence of vector fields using visual methods (DS or IS), either in class or during homework.

### B. Study design and procedure

The study design of the experiment is illustrated in Fig. 3. All students came one by one to our eye-tracking lab where a researcher guided them through the experiment. First, they completed the preinstruction survey (approximately 10 min) to ensure that they had sufficient prerequisite knowledge of the concepts required to understand the instructions.

At the beginning of the experiment, students were placed in front of a computer screen, and a nine-point calibration



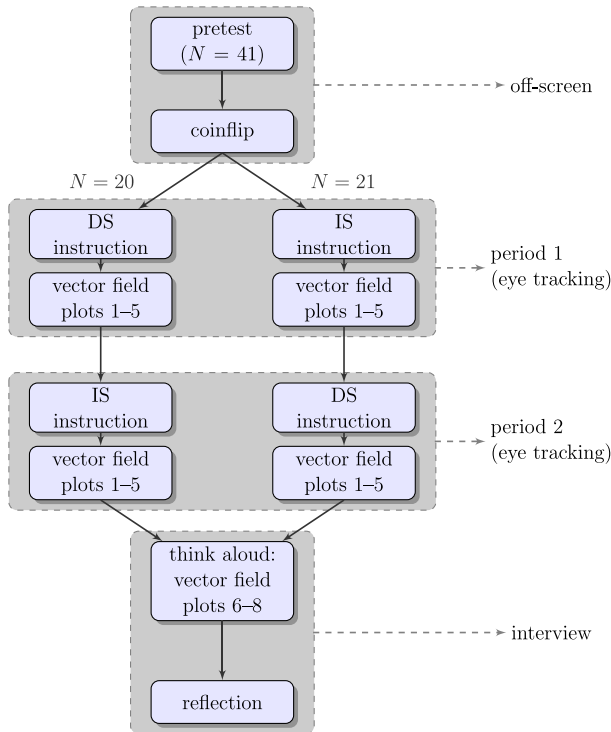


FIG. 3. Study design. During the instruction periods, students first learned one strategy, applied it to five vector fields, and then repeated this procedure with the other strategy. At the end, each student was free to choose one strategy (or both) to solve vector fields V6–V8 while thinking aloud and commenting about difficulties.

and validation procedure was used. To obtain reliable data on students' eye movements, we conducted the clinical study in a darkened room. As shown in Fig. 3, each student randomly started with either the DS ( $N = 20$ ) or the IS ( $N = 21$ ). The experiment (period 1) started with the instruction page (DS or IS). Students read the material without any teaching by the researcher. They were free to use as much time as needed before proceeding to the first task. Students then applied the strategy to five vector fields that were presented one after another (see Sec. II C for more details). The students were not allowed to take notes or use any aids. Whenever a student was ready to give an answer, they pressed a button and gave their answer (divergence is zero or nonzero). The students were instructed to avoid speaking during decision making because this would influence eye movements. They did not receive any feedback after completing a task and were unable to skip back to the instruction page. Afterwards, the procedure was repeated with the other strategy (period 2). The students were told to apply only the current strategy.

As can be seen in Fig. 3, each student applied both strategies to five different vector field plots. To balance recognition and group effects, half of the students started with the DS, while the other half started with the IS. Almost equal numbers of participants were assigned to each

condition. After completing both instruction periods, a short think-aloud interview study was conducted. Students were given three more vector field plots and were encouraged to use one strategy they prefer or both strategies to decide whether divergence is zero or not. Participants' verbal explanations were recorded using the built-in microphone of the computer. Last, they reflected about their understanding of both strategies. The whole procedure was completed within 30 min.

## C. Materials

All study material is included as Supplemental Material (translated from the original German into English) [31]. It consisted of the pretest, two instruction pages, five vector field plots (V1–V5) presented during the eye-tracking study, three similar vector field plots (V6–V8) presented during student interviews, and a short interview protocol.

### 1. Pretest

The pretest consisted of several true-false and open-response items aiming at determining prerequisite knowledge of differential and integral approaches to divergence and requiring students to express their response confidence on a rating scale. The main purpose of the pretest was to ensure that students had sufficient prerequisite knowledge of the concepts required to understand the instructions.

### 2. Instruction pages

Each instruction page covered 250 written words and two figures. It included an introduction to the problem and a step-by-step description of how the strategy has to be applied (worked-out example). The intention of the instructions was to give a standardized explanation of both visual strategies to all students. The instruction for the DS focused on the vector decomposition and the interpretation of the partial derivatives, while the IS instruction explained balancing the flux through a rectangle; cf. Fig. 1. Both instructions used the same vector field plot to explain the different strategies. This vector field plot was one dimensional and different from those used to measure student performances. The instruction pages were checked by four faculty members (theoretical physicists) and five physics education researchers, who can all be considered discipline experts and who have several years of teaching experience.

### 3. Vector field plots

The vector field plots V1–V8 were presented next to a two-dimensional  $x$ - $y$ -coordinate system and had either zero or constant nonzero divergence. Similar problems have been used in the PER community before [22,24]. The mathematical features of the vector fields are summarized in Table I. Vector fields V1–V3 and V6 only have one nonzero field component (one-dimensional fields), while the other vector fields are two dimensional. In each case,

TABLE I. Features of vector field plots used in the eye-tracking study.

Vector field	Nonzero field components	Direction of change
V1	$F_y$	$y$
V2	$F_y$	$x$
V3	$F_x$	$x$
V4	$F_x, F_y$	$y$
V5	$F_x, F_y$	$x$
V6	$F_x$	$y$
V7	$F_x, F_y$	$y$
V8	$F_x, F_y$	$x, y$

vector field arrows change their length or orientation with respect to one (or two) coordinate(s). Vector fields V1, V3, V5, and V7 have nonzero divergence. The vector field plot V5 is also presented in Fig. 1.

#### 4. Interview protocol

The interview protocol covered questions about the students' understanding of each strategy, their preferences, and any difficulties they encountered during problem solving.

#### D. Eye-tracking apparatus

Participants were provided with instructions and tasks on a 22-inch computer screen. The resolution of the computer screen was set to  $1280 \times 960$  pixels with a refresh rate of 75 Hz. The average distance from the participant to the monitor was  $63.1 \pm 8.7$  cm. Each vector field plot covered an area of  $400 \times 400$  pixels corresponding to  $12 \times 12$  degrees of visual angle. Eye movements were recorded with a Tobii X3-120 eye-tracking system [47], which had an accuracy of less than  $0.40^\circ$  of visual angle (as reported by the manufacturer) and a sampling frequency of 120 Hz. The system allows a relatively high degree of freedom in terms of head movement (no chin rest was used).

To detect fixations and saccades, we adopted an I-VT algorithm [48]. An eye movement was classified as a saccade (i.e., in motion) if the eyes' acceleration exceeded  $8500^\circ/s^2$  and the velocity exceeded  $30^\circ/s^2$ .

#### E. Data analysis methods

As stated in the last paragraph of Sec. IB, a correct application of the DS implies vertical and horizontal saccades when the vector field is being looked at. A suitable method to obtain quantitative information about the saccadic directions of a student's eye movements is given by *saccadic direction analysis* [49] (p. 303). Technical details are given in the Supplemental Material [31].

Concerning the IS, we constructed AOIs defined by the symmetry of vector fields. Based on the number of fixations in symmetrical AOIs, a symmetry score, defined by

$$s = \left| \frac{n_1 - n_2}{n_1 + n_2} \right|, \quad (3)$$

provides information about a student's focus within the vector field plot. It compares the relative frequencies of hitting both AOIs, where  $n_1$  and  $n_2$  are total numbers of fixations in two AOIs covering mirror-symmetrical parts of the vector field, respectively. It is  $s = 1$  if a student only looks at one AOI,  $s = 0$  if students look at both AOIs equally often, and  $s$  increases, the more students tend to focus on one AOI. More technical details are given in the Supplemental Material [31].

### III. RESULTS

#### A. Pretest

In summary, most participants achieved a high score on the pretest questions, demonstrating knowledge of the concepts required to follow the instructions (average score and standard deviation of all questions:  $0.84 \pm 0.09$ , range 0.60–0.95). Only three students achieved an average score below 65%. We refer to these students as L1, L2, and L3, respectively, in the sections below. It is worth noting that performance did not differ significantly between both groups defined by the experimental condition (cf. Fig. 3). Interested readers find an item-based discussion of the pretest results in the Supplemental Material [31].

#### B. Student performance on vector fields V1–V5 (RQ 1)

Figure 4 presents the distribution of total scores after completing both test periods. For each student, the performance score reflects the number of correct answers (divergence is zero or non-zero) achieved in the experiment. The maximum score of 10 was achieved by only two out of 41 students, while three students scored 4 points or less. The lowest scores were obtained by those students who also achieved the lowest scores in the pretest (L1, L2, L3), while the students with maximum scores also scored high on the pretest (average 0.90 [50]). We refer to these

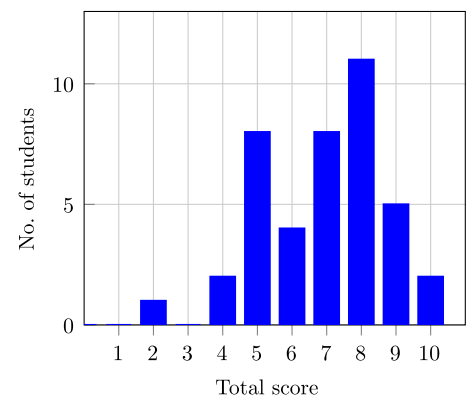


FIG. 4. Distribution of total scores after completing both periods.

extreme performance groups as *best performers* (B1, B2) and *worst performers* (L1, L2, L3) in later sections, respectively. The mean score is  $6.9 \pm 1.8$  points what can be considered as mediocre given that the chance of guessing correct is 50%, that means, guessing every question yields a score of 5.0.

Table II shows the average performance scores for each strategy and each vector field V1–V5. In sum, the DS yielded a correct response in 66% of the cases, and the IS in 72% of the cases, without significant differences overall. However, the effectiveness of each strategy varies across specific vector field representations. It is notable that the DS results in lower scores if the vector field has zero divergence (cases V2 and V4) but is more successful when divergence is nonzero (cases V1 and V5). In other words, students are more inclined to accredit divergence to vector fields when using the DS. If this assumption holds, students might have answered correctly for the wrong reason in some case. This will be investigated further (Secs. III D 2 and III F).

Considering instruction periods 1 and 2 separately, we observe similar trends and a repetition effect. There are no differences on test level (all items) but on individual item level between strategies. Total DS and IS scores increased significantly between both test periods (DS: from 0.61 to 0.71; IS: from 0.68 to 0.76;  $p = 0.02$ ,  $d = 0.42$ ). We

TABLE II. Average scores and standard deviation for all vector field plots V1–V5 (zero or non-zero divergence is indicated in parentheses) and both strategies collected from both instruction periods (DS = differential strategy; IS = integral strategy). Statistics refer to a two-tailed paired  $t$ -test ( $N = 41$ ).

Vector field	DS	IS	$p$ value	Cohen's $d$
<u>Both periods</u>				
V1 (nonzero)	0.85 (0.36)	0.54 (0.50)	<0.001	0.72
V2 (zero)	0.56 (0.50)	0.88 (0.33)	<0.001	-0.75
V3 (nonzero)	0.80 (0.40)	0.78 (0.42)	0.74	...
V4 (zero)	0.29 (0.46)	0.71 (0.46)	<0.001	-0.90
V5 (nonzero)	0.80 (0.40)	0.68 (0.47)	0.10	0.28
Total	0.66 (0.47)	0.72 (0.45)	0.21	...
<u>Only period 1</u>				
V1 (nonzero)	0.85 (0.37)	0.48 (0.51)	0.01	0.84
V2 (zero)	0.30 (0.47)	0.81 (0.40)	<0.001	-1.16
V3 (nonzero)	0.80 (0.41)	0.76 (0.44)	0.78	...
V4 (zero)	0.25 (0.44)	0.76 (0.44)	<0.001	-1.16
V5 (nonzero)	0.85 (0.37)	0.57 (0.51)	0.05	0.63
Total	0.61 (0.49)	0.68 (0.47)	0.32	...
<u>Only period 2</u>				
V1 (nonzero)	0.86 (0.36)	0.60 (0.50)	0.08	0.59
V2 (zero)	0.81 (0.40)	0.95 (0.22)	0.17	...
V3 (nonzero)	0.81 (0.40)	0.80 (0.41)	0.95	...
V4 (zero)	0.33 (0.48)	0.65 (0.49)	0.05	-0.65
V5 (nonzero)	0.76 (0.44)	0.80 (0.41)	0.72	...
Total	0.71 (0.43)	0.76 (0.45)	0.46	...

conclude that students improved their performance in the second instruction period, independent of the specific instruction. This may be due to repetition effects or reconsideration of the first strategy, unless students were told not to do so (student interviews shed light on this issue, cf. Sec. III F). Thus, data from the first instruction period are more reliable to answer research question 4.

Last, we found that both groups achieved similar results after completing both test periods. The left group in Fig. 3 (starting with DS) achieved a total average score of  $6.9 \pm 4.6$  points, and the right group (starting with IS) achieved  $7.0 \pm 4.6$  points ( $p = 0.83$ ).

### C. Visual processing of both strategies (RQ 3)

Next, we investigate differences in eye movements between both strategies. First, we considered the distribution of visual attention (heat map) on each vector field plot when the students used either the DS or the IS. The result is illustrated in Fig. 5 for vector field V1:

- Employing the DS, students made more fixations on the coordinate system ( $x$ - $y$  labels) compared to the IS.
- Students focused on a smaller region when using the IS compared to the DS.
- Individual students have concentrated on similar positions using the IS (center of field plot), whereas they focused on different positions using the DS (broad distribution).

Heat maps of the other vector fields V2–V5 are not demonstrated but they show similar trends.

Second, in addition to the qualitative observations provided by heat maps, classical eye-tracking measures provide a quantitative approach to data analysis, cf. Table III.

Both qualitative and quantitative results provide evidence that the strategies induce a different visual behavior (and therefore a different cognitive processing) when students inspect the vector field plot. When students apply the IS, they made significantly more fixations on the visual representation ( $p = 0.03$ ,  $d = -0.42$ ) of longer duration ( $p = 0.01$ ,  $d = -0.25$ ) compared to the DS. Students need to construct a mental representation of a test volume in order to balance the flux through its boundary demanding higher cognitive effort.

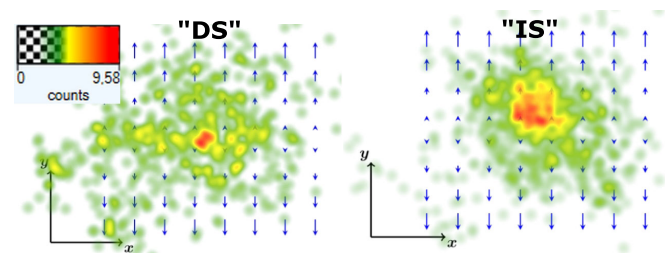


FIG. 5. Heat maps of the first vector field (V1) when students applied the differential strategy (left) or the integral strategy (right). Data were taken from all students.

TABLE III. Comparison of different eye-tracking measures (mean and standard deviation) between both strategies (in both test periods). The meaning of the  $p$  values is the usual statistical significance level, i.e. the greater  $p$ , the higher the probability that the differences between the means were found due to chance. The magnitude of the difference is quantified by Cohen's  $d$ , a standard effect size measure.

Measure	DS	IS	$p$	$d$
Number of fixations	37.1 (13.4)	43.1 (15.5)	0.03	-0.42
Fixation duration	0.29 (0.08)	0.31 (0.09)	0.01	-0.25
Saccade length	4.05 (1.04)	3.15 (0.93)	<0.001	0.91
Number of fixations (axes)	3.32 (4.89)	1.85 (3.34)	0.01	0.49

Concerning the DS, students were required to decompose the vector field arrows in the  $x$  and  $y$  components before looking for changes along horizontal and vertical directions. When students explored the change of the field, they extract information from widely distanced and parafoveal regions (surrounding the fovea, that is a small depression in the retina of the eye where visual acuity is highest), resulting in greater saccade lengths ( $p < 0.001$ ,  $d = 0.91$ ) and a broader attention distribution. Furthermore, we found that students spent more attention to the coordinate axes using DS as compared to IS ( $p = 0.01$ ,  $d = 0.49$ ). This can be explained by the fact that the DS uses coordinates explicitly, whereas the IS is coordinate-free; cf. Eqs. (1) and (2), respectively.

However, the DS did not require decomposing the vector field in all cases (one-dimensional fields). If the analysis is restricted to V1–V3, the difference between numbers of fixations is more pronounced ( $p = 0.002$ ,  $d = -0.61$ ) between DS and IS (data on item level are provided in Ref. [51]). We obtain evidence that vector decomposition (DS) and mental construction of rectangles or balancing flux (IS) causes cognitive load.

#### D. Discrimination between correct and incorrect responses (RQ 4—part 1)

Next, we split our sample in those students who answered correctly and incorrectly. Since the chance of guessing correct is 50%, performance data might contain noise (correct answers for incorrect reasons). Therefore, we conducted this analysis both on the item level and on the test level (collecting values over all vector fields for each strategy), and we compared eye-tracking measures between best and worst performers. For better readability, results concerning best and worst performers are provided separately (Sec. III E). We restricted our analysis to the first test period (cf. Fig. 3) because the students' visual behavior might be biased in the second period as they have already seen the field plots before (cf. Sec. III B).

As seen in the previous section, eye-tracking measures differ between both strategies. Therefore, we treat both

TABLE IV. Comparison of classical eye-tracking measures between correct and incorrect responses for each strategy.

Measure	Correct	Incorrect	$p$	$d$
<u>Differential strategy</u>				
Number of responses	61	39		
Number of fixations	30.5 (17.0)	38.8 (24.2)	0.05	0.41
Fixation duration	0.29 (0.09)	0.29 (0.08)	0.82	...
Saccade length	4.54 (1.26)	4.30 (1.22)	0.38	...
<u>Integral strategy</u>				
Number of responses	70	35		
Number of fixations	41.7 (25.2)	44.9 (29.4)	0.57	...
Fixation duration	0.30 (0.08)	0.27 (0.09)	0.12	...
Saccade length	3.69 (1.01)	3.81 (1.33)	0.60	...

strategies separately. This and the next section are structured as follows: First, we present results regarding classical eye-tracking measures, followed by strategy-specific analyses (saccadic direction analysis and symmetry analysis).

#### 1. Classical measures

Table IV shows the results for each strategy. For the IS, we found no significant differences in eye-tracking measures between all correct and incorrect responses. There were also no significant differences on item level. Hence, classical measures were not able to distinguish whether students can correctly judge the divergence of vector fields when they employ the IS.

For the DS, we found a significant difference in the number of fixations between all correct and all incorrect answers ( $p = 0.05$ ,  $d = 0.41$ ). On average, correct answers were given with less fixations. Since there were no differences in the average fixation duration, this result also means that students who answered correctly needed less time for responding. There were no systematic differences between correct and incorrect answers in terms of fixation duration or saccade length. Since fixation duration is an indicator of cognitive effort, we can assume that wrong answers do not result from students' unwillingness to come to a right solution. In other words, both groups of students showed the same mental effort. This claim can be supported by the finding that both correct and incorrect answers show saccade lengths of characteristic magnitude for this strategy (cf. Table III). We observed similar trends on the individual item level (less number of fixations for correct responses but no other effect).

#### 2. Saccadic direction analysis (differential strategy)

To calculate saccadic directions (cf. Sec. II E and details in the Supplemental Material [31]), we considered two consecutive fixations located within the vector field plot. By doing so, we obtained  $N = 151\,04$  saccadic directions in total which have been converted from radians to degrees. The data set was further restricted to the DS ( $N = 6977$ ) and



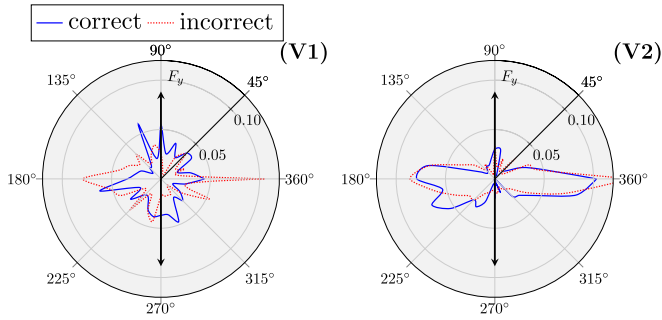


FIG. 6. Polar distributions of saccadic directions with regards to vector fields V1 and V2. The direction of the nonzero field vector  $F_y$  is highlighted.

finally to those students who started with the DS ( $N = 2984$ ). Figure 6 shows the polar distribution of saccadic directions for vector field plots V1 and V2, separated by correct and incorrect responses. The radial axis represents the relative frequency of each saccadic direction.

The first vector field (V1) is one dimensional. The field vector only has a vertical component and this component is also changing with respect to the vertical direction. Figure 6 (V1) indicates that students performed more saccades in the horizontal direction and less in the vertical direction when responding incorrectly compared to responding correctly. Moreover, correct answers correlate with more saccades in the vertical direction in line with what can be expected from correct application of the partial derivative  $\partial F_y / \partial y$ .

The second vector field (V2) is also one dimensional and provides a useful example to test whether students are paying attention to the direction of the field vector or the direction of change. The direction of change is horizontal, whereas the field vector only has a vertical component. Figure 6 (V2) shows that saccadic direction distributions are not different between correct and incorrect answers. The horizontal direction coincides with the direction of change and might therefore attract the students' attention. As seen in Sec. III B, most students thought this field has nonzero divergence, most likely because they noted changes in the orientation of vector field arrows. Student interviews support this interpretation; see Sec. III F. This indicates a poor understanding of partial vector derivatives and the saccade direction distribution is reflecting students' misinterpretation. However, most correct problem solvers do not seem to follow the appropriate strategy which would induce looking in the direction of the field vector ( $y$  direction).

Results from vector fields V3–V5 are provided as Supplemental Material [31]. In summary, saccade direction analysis revealed that (i) horizontal directions are predominant in general and (ii) students are predominantly performing eye movements in that direction in which the field vector is changing its length or orientation (that is  $y$ ,  $x$ ,  $x$ ,  $\{x$  and  $y\}$ , and  $x$  for vector fields V1, V2, V3, V4, and V5, respectively). However, a correct interpretation of divergence must

particularly take the direction of the field vector into account. With the exception of the first vector field (V1), we did not find significant differences between correct and incorrect answers with regards to the saccadic distribution. As discussed in the introduction of this section, the group of correct answers might include false positives (correct answers for incorrect reasons). This clearly limits the validity of this analysis on individual item level. We refer to the comparison of best and worst performers to resolve this limitation (Sec. III E).

### 3. Symmetry analysis (integral strategy)

As described in Sec. II E, we calculated the numbers of fixations within two symmetrical AOIs for each vector field and each participant using the symmetry score; cf. Eq. (3). We restricted the analysis to data of the IS from test period 1 yielding a total of 2594 fixations. We did not include data from vector field V4 since no symmetry axis could be defined. As shown in Table V, average symmetry scores are not different between correct and incorrect answers on test level. On an individual item level, correct answers showed a significantly higher symmetry score than incorrect answers concerning the first vector field (V1) ( $p = 0.003$ ,  $d = 1.50$ ). Students who answered correctly tend to focus at one symmetrical subarea of the field plot. Within one symmetrical subarea, vector arrows do not change its orientation, and hence balancing the flux through an integration area is easier. As seen in Sec. III B (Table II), this task was the most difficult task for students using the IS (period 1, performance score = 0.48). Hence, we assume that conserving mental resources by focusing on one symmetrical subarea helped students to answer correctly. However, there were no similar effects concerning other items.

### E. Discrimination between best and worst performing students (RQ 4—part 2)

Apart from comparing all right vs all wrong answers, we have also compared the best performers (B1, B2; all answers were correct,  $N = 2$ ) with the worst performers (L1, L2, L3; 2–4 out of ten answers were correct,  $N = 3$ ). One best performer and two of the worst performers were exposed to the DS first. On the one hand, only two vs three students are being compared and hence we need to be careful to generalize results. On the other hand,

TABLE V. Average symmetry scores  $s$  (with standard deviation in parenthesis) for correct and incorrect responses by vector field.

Vector field	$s$ (correct)	$s$ (incorrect)	$p$	$d$
V1	0.66 (0.19)	0.32 (0.26)	0.003	1.50
V2	0.28 (0.23)	0.28 (0.14)	0.97	...
V3	0.26 (0.24)	0.43 (0.40)	0.24	...
V5	0.25 (0.19)	0.26 (0.20)	0.90	...
Total	0.34 (0.26)	0.31 (0.25)	0.73	...

TABLE VI. Comparison of classical eye-tracking measures between best (B1, B2) and worst performers (L1, L2, L3) for each strategy.

Measure	B1, B2	L1, L2, L3	$p$	$d$
<u>Differential strategy</u>				
Number of responses	10	15		
Number of fixations	32.5 (15.1)	41.13 (22.0)	0.29	...
Fixation duration	0.35 (0.15)	0.25 (0.07)	0.03	0.80
Saccade length	2.62 (0.88)	3.39 (1.03)	0.03	-0.79
<u>Integral strategy</u>				
Number of responses	10	15		
Number of fixations	27.2 (9.50)	42.9 (33.4)	0.16	...
Fixation duration	0.37 (0.09)	0.23 (0.07)	<0.001	1.84
Saccade length	3.58 (0.91)	3.41 (1.19)	0.71	...

these groups most likely represent students who did understand the strategies very well or failed to understand them, respectively. This will become more evident during student interviews (Sec. III F).

### 1. Classical measures

Concerning the IS, average fixation duration was longer for the best performing students compared to the worst performers ( $p < 0.001$ ,  $d = 1.84$ ), cf. Table VI. The best performers have engaged more deeply with more cognitive effort than the worst performers.

For the DS, we also found that average fixation duration was longer for the best performers compared to the worst performers ( $p = 0.03$ ,  $d = 0.80$ ). Furthermore, average saccade length was shorter for the best performers ( $p = 0.03$ ,  $d = -0.79$ ). Figure 7 compares the gaze paths of the best performing students with two worst performers. First, we can observe that the best performers looked at the vector field plot in a very systematic way, concentrating on a subspace of the field plot. Second, the best performers followed the field vectors (y direction) systematically, looking for changes in the vertical direction, and therefore reflecting a correct interpretation of the partial derivative  $\partial F_y / \partial y$ . They have made very few horizontal saccades compared to the worst performers. This qualitative impression will be quantified in the next section. Last, we note that

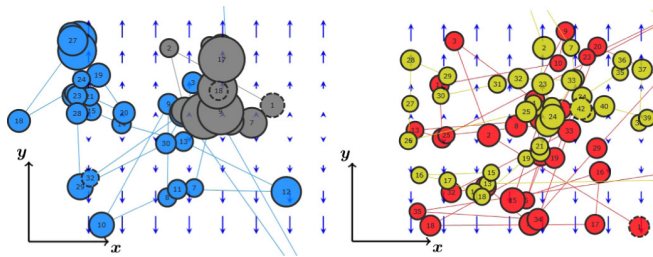


FIG. 7. Gaze plots of the two best performing students (left) and of two worst performing students (right). In both cases, students applied the differential strategy to vector field V1.

the fixation duration varies between different fixations, illustrating cognitive processing after information intake. In contrast, the gaze plot of the two worst performers demonstrates a nonsystematic skimming over the field with almost equal duration of each fixation and large saccade lengths.

### 2. Saccadic direction analysis (DS)

Following the analysis procedure applied in Sec. III D 2, polar distributions of saccade directions have been calculated for best (B1, B2) and worst performers (L1, L2, L3). While detailed figures for each vector field are provided as Supplemental Material [31], we restrict this paragraph to general trends and provide a summary of results.

We observe characteristic differences in saccadic-direction distributions between the best and worst performers regarding vector fields V1, V2, V4, and V5. Best performing students focus on the vertical and horizontal directions and neglect oblique directions. For the one-dimensional fields (V1–V3), they perform a significant amount of saccades in the direction of the field arrows and in direction of change. The worst performers, in contrast, do not perform any vertical saccades when engaging with the second vector field (V2).

Figure 8 demonstrates collected data from all vector fields comparing best performers, worst performers, and all students (except extreme performers). The distribution of best performers is highly symmetric. Both the horizontal and the vertical direction is pronounced, whereas oblique directions are avoided. This result reflects the focus on the directions of the coordinate directions  $x$  and  $y$ . In contrast,

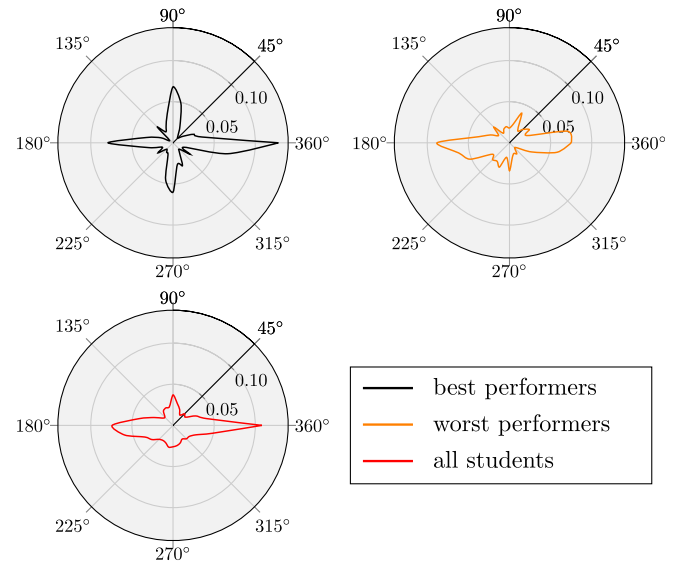


FIG. 8. Polar distribution of saccadic directions for the best performing students (upper left), worst performing students (upper right) and all students (extreme performers excluded; bottom) when they applied the differential strategy. Fixation data from all five vector fields were included.

the upper right part of Fig. 8 shows the saccadic direction distribution of the worst performers. The vertical direction is much less pronounced and oblique saccades occur more often. There is little difference between all students and the worst performers.

### 3. Symmetry analysis (integral strategy)

Last, we calculated the symmetry score  $s$  for the best and worst performing students by vector field. Since only 2 vs 3 values can be compared on the item level, no statistical tests were used on individual item level. Interested readers might refer to descriptive statistics for each student (B1, B2, L1, L2, and L3) provided as Supplemental Material [31]. For total average symmetry scores, we found a significant difference between best ( $s = 0.68 \pm 0.20$ ) and worst performers ( $s = 0.27 \pm 0.24$ ) with a large effect size ( $p < 0.001$ ,  $d = 1.75$ ). Hence, best performers were more inclined to focus on one symmetrical half of the vector field plot, whereas worst performers were looking at both symmetrical parts. As discussed above, one explanation is that students' conserve mental resources by focusing on one symmetrical half since the vector arrows do not change its orientation. However, it is also possible that expert students focus on smaller regions of the field plot in general compared to novices. Additional research is needed to investigate the reason of this effect further.

### F. Student interviews (RQ 1 and RQ 2)

Finally, we present results concerning the interview study. After completing both test periods, students were given three more vector field plots and were allowed to apply the strategy they prefer to judge whether divergence was zero or nonzero. During their performance, students were encouraged to comment on their thoughts. These interviews provide information whether the instructions were understood in the way to what was intended and reveal student difficulties with both strategies.

In the following, we will first present performance and preference data with respect to the vector fields V6–V8. Then, we will provide a full transcript of one best performing student (B1) and of one worst performing student (L1). We expect different reasoning between these students presenting appropriate and inappropriate application of both strategies. We will complement the results with interview excerpts of other students who apply correct and incorrect reasoning. Students who are not part of the extreme performance groups are referred to as S1–S8 in the order of appearance in this text. The researcher conducting the interview is referred to as R.

#### 1. Student performances, strategy preferences, and difficulties

In total, students achieved an average score of 2.65 points out of the maximum score of 3.00 points (88.4%)

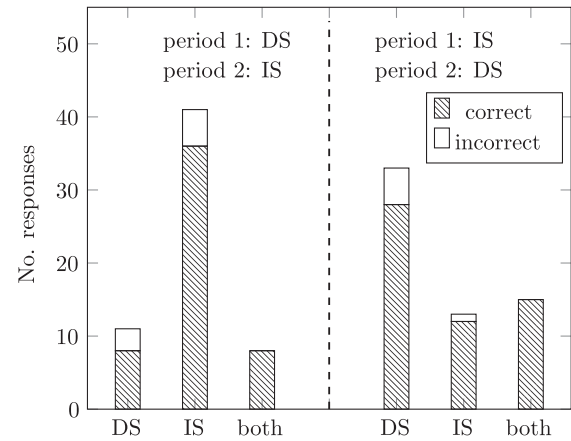


FIG. 9. Number of correct and incorrect responses per strategy summarized for vector field plots V6–V8. The left part of the figure refers to those students which were first instructed by the differential strategy and then by the integral strategy, and the right part refers to the reverse order; cf. Fig. 3.

regarding the vector fields V6, V7, and V8. This can be considered as a high achievement compared to the test periods 1 and 2 (cf. Sec. III B). Students have used the DS in 36% of cases, the IS in 45% of cases, and in 19% of cases they used both strategies to cross-check their answer before giving a final response. Figure 9 presents the performance scores per strategy; the sample was split according to the experimental condition (which strategy was instructed first). First, we notice that students made no errors when they used both strategies before giving their answer. One could assume that only the best students cross-checked their answers with both strategies but that was not the case. We found that ten students used both strategies at least once, and the performance scores of these students during the experiment periods ranged from 5–10 points (average 7.2). One student (S1) corrected himself when considering the vector field V6: (translated from original German)

S1: *So I try to use the differential strategy. In this case the y component remains zero but the x component is always positive when I have a look on different columns, mmmh,..., wait it changes direction,..., I don't know for sure, can I check the other strategy?*

R: *Sure, go on please.*

S1: *I will consider a rectangle in the middle of the field. Then they [the arrows] are entering the region and leaving it, hence divergence is zero.*

Another student (S2) explicitly mentioned that both strategies yield the same result:

S2: *In this case [V6], both strategies could easily be applied. If I consider an area then I see that everything that enters from the left will go out to the right. However, this is equal to the finding that there is no change of this*



*field vector in x direction. When I walk from left to right, nothing changes. Both strategies are connected in this way. The field is divergence-free.*

This student refers to the theorem of Gauss (without mentioning it) and demonstrates a deep conceptual understanding of both strategies.

Second, we asked every student to comment on the difficulties of each strategy and which one was easier for them after completing vector fields V6–V8. In the following, we summarize the responses (numbers in brackets refer to the number of students who made similar statements):

- (1) I would prefer the differential strategy for one-dimensional fields because I found this one easier or faster than the integral strategy. (16)
- (2) In my opinion, the differential strategy was easier *per se* because one can concentrate on a row or a column of arrows and look for changes. (2)
- (3) I would prefer the differential strategy because visualizing areas is exhausting. (3)
- (4) Concerning the differential strategy, the decomposition of components was very hard for me. (8)
- (5) I would always prefer the integral strategy because this one is more appealing to me (2), seems more reliable to me (3) or makes more sense for me. (3)
- (6) The integral strategy was more useful when the vector field arrows had  $x$  and  $y$  components. (11)
- (7) The integral strategy was more difficult to apply because I had to imagine a test volume and focus on the arrows simultaneously. (6)
- (8) If I discovered some kind of symmetry I would prefer the integral strategy. (3)
- (9) Concerning the integral strategy, I had problems to identify the shape of a suitable area (5), to choose the right location for the area (4), to estimate positive and negative flux (6) or to decompose the components in order to estimate the scalar product  $\vec{F} \cdot d\vec{n}$ . (11)

Furthermore, several students claimed that their performance would increase if they were allowed to draw the integration area in case of the IS (7) or to sketch the vector decomposition in case of the DS (4). This fits well into student statements 1, 3, 4, and 9 above and extends previous findings on student difficulties with vector calculus concepts (cf. Sec. I A 2). This finding also indicates that students' spatial ability influences student performances and understanding of the problem-solving strategies—an issue that will be investigated in further studies. Moreover, some students admitted that they did not really understand the DS. This will be presented and discussed in the next paragraph.

Last, we observe a remarkable trend in Fig. 9. Students have mostly chosen the strategy that was instructed during the second test period (IS in the left part of the figure, DS in the right part). A  $t$  test revealed that students have chosen the period-2 strategy significantly more often than

the period-1 strategy ( $p < 0.001$ ). We did not recognize this trend during the interviews, but in hindsight two students provided information about this phenomenon:

S3: *When I started to work with the second strategy [the integral strategy], I felt like I was understanding everything better. I knew about the tasks and what followed and therefore, I was able to concentrate better on the second strategy.*

R: *So, these two test periods have improved your understanding?*

S3: *Exactly. I would even say that I could now work better with the first strategy than before.*

Student S3 benefited from the repetition of the vector fields and claimed that understanding of one strategy regulates learning of the other. It could be assumed that other students were also able to better concentrate during the second instruction period as they knew what followed. This assumption could explain students' better achievement in the second period (cf. Sec. III B) and why they preferred the period-2 strategy. However, a memory effect could also be important as student S4 stated:

R: *I noticed that you only used the integral strategy [the period-2 strategy] during the last vector fields. Please explain why you did so.*

S4: *Oh, I knew pretty much how that works because I've applied it five times before. I cannot remember exactly the first strategy so I stuck to the second strategy.*

Student S4 was using the IS since he has used it directly before.

## **2. Interview transcripts contrasting one best and one worst performers**

Last, we provide a full transcript of one best performing student and one worst performer. Student B1 provided correct answers on vector fields V6–V8. He explained,

B1: *If I consider the first vector field [V6] then the first thing that comes to my mind is that this field is one dimensional. There is no  $y$  component. Therefore, I only need to consider the flux with respect to vertical edges of a rectangle. It's the same arrows on both sides, hence divergence is zero.*

R: *So you are using the integral strategy?*

B1: *Yes, I do. In this case it doesn't matter. Everything that goes in from left will go out on the right side. There is no change in horizontal position; it also works with the other strategy.*

Student B1 started his reasoning with the IS and used a symmetry argument to come to the right conclusion. He also supported his statement with the DS explicitly.



B1: *The next field has two components, but... the x component is not changing. Oh, the y component does. Hence, this one is not source-free; it has some source.*  
 R: *Now, you have used the differential strategy. Can you explain why, please?*

B1: *I think that the differential strategy was easier in general.*

R: *Why do you think so, and what makes the integral strategy difficult?*

B1: *Well, especially in this case [V7], it's hard to see if the field arrows are really perpendicular to the surface normal or not. And if not, one must count how much is going in or out. Especially, if the field is confusing like this one, one can easily make mistakes. The other strategy is way easier. When I go along the axis, I can judge whether the arrows are changing accordingly or not.*

R: *I see. Could you please describe in more detail how you applied the differential strategy in this case?*

B1: *In my mind, I went along the x direction and was looking whether the length of the projected arrows are changing with this direction. I did the same with the vertical direction.*

R: *Ok, great. There is one last field to discuss.*

B1: *Well, this is hard [the last vector field]. I could use again the differential strategy but on the screen this looks somehow skewed... In this case it's better to imagine a volume with spherical symmetry. Then there is nothing coming in or out because the field vectors are all perpendicular to the surface normal. Hence, this field has no source.*

Student B1 was able to switch back and forth between both strategies. We obtain evidence that this student applied both strategies correctly, making use of symmetry arguments and vector decomposition. For instance, he considered the projected field arrow and was looking for changes in horizontal (or vertical) direction, which is also confirmed by the saccadic direction distribution (Fig. 8).

Next, the transcripts of one poor performer gives insight into misunderstandings and confusions with both strategies:

L1: *In the first field, the y direction does not change, but the x direction is different, hence there must be a contribution in x direction. Then there should be a source if I got this right.*

R: *What do you mean with contribution in the x direction. Please explain that.*

L1: *I think we have a change there, and if it changes, then it is nonzero. [...]*

L1: *The second field... we have this case in which it is nonzero. I mean if I decompose the field, then nothing is zero. Now it changes from both directions. This should also have a source.*

L1: *Ok, now I'm totally confused [field V8]. This field obviously has some curl, but it does also have a contribution to the x and y directions. I fear I totally misunderstood this.*

R: *Ok, don't worry. Maybe you should try the other strategy? How would you do that?*

L1: *I would imagine a volume and balance what is going in and out. In this case I would say it's the same, hence it is source free.*

R: *Could you explain at what position you imagined the volume?*

L1: *Exactly in the center. Those four arrows. Then I will get these four arrows in, hence there should also be four arrows going out.*

This student obviously confuses the primary quantity (the arrow) with the change of the quantity. L1 claims, e.g., that there is a contribution in the x direction and in both directions concerning field V6 and V7. The student shows no appropriate understanding of partial vector derivative at all but assigned divergence to vector fields whenever there is something present or changing. Concerning the last field, L1 fails to apply the DS and recognizes that he misunderstood something because he knew that field was divergence-free. When applying the IS, L1 was not able to provide a reasonable explanation of his conclusion but rather remembers this field (*It obviously has some curl*) from prior learning experiences.

Statements of other students reveal similar misunderstandings:

L2: *The first vector field [V6] is source-free since the arrows cancel each other out. I used the differential strategy, and when I look in a row, for each arrow going left, there is one going to the right.*

L2: *This vector field [V7] has no divergence. Like the one before. I decompose the field in x and y components and they are the same.*

R: *Could you explain which strategy you use and what you mean by saying that they are the same?*

L2: *I use the differential strategy. I try to compare the x direction with the x direction above, and the same with the y direction. I don't see anything changing.*

Student L2 confuses both strategies concerning V6, fails to decompose the field arrows concerning V7, and compares columns of errors demonstrating no understanding of the DS at all. We close with quotes of four more students (S5–S8) who also struggled with correct application of the DS.

S5: *I understood the differential strategy in that way: If I decompose the arrow and the arrow is not zero, then divergence is nonzero. I think I understood this completely wrong. The errors cannot cancel each other out.*

S6: *Sometimes I told you that there is divergence but in hindsight I think there was none. I had a look at the arcs and I saw that they were smaller in the middle of the field but larger somewhere else and I concluded that there is divergence. I did not understand what was about the direction. That was confusing to me.*

S7: *Well this field [V6] has only one direction and this one is not the same at different locations. Hence, there must be some divergence.*

S8: *This field is easy [V6]. We see that the y direction is zero hence the derivative is zero as well. But the x direction is nonzero so the derivative is nonzero as well. Therefore, this field has divergence.*

Student S5 and S8 confused the primary quantity (vector arrow) with its derivative (change of vector arrow in terms of length or orientation), whereas students S6 and S7 did not understand the concept of direction. They were comparing the size or orientations of different arrows at arbitrary positions to decide about divergence instead of looking in coordinate directions.

#### IV. CONCLUSION AND FUTURE DIRECTIONS

In this paper, we presented an instruction-based study on students' visual understanding of vector field plots with respect to divergence. We used multiple data sources and a mixed-method approach to investigate students' understanding and difficulties concerning two qualitative problem-solving strategies, including performance measures (correct or incorrect), eye-gaze measures (e.g., fixations, saccadic directions), and interview transcripts.

##### A. Students' understanding of both strategies and students' difficulties (RQ 1 and RQ 2)

First, we can confirm the results of previous research [22,24–26], that is, that students struggle to determine whether vector fields have zero or nonzero divergence. Even after explicit instruction on the DS and the IS, students' decisions on five vector field plots were correct in only 65% of cases, with no significant differences between both strategies. After a second instruction period, overall performance increased to 73%, again with no differences between strategies. During the postinterviews, students named several sources of difficulties for both strategies, such as decomposition of vector field arrows, visualizing an integration area, and balancing positive and negative flux through an imaginary test volume. Several students did not show an appropriate understanding of partial vector derivatives, a result that has also been documented elsewhere [26] and that can be further validated with the eye-tracking methodology (see below). We found that these difficulties arise from confusion between the primary quantity and its derivative (what has also been documented in another context [52]), not paying attention

to the direction of change (supported by eye-tracking data), and difficulties with the concept of components and decomposition of components. Many students claimed that additional visual aids, such as sketches of the vector decomposition or the integration area, would have improved their performance. This indicates that students' spatial ability might influence their performance and understanding of the instructions. In upcoming studies in this field, we suggest to measure students' spatial ability and explore correlations to performance measures.

However, when the students were free to choose the problem-solving strategy (V6–V8), their answer was correct in 88% of the cases. Even though they actually worked with slightly different vector field plots, this result indicates that the availability of both strategies supports student performance. This fits into the motivation of this article, that is, supporting students' understanding of physical concepts with various tools and representations of physical laws (in the sense of Feynman; cf. Sec. IA 1). We conclude with a short answer to RQ 1 and RQ 2:

- No strategy (DS or IS) is *best* to qualitatively judge the divergence of vector fields. Each strategy has characteristic difficulties, which have been identified above.
- Knowledge of both strategies is complementary and yields the best performance.

##### B. Differences in visual processing of both strategies (RQ 3)

Second, eye-tracking methodologies were able to underline the conceptual differences between both strategies quantitatively. We found that the IS causes more fixations of longer duration, reflecting a higher cognitive effort before decision making. This finding was validated by student statements, as most of them indicated that the IS was more effortful than the DS. Some students suggested sketching an integration region to preserve mental resources. In contrast, the DS causes longer saccades, indicating a seeking behavior after changes of the field arrows. The students also paid more attention to the coordinate axes because this strategy is closely connected to the concept of coordinates. During interviews, we observed that most students argued with the coordinate directions when applying the DS (cf. Sec. III F), confirming this finding. We conclude with a short answer to RQ 3:

- Both strategies are processed differently. Differences in eye-tracking measures reflect conceptual differences between the strategies.

##### C. Gaze-based discrimination between correct and incorrect responses and between extreme performance groups (RQ 4)

Concerning classical eye-tracking measures, we found that correct answers were given with fewer fixations corresponding to a shorter time on task. While this reproduces the

findings of expertise research [53], we want to treat this result carefully, as the effect size is small and no other between-groups effects were found. Given that the chance of guessing correctly is 50% and that some students might have answered correctly for incorrect reasons (cf. statements of students L2, S5–S8), these data contain noise, and the groups might not reflect correct or incorrect application of the strategies. In contrast, when more information about student reasoning is given (e.g., when comparing best and worst performers), we found significant differences in the number of fixations and did not observe that greater expertise correlates with shorter response times. Analyzing saccade directions or symmetry scores on the item level also did not yield a positive discrimination between all correct and all incorrect responses in all cases, possibly for the same reason as given above. This result underlines the importance of taking multiple data sources into account when interpreting eye-tracking data. Nevertheless, for extreme performance groups, saccade direction analysis discriminates reliably.

Finally, we found that eye-tracking measures are significantly different between best and worst performers (as identified *post hoc* by pretest scores and by performance scores during the experiment). In particular, the best performers had fewer fixations but of longer duration than the worst performers, confirming a similar finding by van Gog *et al.* [40]. Gaze plots and symmetry analysis revealed that the best performing students focused on a small subarea of the vector field plot and performed systematic eye movements in horizontal and vertical directions when the DS was considered. In contrast, worst performers apparently skimmed over the vector field plot nonsystematically. This qualitative impression was quantified using saccade direction analysis and symmetry analysis. Using saccadic direction analysis, we obtained evidence that the best performing students concentrated on horizontal and vertical eye movements and ignored oblique saccades. On the item level, we also confirmed that the best students performed significantly more saccades in the relevant direction of the field arrow compared to the worst performers, indicating expert understanding of partial derivatives. This was confirmed during student interviews. We conclude with a short answer to RQ 4:

- Eye-tracking methods were not sufficiently discriminatory between correct and incorrect answers, but they reveal significant differences between extreme performing groups.

We also noted a general trend of students to perform horizontal eye movements. In recent studies, it has been found that horizontal eye movements are dominant compared to others, for example, when viewing pictures or landscape scenes [45]. This can be explained by oculomotor factors of the eye and cultural habits of reading. Therefore, if the relevant direction of a vector field corresponds to the  $x$  direction, the method of saccade direction analysis is not appropriate to distinguish between

the best and the worst performers because this direction reflects both top-down and bottom-up processes of viewing. In future studies, we will explore this methodology in the framework of radial vector fields that require, for instance, oblique saccades for correct problem solving. In this case, we expect an even more accurate and reliable discrimination between different levels of expertise.

#### D. Implications for teaching

As this study shows, connecting mathematical concepts to graphical representations of vector fields is difficult for students. It is hard to overestimate the importance of representational competence in physics learning and problem-solving; therefore, we advocate incorporating qualitative exercises like those presented here into lectures or tutorials.

Lecturers are encouraged to demonstrate correct qualitative interpretation of the partial vector derivative using various examples, stressing the focus on horizontal and vertical directions. For the IS, symmetry arguments and different integration areas should be considered explicitly to foster mental flexibility.

All vector fields used in this study had either zero or nonzero divergence everywhere. While this restriction is useful for teaching and learning, it should be made explicit. The drawback of this restriction consists of reinforcing a common misconception that a field has only a single value for divergence rather than divergence being a local property of the field that can change value at any location. To avoid this misconception, students must be told about this special case of the instructional material. Once students understand the strategies, they can be applied to any specific point within the field considering an infinitesimal neighborhood of the point in order to evaluate divergence. In this way, students get to know the meaning of an operator. Because there is an undesired tendency that students (develop) believe that divergence is a global property [25], examples of nonconstant divergence should be discussed subsequently.

To our surprise, many students were thankful for the chance to take part in this short instruction sequence after the completion of the experiment and requested the instruction material. On one hand, this underlines students' need for instructional support in developing connective understanding between abstract concepts and visual representations. The experiment helped students develop a rigorous procedure for interpreting and determining divergence rather than thinking about the common meaning of the word divergence or causing other imperfect associations [24]. On the other hand, it shows that such learning materials are still underrepresented in the physics curriculum.

#### E. Outlook

In future studies, we will extend this work on the concept of curl (where similar strategies can be applied) and



investigate whether student performance can be improved based on various instructional material, including, for example, dynamic visualizations of both strategies, different worked-out examples, or preinstructions focusing on the student difficulties described above. In future studies, our interview findings relating to student difficulties will be used to design a questionnaire to relate eye-tracking

measures to a quantitative measure of students' cognitive load during problem solving.

### ACKNOWLEDGMENTS

We would like to thank TUK physics faculty members Professor Georg von Freymann and Dr. Imke Schneider for welcoming this research in their courses.

- 
- [1] E. M. Smith, Master's thesis, Oregon State University, Corvallis, Oregon, USA, 2014.
- [2] L. Bollen, P. van Kampen, C. Baily, M.-Kelly, and M. De Cock, Student difficulties regarding symbolic and graphical representations of vector fields, *Phys. Rev. Phys. Educ. Res.* **13**, 020109 (2017).
- [3] The drawbacks of this restriction are discussed in Sec. IV D.
- [4] S. E. Ainsworth, The functions of multiple representations, *Comput. Educ.* **33**, 131 (1999).
- [5] D. E. Meltzer, Relation between students' problem-solving performance and representational format, *Am. J. Phys.* **73**, 463 (2005).
- [6] A. Van Heuvelen, Learning to think like a physicist: A review of research-based instructional strategies, *Am. J. Phys.* **59**, 891 (1991).
- [7] P. B. Kohl and N. D. Finkelstein, Patterns of multiple representation use by experts and novices during physics problem solving, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010111 (2008).
- [8] P. Nieminen, A. Savinainen, and J. Viiri, Relations between representational consistency, conceptual understanding of the force concept, and scientific reasoning, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010123 (2012).
- [9] S. E. Ainsworth, P. A. Bibby, and D. J. Wood, Analyzing the costs and benefits of multi-representational learning environments, in *Learning with Multiple Representations*, edited by M. W. van Someren, P. Reimann, H. P. A. Boshuizen, and T. de Jong (Pergamon, Oxford, England, 1998).
- [10] T. Rasch and W. Schnotz, Interactive and non-interactive pictures in multimedia learning environments: Effects on learning outcomes and learning efficiency, *Learn. Instr.* **19**, 411 (2009).
- [11] W. Schnotz and M. Bannert, Construction and interference in learning from multiple representations, *Learn. Instr.* **13**, 141 (2003).
- [12] S. E. Ainsworth, P. A. Bibby, and D. J. Wood, Examining the effects of different multiple representational systems in learning primary mathematics, *J. Learn. Sci.* **11**, 25 (2002).
- [13] M. A. Rau, V. Alevan, N. Rummel, and S. Rohrbach, Sense making alone doesn't do it: fluency matters too! ITS support for robust learning with multiple representations, in *Intelligent Tutoring Systems*, edited by S. Cerri, W. Clancey, G. Papadourakis, and K. Panourgia (Springer, Berlin, Heidelberg, 2012), Vol. 7315, pp. 174–184.
- [14] R. J. Dufresne, W. J. Gerace, and W. J. Leonard, Solving physics problems with multiple representations, *Phys. Teach.* **35**, 270 (1997).
- [15] P. C.-H. Cheng, Unlocking conceptual learning in mathematics and science with effective representational systems, *Comput. Educ.* **33**, 109 (1999).
- [16] D. Rosengrant, Case Study: Students' use of multiple representations in problem solving, *AIP Conf. Proc.* **818**, 49 (2006).
- [17] C. J. De Leone and E. Gire, Is instructional emphasis on the use of non-mathematical representations worth the effort?, *AIP Conf. Proc.* **818**, 45 (2006).
- [18] P. B. Kohl and N. D. Finkelstein, Effect of instructional environment on physics students' representational skills, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010102 (2006).
- [19] R. Kozma, E. Chin, J. Russell, and N. Marx, The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning, *J. Learn. Sci.* **9**, 105 (2000).
- [20] R. P. Feynman, *The Character of Physical Law* (MIT Press, Cambridge, MA, 1967).
- [21] It is worth noting that it is no foregone conclusion that using multiple representations leads to better learning in general. In contrast, using multiple representations can also have detrimental effects. This occurs, for example, when learners are unable to translate between representations or lack a visual understanding of single representations (see, e.g., S. E. Ainsworth, P. A. Bibby, and D. J. Wood, Examining the effects of different multiple representational systems in learning primary mathematics, *J. Learn. Sci.* **11**, 25 (2002).
- [22] C. Singh and A. Maries, Core graduate courses: A missed learning opportunity?, *AIP Conf. Proc.* **1513**, 382 (2013).
- [23] L. Bollen, P. Van Kampen, and M. De Cock, Students' difficulties with vector calculus in electrodynamics, *Phys. Rev. ST Phys. Educ. Res.* **11**, 020129 (2015).
- [24] L. Bollen, P. Van Kampen, C. Baily, and M. De Cock, Qualitative investigation into students' use of divergence and curl in electromagnetism, *Phys. Rev. Phys. Educ. Res.* **12**, 020134 (2016).
- [25] C. R. Baily, L. Bollen, A. Pattie, P. van Kampen, and M. De Cock, Student thinking about the divergence and curl in mathematics and physics contexts, *Proceedings of the Physics Education Research Conference 2016, College Park, MD*, edited by A. D. Churukian, D. Jones, and L. Ding (AIP, New York, 2016), pp. 51–54.



- [26] R. E. Pepper, S. V. Chasteen, S. J. Pollock, and K. K. Perkins, Observations on student difficulties with mathematics in upper-division electricity and magnetism, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010111 (2012).
- [27] J. E. Hoffman and B. Subramaniam, The role of visual attention in saccadic eye movements, *Perception & Psychophysics* **57**, 787 (1995).
- [28] D. D. Salvucci and J. R. Anderson, Automated eye-movement protocol analysis, *Human-Computer Interactions* **16**, 39 (2001).
- [29] 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).
- [30] M. A. Just and P. A. Carpenter, Eye fixations and cognitive processes, *Cogn. Psychol.* **8**, 441 (1976).
- [31] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.14.010116> for the video sequence, methodological details, study materials, and additional data.
- [32] 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).
- [33] 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).
- [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] 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).
- [36] L. Mason, P. Pluchino, M. C. Tornatora, and N. Ariasi, An eye-tracking study of learning from science text with concrete and abstract illustrations, *J. Exp. Educ.* **81**, 356 (2013).
- [37] P. A. O'Keefe, S. M. Letourneau, B. D. Homer, R. N. Schwartz, and J. L. Plass, Learning from multiple representations: An examination of fixation patterns in a science simulation, *Comput. Hum. Behav.* **35**, 234 (2014).
- [38] S. S. Mozaffari, J. Kuhn, P. Klein, A. Dengel, and S. S. Bukhari, Entropy Based Transition Analysis of Eye Movement on Physics Representational Competence, *UbiComp'16 Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany* (ACM, New York, USA, 2016), pp. 1027–1034.
- [39] 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, *Computer Education* **74**, 61 (2014).
- [40] 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).
- [41] A. C. Graesser, S. Lu, B. A. Olde, E. Cooper-Pye, and S. Whitten, Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down, *Mem. Cogn.* **33**, 1235 (2005).
- [42] M. Hegarty, in *Eye Movements and Visual Cognition*, edited by K. Rayner (Springer-Verlag, New York, 1992), pp. 428–443.
- [43] M. Kozhevnikov, M. A. Motes, and M. Hegarty, Spatial Visualization in Physics Problem Solving, *Cogn. Sci.* **31**, 549 (2007).
- [44] K. Rayner, Eye movements in reading and information processing: 20 years of research, *Psychol. Bull.* **124**, 372 (1998).
- [45] T. Foulsham, A. Kingstone, and G. Underwood, Turning the world around: Patterns in saccade direction vary with picture orientation, *Vision Res.* **48**, 1777 (2008).
- [46] H. Collewijn, C. J. Erkelens, and R. M. Steinman, Binocular coordination of human horizontal saccadic eye movements, *J. Physiol.* **404**, 157 (1988).
- [47] More specifications can be found on the product website <https://www.tobiipro.com>.
- [48] D. D. Salvucci and J. H. Goldberg, Identifying fixations and saccades in eye-tracking protocols, *Proceedings of the 2000 Symposium on Eye Tracking Research and Applications* (ACM, New York, USA, 2000), pp. 71–78.
- [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] Correlation between pretest and students performance during the experiment was not significant,  $r = 0.23$ ,  $p = 0.13$ .
- [51] Number of fixations  $NF$  (mean and standard deviation):  $NF_{V1DS} = 34.8(18.3)$ ,  $NF_{V1IS} = 45.9(27.7)$ ,  $NF_{V2DS} = 37.1(22.9)$ ,  $NF_{V2IS} = 46.3(23.8)$ , and  $NF_{V3DS} = 30.0(20.1)$ ,  $NF_{V3IS} = 35.6(22.7)$ .
- [52] L. C. McDermott, M. Rosenquist, and E. van Zee, Student difficulties in connecting graphs and physics: Examples from kinematics, *Am. J. Phys.* **55**, 503 (1987).
- [53] M. T. H. Chi, Two approaches to the study of experts' characteristics, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, Cambridge, 2006), pp. 21–30.