Designing e-learning courses for classroom and distance learning in physics: The role of learning tasks

Daniel Laumann,[1,2] Julian Alexander Fischer,[3] Tatjana K. Stürmer-Steinmann,[4,5] Julia Welberg,[1] Susanne Weßnigk,[6] and Knut Neumann[3]

[1]*Institute of Physics Education, University of Münster, 48149 Münster, Germany*
[2]*Institute of Physics, University of Münster, 48149 Münster, Germany*
[3]*IPN—Leibniz-Institute for Science and Mathematics Education, 24118 Kiel, Germany*
[4]*Leonore-Goldschmidt-School, IGS Hannover-Mühlenberg, 30457 Hannover, Germany*
[5]*Institute for Didactics of Mathematics and Physics, Leibniz University Hannover, 30167 Hannover, Germany*
[6]*QUEST Leibniz Research School, Leibniz University Hannover, 30167 Hannover, Germany*

(Received 26 March 2023; accepted 4 January 2024; published 16 February 2024)

Digital learning technologies have grown increasingly important in physics education, partly enforced through the COVID-19 pandemic. During the pandemic, digital technologies allowed for continued teaching and learning of students even when schools were closed. While research in psychology and educational technology has yielded many insights into the effectiveness of e-learning courses, fewer studies have examined the design of e-learning courses. Few studies have empirically investigated the design of learning tasks as a central element of e-learning courses. The present study analyzes how the design of tasks in e-learning courses, specifically with respect to their degree of openness as well as the relevance of their contexts, influences students' behavioral engagement, learning outcomes, and situational interest. Due to the importance of e-learning courses during the COVID-19 pandemic, we also analyzed the extent to which specific learning settings (classroom learning, distance learning) influence the effects of e-learning course design on students' behavioral engagement, learning outcomes, and situational interest. To investigate the research questions, we analyzed a total of $N = 1060$ datasets for 12 different e-learning courses (3 to 5 lessons, middle school physics), of which $n = 557$ were completed before and $n = 503$ during the COVID-19 pandemic. The results suggest that e-learning courses with a high proportion of learning tasks that relate to meaningful real-world contexts appear to be more conducive to behavioral engagement, learning outcomes, and situational interest. Regarding the consideration of open-ended tasks, the results suggest that these appear to be more useful for classroom learning but should be used in a limited way when designing e-learning courses for distance education.

DOI: 10.1103/PhysRevPhysEducRes.20.010107

I. INTRODUCTION

The COVID-19 pandemic has highlighted the need for a better understanding of the role that digital technologies can play in modern education. Physics education research has provided insights into the use of specific digital technologies, such as digital measurement acquisition or interactive simulations, in specific teaching-learning contexts. However, there is little insight on how to design e-learning courses, typically in the form of, for example, *moodle* courses being accessed online or offline. E-learning courses seem particularly interesting because they can be used for classroom learning as well as distance learning. The design of e-learning courses requires specific decisions on the technological side regarding how the educational features are presented and how they appear. From this technological perspective, there are already numerous findings from research and practice regarding the suitability of specific platforms but also regarding more general design elements of e-learning courses, such as features of course structuring, consideration of self-assessments, and clarity of instructions [1–3]. With reference to the technological design perspective, insights from the cognitive theory of multimedia learning (CTML) [4], which deals with the arrangement and design of different representations, are particularly worth mentioning.

Significantly fewer insights exist on the educational design and specifically on the design of learning tasks within e-learning courses. Up to now, very little research has empirically investigated the influence of the design of learning tasks on learning and learning motivation in e-learning courses. This seems particularly important for physics education where digital learning technologies already possess great importance. The aim of the current study hence was to empirically investigate the design of e-learning courses with respect to learning tasks, which are important elements of these courses. One particular strength of e-learning courses is that they can be used in both classroom learning as well as distance learning settings. Due to the popularity of e-learning courses during the COVID-19 pandemic, we also considered the extent to which the learning setting (classroom learning and distance learning) influences the effects of the design of e-learning courses on learning processes.

A. Digital technologies in physics education

Digital technologies have played a significant role in physics education for decades. Personal computers with various software have been an important tool since 1970s [5], for example, for experimental data acquisition [6] or to simulate and model physical processes [7]. Nowadays, mobile digital technologies, such as smartphones or tablet computers, have become core technologies due to their high availability as well as their versatile applications [8]; for example, for data acquisition (e.g., digital measurement systems) or the (computational) modeling activities (e.g., simulations). In the not-too-distant future, augmented as well as virtual reality will further extend the range of technologies used in physics education [9].

The importance of digital technologies for the individualization of learning processes and training of specific competencies in physics has been shown by empirical studies in physics education and science education, which have focused on the analysis of individual technologies and their effectiveness compared to traditional teaching elements [10–14]. Now, however, it no longer seems appropriate to consider individual digital technologies in an isolated manner. Increasingly, e-learning courses are becoming more and more relevant for providing “a learning space integrating different technologies and associated curricular and pedagogical practices that frame the context of learners and the learning environment” [15] (pp. 1520–1521).

In this respect, e-learning courses, for example, provide information through various media, integrate assessment tools, and support communication between students and teachers. They can be offered online through different learning platforms, such as the *Khan Academy or Coursera*, but they can also be offered offline via school servers. E-learning courses are realized through digital learning management systems and appear fruitful for innovating teaching-learning processes: Digital technologies of appropriate design can contribute to lifelong learning, due to “dissolving of the boundaries between formal, informal, and non-formal learning” [16] (p. 31). Furthermore, appropriate technologies also appear to be a key to providing access to science learning content for specific communities against the background of race, ethnicity, gender, or language [17–19] if learners have access to appropriate (mobile) digital devices. E-learning courses also offer the advantage that they can be used for both classroom learning and distance learning. In terms of classroom learning, they appear to be a promising technology for supporting, for example, individualized learning or individual feedback, with teachers acting as learning guides and facilitators to support students. When used in distance learning, e-learning courses can also be used to combine learning materials from a series of lessons and integrate learning activities and assessments, for example.

B. Digital learning during the COVID-19 pandemic

During the COVID-19 pandemic, an acceleration of developments toward digital technologies occurred [20]. In most countries around the world [21–22], including the United States [23], online learning environments and e-learning courses took on a special role. Thus, it became apparent that instruction was essentially no longer conducted as classroom learning but as distance learning. Only through e-learning courses could school learning continue during the COVID-19 pandemic. Despite teachers expressing significant challenges in transforming instructional processes [24], studies found that e-learning courses in distance learning could certainly teach similar competencies compared to classroom learning [23]. This even applied to experimental lab courses in physics, whose transformation into a digital format is a particular challenge for learners and teachers [25–26].

Nevertheless, teachers perceived themselves as insufficiently supported and prepared to use digital learning technologies during the COVID-19 pandemic [27]. The teachers’ impression can be attributed, among other things, to a lack of guidance on the design of e-learning courses. Teachers have little influence on the design of e-learning courses in terms of technological implementation, as the prerequisites are generally predefined for each school. Surface features for the design of multimedia also appear to be less significant for teachers. However, it is important to support teachers in making the best possible decisions about the educational design of the courses, for example, about how learning tasks can be best designed as part of e-learning courses.

C. Design of e-learning courses

There are several lines of research that aim to identify features of how to effectively use e-learning courses for supporting students’ learning and learning motivation. These include findings from cognitive psychology related to learning as well as research on educational technologies and ideas of educational sciences.

Many findings from learning psychology can be used to design e-learning courses since these can be traced back to multiple representations as "fundamental building blocks." Therefore, the design of e-learning courses is supported, among other theories, by the cognitive theory of multimedia learning (CTML) [4]. CTML describes cognitive learning processes based on multiple representations (e.g., texts and pictures) and modalities (e.g., seeing and hearing) and provides more than ten design principles (e.g., multimedia principle and multimodality principle) that describe how to design multimedia learning content conducive to learning and that have already been tested many times in the form of meta-analyses [28–32]. Thus, according to the spatial contiguity principle, as one example of the multimedia design principles, it seems advantageous for learning to present textual and pictorial elements that belong together in as close spatial proximity as possible [27,33–35]. Considering the multimodality principle as another example, learning is supported by addressing different sensory channels simultaneously; therefore, an illustration should be complemented by spoken, not written text [25,36–39]. Similarly, research on the influence of digital media on affective characteristics shows, for example, how the shape and color of the design affect learners' emotions, such as their situational interest, and, thus, indirectly influence cognitive processes [40–41]. Conclusions from findings in CTML can also be found for the field of e-learning and, thus, for the design of e-learning courses [42–43]. The literature review by Oh *et al.* (2020) emphasizes that design principles for organization and presentation are analyzed in previous studies but notes on pedagogical design need to be more fully developed [43].

Much research on e-learning courses has focused on online learning courses or, more specifically, massive open online courses (MOOCs). Successful MOOCs are characterized by certain features, for example, course structuring, consideration of self-assessments, or form of instructions [1–3]. With respect to tasks and content, two findings seem to be of particular interest: First, the perceived task value of a learning task [44] is important to positively influence learners' behavioral engagement and their situational interest in MOOCs. Behavioral engagement, in this sense, "can be observed when students […] attend to an academic task" [45]. To increase both learners' behavioral engagement and their perceived relevance of a task, it seems possible to include relevant and meaningful contexts [46–49], according to studies within physics education research. Second, the type of learning task not only influences learners' behavioral engagement but also influences the quality of what students learn from it [50]. This seems important because it means the learning tasks have a key function in the general teaching-learning process: "They [learning tasks] may be characterized as an interface between the learners and the information offered in the learning environment. They serve to activate and control learning processes to facilitate successful learning." [51] (p. 1976).

Regarding the first point, to increase the perceived task value, the design of e-learning courses could integrate meaningful (relevant) contexts for learning and pay attention to the nature (form) of the learning tasks. Notably, based on the results of research in science education, the perceived contextual relevance of a learning task seems crucial for cognitive and affective outcomes, as it is assigned a high activation potential [46–49]. Meaningful context can support students' behavioral engagement and raise their situational interest [52–53]. Although there is already ample evidence that context-based learning is conducive to learning, Sevian *et al.* (2018), in a special issue on context-based learning, point out as a perspective for future research, "We need more studies that compare how students learn in different context-based learning classrooms" [48] (p. 1105). Against the background of the previous considerations, e-learning courses seem to represent such a special setting for the application of context-based learning for which there is still little empirical evidence. This is especially true when comparing the application of e-learning courses in different learning settings (classroom learning and distance learning).

With regard to the typification of learning tasks, it can be considered, among other approaches to classification, whether they are open-ended learning tasks or, in contrast, closed-ended learning tasks. Open-ended tasks, e.g., free-text questions, require more effort from learners to answer the task but are considered effective for in-depth learning [54]. Closed-ended tasks, e.g., multiple-choice tasks, true or false questions, or matching tasks, can be answered in a low-threshold manner and, thus, initially activate learners but are often not assigned a deeper activation of thought processes [55]. In addition to these two types of learning tasks, there are also tasks called semi-open-ended tasks, where parts of the solution path are predefined, while other parts of the problem-solving process are up to the learners, such as semistructured worksheets [56]. As a summarizing characteristic of tasks, the term "degree of openness" [57] (p. 1163) can be found in the literature for the spectrum of tasks mentioned from closed-ended to semi-open-ended to open-ended learning tasks. As studies show, the "degree of openness" is a relevant characteristic for success in teaching-learning processes [58,59] as well as for the processing quality of online surveys [60]. Based on the previous findings, it can be hypothesized that open tasks may positively influence learning effectiveness [57], but the probability that tasks are processed at all decreases with the degree of openness [60]. Since learning tasks play a central role in controlling learning processes, especially in e-learning courses, it seems necessary to test this hypothesis empirically.

In addition, previous studies have not considered the learning settings of e-learning courses. It seems feasible that the use of e-learning courses in the classroom, in the presence of the teacher, places different requirements on the design of the courses than when e-learning courses are used in a distance learning setting. This assumption is supported by the fact that studies already point to the influence of self-directed learning skills and self-regulated learning on the processing of e-learning courses [61]. Thus, a research deficit seems to exist with regard to the question of how the design of learning tasks in e-learning courses with regard to relevant contexts and the openness of tasks influences the processing of e-learning courses in different learning settings.

D. Research questions

Research has repeatedly shown the potential of e-learning courses. In particular, if e-learning courses are designed to include meaningful tasks and contexts, they support students' behavioral engagement, situational interest, and learning. This seems particularly relevant for physics, where numerous studies point to the importance of meaningful contexts and, thus, the contextual relevance of learning tasks. However, so far, there is a lack of knowledge on how specific design features of learning tasks (degree of openness and contextual relevance) influence students' behavioral engagement, situational interest, and learning. Furthermore, no studies exist that include the learning setting of e-learning courses; for example, we do not yet know the extent to which classroom learning requires a different design of learning tasks compared to distance learning. Therefore, this study addressed two research questions:

RQ1: How do the degree of openness and the contextual relevance of learning tasks influence students' behavioral engagement, learning outcomes, and situational interest in e-learning courses?

RQ2: To what extent does the effect of openness and contextual relevance of learning tasks on students' behavioral engagement, learning outcomes, and situational interest in e-learning courses differ for classroom learning and distance learning?

II. METHODS

The study presented here draws on data from a larger project that was conducted with students in grades 7 to 11 in 3- to 5-h teaching units from January 2020 until June 2021. Within this project, 12 e-learning courses were developed, which were analyzed with regard to their use in digitally supported classroom teaching. All units were administered as Moodle courses. The courses included learning content (videos, texts, illustrations related to phenomena, experiments, or models, etc.) as well as activities and learning tasks (free-text responses, multiple choice, experimental tasks, etc.).

The courses took place in the classroom in the presence of the teacher, who guided students through the activities, facilitating discussions, and responding to questions. In addition, parts of the activities were done in the form of homework. In about the middle of the study, due to the COVID-19 pandemic, it was no longer possible to use the e-learning courses for classroom learning; they instead had to be used for distance learning. Therefore, minor adjustments were made to the courses to allow students to participate in e-learning courses in distance learning. However, the adjustments did not affect the structure and learning content of the courses. In addition, the teacher continued to guide students through the activities (online), being available (online) to facilitate discussion and clarify questions, analogous to regular classroom learning. As with regular classroom learning, parts of the material were

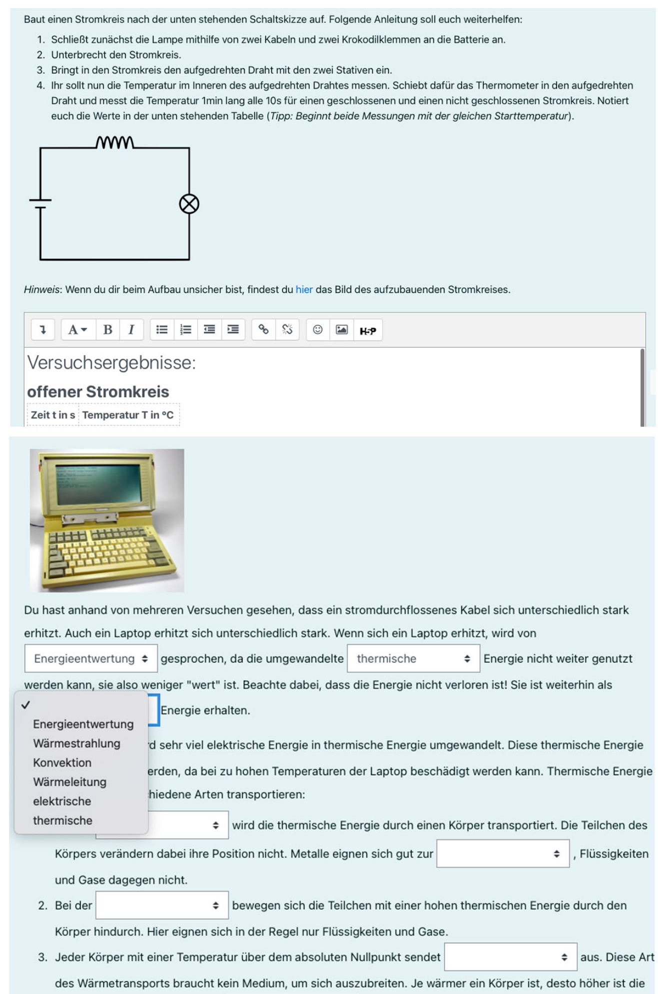

FIG. 1. Two sample activities typical for the e-learning courses. Left: Experimental task—The experimental task is to set up an electric circuit with a battery, cables, and a wire and to determine the temperature of the wire at given time intervals. Based on this, the devaluation of energy during the conversion of electrical energy into thermal energy is discussed. Right: Cloze or gap-fill task—The cloze or gap-fill task deals with the conversion and devaluation of energy in a laptop. Learners have to identify the forms of energy involved in the conversion and describe the processes taking place in the laptop.

processed synchronously in the (online) presence of the teacher, while other parts were completed asynchronously, analogous to homework.

The 12 e-learning courses (units) used in this study aimed to support the development of an increasingly connected knowledge base on the scientific core idea of "energy" as the foundation of competence, and these courses were developed within the framework of another project (see Fig. 1 for an example). The 12 units linked content from different areas of middle school physics and were designed such that they seamlessly integrated with the existing curriculum [62] and related to a meaningful driving question. Table I presents the driving questions, while Fig. 2 illustrates the grade levels and content (energy forms) to which the respective driving questions are assigned.

For example, the unit "Why does a laptop sometimes get hot?" linked the area of electricity with the area of thermodynamics and was designed to align with the grade 7 curriculum in German high schools. Such units were also referred to as curriculum replacement units [63]. To ensure the comparability of the content of the units, all e-learning courses followed a uniform structure. According to this structure, each e-learning course could be divided into five sections (one section each for introduction and reflection as entry and conclusion, and three sections for the elaboration of the driving question). An overview of the structure of the courses with the indication of exemplary questions can be found in Fig. 3.

To ensure that the contexts of the e-learning courses and driving questions are in fact perceived as meaningful by the learners, a preliminary study with $N = 285$ learners was carried out in which it was investigated for a total of 30 driving questions on different contexts and how these are assessed by learners with regard to (a) interestingness and (b) importance. Considering the results of this preliminary study, 12 driving questions that were selected by learners as being significantly meaningful and having the widest possible range of subject content to ensure applicability to different teachers and age levels were selected.

TABLE I. Overview of the driving questions of the e-learning courses.

Course	Driving question
1	Why are there seasons on earth?
2	What color should you choose for your clothes in summer to avoid sweating?
3	How much energy is in a "stick bomb"?
4	Why does a laptop sometimes get hot?
5	How should solar cells be attached to a house to convert as much energy as possible?
6	How long would you have to ride a bicycle to charge a smartphone?
7	Why does a roller coaster go at breakneck speed even without its own drive?
8	How does a skateboard constantly reach the same height in a half pipe?
9	How can you charge a smartphone without a power outlet or power bank?
10	How does a microphone work?
11	How high can you jump on a trampoline?
12	Why do you burn your skin when you slide across the gym floor?

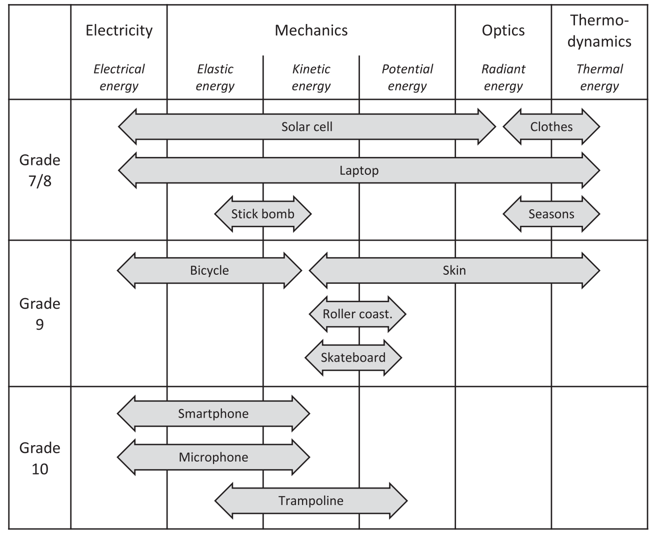

FIG. 2. Overview of the subject content of the e-learning courses by specifying the forms of energy considered, supplemented by the context of the driving questions (see also Table I). Each gray arrow in the figure corresponds to one e-learning course.

The e-learning courses then each contained different types of learning tasks, with the number of learning tasks varying between 21 and 39 learning tasks per e-learning course, on average 27.6 ± 5.1 learning tasks per e-learning course.

A. Research design

Since the COVID-19 pandemic started during data collection and teachers had to move their instruction from classroom learning to distance learning, a situation resulted in which about half of the courses examined were still held in a classroom learning setting and the other half in a distance learning setting. That is, the data structure corresponds with a quasiexperimental group comparison design to investigate the influence of the design of learning tasks in e-learning courses on learning as well as on learning motivation (RQ1), also considering the specific constraints of classroom learning and distance learning according to the form of instruction (RQ2). For the analyses, the average degree of openness and the average contextual relevance of each e-learning course were considered independent variables, and students' behavioral engagement, situational interest, and learning outcomes as dependent variables. While for RQ1 all data were considered regardless of the

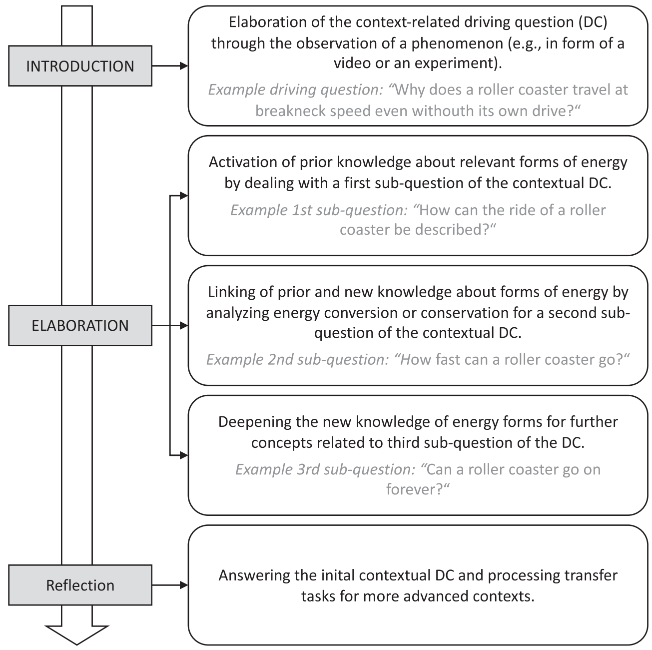

FIG. 3. Presentation of the structuring of introduction, elaboration, and reflection, which is uniform for all 12 courses, with an indication of exemplary questions.

learning setting (classroom learning and distance learning), the analyses for RQ2 were performed separately by learning setting.

When analyzing the research questions, we were aware that, in addition to the characteristics of the learning tasks, numerous other factors influence students' behavioral engagement, learning outcomes, and situational interest in e-learning courses. These include each learner's individual preconditions (e.g., individual interest and prior knowledge) as well as characteristics of the e-learning courses (e.g., complexity and choice of context) and features of implementation (e.g., class atmosphere and environmental variables) so that we expected to find strong variances in the variables mentioned. However, in order to conduct a study with authentic data during the COVID-19 pandemic, we accepted the limitations of the effects given by the expected variances.

1. Design features of e-learning courses—Learning tasks

Of particular importance for investigating both research questions was classifying the e-learning courses in terms of the degree of openness and the contextual relevance of their learning tasks. For this purpose, each e-learning course was first divided into its learning tasks.

To describe the degree of openness of an e-learning course, each learning task was first assigned a value according to the following assignment: 1 = open-ended task (open-ended questions), 0.5 = semi-open-ended task (experimental measurement tasks with semi-structured digital worksheets), and 0 = closed-ended task (multiple-choice tasks, true or false questions, matching tasks, cloze or gap-fill tasks). It should be considered that in the analyzed learning courses, only the mentioned task types can be found. In this way, a relative value for the degree of openness of an e-learning course could be determined as the sum of the values over each learning task of the respective course divided by the total number of learning tasks of an e-learning course. The degree of openness varied between 45.2% (minimum) and 87.1% (maximum) open-ended learning tasks per e-learning course, on average $64.4 \pm 9.9\%$ open-ended learning tasks per e-learning course. For validation purposes, two independent raters evaluated a subsample of about 15% of the learning tasks regarding the "degree of openness." Due to the (almost) unambiguous assignment of the learning tasks to the three expressions of the "degree of openness," almost perfect interrater reliability of Cohen's $\kappa = 0.91$ between two independent raters resulted [64].

To describe the contextual relevance of an e-learning course via the respective learning tasks, a similar approach was taken. First, for each learning task, we determined whether it referenced the overarching meaningful context of the e-learning course. A value was assigned: 1 = contextual reference (e.g., "Describe the energy transformations of a skateboard in a half pipe." or "Why does a laptop heat up during intensive use?"), 0 = no contextual reference (e.g., "Describe the energy transformations of a pendulum." or "Why does a resistor of an electrical circuit heat up (depending on the electrical voltage and current)?"). Here, the relative contextual relevance of the e-learning course could be determined as the sum of the values over each learning task of the respective course divided by the total number of learning tasks of an e-learning course. The contextual relevance varied between 0% (minimum) and 80% (maximum) context-related learning tasks per e-learning course, on average $26.2 \pm 18.5\%$ context-related learning tasks per e-learning course. For validation purposes, two independent raters evaluated a subsample of about 15% of the learning tasks regarding the "contextual relevant." An almost perfect interrater reliability of Cohen's $\kappa = 0.86$ between two independent raters resulted [64].

2. Behavioral engagement, learning outcomes, and situational interest

The research questions consider three variables relevant to students' learning: behavioral engagement, learning outcomes, and situational interest.

To determine students' behavioral engagement, all learning tasks of every learner were assessed individually based on the work products available in Moodle after the processing by the learners was finished according to the definition that behavioral engagement "can be observed when students […] attend to an academic task" [45]. Therefore, a specific value was assigned according to the degree of engagement: 0 = “task ignored,” 1 = “task opened but not completed,” 2 = “task completed.” In addition, a mean value of behavioral engagement was determined for each learner and the entire e-learning course but also for each of the five sections of the e-learning course, providing longitudinal data of task processing.

To analyze students’ learning outcomes, we used the input that students entered into Moodle in response to each learning task. The learning tasks were divided into (type 1) scientific explanation and (type 2) scientific inquiry. Scientific explanation refers to students’ ability of using scientific concepts to explain phenomena while scientific inquiry refers to the meaningful use of scientific practices for scientific inquiry [65]. Different models exist describing students’ competence in scientific explanations and scientific inquiry [66]. These models delineate different levels of students’ competence based on the notion of hierarchical complexity [67,68]. Based on this approach, both types of tasks were scored on a five-level scale adapted from a coding book taken from existing frameworks [69,70]. For the scientific explanation tasks (type 1), each task was scored on a five-level scale (0 = “No text,” 1 = “Off-task,” 2 = “Non-normative ideas,” 3 = “Normative ideas,” 4 = “Connections between normative ideas”). For example, to explain the heating up of a laptop, students use elements of a laptop to explain the heating up (level 2), single and unconnected normative ideas like electric energy (level 3), or connected normative ideas like energy transformation from electric into thermal energy (level 4). For the scientific practice tasks (type 2), tasks were scored on a five-level scale (0 = “No text,” 1 = “Off-task,” 2 = “Non-scientific practice,” 3 = “Non-systematic approach,” 4 = “Systematic approach”). For example, to compare pictures including solar cells, students mention single elements like “each picture shows solar cells” (level 2), state unrelated facts like “some solar cells are blue, most solar cells are on rooftops” (level 3), or connect these facts systematically like “Most solar cells are on rooftops, but each solar cell is positioned in different directions” (level 4). Both task types were assessed by three experts in physics education in which we found a moderate agreement (Fleiss’ $\kappa = 0.49$) [64]. By analyzing the conflicting scored tasks, the experts could resolve them.

Students’ situational interest [71–72] was assessed before and after as well as 4 times throughout each e-learning course using adapted versions of established instruments on five 5-level Likert scale items (emotion-related valence beliefs [73] and value-related valence beliefs [74]). Thereby, for emotion-related valence beliefs, a high value of situational interest results in a positive evaluation of the joy of the content (“I am looking forward to the content.”) as well as a negative evaluation of frustration by the content (“I am dissatisfied while working on the content.”). Regarding value-related valence beliefs, a high value of situational interest emerges with a positive evaluation of intrinsic value (“I think the content will be important to me.”), usefulness-private (“The content will be useful in my daily life.”), and usefulness-academic (“Working on the unit tasks will be worthwhile because it is expected in school.”). Again, means for entire e-learning courses as well as longitudinal data for each section of the e-learning course are available. When considering the overall reported situational interest across all learners and courses, a value of $M = 3.45$ ($SD = 0.66$) indicates that the courses and contexts are perceived as more likely to be meaningful, supporting the appropriateness of the course selection based on the preliminary study (see Sec. II).

B. Sample

The overall sample included $N_D = 1060$ datasets, each corresponding to the completion of one e-learning course by one student. $N_C = 46$ e-learning courses were completed by the participating classes. The assignment of the students to the experimental conditions (group 1: classroom learning and group 2: distance learning) was not random but resulted from the time of implementation (before or during the COVID-19 pandemic). Further information on the sample can be found in Table II.

C. Analytical approach

The research questions based on the available data were answered by means of linear mixed-effect models. We examined the extent to which two independent variables (IV: CRE and OPE) each influence one of three dependent variables (DV: BE, LO, and SI) with the purpose to identify possible influences of the design characteristics on the learning-related variables and to be able to assess the strength of the effects. Since the present data were from learners clustered into classes, the student’s class was considered as a subject factor in the

TABLE II. Sample description of the study and description of the study groups.

	Total	Classroom learning	Distance learning
Datasets	1,060	557	503
Grade 7	419	216	203
Grade 8	132	0	132
Grade 9	406	297	109
Grade 10	92	33	59
Grade 11	11	11	0
Completed courses	46	23	23
Grade 7	17	8	9
Grade 8	5	0	5
Grade 9	19	13	6
Grade 10	4	1	3
Grade 11	1	1	0

TABLE III. Intraclass correlation coefficients (ICC) and design effects (DEFF) for linear mixed-effect models with respect to behavioral engagement, learning outcome, and situational interest as dependent variables.

Dependent variable	ICC	DEFF
Behavioral engagement (BE)	0.222	5.891
Learning outcome (LO)	0.173	4.808
Situational interest (SI)	0.072	2.585

models to account for the nested data structure [75,76]. Both the intraclass correlation coefficient (ICC) [77] and the design effect (DEFF) [78], representing the degree of variance inflation due to cluster sampling, indicate the multilevel structure must be taken into account (see Table III). Overall, the following models (random intercepts and no random slopes) were determined for the two research questions:

RQ1: $\mathrm{DV} \sim \mathrm{IV} + (1|\mathrm{CLASS})$.

RQ2: $\mathrm{DV} \sim \mathrm{IV} + \mathrm{SETTING} + \mathrm{IV} \times \mathrm{SETTING} + (1|\mathrm{CLASS})$.

III. RESULTS

Our study aimed to investigate the role of design features of learning tasks, namely degree of openness (OPE) and contextual relevance (CRE) on students' behavioral engagement (BE), learning outcomes (LO), and situational interest (SI) in e-learning courses. Furthermore, the effects of the design of learning tasks on students' processing of e-learning courses were analyzed regarding the learning setting—classroom learning versus distance learning. Data were analyzed with IBM SPSS Statistics (Version 29).

Effect sizes are given according to Ref. [79]. The following interpretive ranges for effect size apply to all results. Partial η^2: $\eta^2 \geq .01$ small effect, $\eta^2 \geq .06$ moderate effect, and $\eta^2 \geq .14$ large effect.

A. Effects of design of learning tasks in e-learning courses on students' behavioral engagement, learning outcomes, and situational interest (RQ1)

In order to answer the research question (RQ1), separate linear mixed-effect models (restricted maximum likelihood estimation) were calculated for each dependent variable [DV: behavioral engagement (BE), learning outcomes (LO), situational interest (SI)] and each independent variable [IV: degree of openness (OPE), contextual relevance (CRE)] as well as students' class as a subject factor according to the following models: $\mathrm{DV} \sim \mathrm{IV} + (1|\mathrm{CLASS})$. All data were considered, regardless of the learning setting. The model estimates for the linear mixed-effect models are reported in Table IV (as a measure of effect size, partial η^2 was calculated [80].

The results show that all three dependent variables are influenced by the degree of openness of the learning tasks, as all standardized coefficients become statistically

TABLE IV. Estimates obtained from linear mixed-effect models regarding the influence of design features (OPE: degree of openness, CRE: contextual relevance) on learning (BE: behavioral engagement, LO: learning outcomes, SI: situational interest) considering students' class as subject factor.

Model	Effect	β	$SE\beta$	t	df	p	Partial η^2
BE	OPE	−0.01	0.07	−0.20	53.60	0.846	0.00
	CRE	0.20	0.07	2.74	62.28	0.008	0.11
LO	OPE	−0.04	0.07	−0.64	51.59	0.528	0.01
	CRE	0.17	0.07	2.63	55.58	0.011	0.11
SI	OPE	−0.06	0.05	1.24	50.51	0.221	0.03
	CRE	0.16	0.05	3.40	44.25	0.001	0.21

insignificant. This can be observed differently for the influence of contextual relevance. Here, a moderate positive effect on the behavioral engagement of the learners ($p = 0.008$ and partial $\eta^2 = 0.11$) and also on the learning outcomes ($p = 0.011$ and partial $\eta^2 = 0.11$) can be determined. The situational interest of the learners is even positively influenced by the contextual relevance with a large effect size ($p = 0.001$ and partial $\eta^2 = 0.21$).

B. Effects of design of learning tasks in e-learning courses on students' behavioral engagement, learning outcomes, and situational interest in comparison between classroom learning and distance learning (RQ2)

At first, due to the analysis of differential effects of the learning session (classroom learning and distance learning), it seems necessary to analyze whether both groups are comparable regarding the dependent variables (BE, LO, and SI) and the independent variables as design characteristics (OPE and CRE). Performing t tests for independent samples showed that conducted courses differed significantly with respect to the dependent variable only with respect to SI [CL: $M = 3.37$, $SD = 0.68$; DL: $M = 3.55$, $SD = 0.62$; $t(964) = -4.21$, $p < 0.001$, Pearson's $r = 0.13$]. Therefore, students showed significantly higher situational interest in distance learning compared to classroom learning. With respect to the independent variables, a significant difference was found for CRE [CL: $M = 0.22$, $SD = 0.11$; DL: $M = 0.31$, $SD = 0.23$; $t(692.25) = -7.92$, $p < 0.001$, Pearson's $r = 0.29$] indicating that courses taught during distance learning had a stronger contextual relevance. However, respective differences showed only small effect sizes, and both learning settings appear comparable, although this aspect will be considered in the final discussion.

Finally, to answer this research question (RQ2), separate linear mixed-effect models (restricted maximum likelihood estimation) were calculated for each dependent variable [DV: behavioral engagement (BE), learning outcomes (LO), and situational interest (SI)] and each independent

TABLE V. Estimates obtained from linear mixed-effect models regarding degree of openness (OPE) on learning (BE: behavioral engagement, LO: learning outcomes, SI: situational interest) in varying learning setting (CL: classroom learning, DL: distance learning) considering students' class as subject factor.

Model	Setting	Effect	β	$SE\beta$	t	d.o.f.	p	Partial η^2
BE	CL	OPE	0.14	0.11	1.21	20.58	0.240	0.07
	DL	OPE	0.02	0.10	0.17	21.15	0.868	0.00
		OPE × SET	···	···	−0.92	36.85	0.362	0.02
LO	CL	OPE	0.03	0.11	0.31	20.52	0.759	0.00
	DL	OPE	−0.01	0.09	−0.05	20.62	0.957	0.00
		OPE × SET	···	···	−0.46	37.78	0.647	0.01
SI	CL	OPE	0.01	0.06	0.11	18.77	0.914	0.00
	DL	OPE	−0.17	0.07	−2.45	24.63	0.022	0.20
		OPE × SET	···	···	−1.96	43.63	0.056	0.08

variable [IV: degree of openness (OPE) and contextual relevant (CRE)] as well as students' class as subject factor. However, due to the research question, the setting as well as the moderation effects of the setting on the respective independent variable were considered. This resulted in the following models: DV $\sim$ IV $+$ SETTING $+$ IV $\times$ SETTING $+$ $(1|\text{CLASS})$.

The model estimates for the linear mixed-effect models are reported in Table V for the degree of openness and Table VI for contextual relevance as independent variables. As a measure of effect size, partial η^2 was calculated [80].

Regarding the degree of openness as a design feature, the analysis yields one statistically significant effect. This large effect is found on SI in distance learning ($\beta = -0.17$, $p < 0.022$, and partial $\eta^2 = 0.20$), suggesting that for e-learning units with a larger share of open learning tasks, students' situational interest ends up being lower. If we accept effects with type I error rates of less than 10% ($\alpha < 0.10$), there would be a difference in how the openness affects students' situational interest across the two learning settings with moderate effect size ($p = 0.056$ and partial $\eta^2 = 0.08$).

Regarding contextual relevance also only one statistically significant effect can be found; again, a large effect on SI in distance learning ($\beta = 0.13$, $p < 0.026$, and partial $\eta^2 = 0.24$) so that for learning units with a larger share of context-related learning tasks, a higher situational interest in the learners' results is found. Also, for contextual relevance, we find two other effects if considering type I error rates up to 10%. These two effects relate to LO. The analyses reveal a large effect on classroom learning ($\beta = 0.32$, $p = 0.072$, and partial $\eta^2 = 0.14$), where better learning outcomes are thus observed for learning units with more contextual tasks. The corresponding comparison of both learning settings reveals a moderate effect ($p \leq 0.051$ and partial $\eta^2 = 0.05$) in the difference of the effect of contextual relevance depending on the learning setting.

IV. LIMITATIONS

The present data provide insight into how the design of learning tasks for e-learning courses in authentic learning settings before and during the COVID-19 pandemic affected student learning. However, the following limitations must be considered when interpreting the data.

TABLE VI. Estimates obtained from linear mixed-effect models regarding contextual relevance (CRE) on learning (BE: behavioral engagement, LO: learning outcomes, SI: situational interest) in varying learning setting (CL: classroom learning, DL: distance learning) considering students's class as subject factor.

Model	Setting	Effect	β	$SE\beta$	t	d.o.f.	p	Partial η^2
BE	CL	CRE	0.18	0.19	0.97	21.71	0.345	0.04
	DL	CRE	0.12	0.08	1.48	18.24	0.155	0.11
		CRE × SET	···	···	−1.23	70.62	0.223	0.02
LO	CL	CRE	0.32	0.17	1.89	21.96	0.072	0.14
	DL	CRE	0.08	0.07	1.10	17.39	0.287	0.07
		CRE × SET	···	···	−1.98	69.35	0.051	0.05
SI	CL	CRE	0.16	0.10	1.61	21.81	0.122	0.11
	DL	CRE	0.13	0.05	2.42	18.40	0.026	0.24
		CRE × SET	···	···	−0.30	49.69	0.766	0.00

A central limitation of the present study arises from the clustered data structure for a given sample size. As the introductory calculations of the intraclass correlation coefficients show, the consideration of the multilevel structure is necessary for the intended analyses. However, when looking at potential design effects, it becomes apparent that, given the study design, a 2 to -6 times larger sample size, depending on the dependent variables, would have been necessary to adequately analyze significant effects using linear regressions. Due to the complex models required, only particularly robust effects were found. However, it can be assumed that further effects exist that are no longer significant due to the complexity of the models and are below the detection limit. For this reason, effects with type I error rates less than 10%, often only slightly above the commonly used 5% were also reported in some cases.

Another limitation is that effects are limited by the fact that the variables mentioned are influenced by numerous other aspects in addition to the design features, which could not be controlled due to the use of authentic classroom settings. These additional aspects include individual preconditions of the learners (e.g., individual interest and prior knowledge) as well as characteristics of the e-learning courses (e.g., complexity and choice of context) and features of the implementation (e.g., class atmosphere and environmental variables). Consequently, large variances were found in the behavioral engagement, learning outcomes, and situational interest of the learners.

Despite these large variances, we did observe significant influences of the design features of the learning tasks, as elements of e-learning courses, and we did find differences between the learning settings slightly above an alpha level of 5%; these findings show that appropriately designing learning tasks can contribute to the optimization of the teaching-learning processes.

Then again, it must be mentioned as a limitation that to obtain clearer effects in this sense, the design features degree of openness and contextual relevance would have had to be adapted in a certain way to highlight the largest possible differences between the e-learning courses. In the most extreme case, it would have been conceivable to adjust the two design features as much as possible in the sense of a 2×2 research design. This would have resulted in e-learning courses with exclusively open-ended or closed-ended tasks as well as completely with or without context-related learning tasks. However, since the design of the e-learning courses was oriented toward real teaching to derive realistic and practical findings, designing the e-learning courses in this way did not appear to make sense for the study.

A further limitation arises with reference to the sample and the distribution of the courses in classroom learning and distance learning. Here, minor differences between the groups in situational interest and in the design of the e-learning courses regarding contextual relevance were observed. However, due to the small effect sizes, only small influences on the results are expected.

V. DISCUSSION AND CONCLUSIONS

The following discussion and derivation of conclusions are done with reference to the two research questions of the study. With respect to RQ1, which aimed at analyzing the effects of learning task design (openness and contextual relevance) in e-learning courses on students' behavioral engagement, learning outcomes, and situational interest, the results can be summarized in the following way:

The results suggest that the learning tasks' contextual relevance is important, as contextual relevance influenced both the number and the quality of the processed tasks (expressed as behavioral engagement and learning outcomes) and was positively related to situational interest. Results of the linear mixed-effect models showed, independent of the learning setting that e-learning courses with a larger share of learning tasks with context relevance were positively related to all considered dependent variables (behavioral engagement, learning outcomes, and situational interest), with the influence of context relevance on situational interest appearing to be the strongest. This is consistent with findings in the literature that often attribute the importance of meaningful contextualization to positive effects on learner interest and motivation [81,82] and confirms the impression of previous studies, especially for the learning of abstract concepts in physics, which is difficult for many learners and meets with little interest, references to the life world of the learners and related contexts seem to be beneficial. "The concept of interest seemed to be appropriate to understand tendencies of students or adults to engage in certain themes or contexts" [83] (p. 29). In this context, the positive relation between context relevance and behavioral engagement also seems plausible, since it can be assumed that e-learning courses with context-related tasks that positively influence the learners' interest are also worked on more frequently by the learners. This is supported by the fact that regardless of the learning setting, a higher proportion of context-related tasks was associated with greater behavioral engagement, indicating a higher number of completed tasks (see Table IV). However, no corresponding studies have been done in the field of science education on the connection between the use of meaningful contexts and the specific construct behavioral engagement.

As previous studies [84] suggested, the context relevance of the e-learning courses, determined by the proportion of context-related learning tasks, was also positively related to learning outcomes but to a lesser extent compared to the effect on situational interest. This result seems possibly transferable to other subjects in which lifeworld contexts have a similar importance as a complement to more abstract concepts, for example, in chemistry, although empirical tests seem necessary.

Regarding the openness of the e-learning courses, measured by the proportion of open-ended learning tasks, we found that openness did not influence any of the observed dependent variables. This applied both with respect to the alpha levels of 5% typically used and also with respect to effect sizes that might indicate less robust effects, regardless of significance (see Table IV).

It seems interesting and noteworthy that learners in e-learning units with more open and in e-learning units with more closed learning tasks work with similar engagement and interest and achieve comparable learning outcomes. There are at least two possible explanations for this. The first relates to physics as the specific discipline of the study. In contrast to some other subjects, in physics in general and also in the studied learning courses many open-ended tasks are very concretely related to learning objects, such as interpreting phenomena in videos or performing experiments, the open-ended tasks might have subject-specific peculiarities that make the results not directly transferable. The second relates to the learners' possible prior experience with open-ended tasks. Studies suggest that open-ended tasks place rather high demands on learners, but these can be reduced by appropriate prior experience, practical knowledge in the area of self-regulation, and appropriate advice from teachers [85,86]. The results can therefore possibly be interpreted to mean that learners have previous experience with open-ended tasks and that they therefore do not represent a fundamental challenge within the present study.

All in all, the data of the present study suggest that it makes sense to develop e-learning courses in physics with context-related tasks regardless of the learning setting, as these seem to have a positive influence on central variables related to learning and students' motivation to learn, just as they do in traditional teaching and learning materials.

Regarding RQ2, which focused on the effects of learning task design (openness and context relevance) in e-learning courses on students' behavioral engagement, learning outcomes, and situational interest depending on the learning setting (classroom or distance), the results can be summarized as follows:

First, the findings of RQ2 regarding the influence of contextual relevance of learning tasks on students' behavioral engagement, learning outcomes, and situational interest in physics e-learning provide a more differentiated picture (see Table VI). Only one statistically significant effect was found, indicating a positive influence of contextual relevance on situational interest in distance learning. Effects related to the two other dependent variables, behavioral engagement and situational interest, and the different learning settings were suggested by the corresponding moderate to large effect sizes, but not significant. Further studies would have to clarify whether the lack of significance can be attributed to the effects themselves or to the study design and the complex models as explained in Sec. IV. Considering the results on RQ1, it can be assumed that with a larger sample, further effects with appreciable effect sizes may have been found to be significant. Regarding the difference between the two learning settings, the significance, which is slightly above the alpha level of 5%, and the moderate effect size suggest that there could only be a difference with regard to the learning outcomes, whereby integrating context-related tasks in e-learning courses for classroom learning fostered students' learning outcomes more than when doing so for distance learning. Again, against the background of the particular importance of contexts from the everyday life of learners, the findings appear consistent with previous studies [81,82] when transferred to learning in physics in e-learning courses but require further research.

The analysis of the e-learning courses with respect to the proportion of open-ended learning tasks provided a more differentiated picture of the findings for RQ1. The results for RQ1 showed that regardless of the learning setting, the openness of the learning courses, specifically the degree of openness of the learning tasks, did not affect behavioral engagement, learning outcomes, or situational interest (see Table IV). However, as the results for RQ2 show (see Table V), a negative influence on situational interest with a large effect size was present for distance learning as a learning setting. In this setting, students who worked on e-learning courses with a larger share of open-ended tasks were substantially less interested. In classroom learning, however, situational interest was statistically independent of the degree of openness. This may be interpreted in a way that especially in distance learning the open learning tasks, which are often perceived as more demanding, have a negative effect on the interest of the learners because in distance learning, the impression arises that they must force themselves to complete tasks, although more interesting activities are available as an alternative. In classroom learning, on the contrary, time must be used for learning anyway so that the openness of the tasks has less of a negative effect due to the lack of alternatives. Considering this result, we get the impression that especially without the direct support of peers or teachers, many open-ended tasks may have a negative effect on the situational interest of the learners in working through e-learning courses. This finding seems comprehensible in view of the literature that suggests open-ended learning tasks are better addressed through collaboration; however, such collaboration requires further competencies and represents a challenge, especially in new e-learning settings [87]. It is possible that this effect is counteracted in the classroom learning setting, where direct communication is available.

Regarding behavioral engagement, no statistically significant effect was observed. Therefore, e-learning courses with a larger share of open-ended tasks were not processed significantly more or less often. This appears contradictory with prior research, where more open-ended questions in classroom education led to greater participation by students [83]. Also, in terms of learning outcomes, the results did not indicate any influence on the degree of openness.

Overall, the results for RQ2 show few statistically significant findings, which are particularly evident for situational interest as a dependent variable in distance learning as a learning setting. Particularly regarding the degree of openness, both learning settings prove to be astonishing and, in part, contrary to the literature, for example, on the engagement of learners in the processing of open tasks [83], independent of the design of the learning courses with rather open or rather closed tasks.

The present study analyzes the influence of two design features, the degree of openness and contextual relevance, of learning tasks as part of e-learning courses on different learning-related variables, such as behavioral engagement, learning outcomes, or situational interest. Looking at the results of the present study on RQ1 and RQ2 as a whole, contextual relevance proves to be a much more influential design feature than the degree of openness, particularly in RQ1. While contextual relevance influences all dependent variables, a corresponding influence of learning courses with more open versus more closed tasks is only found for the situational interest of the learners. The results for RQ2, which only demonstrate particularly robust effects due to the limitations described below, emphasize the influence of the design features on the situational interest of the learners.

Furthermore, they only indicate statistically significant differences between the learning settings in two cases (degree of openness and learning outcomes; contextual relevance and situational interest). The main limitations to be considered here are that, given the multilevel analysis with complex models, a larger overall sample would have been necessary to detect further effects, especially for the sample size, which was reduced once again when divided according to learning settings. A targeted variation of the design features and the construction of "extreme" learning courses with, for example, exclusively open or exclusively closed tasks, would probably also have shown further effects, although their practical benefit would have been questionable.

All in all, both influencing factors, namely the degree of openness and integration of meaningful contexts expressed as contextual relevance, do not claim to extensively predict the dependent variables (behavioral engagement, learning outcomes, and situational interest). Rather, the aim was to find indications of optimal design for these two design features. These findings can thus be interpreted as implications for the two design features of learning tasks, namely context relevance and openness when designing e-learning courses for physics education.

ACKNOWLEDGMENTS

The study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 470250834. We acknowledge support from the Open Access Publication Fund of the University of Münster.

[1] S. Baldwin and Y.-H. Ching, Online course design: A review of the canvas course evaluation checklist, Int. Rev. Res. Open Dist. Learn. **20**, 269 (2019), https://scholarworks.boisestate.edu/edtech_facpubs/222/.
[2] S. Warburton and Y. Mor, A set of patterns for the structured design of MOOCs, Open Learn. **30**, 206 (2015).
[3] J. Kasch, P. van Rosmalen, and M. Kalz, Educational scalability in MOOCs: Analysing instructional designs to find best practices, Comp. Educ. **161**, 104054 (2021).
[4] R. E. Mayer, *Multimedia Learning* (Cambridge University Press, Cambridge, England, 2009).
[5] A. M. Bork, Effective computer use in physics education, Am. J. Phys. **43**, 81 (1975).
[6] S. Staacks, S. Hütz, H. Heinke, and C. Stampfer, Advanced tools for smartphone-based experiments: Phyphox, Phys. Educ. **53**, 045009 (2018).
[7] C. E. Wieman, W. K. Adams, and K. K. Perkins, PhET: Simulations that enhance learning, Science **322**, 682 (2008).
[8] R. Girwidz, L.-J. Thoms, H. Pol, V. López, M. Michelini, A. Stefanel, T. Greczyło, A. Müller, B. Gregorcic, and M. Hömöstrei, Physics teaching and learning with multimedia applications: A review of teacher-oriented literature in 34 local language journals from 2006 to 2015, Int. J. Sci. Educ. **41**, 1181 (2019).
[9] M.-B. Ibáñez and C. Delgado-Kloos, Augmented reality for STEM learning: A systematic review, Comp. Educ. **123**, 109 (2018).
[10] N. Rutten, W. R. van Joolingen, and J. T. van der Veen, The learning effects of computer simulations in science education, Comp. Educ. **58**, 136 (2012).
[11] T. de Jong, M. C. Linn, and Z. C. Zacharia, Physical and virtual laboratories in science and engineering education, Science **340**, 305 (2013).
[12] S.-C. Fang and Y.-S. Hsu, Understanding science teachers' enactments of a computer-based inquiry curriculum, Comp. Educ. **112**, 69 (2017).
[13] C. D. Porter, J. R. H. Smith, E. M. Stagar, A. Simmons, M. Nieberding, C. M. Orban, J. Brown, and A. Ayers,

Using virtual reality in electrostatics instruction: The impact of training, Phys. Rev. Phys. Educ. Res. **16**, 020119 (2020).

[14] D. Sonntag and O. Bodensiek, How mixed reality shifts visual attention and success in experimental problem solving, Phys. Rev. Phys. Educ. Res. **18**, 023101 (2022).

[15] K. Neumann and N. Waight, The digitalization of science education: Déjà vu all over again?, J. Res. Sci. Teach. **57**, 1519 (2020).

[16] J. Jaldemark, Formal and informal paths of lifelong learning: Hybrid distance educational settings for the digital era, in *An introduction to distance education* (Routledge, New York, 2021), pp. 25–42.

[17] B. Mehra, C. Merkel, and A. Peterson Bishop, The internet for empowerment of minority and marginalized users, New Media Soc. **6**, 781 (2004).

[18] R. B. Maier and M. Fisher, Strategies for digital storytelling via tabletop video: Building decision making skills in middle school students in marginalized communities. J. Educ. Technol. Syst. **35**, 175 (2006).

[19] B. O'Mara and A. Harris, Intercultural crossings in a digital age: ICT pathways with migrant and refugee-background youth, Race Ethn. Educ. **19**, 639 (2016).

[20] D. Vargo, L. Zhu, B. Benwell *et al.*, Digital technology use during COVID-19 pandemic: A rapid review, Human Behav. Emerging Technol. **3**, 13 (2020).

[21] B. Williamson, R. Eynon, and J. Potter, Pandemic politics, pedagogies and practices: Digital technologies and distance education during the coronavirus emergency, Learn. Media Technol. **45**, 107 (2020).

[22] P. Klein, L. Ivanjek, M. N. Dahlkemper, K. Jeličić, M.-A. Geyer, S. Küchemann, and A. Susac, Studying physics during the COVID-19 pandemic: Student assessments of learning achievement, perceived effectiveness of online recitations, and online laboratories, Phys. Rev. Phys. Educ. Res. **17**, 010117 (2021).

[23] E. W. Burkholder and C. E. Wieman, Absence of a COVID-induced academic drop in high-school physics learning, Phys. Rev. Phys. Educ. Res. **18**, 023102 (2022).

[24] E. Brewe, A. Traxler, and S. Scanlin, Transitioning to online instruction: Strong ties and anxiety, Phys. Rev. Phys. Educ. Res. **17**, 023103 (2021).

[25] V. Borish, A. Werth, N. Sulaiman, M. F. J. Fox, J. R. Hoehn, and H. J. Lewandowski, Undergraduate student experiences in remote lab courses during the COVID-19 pandemic, Phys. Rev. Phys. Educ. Res. **18**, 020105 (2022).

[26] A. Wert, C. G. West, and H. J. Lewandowski, Impacts on student learning, confidence, and affect in a remote, large-enrollment, course-based undergraduate research experience in physics, Phys. Rev. Phys. Educ. Res. **18**, 010129 (2022).

[27] I. DeCoito and M. Estaiteyeh, Transitioning to online teaching during the COVID-19 pandemic: An exploration of STEM teachers' views, success, and challenges, J. Sci. Educ. Technol. **31**, 340 (2022).

[28] P. Ginns, Meta-analysis of the modality effect, Learn. Instr. **15**, 313 (2005).

[29] P. Ginns, Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects, Learn. Instr. **16**, 511 (2006).

[30] N. L. Schroeder and A. T. Cenkci, Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis, Educ. Psychol. Rev. **30**, 679 (2018).

[31] D. Alpizar, O. O. Adesope, and R. M. Wong, A meta-analysis of signaling principle in multimedia learning environments, Educ. Technol. Res. Dev. **68**, 2095 (2020).

[32] N. Sundararajan and O. O. Adesope, Keep it coherent: A meta-analysis of the seductive details effect, Educ. Psychol. Rev. **32**, 707 (2020).

[33] R. E. Mayer, K. Steinhoff, G. Bower, and R. Mars, A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text, Educ. Technol. Res. Dev. **43**, 31 (1995).

[34] J. Holsanova, N. Holmberg, and K. Holmqvist, Reading information graphics: The role of spatial contiguity and dual attentional guidance, Appl. Cogn. Psychol. **23**, 1215 (2009).

[35] R. E. Mayer and L. Fiorella, 12—Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles, in *The Cambridge Handbook of Multimedia Learning* (Cambridge University Press, Cambridge, England, 2014), pp. 279–315.

[36] S. Mousavi, R. Low, and J. Sweller, Reducing cognitive load by mixing auditory and visual presentation modes, J. Educ. Psychol. **87**, 319 (1995).

[37] R. E. Mayer and R. Moreno, A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory, J. Educ. Psychol. **90**, 312 (1998).

[38] R. Moreno and R. E. Mayer, Cognitive principles of multimedia learning: The role of modality and contiguity, J. Educ. Psychol. **91**, 358 (1999).

[39] W. Schnotz, Colorful bouquets in multimedia research: A closer look at the modality effect, Z. Pädagog. Psychol. **25**, 269 (2011).

[40] J. L. Plass, S. Heidig, E. O. Hayward, B. D. Homer, and E. Um, Emotional design in multimedia learning: Effects of shape and color on affect and learning, Learn. Instr. **29**, 128 (2014).

[41] J. L. Plass and U. Kaplan, Emotional design in digital media for learning, *Emotions, Technology, Design, and Learning* (Academic Press, London, 2016), pp. 131–161.

[42] R. C. Clark and R. E. Mayer, *E-learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning* (Pfeiffer, San Francisco, 2011).

[43] E. G. Oh, Y. Chang, and S. W. Park, Design review of MOOCs: Application of e-learning design principles, J. Comput. Higher Educ. **32**, 455 (2020).

[44] D. Lee, S. L. Watson, and W. R. Watson, The influence of successful MOOC learners' self-regulated learning strategies, self-efficacy, and task value on their perceived effectiveness of a massive open online course, Int. Rev. Res. Open Dist. Learn. **21**, 82 (2020).

[45] A. Gregory, J. P. Allen, A. Y. Mikami *et al.*, Effects of a professional development program on behavioral engagement of students in middle and high school, Psychol. Schools **51**, 143 (2013).

[46] J. Kuhn and A. Müller, Context-based science education by newspaper story problems: A study on motivation and learning effects, Perspect. Sci. **2**, 5 (2014).

[47] K.-C. Yu, S.-C. Fan, and K.-Y. Lin, Enhancing students' problem-solving skills through context-based learning, Int. J. Sci. Math. Educ. **13**, 1377 (2015).

[48] H. Sevian, Y. J. Dori, and I. Parchmann, How does STEM context-based learning work: What we know and what we still do not know, Int. J. Sci. Math. Educ. **40**, 1095 (2018).

[49] W. R. Penuel, M. L. Turner, J. K. Jacobs, K. Van Horne, and T. Sumner, Developing tasks to assess phenomenon-based science learning: Challenges and lessons learned from building proximal transfer tasks, Sci. Educ. **103**, 1367 (2019).

[50] N. Nordin, H. Norman, M. A. Embi *et al.*, Factors for development of learning content and task for MOOCs in an Asian context, Int. Educ. Stud. **9**, 48 (2016).

[51] S. Richter, *Learning Tasks, Encyclopedia of the Sciences of Learning* (Springer, New York, 2012), pp. 1975–1979.

[52] S. Habig, J. Blankenburg, H. van Vorst, S. Fechner, I. Parchmann, and Elke Sumfleth, Context characteristics and their effects on students' situational interest in chemistry, Int. J. Sci. Educ. **40**, 1154 (2018).

[53] J. Lavonen, B. Ávalos, K. Upadyaya, S. Araneda, K. Juuti, P. Cumsille, J. Inkinen, and K. Salmela-Aro, Upper secondary students' situational interest in physics learning in Finland and Chile, Int. J. Sci. Educ. **43**, 2577 (2021).

[54] C. A. Melovitz Vasan, D. O. DeFouw, B. K. Holland *et al.*, Analysis of testing with multiple choice versus open-ended questions: Outcome-based observations in an anatomy course, Anat. Sci. Educ. **11**, 254 (2017).

[55] S. H. S. Tan, G. Thibault, A. C. Y. Chew, and P. Rajalingam, Enabling open-ended questions in team-based learning using automated marking: Impact on student achievement, learning and engagement, J. Comput. Assist. Learn. **38**, 1347 (2022).

[56] L. Zhang, Y. Huang, X. Yang, S. Yu, and F. Zhuang, An automatic short-answer grading model for semi-open-ended questions, Interact. Learn. Environ. **30**, 177 (2022).

[57] B. A. S. Buldur and S. Kartal, The effect of open-ended science experiments based on scenarios on the science process skills of the pre-service teachers, Procedia **93**, 1162 (2013).

[58] B. Štěpánková and P. Emanovský, On open-ended and closed-ended questions in didactics tests of mathematics, Probl. Educ. 21st Century **28**, 114 (2011), http://www.scientiasocialis.lt/pec/node/527.

[59] W.-L. Lin and Y.-W. Lien, The different role of working memory in open-ended versus closed-ended creative problem solving: A dual-process theory account, Creativity Res. J. **25**, 85 (2013).

[60] R. Zhou, X. Wang, L. Zhang, and H. Guo, Who tends to answer open-ended questions in an e-service survey? The contribution of closed-ended answers, Behav. Inf. Tech. **36**, 1274 (2017).

[61] M. Zhu, C. K. Bonk, and S. Berri, Fostering self-directed learning in MOOCs: Motivation, learning strategies, and instruction, Online Learn. **26**, 153 (2022).

[62] S. Weßnigk, K. Neumann, and M. Kerres, Energie unterrichten über eine digitale Lehr-Lernplattform, Naturwiss. Unterricht Phys. **31**, 31 (2020).

[63] S. Marion and L. Shepard, *Let's Not Forget About Opportunity to Learn: Curricular Supports for Innovative Assessments* (Center for Assessment, Dover, 2010).

[64] J. R. Landis and G. G. Koch, The measurement of observer agreement for categorical data, Biometrics **33**, 159 (1977).

[65] National Research Council, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (The National Academies Press, Washington, DC; 2012).

[66] R. Lehrer and L. Schauble, Learning progressions: The whole world is not a stage, Sci. Educ. **99**, 432 (2015).

[67] L. K. Berland and K. L. McNeill, A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts: Learning progression for scientific argumentation, Sci. Educ. **94**, 765 (2010).

[68] C. F. Herrmann-Abell and G. E. DeBoer, Investigating a learning progression for energy ideas from upper elementary through high school, J. Res. Sci. Teach. **55**, 68 (2018).

[69] H.-S. Lee and O. L. Liu, Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective, Sci. Educ. **94**, 665 (2010).

[70] S. Bernholt and I. Parchmann, Assessing the complexity of students' knowledge in chemistry, Chem. Educ. Res. Pract. **12**, 167 (2011).

[71] A. Krapp, Interesse, Lernen und Leistung—Neue Forschungsansätze in der Pädagogischen Psychologie, Z. Pädagogik **38**, 747 (1992).

[72] U. Schiefele, Interest and learning from text, Sci. Stud. Read. **3**, 257 (1999).

[73] R. Pekrun, E. Vogl, K. R. Muis *et al.*, Measuring emotions during epistemic activities: The epistemically-related emotion scales, Cogn. Emot. **31**, 1268 (2016).

[74] H. Gaspard, A.-L. Dicke, B. Flunger, B. Schreier, I. Häfner, U. Trautwein, and B. Nagengast, More value through greater differentiation: Gender differences in value beliefs about math, J. Educ. Psychol. **107**, 663 (2015).

[75] J. J. Hox, M. Moerbeek, and R. van de Schoot, *Multilevel Analysis: Techniques and Applications* (Routledge, New York, 2017).

[76] F. L. Huang, Multilevel modeling myths, Sch. Psychol. Q. **33**, 492 (2018).

[77] A. Donner and J. J. Koval, The estimation of intraclass correlation in the analysis of family data, Biometrics **36**, 19 (1980).

[78] L. Kish, *Survey Sampling* (John Wiley & Sons, New York, 1965).

[79] J. Cohen and P. Cohen, *Applied Multiple Regression/Correlation Analysis for the Behavioral Science* (Erlbaum, Hillsdale, 1983).

[80] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Erlbaum, Hillsdale, 1988).

[81] P. Häussler and L. Hoffmann, A curricular frame for physics education: Development comparison with students' interests, and impact on students' achievement and self-concept, Sci. Educ. **84**, 689 (2000).

[82] P. Potvin and A. Hasni, Interest, motivation and attitude towards science and technology at K-12 levels:

A systematic review of 12 years of educational research, Stud. Sci. Educ. **50**, 85 (2014).

[83] A. Krapp and M. Prenzel, Research on interest in science: Theories, methods, and findings, Int. J. Sci. Educ. **33**, 27 (2011).

[84] H. Çakir and Ö. Cengiz, The use of open ended versus closed ended questions in Turkish classrooms, Open J. Mod. Linguist. **6**, 60 (2016).

[85] J. S. Kinnebrew, J. R. Segedy, and G. Biswas, Analyzing the temporal evolution of students' behaviors in open-ended learning environments, Metacogn. Learn. **9**, 187 (2014).

[86] K. Beckmann, T. Apps, S. Bennett *et al.*, Self-regulation in open-ended online assignment tasks: The importance of initial task interpretation and goal setting, Stud. Higher Educ. **46**, 821 (2019).

[87] D. Koert, S. Trick, M. Ewerton, M. Lutter, and J. Peters, Incremental learning of an open-ended collaborative skill library, Int. J. Humanoid Robot. **17**, 2050001 (2020).