Received: 31 July 2023 | Revised: 28 March 2024 | Accepted: 12 April 2024
DOI: 10.1002/tea.21947

JRST | WILEY

RESEARCH ARTICLE

Learning with multiple external representations in physics: Concreteness fading versus simultaneous presentation

Andreas Lichtenberger[1] | **Tommi Kokkonen**[2] | **Lennart Schalk**[3]

[1]Laboratory for Solid State Physics, ETH Zurich, Zurich, Switzerland
[2]Department of Physics and Astronomy, University of Turku, Turku, Finland
[3]Pädagogische Hochschule Schwyz, Goldau, Switzerland

Correspondence
Andreas Lichtenberger, Laboratory for Solid State Physics, ETH Zurich, John-von-Neumann-Weg 9, 8093 Zurich, Switzerland.
Email: lichtenberger@phys.ethz.ch

Abstract
Multiple external representations (MERs) are useful for teaching complex content in science education. An open question is whether there is an especially effective way to sequence MERs. On the one hand, the so-called concreteness fading approach suggests starting instruction with more concrete representations and proceeding stepwise to more idealized representations. The effectiveness of this fading approach is, however, supported mainly by studies in mathematics education, while the results in physics are equivocal. On the other hand, presenting different representations simultaneously may support linking, that is, the comparison and contrast of representations, which may benefit learning. In an experimental classroom study ($N = 187$), we compared concreteness fading and simultaneous presentation of MERs for learning a challenging physics content in high school, namely, Faraday's law. We found no significant differences between conditions in posttest performance, and an equivalence test with bounds $d = -0.5$ to 0.5 showed that both approaches performed equally. The results align with previous findings questioning the superiority of concreteness fading over other ways of sequencing MERs.

Therefore, facilitating students' understanding of a complex physics content may involve more than determining the optimal order of presenting MERs. We discuss limitations of the present study and implications for future research and practice.

KEYWORDS

concreteness fading, multiple external representations, physics instruction, science education

1 | INTRODUCTION

A central aim but also a major challenge in science education is to foster understanding of scientific concepts, that is, supporting the construction of relational knowledge of the concepts in a domain (Goldwater & Schalk, 2016). Learning scientific concepts is typically supported using *multiple external representations* (MERs), such as manipulatives, figures, diagrams, graphs, and equations (e.g., Ainsworth, 2008; Corradi et al., 2012; Treagust et al., 2017), for several reasons. First, each representation brings unique advantages and benefits for conceptual understanding (e.g., Fredlund et al., 2014; Lampinen & McClelland, 2018). For instance, a graph of an object's velocity or acceleration is effective for inferring the kinematics over time, but tables are often better for reading exact values or finding maximum/minimum values (Ainsworth, 2014, p. 468). Second, MERs convey complementary information about scientific concepts. Thereby, they might foster students' understanding (Kozma et al., 1996) and support them in recognizing what is essential in the various representations. That is, multiple representations complement but also constrain each other: A formula may focus the interpretation of the outcome of an experiment directly on the relevant variables; a graph may focus the interpretation on the dynamicity of the relationship between variables; and a formula may constrain the interpretation of a graph by determining, for example, the scaling of the axes. Third, using MERs can support abstraction across representations, that is, help learners construct relational knowledge and thus generalizable and transferable schemata. There is now clear evidence that learning with MERs is more effective than learning with single external representations (e.g., Malone et al., 2020; Müller et al., 2017; Treagust et al., 2002). However, how to sequence or combine MERs to support learners' understanding of scientific concepts is less clear, with several approaches, sometimes contradictory, being proposed in the research literature. In the present study, we compared two approaches, concreteness fading versus simultaneous presentation of representations, in a physics classroom experiment with upper secondary school students learning about Faraday's law.

Concreteness fading (Fyfe & Nathan, 2019) is a widely known approach to promote students' understanding of concepts. Accordingly, MERs should be presented in a stepwise sequence from more concrete to more idealized external representations. This approach builds on Bruner's (1966) theoretical suggestion to sequence enactive, iconic, and symbolic representations and on empirical evidence according to which concrete representations help students activate prior knowledge and ground the following more idealized representations in the initial concrete learning experiences (Fyfe & Nathan, 2019). Several implementations of concreteness fading exist. For example, implementations differ regarding the number of steps (two vs. three) or the types of representations used (pictures vs. physical objects). However, all implementations share the idea that learning should start from concrete representations and proceed stepwise, presenting one representation after another to more idealized representations (for an overview, see Kokkonen & Schalk, 2021; Suh et al., 2020). Note that some researchers refer to the more idealized representations as "abstract" representations (e.g., Ainsworth, 2006). We chose to follow Fyfe and Nathan's (2019) terminology in which abstractness is used to describe ideas (e.g., the triangle-ness of triangles), and representations represent ideas in more or less concrete or idealized ways. Specifically, "the term 'idealized' is relative, not dichotomous [...], as some representations can be more idealized than others" (Fyfe & Nathan, 2019, p. 407). There is some encouraging evidence for concreteness fading in mathematics education, but the evidence from physics education is mixed (Jaakkola & Veermans, 2018; Kokkonen et al., 2022). The generalizability of the approach across domains has also been questioned in a conceptual analysis (Kokkonen & Schalk, 2021).

Besides the sequential introduction of different external representations, there is also evidence that simultaneous presentation of MERs may be beneficial. It has been shown repeatedly that simultaneous presentation prompts processes of comparing and contrasting, which in turn support the understanding of concepts and the ability to successfully transfer knowledge to novel problems (for a meta-analysis, see Alfieri et al., 2013). Likewise, theories of multimedia learning suggest that people learn more when corresponding representations (e.g., text, words, and pictures) are presented according to the spatial and temporal contiguity principle; that is, different representations are presented at the same time, close to each other, and thus aligned in time and space (Mayer & Fiorella, 2014). For example, Mayer (2009) reported meta analytic results which show strong positive effects on transfer performance for the spatial contiguity principle ($d = 1.09$). That is, "learners performed better [...] when corresponding text and illustrations were placed near each other on the page (or when corresponding on-screen text and animation segments were placed near each other on the screen) than when they were placed far away from each other" (Mayer, 2009, p. 135). A similar effect was reported for the temporal contiguity principle ($d = 1.31$), that is, "when corresponding portions of animation and narration were presented simultaneously rather than successively" (Mayer, 2009, p. 153). Notably, these meta analytic results are based on rather "simple" learning materials: the combination of text/narration with illustration/animation. Ginns (2006) conducted a meta-analysis on the spatial and temporal contiguity principle, too, and identified the complexity of the learning materials as a moderator. That is, for more complex materials, increasing the spatial and temporal contiguity may be even more effective. These meta-analytic results provide initial additional evidence for why a simultaneous presentation of representations such as manipulatives, diagrams, and graphs in more "complex" learning materials may be more beneficial than following a concreteness fading sequence. Especially in physics instruction, where experimental observations are a key ingredient of instruction (Börlin, 2012), essential information is often distributed across MERs. More idealized external representations such as visual-graphical models or mathematical formalisms are typically presented before or after experimentation. This form of sequencing creates spatial and temporal discontiguity (Schroeder & Cenkci, 2018), as the external representations cannot be directly compared to or contrasted with the observations nor do they adapt dynamically to changes in the experimental manipulations. When essential information is distributed across external representations, learners need to establish links between MERs to derive an integrated understanding (Ainsworth, 1999; Mayer, 2014). Put differently, prompting students to make connections across MERs is crucial, and simultaneous presentation has been suggested to support students in recognizing such connections (Goldwater & Gentner, 2015; Johnson & Mayer, 2012; Mason et al., 2013; Son et al., 2011).

Thus, while proponents of the concreteness fading approach explicitly hypothesize that sequential presentation is more beneficial than presenting several representations simultaneously (Fyfe & Nathan, 2019, Hypothesis 5, p. 413), research on the spatial and temporal contiguity principle and on learning with MERs in science education suggests that simultaneous presentation may have advantages (e.g., Alfieri et al., 2013; Goldwater & Schalk, 2016). Given this conflicting evidence, it seems likely that these two approaches may have specific benefits (and detriments) that could also cancel each other out.

In the present study, we compared concreteness fading to simultaneous presentation of MERs with a fine-grained assessment to identify possible specific benefits of either of the two approaches. Students learned about an advanced content in high school physics, namely, Faraday's law. The law is at the core of electromagnetic induction, which is considered an important but challenging topic (Bagno & Eylon, 1997). Faraday's law is especially demanding because of the abstract nature of the involved concepts of magnetic field and magnetic flux and their difficult relational structure. Based on the argumentation of its proponents, concreteness fading seems to be an apt approach to learn Faraday's law because concrete hands-on experiments involving familiar manipulatives (e.g., bar magnets and coils) may facilitate students' access to this content. However, starting with experiments has also long been criticized in physics education since students often miss the relevant features of the experiments (Chi & VanLehn, 2012; Driver, 1983; Hodson, 1993). In the present case of learning about Faraday's law, a possible solution for focusing students on the relevant features is to support them in relating the hands-on experiments to more idealized representations such as magnetic field line diagrams. Hence, presenting multiple representations simultaneously might be beneficial. Before going into the details of the present study, we review why and how MERs have been used in science education and provide more detail on stepwise sequencing as well as simultaneously presenting MERs.

2 | MULTIPLE EXTERNAL REPRESENTATIONS IN SCIENCE EDUCATION

An *external representation* denotes purposefully constructed objects that stand for a concept (cf. Zhang, 1997). As educators introduce new content (e.g., a law) to students in science education, they routinely make use of MERs. For example, electromagnetic induction may be demonstrated by moving a bar magnet near a coil, which is connected to a voltmeter to indicate whether a voltage is induced (and, hence, electromagnetic induction has occurred). The experiment may be followed by a diagrammatic or pictorial representation of the experimental setup accompanied by corresponding magnetic field lines. Moreover, the symbolic mathematical representation of Faraday's law must also be introduced. Using these MERs serves various aims and reasons. One reason is the nature of physics itself: an important goal of physics education is to be able to understand highly idealized representations, which can subsequently be applied to various concrete instantiations of a concept (Kokkonen & Schalk, 2021; Opfermann et al., 2017).

Another reason for introducing MERs is that their combination can benefit the learning and understanding of new contents. In the above example, each of the representations may highlight a different aspect of the target, Faraday's law, or allow for different reasoning so that none of them alone is sufficient. The concrete, hands-on experiment helps students understand the phenomenological context in which Faraday's law applies (Girwidz et al., 2021). The mathematical symbolism allows for precise reasoning about the dynamics of the experiment–for example, predicting the magnitude of the induced voltage as a function of time. Moreover, mathematical symbolism, as the most idealized representation, can support knowledge transfer to novel contexts if understood correctly. Diagrammatical or pictorial representations such as the field line representation may serve as an important link between the experiment and mathematical symbolism (Küchemann et al., 2021). They may also allow for qualitative reasoning regarding the dynamics of the situation. Thus, seeing Faraday's law represented differently and having to work out the connections, similarities, and differences between the representations may support students in developing deep and transferable understanding.

In the design, function, tasks (DeFT) framework, Ainsworth (2006) identified several factors that affect how presenting MERs influences learning and understanding of a target concept. Among the factors are design principles such as the number, type, and sequences of MERs, the functions that MERs play (e.g., mutually constraining interpretation, providing complementary information), and the cognitive tasks involved in learning with MERs. Deciding how to sequence MERs or whether to present them simultaneously is often accompanied by conceptualizing the representations. For example, according to Ainsworth (2006), representations can be categorized based on their type (e.g., histogram, table, equation, picture, etc.), modality (e.g., text-based or graphical), or specificity (e.g., domain-general vs. domain-specific). One conceptualization that has garnered interest during recent years is the level of concreteness of MERs (e.g., Belenky & Schalk, 2014; Fyfe & Nathan, 2019). Fyfe and Nathan (2019) argue that all external representations are concrete in that they are necessarily depicted "in this physical world" (p. 407)—even mathematical symbolism like, for example, equations become concrete by writing them on a piece of paper. However, some representations are less concrete, or put differently, more idealized than others. Accordingly, hands-on materials such as the tools used in an experiment would be considered less idealized (or more concrete) and equations more idealized (or less concrete). While hands-on materials have the benefit of being physically manipulable, they also contain irrelevant information such as color or additional functionalities. Thus, learners face the challenge of understanding what is relevant and what is irrelevant. In contrast, highly idealized representations such as mathematical symbolisms precisely inform students about what aspects are relevant (i.e., exactly those variables represented by the symbolism). Thus, a more idealized external representation makes the to-be-learned information more salient.

However, students often struggle to grasp the generality of highly idealized representations such as mathematical symbolisms and how such representations map to their targets, that is, to the "concrete" phenomena. Science education faces this dual challenge: learners need to know how different representations highlight different aspects of a concept, and they also need to understand the concept itself and how to use it in novel contexts, that is, to transfer their knowledge.

2.1 | Sequences of representations

The previous section highlighted that every representation offers its own advantages. Consequently, sequences of MERs can be developed based on these particular advantages. The concreteness fading approach suggests that learning should begin with less idealized (i.e., more concrete) external representations such as manipulatives or hands-on experiments. These representations are supposed to embed the target concept in a grounded and familiar context, which can trigger learners' prior knowledge or support their reasoning (Cobb et al., 1992; Donovan & Fyfe, 2022; Martin & Schwartz, 2005). Subsequently, learning should proceed to more idealized representations. Concreteness fading aligns with various other approaches, such as the Concrete-Representational-Abstract (CRA, e.g., Bouck et al., 2017) approach. There are, however, different ways of implementing concreteness fading. They vary, for instance, in the number (e.g., two vs. three) or in the types of representations used. While some advocate the use of physical manipulatives in the first step (followed by pictorial representations, for example), others implement virtual or pictorial representations of varying concreteness in each step.

Fyfe and Nathan (2019, p. 411) were the first to provide a rigorous definition of concreteness fading as the "three-step progression by which a concrete representation of a concept is explicitly faded into a generic, idealized representation of that same concept." To foreshadow, our implementation of concreteness fading for learning about Faraday's law will follow this definition. Fyfe and Nathan (2019) further argued that concreteness fading is a general framework for teaching concepts across disciplines. From reviewing the literature, they derived several hypotheses regarding the effectiveness of concreteness fading. For instance, they hypothesized that concreteness fading is more effective than the reversed sequence (*concreteness introduction*), in which more idealized external representations are followed by more concrete ones (Fyfe & Nathan, 2019, pp. 414–415). This hypothesis has received some encouraging support, mostly from mathematics education (e.g., Ching & Wu, 2019; Fyfe et al., 2015; Ottmar & Landy, 2016).

However, in physics education, the empirical evidence for concreteness fading is equivocal. In a study closely following Fyfe and Nathan's (2019) definition of concreteness fading, high school students learned about Faraday's law with three external representations differing in concreteness: physical manipulatives (bar magnets and coils) used in hands-on experiments, field line representations, and mathematical symbolism (Kokkonen et al., 2022). The students learned either in a three-step concreteness fading or concreteness introduction sequence. The results indicate equal effectiveness of both sequences for conceptual understanding. Johnson et al. (2014) even found evidence for the superiority of concreteness introduction when they examined learning of DC circuit problems with circuit diagrams embedded in realistic pictorial illustrations (referred to as "concrete" by the authors) and conventional circuit diagrams with no contextualized illustration (referred to as "abstract" by the authors); specifically, sequencing the representations from abstract to concrete produced better performance in the transfer tests than concrete-to-abstract, concrete-only or abstract-only conditions.

The seemingly discordant evidence between mathematics and physics may be due to the different nature of the domains and the complexity of the target knowledge (Kokkonen & Schalk, 2021). Moreover, repeated critique against hands-on practical work as a starting point (i.e., as the first representation) has been expressed in science education research because students often do not observe the aspects that the teacher wanted to exemplify: the students tend to "see what they expect to see" (Driver, 1983, p. 2; see also Chi & VanLehn, 2012; Gunstone, 1991; Hodson, 1993). Thus, concrete representations may become relevant only after the learner has some knowledge of the concept that the concrete representation instantiates, especially if the target knowledge comprises complex relational structures (Kokkonen & Schalk, 2021).

Furthermore, the way in which more and less idealized representations relate to each other in mathematics and physics may differ. In mathematics, the more concrete representations are mainly used to make the more idealized representations understandable. A teacher is quite free to construct more concrete representations and make them closely analogous to the more idealized representation (e.g., use 5 apples or 5 oranges to represent a set of 5). Conversely, in physics, a teacher has less freedom in constructing the more concrete instantiations (e.g., hands-on-experiments) to represent the more idealized laws, and the different representations may be less analogous than in mathematics (for a detailed discussion, see Kokkonen & Schalk, 2021). For example, pictures with field-like diagrams of electromagnetic experiments do not only shed irrelevant details (as the more idealized representations in mathematics do) but also introduce relevant conceptual information (i.e., the field lines).

2.2 | Simultaneous presentation of MERs

In physics education, a major focus is typically on understanding the generality of a concept that is often best represented by mathematical formalism. However, as we already pointed out, students often have difficulties relating such formalism to concrete phenomena. A simultaneous presentation of the formalism with other less idealized representations may help students understand what the formalism stands for. An explanation for this benefit is that simultaneous presentation creates spatial and temporal contiguity (Ginns, 2006; Mayer, 2009; Mayer & Fiorella, 2014) and prompts processes of comparing and contrasting, a specific form of analogical learning (e.g., Gentner, 2010). There exists a vast body of empirical studies whose results empirically support simultaneous over sequential presentation (Alfieri et al., 2013). Crucially for the present study, simultaneous presentation can support learning of relational, law-like schemata underlying scientific knowledge by triggering processes of structural alignment (Gentner, 2010; Goldwater et al., 2011; Goldwater & Schalk, 2016). The processes of structural alignment according to Gentner (2010) are schema abstraction, difference detection, re-representation, and inference projection. These processes fit the proposed beneficial functions of MERs suggested by Ainsworth (2008) well: MERs can play complementary roles (difference detection, inference projection), constrain the interpretation of each other (difference detection, re-representation, inference projection), and help construct a deeper understanding (schema abstraction, re-representation); for details, see Kokkonen and Schalk (2021). Moreover, via the processes of structural alignment, students might not only learn about the physical concept but also improve their ability to interpret, generate, and switch between representations (e.g., Kohl & Finkelstein, 2005; Kozma & Russell, 2005).

There is also empirical support for the benefits of simultaneous over sequential presentation of MERs in the context of physics education (Gadgil & Nokes, 2009; Kurtz et al., 2001; Mason, 2004). In these studies, the examples were typically in similar representational format. For example, Kurtz and colleagues (2001; see also Mason, 2004) presented line drawings representing different concrete phenomena of heat transfer either sequentially or simultaneously. However, in physics education, students are often asked to learn from and with different representational formats. For instance, in the context of kinematics, an object's motion (displacement, velocity, and/or acceleration) has to be graphically represented. Students typically have difficulties in interpreting graphs and/or in translating between a concrete hands-on experiment and a more idealized graphical representation (McDermott et al., 1987). Seeing a graph of an object's motion unfold simultaneously with the perceived real motion of the object instead of viewing them sequentially has been shown to support students graphing skills and conceptual understanding of the related concepts of velocity and acceleration (Becker et al., 2020; Beichner, 1996; Brasell, 1987; Thornton & Sokoloff, 1990). For example, Becker et al. (2020)

found that a video analysis approach to teaching kinematics concepts, which was designed accordingly, was more beneficial than traditional lab work where simultaneous graphing was not available. However, earlier studies directly comparing simultaneous graphing to delayed viewing of the graphs showed no statistically significant benefits regarding students' graphing skills (Beichner, 1990; Brungardt & Zollmann, 1995). Thus, similar to the evidence for concreteness fading in physics, the empirical support for simultaneous presentation (compared to sequential presentation) is also equivocal.

To summarize the previous two sections, the research on simultaneous presentation of MERs and on concreteness fading (i.e., the sequential presentation of MERs) do not fit together well. While the research on the benefits of simultaneous presentation typically pits the performance of learners in a simultaneous condition against the performance in a sequential condition, Fyfe and Nathan (2019) hypothesized that a concreteness fading sequence should work better than presenting the representations all at once (p. 417). However, Fyfe and Nathan (2019) also argued that concreteness fading should support thinking of the "representations as one and the same, each an instantiation of a common invariant set of relations" (p. 417), thereby supporting learning the meaning of the most idealized representations. This aim is similar to the proposed processes in analogical learning, where learners are confronted with several representations simultaneously to help them construct a general schema that aligns the representations. Thus, either sequential presentation in the form of concreteness fading or simultaneous presentation in the vein of analogical learning are suggested to be superior, depending on the respective research literature.

3 | THE PRESENT STUDY

We took up the conflicting predictions from concreteness fading research and research on the simultaneous presentation of MERs. Specifically, based on the hypothesis presented by Fyfe and Nathan (2019, Hypothesis 5, p. 417), we examined whether concreteness fading is more beneficial than the simultaneous presentation of MERs. We tested this hypothesis in an experimental classroom study situated in upper secondary school physics education. Students learned about Faraday's law with three representations: manipulatives in a hands-on experiment, iconic field lines, and mathematical graphs together with a formula. These three representations are complementary; that is, they all represent the same content—Faraday's law—but differ in the supported processes (i.e., hands-on experimentation supports enactment, field lines highlight dynamic relationships, and formulas support effective computation). To simplify our subsequent methodological description of the study and stick to the terminology used in previous research on concreteness fading (e.g., Fyfe & Nathan, 2019; Kokkonen et al., 2022), we refer to these three representations as concrete, intermediate, and idealized, respectively. One condition implemented the three-step concreteness fading sequence. In the other condition, the three representations were presented simultaneously. Before receiving the intervention on Faraday's law, the students underwent a self-learning sequence on the concept of magnetic flux to ensure that students had the relevant prior knowledge for learning about Faraday's law.

To assess learning, we adapted and applied two detailed multiple-choice assessments developed by Kokkonen et al. (2022) as pre- and posttests. We enriched the posttest with two open questions that involved a transfer task to gain further insight into students' thinking. To better understand the underlying learning processes, we explored how students made use of the multiple representations in both conditions. For this purpose, we (1) investigated students' ability to translate between representations with dedicated tasks, (2) included a specific prompt asking students to explain Faraday's law in their own words to determine how many and which representations students spontaneously used, and (3) explicitly asked students about the usefulness of the different representations in solving the transfer task mentioned above.

The topic of electromagnetic induction, and Faraday's law in particular, is challenging for students. Faraday's law governs the generation of an electromotive force (i.e., a voltage) in an electric circuit due to a changing magnetic flux through the circuit. At the upper secondary school level, magnetic flux through a surface is usually formally defined as the strength of the magnetic field times the area of that surface perpendicular to the field. When using the magnetic field line representation, it can also be conceptualized as the number of field lines through the surface. Faraday's law is typically experimentally demonstrated by placing a coil (which constitutes the electric circuit) in a magnetic field and changing either the orientation of the coil or the magnetic field, which generates a voltage in the coil due to changing magnetic flux through the surface of the coil. In this hands-on experiment, Faraday's law implies that the voltage is equal to the negative of the rate of change of the magnetic flux through the coil. This is equivalent to the voltage being the negative time derivative of the magnetic flux as a function of time. Written as an equation, this reads: $U = -\Phi'(t)$, with U denoting the voltage, Φ the magnetic flux, and t the time.

The difficulty of learning Faraday's law is partly associated with its relational complexity: it requires understanding the quantity voltage as the rate of change of the quantity magnetic flux, which itself is related to the magnetic field strength and surface area. Students often struggle to correctly relate the quantities (Jelicic et al., 2017). For example, they think that the voltage is related to the magnetic flux instead of the rate of change of the magnetic flux, or they confuse magnetic effects with electric effects (Bagno & Eylon, 1997; Jelicic et al., 2017; Maloney et al., 2001). Another difficulty is related to recognizing when and how Faraday's law should be applied in concrete contexts (e.g., experimental setups). For instance, students may think that movement of the magnetic field or the circuit/coil is necessary to generate voltage, thereby mistaking “change” as change in position or orientation instead of the change in the magnetic flux and failing to apply the concept of magnetic flux altogether (Jelicic et al., 2017; Maloney et al., 2001). The relational complexity of this content and the challenges in learning and understanding the respective concepts make it a well-suited domain to compare learning via a concreteness fading sequence vs. learning by simultaneous presentation.

4 | METHODS

4.1 | Participants

To plan the sample size, we took the previous concreteness fading studies as starting points. Fyfe et al. (2015) and McNeil and Fyfe (2012) reported effect sizes of $d = 0.76$ and $d = 0.83$, respectively, for comparing concreteness fading against other approaches (e.g., concreteness introduction, in which the concreteness fading sequence is reversed) in an immediate posttest but not against simultaneous presentation. We performed a power analysis with the software G*Power (version 3.1.9.6; Faul et al., 2009) for an analysis of covariance (ANCOVA) design with one covariate, alpha $= 0.05$ and power of 0.9. Because of the contradictory evidence regarding our conditions, we chose a more conservative effect size estimate of $d = 0.50$ (corresponding to

$f = 0.25$). The analysis yielded a minimum total sample size of 108 (assuming $R^2 = 0.35$ for the covariate).

We recruited 11th grade students (mean age $M = 17.8$ years, $SD = 0.7$ years) from 12 school classes (six different teachers) in three upper secondary schools in the German-speaking part of Switzerland. In the analytic sample, we included only students who completed all tasks and tests, gave consent to the use of their data in the present study, and indicated on a validation question that their data were trustworthy. Due to a very small attrition rate, we obtained a sample size of $N = 187$, which was more than adequate for detecting an effect as specified in the power analysis.

All students were given anonymized codes for study participation. They were ranked according to their most recent physics grades by their teachers in advance and then pairwise randomly split between conditions by the study investigators. Thus, the assignment was random but balanced regarding previous performance in physics to ensure a fair comparison between conditions. The group sizes for the concreteness fading (CF) condition and the simultaneous presentation (SIM) condition were $N = 98$ (63% female) and $N = 89$ (54% female), respectively. Unequal group sizes are due to the exclusion of students who did not fulfill the criteria mentioned above.

4.2 | Study design and procedure

We implemented an experimental within-classroom between-subjects design with condition (CF or SIM) as the independent variable, the pretest score as the covariate, and the posttest score as the dependent variable. The whole study took place during the three normal weekly physics lessons at the participants' schools, which were divided into a single lesson (45 min) and a double lesson (90 min). The study comprised four phases: pre-instruction (in the single lesson), pretest, learning phase (also referred to as intervention), and posttest (all in the double lesson). The study was coordinated with the curriculum of all participating classes. Physics lessons prior to the study covered basic knowledge on magnetism. However, magnetic flux and magnetic induction were not taught before the intervention. Figure 1 shows the procedure of the study for both conditions.

In the pre-instruction, the students learned about magnetic flux to ensure that all students had the relevant prior knowledge to learn about Faraday's law. To minimize the influence of the teacher, we prepared self-learning materials. This material was identical for both conditions. Teachers were advised not to intervene. They only attended to the students in administrative matters (e.g., preventing student interactions, distributing and collecting materials). Pre-instruction occurred in the regular single physics lesson prior to the intervention, which was between 0 and 3 days before the intervention (depending on the schools' timetables).

To control for the varying time span, the pretest was administered immediately before the intervention at the beginning of the double lesson in all courses. The pretest assessed knowledge of magnetic flux and necessary prior mathematics knowledge. During the intervention, participants learned about Faraday's law using materials designed for a single-lesson period (for details, see Materials). Students from both conditions worked in the same classroom. To prevent cross-condition contamination and student interactions, they were spread out so that they could not talk or see each other's answers. During the entire intervention, students worked individually without help from other students, teachers, or study administrators. Directly after the intervention, students completed the posttest. Both pre- and posttests were administered via an online survey tool. Students were allowed to go back and forth in all online tests.

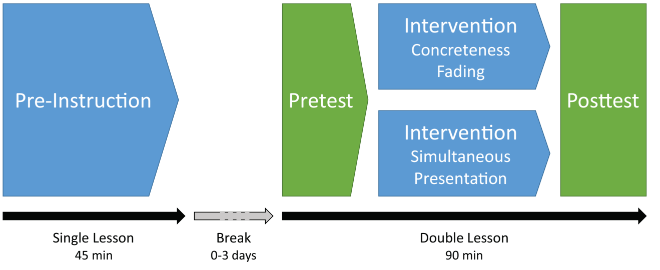


FIGURE 1 Study procedure. All students received a pre-instruction on magnetic flux in a single lesson. In the subsequent double lesson 0–3 days later, students first solved a pretest. For the intervention about Faraday's law, students were randomly assigned to the two conditions (concreteness fading or simultaneous presentation) within classrooms. After completing the intervention, they solved a posttest.

5 | MATERIALS

5.1 | Pre-instruction

The pre-instruction phase was designed with the goal of students acquiring the necessary prior knowledge to learn Faraday's law. Students were given a self-learning booklet that introduced the concept of magnetic flux using three representations: pictures of a real magnet and coil (concrete), magnetic field line diagrams (intermediate), and a formula (idealized). While students were already familiar with the types of representations used in the booklet, the concept of magnetic flux was new to them.

Magnetic flux was first described and illustrated using the concrete and intermediate representations simultaneously: a picture of a magnet with a conductor loop ("magnetic flux as the amount of magnetic field through the surface of the conductor loop") was placed next to an illustration with magnetic field lines ("magnetic field as the number of field lines through the surface"). Afterward, students elaborated on the dependence of magnetic flux on the strength of the magnetic field, on the area of the loop, and on the position of the loop relative to the magnetic field direction using field line diagrams. Eventually, the formula for the magnetic flux was introduced. Finally, the students solved three tasks (one per representation) with prepared solutions to obtain feedback on their understanding of magnetic flux. We used this approach because, first, in our opinion it did not favor one of our conditions in the intervention. The first two representations were introduced simultaneously to support students linking the experimental components and the field line diagrams (closer to the SIM condition). The most idealized representation was presented last (closer to the CF condition), after students had elaborated the relation of the magnetic flux to the surface and magnetic field qualitatively using the intermediate representation. Second, this approach with a strong focus on the intermediate representation is also suggested in the physics education literature (e.g., Erfmann & Berger, 2015).

5.2 | Pretest

The pretest consisted of 25 multiple-choice items, each with one correct answer and four distractors. Correct answers were awarded 1 point, and incorrect responses were awarded 0 points. Six questions asked about the derivatives of linear and sinusoidal functions. They were taken and slightly adapted from Planinic et al. (2013). Understanding the derivative of a function is a mathematical prerequisite to learn Faraday's law because of the corresponding relation between voltage and magnetic flux. In addition, we included an item from the Representational Competence of Fields Inventory (Küchemann et al., 2021) to check students' understanding of the field line representation. Furthermore, the test comprised 18 questions on magnetic flux to control for the effect of the pre-instruction phase. Twelve items made use of only one representation (concrete, intermediate, idealized; four per representation), whereas three items included quantitative calculations (one per representation). Six items required translation between two representations (e.g., from concrete in the question to intermediate in the selection of answers; one per combination of representations). These items could also be solved without a deep understanding of magnetic flux but by identifying similar situations in the various representations. Eleven questions were taken from Kokkonen et al. (2022) and complemented with seven newly developed items. Table 1 gives an overview of the pretest structure. Items 1–15 were presented in individually randomized order to the students. Items 16–18, which required quantitative calculations, were presented in fixed order at the end of the test. An English translation of the 18 multiple-choice items is available as Supporting Information accompanying the online article.

To assess the psychometric properties of the pretest, we calculated the item difficulties (mean value $M = 0.82$), the item discrimination indices ($M = 0.32$), and the point biserial correlations ($M = 0.36$) according to Ding and Beichner (2009), which are in the desired range. The relatively low mean values of the item discrimination indices and point biserial correlations mainly result from the items with high solution rates. Items that are solved correctly by nearly all participants do not discriminate well nor do they highly correlate with other items. To estimate the internal consistency of the pretest, we calculated McDonald's omega for the present sample, $\omega = 0.75$, indicating acceptable consistency. Based on these analyses, we regard the pretest as a reliable and valid instrument to assess students' knowledge prerequisites for the intervention.

We refrained from including items on Faraday's law in the pretest. The students did not learn about Faraday's law in physics at school, and according to Bagno and Eylon (1997), they lack everyday experiences with this topic. We therefore assumed that students had no significant prior knowledge of Faraday's law, and that their knowledge development could thus be largely attributed to the intervention. Furthermore, including items on Faraday's law could have negative effects on the intervention. First, students would have been exposed to all representations regarding Faraday's law. This initial introduction to all representations would have confounded the learning phase, especially in the concreteness fading condition. Second, we did not want to target student's learning to specific facets of problems they would only later investigate in the intervention. Third, we assumed that exposing students to a large set of items they

TABLE 1 The 18 magnetic flux items in the pretest categorized based on the representation(s) used.

	Single				Translation					
Concrete	1	2	3	16^{C}	10^{o}	11^{X}	12^{o}	13^{X}		
Intermediate	4	5	6	17^{C}	10^{X}	11^{o}			14^{o}	15^{X}
Idealized	7	8	9	18^{C}			12^{X}	13^{o}	14^{X}	15^{o}

Note: Items marked with a superscript ^{C} include quantitative calculations. Items requiring translation between representations are denoted either with superscript ^{o} (origin representation in question) or ^{x} (target representation in the selection of answers).

are not able to solve could lead to frustration and cause a negative attitude toward the learning phase. Moreover, the items of the pre- and posttest cover similar situations but ask about different physical quantities (magnetic flux and induced voltage, respectively). Without knowledge about the induced voltage, items on Faraday's law could even trigger a confusion of magnetic flux and induced voltage and thus increase the learning difficulty to correctly relate the two quantities (Jelicic et al., 2017).

5.3 | Learning phase

In the intervention, we also provided a self-learning booklet designed for the purposes of the present study. The booklet covered Faraday's law using three representations with varying degree of idealization, that is, ranging from concrete (hands-on experiments) through intermediate (field line diagrams) to idealized (graphs depicting magnetic flux combined with a formula). There were four tasks for each representation (total: 12 tasks), depicting the following four situations: (1) a magnet resting next to the coil, (2) a magnet moving backward and forward toward the front of the coil, (3) a magnet moving backward and forward toward the side of the coil, and (4) the coil spinning around next to the magnet. In every task, participants had to investigate or deduce whether a voltage was induced in the coil. The four situations were designed to support a deeper understanding of Faraday's law by focusing on its important features. For example, by contrasting (1) and (2), one can see that the induced voltage is not related to the magnetic flux but to the rate of change of the magnetic flux. Furthermore, (2) and (3) show that the relative movement of the magnet and coil alone is not sufficient for an induced voltage; it depends on the relative position of both objects. Finally, linear and rotational movements can be contrasted by (2) and (4). An overview of all tasks is presented in Figure 2.

In the CF condition, learners first received the four tasks dealing with the concrete representation and proceeded stepwise to tasks with the intermediate and the idealized representations (see Figure 2). Within each step, the tasks were preceded by a short introductory text that explained how Faraday's law is illustrated by the specific representation. In the concrete representation step, the four tasks were framed as predict-observe-explain experiments (White & Gunstone, 1992), a well-established effective and motivating way to design experiments in science education (Hong et al., 2021; Özcan & Uyanık, 2022; Rivera Gavidia & Marrero Galván, 2023). Students were first given an explanation of the task and asked whether there would be a voltage induced (predict). Then, they had to conduct the experiment (observe) with a bar magnet and a coil that was connected to a voltmeter, write down the observation and finally explain it. Figure 3 shows an excerpt of the learning materials in the CF condition that covers the four situations in the concrete step. In the intermediate step, four tasks similar to the concrete tasks were described but were represented by field line diagrams. Students were explicitly told that the described situations were similar to the ones they investigated with the experiments (as suggested by proponents of the concreteness fading approach, see e.g., Fyfe & Nathan, 2019). Again, students had to state and explain for every situation whether a voltage was induced. Finally, in the idealized phase, students were presented four graphs depicting the flux through a coil as a function of time. Again, the materials stated that the four tasks were similar to the ones in the other representations. Students had to draw the four corresponding graphs to depict the induced voltage as a function of time as well as to calculate the voltages quantitatively.

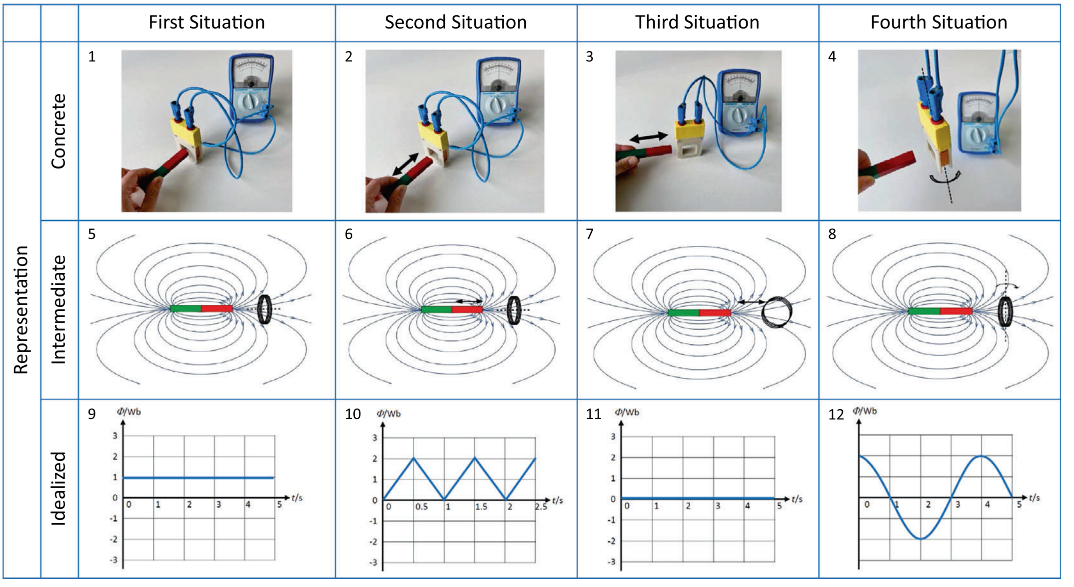
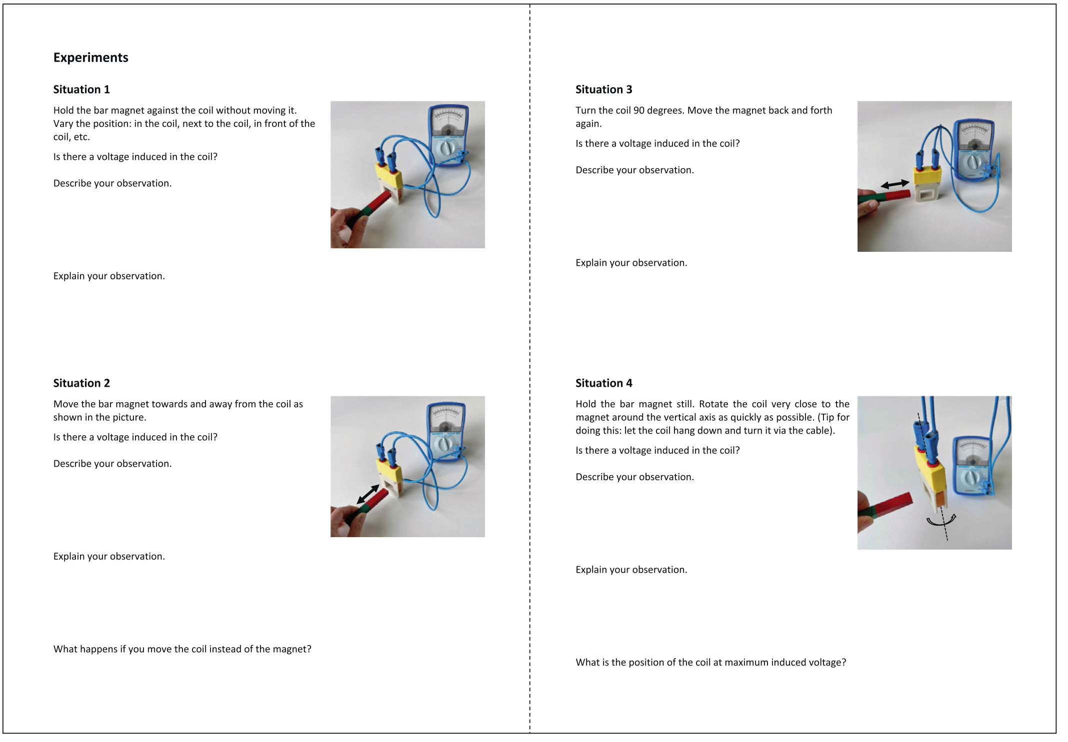


FIGURE 2 Overview of all the situations, representations, and tasks used in the learning phase. In the concreteness fading condition, students encountered the 12 tasks grouped by the type of representation. They started with tasks 1–4 presented in the concrete representation, followed by tasks 5–8 in the intermediate representation, and concluded with the 9–12 tasks in the idealized representation. Conversely, in the simultaneous presentation condition, tasks were grouped by the situation. That is, students first worked on tasks 1, 5, and 9, which all related to the first situation. This pattern continued with the next group of tasks (2, 6, 10) associated with the second situation, and so on.

In the SIM condition, Faraday's law was introduced by presenting all three representations (concrete, intermediate, idealized) simultaneously creating spatial and temporal contiguity. Afterward, the students worked on the three tasks depicting the first situation in all three representations (see Figure 2). As in the CF condition, a text described that the tasks showed a similar situation. The students conducted the first experiment in the predict-observe-explain manner, explained the situation in the intermediate representation, drew the graphs and executed the calculations in the idealized representation. This set of tasks was followed by the second set of tasks referring to the second situation and so forth. Figure 4 provides an excerpt of the learning materials in the SIM condition. It exemplarily shows the tasks that relate to the second situation (i.e., tasks 2, 6, and 10 in Figure 2). Note the spatial and temporal proximity of the three different representations in the excerpt from the SIM learning materials (Figure 4) in contrast to the excerpt from the CF learning materials (Figure 3), where four different situations in the same representation are presented side by side.

All representations were identical in both conditions, only the order differed. At the end of the booklet, all students were given very brief solutions (without explanations) to the four sets of tasks in text form (without using any of the representations of the learning materials). As a final task in both conditions, students responded to the following prompt: “You have studied four tasks in different representations [German word used in the materials: Darstellungen]. Please summarize: How can you determine whether there is a voltage induced?” These prompted self-explanations on Faraday's law served to, first, support students' active and constructive processing (Chi et al., 1994; Chi & Wylie, 2014) and, second, to explore which


Experiments

Situation 1

Hold the bar magnet against the coil without moving it. Vary the position: in the coil, next to the coil, in front of the coil, etc.

Is there a voltage induced in the coil?

Describe your observation.

Explain your observation.

Situation 2

Move the bar magnet towards and away from the coil as shown in the picture.

Is there a voltage induced in the coil?

Describe your observation.

Explain your observation.

What happens if you move the coil instead of the magnet?

Situation 3

Turn the coil 90 degrees. Move the magnet back and forth again.

Is there a voltage induced in the coil?

Describe your observation.

Explain your observation.

Situation 4

Hold the bar magnet still. Rotate the coil very close to the magnet around the vertical axis as quickly as possible. (Tip for doing this: let the coil hang down and turn it via the cable).

Is there a voltage induced in the coil?

Describe your observation.

Explain your observation.

What is the position of the coil at maximum induced voltage?

FIGURE 3 Excerpt from the concreteness fading learning materials. The double-page excerpt shows the combination of tasks and situations as they were presented in the first step (i.e., concrete representation) of the concreteness fading learning materials (see also Figure 1). In the second (intermediate representation) and third step (idealized representation), analogous combinations (4 tasks, 4 situations) were presented. The excerpt is a translated version of the learning materials that were originally in German.

representations students used in their explanations so that we could gain deeper insights into their learning process.

5.4 | Posttest

The posttest consisted of a first part with multiple-choice items and a second part with open questions. For the first part, we used and adapted a test developed by Kokkonen et al. (2022) to assess the learning of Faraday's law. All multiple-choice items had one correct answer and four distractors. Correct answers were awarded 1 point, and false responses were awarded 0 points. The multiple-choice part of the posttest comprised 16 items that assessed the understanding of Faraday's law. In addition, this part included the six items from the pretest about magnetic flux that required translation between representations. Regarding the multiple-choice items on Faraday's law, the posttest structure was identical to that of the pretest (see Table 1) with one exception: a meaningful task on Faraday's law involving the translation between the concrete and intermediate representation could not be realized (i.e., items 10 and 11 in Table 1). While concrete and intermediate representation can provide information about whether voltage is induced in a depicted situation, the translation between these representations can be realized

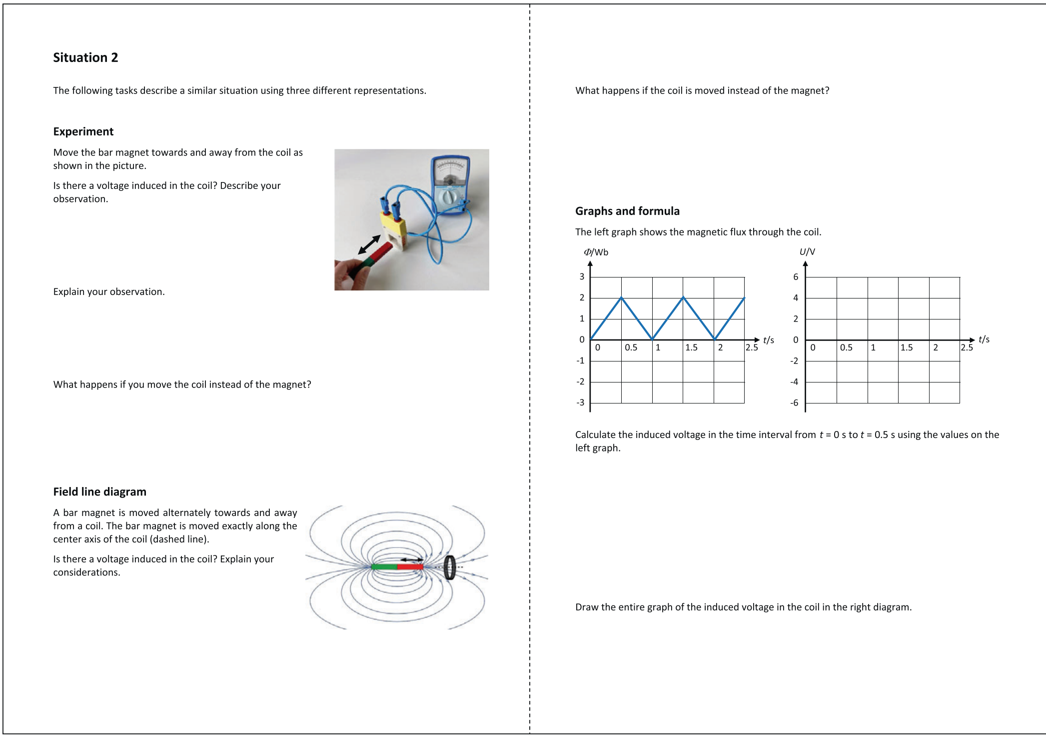


FIGURE 4 Excerpt from the simultaneous presentation learning materials. The double-page excerpt shows the combination of tasks and representations as they were presented with spatial and temporal contiguity in the simultaneous presentation learning materials (here for situation 2; see also Figure 1). Analogous combinations of tasks (i.e., combinations of different representations) were presented for situations 1, 3, and 4. The excerpt is a translated version of the learning materials that were originally in German.

only by comparing superficial features of the illustrations and therefore without understanding the content. Hence, the 16 items on Faraday's law comprised nine items that were presented in one representation (three per representation: concrete, intermediate, and idealized), three items that made use of one representation and included a quantitative calculation, and four items that required translation from one representation to another. As in the pretest, the items were presented in individual random order. Again, the three items that included quantitative calculations were placed at the end of the multiple-choice part of the test. The students were allowed to go back and forth in the test. An English translation of the multiple-choice posttest items is available as Supporting Information accompanying the online article.

To evaluate the psychometric properties of the multiple-choice items on Faraday's law, we calculated the item difficulties (mean value $M = 0.55$), the item discrimination indices ($M = 0.62$) and the point biserial correlations ($M = 0.51$). In addition, we conducted a confirmatory factor analysis to investigate the structural validity of the test. Both Velicer's (1976) minimum average partial test and Cattell's (1966) scree test suggested a single factor solution. A one-factor model indicated a moderate fit between the model and the observed data. However, Horn's (1965) parallel analysis indicated a three-factor model. Modeling three factors assigned to the items with single representations, the items with quantitative calculations, and the items with a translation between representations, respectively (see Table 1), we obtained slightly better fit indices. Therefore, we also report students' subscores with respect to these three factors. Regarding the internal consistency of the posttest, McDonald's omega for the present sample was $\omega = 0.82$, indicating sufficient consistency. Based on these analyses, the multiple-choice part of the posttest can be considered a valid and reliable assessment instrument. The full results of all item analyses are available as Supporting Information accompanying the online article.

The second part of the posttest was designed to gain deeper insight into students' thinking about Faraday's law and their use of representations. For practical reasons (posttest duration and because the answers required content analysis), we applied these open questions to approximately only half of the classes that participated in the study. We designed three open questions that build upon each other. The first question reads, "A friend claims: *A voltage is induced only if either the magnet or the coil is moved. Without movement, there is no voltage*. Please comment on this statement." As discussed before, there is evidence that many students hold the belief (also referred to as misconception) that movement is necessary for induction (Jelicic et al., 2017). Indeed, during the intervention, induction occurred only when students moved either the magnet or the coil. Therefore, with this question, we wanted to determine whether students were able to spontaneously think of new situations in which voltage is induced without movement. The second task was as follows: "Electromagnets can be used to switch magnetic fields on and off. A coil is placed next to a resting electromagnet. Now the electromagnet is switched on so that its magnetic field partly goes through the coil. Explain whether there is a voltage induced in the coil." This task constitutes an example of an induced voltage with all components (electromagnet, coil) at rest. Therefore, with this item, we tested whether students could transfer Faraday's law to a new situation, in which voltage is induced without movement and whether there was evidence of the mentioned misconception. After these two items, which focus on conceptual knowledge of Faraday's law, we asked students which representation they regarded as useful to solve the previous question. On a five-point Likert scale, students rated the usefulness of every representation, which we explicitly listed ("an experiment with the electromagnet and coil," "the field line diagram of the magnetic field through the coil," "the graph of the magnetic flux in a diagram or the formula to describe Faraday's law"). Different representations may offer complementary information and connect differently to students' prior knowledge (cf. Lampinen & McClelland, 2018). In the present study, students might accordingly find different representations helpful depending on the condition (CF vs. SIM).

6 | IMPLEMENTATION FIDELITY

All teachers were assisted by the first author in conducting the study with their first class. Teachers were, however, allowed to conduct the study themselves if they participated with another class. Their role was restricted to administrative tasks. All students worked individually with the learning materials. Teachers did not give additional instructions or assistance in carrying out the experiments. There were no incidents reported in which students talked to each other, supported or distracted their peers. The experimental equipment worked correctly for all students, and no technical difficulties occurred. To ensure that students proceeded in the correct way (pre-instruction, pretest, learning phase, posttest), they were initially given an overview of the procedure. Pre- and posttests were available only during predefined time windows so that students could not unintentionally access the wrong test. All teachers reported that according to their observations, the students worked actively and seriously through the materials. Also, there were no student questions regarding the content of the materials or the experimental procedure. To corroborate this impression, we analyzed the self-learning materials of all 187 students. The students had to work on 12 tasks (four tasks in three representations), which yielded 2244 instances. A vast majority of the students (89%) wrote down answers or explanations for all 12 problems. We found only 37 instances (1.7%) spread over 20 students in which no answer or explanation was given. However, this does not mean that they did not think about the unanswered tasks. We did not rate the answers according to scientific correctness but evaluated whether students engaged with the problems and experiments and wrote down solutions and explanations. Hence, this gives us a good insight in students' engagement. However, we cannot be sure that all tasks were understood correctly and all of the experiments were carried out accurately by everyone.

Regarding the multiple-choice part of the posttest, we statistically evaluated the response times of students. A series of very quick answers may indicate that a student clicked through the test without seriously thinking about the questions (Lichtenberger et al., 2017). We counted the number of items that were solved in less than 5 or 10 s for every student. We found only seven participants who solved five or more items in less than 10 s. Of these students, only three solved three to five items in less than 5 s. Two students were in the SIM, and one was in the CF condition. We decided that these rare cases of fast answers could be neglected. We did not find evidence for systematic cheating or skipping test items in the response times. Moreover, all students confirmed that they took the tests seriously. Overall, we are confident that the study was implemented in the planned way and that students' data are trustworthy.

7 | RESULTS

In what follows, we first present the data gained from the pretest. The subsequent presentation of results from the posttest is divided into two sections. In the first section, we address the effect of condition on learning (i.e., to test the hypothesis that the CF condition outperforms the SIM condition) by analyzing the 16 multiple-choice posttest items about Faraday's law and presenting the answers to the first two open questions, which focused on the conceptual understanding of Faraday's law. In the second section, we present exploratory analyses based on data from the self-learning materials and the posttest tasks that focus on students' use of the representations.

7.1 | Pretest results

Participants' pretest scores were skewed, so we performed a Mann–Whitney *U* test (see Table 2), which did not reveal significant differences between the conditions regarding the total score or the subscores for the magnetic flux, mathematics, and field line items (see Table 1 for an explanation of these items).

7.2 | Effect of condition on learning Faraday's Law

The following two subsections present the comparison between the CF and SIM conditions regarding students' posttest performance on the multiple-choice items and the open questions on Faraday's law.

7.2.1 | Multiple-choice items

To assess the effects of the intervention on learning, we conducted an ANCOVA with the sum score of the 16 multiple-choice items as the dependent variable, condition (CF vs. SIM) as the dependent variable and pretest sum score as the covariate. We confirmed that the assumptions for the ANCOVA were met: the variables were linearly related with a homogeneous relationship across conditions, and Levene's test indicated no significant differences of variances ($F = 0.719$, $p = 0.398$). The results of the ANCOVA are presented in Table 3. The analysis did not reveal a significant difference between conditions. We also found no significant differences for the subscales consisting of the nine items with one representation, the three items that included quantitative calculations, or the four items with translations between representations (assumptions for the analyses of covariance were also met for the subscales, Levene's tests: $F = 3.341$, $p = 0.069$; $F = 0.059$, $p = 0.808$; $F = 0.00$, $p = 0.981$). Apparently, the small nonsignificant difference in the total score seems to result from the items with only one representation involved, which were slightly better solved in the CF condition.

However, nonsignificant findings do not warrant the conclusion that there are no differences between conditions. To examine whether the posttest scores imply that the conditions were equally effective, we conducted an equivalence test (Lakens et al., 2018) with $d = 0.50$ as our smallest effect size of interest (SESOI). Consequently, if the 90% confidence interval of the difference between the CF and SIM conditions lies fully in the equivalence interval with bounds $\pm d \cdot SD$ (SD = pooled standard deviation), the two conditions are considered statistically equivalent. In our sample, the pooled standard deviation (calculated from the standard error, see Table 3) was $SD = 3.02$ in the posttest score, and the resulting equivalence interval was [−1.51; 1.51]. The 90% confidence interval for the adjusted mean difference listed in Table 3 lies entirely in the specified range. Thus, the performance of both conditions in the posttest was equivalent with respect to the SESOI of $d = 0.5$. Regarding the subscales, the equivalence interval bounds were ±0.90, ±0.48, and ±0.66 for single representation, quantitative calculations, and translation items, respectively. Hence, the two conditions were equivalently effective regarding the subscales as well. Note that the subscales including quantitative calculations and translations only comprised three and four items, respectively. The equivalence interval bounds therefore included about 1/3 of the whole scales.

TABLE 2 Means, standard deviations, medians, and results from the Mann–Whitney *U* test for the pretest sum score as well as flux items, math items, and field line item subscores.

		CF ($N = 98$)		SIM ($N = 89$)				
Pretest	**Max. Score**	***M* (SD)**	***Md***	**M (SD)**	***Md***	***W***	***p***	**Rank-biserial correlation [95% CI]**
Total	25	20.76 (3.26)	21	20.47 (3.49)	21	4518	0.669	0.036 [−0.129; 0.200]
Flux items	18	15.22 (2.19)	16	15.08 (2.48)	15	4399	0.918	0.009 [−0.156; 0.173]
Math items	6	4.62 (1.51)	5	4.45 (1.54)	5	4664	0.397	0.069 [−0.096; 0.231]
Field line item	1	0.91 (0.29)	1	0.94 (0.23)	1	4205	0.358	−0.036 [−0.199; 0.130]

Abbreviations: CF, concreteness fading; SIM, simultaneous presentation.

TABLE 3 Adjusted means, standard errors, results from the analysis of covariance, effect sizes (partial eta squared), and 90% CI for the adjusted mean difference as specification of the equivalence range.

		CF ($N = 98$)	SIM ($N = 89$)				
Multiple-choice posttest	**Max. Score**	***M* (SE) (adjusted)**	***M* (SE) (adjusted)**	***F* (1,184)**	***p***	**η^2**	**90% CI for adj. mean difference**
Total	16	9.08 (0.31)	8.56 (0.32)	1.372	0.243	0.004	[−0.21; 1.25]
Single	9	5.49 (0.18)	5.11 (0.19)	2.146	0.145	0.007	[−0.05; 0.82]
Quant. Cal.	3	1.24 (0.10)	1.31 (0.10)	0.185	0.667	0.001	[−0.30; 0.17]
Translation	4	2.27 (0.13)	2.24 (0.14)	0.023	0.879	0.000	[−0.29; 0.35]

Note: The subscales refer to the sets of items that make use of one representation (Single), include quantitative calculations (Quant. Cal.), or require a translation between representations (Translation).
Abbreviations: CF, concreteness fading; SIM, simultaneous presentation.

7.2.2 | Open questions

In two open questions, students had, first, to explain whether a statement on induction based on a common misconception ("without movement, there is no voltage") was true and, second, to analyze a situation with nonmoving components (electromagnet and coil). Student answers were coded by two researchers, one with a PhD in physics education and the other in theoretical physics, and physics teacher qualifications. We followed the guidelines of O'Connor and Joffe (2020) for qualitative research analysis. First, a coding frame with three categories (Miscon, Unclear, Correct) was developed. Student answers that identified the statement as correct and confirmed that movement is necessary for the induction of a voltage were categorized as "Miscon" (evidence for the *miscon*ception). The category "Unclear" was used when students said the statement was incorrect but gave a wrong or no explanation. Rejection of the statement together with a correct explanation was assigned to "Correct." The two researchers independently coded all answers accordingly. Krippendorff's (1980) α for the intercoder reliability (and the percentages of agreement) regarding the two student tasks were 0.98 (99%) and 0.89 (96%), respectively, indicating a very high level of agreement. A consensus could be found in discussing the four instances of disagreement. Table 4 shows the frequencies of coded student answers to the two open questions. Most students (54% in the CF condition, 60% in the SIM condition) agreed with the statement that movement was necessary to induce a voltage in a coil. This statement might be regarded as correct when considering only the experiments students conducted in the learning phase since all experiments with induced voltages involved movement. However, it cannot be generalized because a voltage can also be induced by a changing magnetic field without macroscopic movement of a component. Nevertheless, approximately 40% of the students in both conditions who stated that movement was necessary for induction answered correctly when exposed to the new situation with a resting electromagnet and coil. Overall, 80% of the students correctly answered this transfer question.

7.3 | Effect of condition on the use of representations

The following sections explore how students were able to translate between representations and how they used and valued the usefulness of different representations.

7.3.1 | Translation of representations

To explore the effects of the intervention on students' ability to translate between representations, we analyzed the six multiple-choice translation items on magnetic flux (see Table 1) that were applied both in the pre- and posttest. As mentioned before, these items were specifically designed to test students' ability to translate between representations. We performed an ANCOVA with the posttest sum score of the six translation items on magnetic flux as the dependent variable, condition (CF vs. SIM) as the independent variable and pretest sum score of the same items as the covariate. We again checked whether the assumptions for ANCOVA were met. There was a violation of normality regarding the residuals of the variables. However, ANCOVA is quite robust against such violations (Knief & Forstmeier, 2021), and as all other assumptions were met, we carried out the analysis. We did not find a significant difference between conditions, CF: $M(\text{adj.}) = 4.56$, $SE = 0.12$, SIM: $M(\text{adj.}) = 4.66$, $SE = 0.13$, $F(1,184) = 0.318$, $p = 0.573$. The effect size between conditions was close to zero, $d = -0.08$, with a 95% confidence interval [−0.37; 0.21]. The means of the gain in the sum scores between the pre- and the posttest were $M = -0.06$ in the CF condition and $M = 0.07$ in the SIM condition, respectively. Thus, students' ability to translate between representations regarding magnetic flux did not change during the intervention; please note that this ability was not the focus of the intervention.

7.3.2 | Spontaneous use of representations

At the end of the learning phase, students were prompted to explain Faraday's law. To explore which representations they spontaneously used and to look for evidence of integrating representations, their explanations were analyzed by means of content analysis (Krippendorff, 1980; O'Connor & Joffe, 2020) as described above. We categorized students' answers with respect to the representation they referred to in their answer. Answers were assigned to the concrete representation when they included words referring to the hands-on experiments (e.g., "magnet," "voltmeter," "movement"). The intermediate representation was characterized by the use of the term "field lines." We assigned answers to the idealized representation category when students used symbols for physical quantities, formulas, mentioned graphs or terms such as "derivative" to describe electromagnetic induction. Additionally, statements of a quantitative nature such as "the induced voltage is *proportional* to the change of magnetic flux" were interpreted as indicating a reference to the idealized representation. A further category, "Generic," encompasses explanations that did not include representation-specific terms as described above and were thus regarded as representation-independent statements (e.g., "Voltage is induced when the magnetic flux changes.") The category "Multiple Reps" comprises explanations that included the use of more than one representation. Giving a generic statement in combination with the use of a representation was assigned to the category of the respective representation, not to "Multiple Reps" or "Generic." Therefore, "Generic" also indicates the absence of the explicit use of a representation. In some cases, the representations were explicitly linked in the explanations. We listed these cases additionally in the category "Integration," which was consequently a subset of "Multiple Reps." Finally, the correctness of the answers was rated in a dichotomous way (correct or incorrect).

The coding of the student answers according to (1) the five main categories (Generic, Concrete, Intermediate, Idealized, or Multiple Reps), (2) the subcategory "Integration" and (3) their

TABLE 4 Categorization (frequencies) of students' answers on the open questions on Faraday's law in both conditions.

CF ($N = 39$)		Question 2			SIM ($N = 35$)		Question 2		
		Unclear	Miscon	Correct			Unclear	Miscon	Correct
Question 1	Unclear	2	0	2	Question 1	Unclear	1	0	3
	Miscon	2	4	15		Miscon	3	3	15
	Correct	1	0	13		Correct	0	0	10

Note: Unclear = ambiguous answers, Miscon = evidence for the misconception, Correct = correct explanation. Note that these questions were only answered by half of the classes.
Abbreviations: CF, concreteness fading; SIM, simultaneous presentation.

TABLE 5 Categorization (frequencies and percentages) of students' answers regarding the representation they used in explaining Faraday's law.

Category of representations	CF ($N = 83$)	SIM ($N = 76$)
Generic	19 (23%)	13 (17%)
Single rep	46 (55%)	50 (66%)
Concrete	7 (8%)	23 (30%)
Intermediate	3 (4%)	12 (16%)
Idealized	36 (43%)	15 (20%)
Multiple reps	18 (22%)	13 (17%)
Integration	5 (6%)	8 (11%)
Correct	65 (78%)	52 (68%)

Note: *N* denotes the number of students who answered the question. Multiple Reps refers to the use of more than one of the single representations. Every answer was assigned to only one of the categories Generic, Concrete, Intermediate, Idealized, or Multiple Reps. Integration is a subcategory of Multiple Reps and was assigned when representations were explicitly linked.

correctness was carried out independently by the same two researchers as in the coding described above. Krippendorff's α for the intercoder reliability (and the percentages of agreement) for the three categorizing tasks were 0.90 (93%), 0.72 (96%), 0.71 (89%) and indicated sufficient to good levels of agreement. The slightly lower intercoder reliability regarding the second task was due to the very few instances of integration. Thus, discrepancies in the coding had a strong impact on the reliability. The definition of correctness was sharpened for a second round of coding. Coders agreed upon the following: answers that contained one correct and one clearly false argument were considered wrong; answers with one correct and one ambiguous statement were categorized as correct; incomplete statements such as “you need magnetic flux to have a voltage” were rated as wrong; and the few remaining unmatched cases in all three coding tasks were finally discussed and categorized through agreement.

Table 5 shows the frequencies of student answers regarding the established categories separately for the two conditions. The percentages indicate the total number of students in each condition who gave a written answer to the question. In total, 28 students (CF: 15, SIM: 13) left the answer field empty. To contrast the use of multiple representations with the use of single representations, a category “Single Rep” is introduced, which simply summarizes the categories “Concrete,” “Intermediate,” and “Idealized.”

Fisher's exact test indicated that the frequencies of the categories “Generic,” “Single Rep” and “Multiple Reps” did not differ between the groups ($p = 0.41$). Generic answers were given slightly more frequently in the CF condition, while students in the SIM condition more often used single representations in their answers. However, within the single representation categories, there were significant differences. Idealized representations were more frequently used in the CF condition ($p < 0.001$). In the SIM condition, the distribution over the representations was quite balanced. Compared to the CF, the concrete representation was more often used in the SIM condition ($p = 0.002$). While the groups had similar percentages of cases with multiple representations, explicit links between representations were more prevalent in the SIM condition (8 out of 13 cases, 62%) than in the CF condition (5 out of 18 cases, 28%). However, these differences were not significant ($p = 0.08$). In the CF condition, representations were mostly listed without references between them. However, spontaneous integration of multiple

TABLE 6 Means and standard deviations for students' ratings of the usefulness of representations.

Representation	CF ($N = 36$) M (SD)	SIM ($N = 29$) M (SD)
Concrete (experiment)	2.4 (1.4)	3.1 (1.4)
Intermediate (field lines)	3.4 (1.3)	3.9 (1.1)
Idealized (graph or formula)	2.9 (1.2)	2.5 (1.1)

representations was rather rare in both groups. Considering the correctness, we again ran Fisher's exact test, which indicated no significant difference between conditions ($p = 0.21$).

7.3.3 | Usefulness rating

In a final task, students rated how useful the representations were for solving the transfer item in which a voltage is induced by an electromagnet. The Likert scale ranged from 1 = "not useful" over 3 = "indifferent" to 5 = "very useful." We obtained complete ratings from 36 students in the CF condition and 29 students in the SIM condition (as explained in the Posttest section above, this task was only presented to half of the classes). Table 6 shows the mean values and standard deviations regarding the three representations. Although Likert scale data may not strictly fulfill the requirements for an analysis of variance (ANOVA), it has been argued that this method is robust against violations and preferable to nonparametric methods (Carifio & Perla, 2008; Norman, 2010). We therefore performed ANOVAs to compare the distributions regarding the various representations within conditions and regarding the same representation between conditions. A comparison of the mean values shows that in both conditions, the intermediate representation, which was rarely used in the open answers previously, was considered the most helpful representation for solving the task. Additionally, only for this representation, the students in both conditions most frequently selected "very useful" (CF: 10 students, SIM: 11 students). In the CF condition, the intermediate representation was rated significantly higher than the concrete representation, $F(1,70) = 9.91$, $p = 0.002$. In the SIM condition, the intermediate representation was regarded as significantly more useful than both other representations; $F(1,56) = 6.14$, $p = 0.016$ (intermediate vs. concrete), $F(1,56) = 22.6$, $p < 0.001$ (intermediate vs. idealized). No other comparisons were significant at the 0.05 level. The usefulness of the concrete representation was rated considerably higher in the SIM than in the CF condition, $F(1,63) = 3.89$, $p = 0.053$. In contrast, the idealized representation was regarded slightly more helpful in the CF than in the SIM condition, $F(1,63) = 2.36$, $p = 0.13$. However, both differences were not significant.

8 | DISCUSSION

In an experimental classroom study, we examined whether concreteness fading (i.e., starting with concrete representations) was more beneficial than the simultaneous presentation of multiple external representations (MERs) in learning Faraday's law. We specifically addressed a hypothesis put forward by Fyfe and Nathan (2019, Hypothesis 5, p. 417) according to which "presenting the three stages one at a time in the specified order will be more effective than presenting the stages all at once or presenting any two of the three simultaneously." An ANCOVA did not indicate significant differences between the CF and SIM conditions in the posttest sum scores of the multiple-choice concept test. Additionally, the subscores (items with single representations, quantitative calculations, and translation between representations) did not reveal any statistically significant differences. Equivalence tests indicated that both conditions performed similarly. Content analysis of two open questions also gave no hints of clear differences between the two conditions. A comparable majority of students from both conditions were able to solve a transfer task, indicating that they were able to build solid conceptual knowledge with respect to Faraday's law.

On the one hand, these results challenge the proposal that concreteness fading is generally more beneficial than other ways of sequencing MERs (Fyfe & Nathan, 2019). Although concreteness fading has received promising evidence from mathematics education research, the results from research on physics education have been equivocal. The results of the present study align with previous findings that did not support concreteness fading in physics (Jaakkola & Veermans, 2018; Kokkonen et al., 2022). On the other hand, the theoretical and empirical literature on learning with MERs underlines that creating connections across MERs is crucial for learning (Goldwater & Gentner, 2015; Johnson & Mayer, 2012; Mason et al., 2013; Son et al., 2011). Simultaneous presentation of MERs creates spatial and temporal contiguity (Mayer, 2009; Mayer & Fiorella, 2014) and can prompt comparing and contrasting, and thus processes of structural alignment, which should support connecting across instances (for a review, see, Alfieri et al., 2013). The supposed benefits of simultaneous presentation of MERs did, however, not result in better learning than concreteness fading in the present study. Both conditions (CF and SIM) were equally effective.

In addition to supporting conceptual knowledge development, a specific aspect of learning with MERs is to learn to link various external representations and translate between them. To explore whether either of the approaches was more beneficial to this end, we used multiple-choice items regarding magnetic flux that required translating across two representations in the pre- and posttest. However, we also did not find a significant difference regarding representation translation. Students' performance on these items was already high after the pre-instruction on magnetic flux and it remained stable during the intervention in both conditions. It might be that students really did not spontaneously increase their ability to translate between representation, notably, increasing this ability specifically was not in the focus of the intervention. However, this finding could also be due to a ceiling effect as there was not much room for improvement regarding the selected items since the pretest performance was already high.

To explore the use of representations in more detail and to identify idiosyncratic advantages of concreteness fading and simultaneous presentation, we analyzed students' answers to the self-explanation prompts presented at the end of the intervention. First, we found weak evidence that students might better link representations spontaneously when they are presented simultaneously, as integration was more frequently observed in the SIM condition. However, the numbers of students who indicated integration in their answers to the self-explanation prompts were low in both groups. Second, students in the CF condition mentioned the idealized representation more often. This finding could be regarded as one of the proposed benefits of concreteness fading (e.g., Fyfe & Nathan, 2019, p. 419): "Concreteness fading is intended to facilitate initial learning by starting with a grounded, meaningful representation and to support transfer by moving toward competency with a decontextualized, idealized representation." However, we did not find a benefit for transfer. Thus, two alternative explanations should be considered for why students in the CF condition more frequently mentioned the idealized representation. On the one hand, this finding could be considered a simple recency effect (e.g., Baddeley & Hitch, 1993), given that the idealized representation is presented in the last step (before learning outcomes are assessed). On the other hand, students might reckon it as the most important representation because the progression ends with it, which could give them the impression that it is the target of the learning sequence. Both explanations could be tested with a concreteness introduction condition in which the sequence starts with the idealized representation and ends with the concrete experiments. If we would find a similar distribution also in the latter condition, the general emphasis on idealized representations (e.g., formulas) in advanced physics courses might be an alternate explanation. In contrast, in the SIM condition, the use of the various representations seemed more balanced. Representations might be regarded as equally important when presented simultaneously with spatial and temporal contiguity. The slightly higher frequency of the concrete representation may stem from the enactment component of the hands-on experiment, which can facilitate retention (e.g., Carbonneau et al., 2013). This tentative finding could be an argument in favor of this condition in the longer run, as it might support enduring integration.

Interestingly, we found a similar tendency when we analyzed student ratings of the usefulness of representation with respect to the selected task. Students in the CF condition seemed to favor the idealized representation over the concrete representation, while it was the opposite in the SIM condition. Nevertheless, both groups rated the intermediate representation the most helpful. From a physics education expert's perspective, this representation can indeed be seen as the most effective for solving the given task. Thus, even though students did not preferably use the intermediate representation spontaneously when describing Faraday's law, students were able to flexibly adapt to the new task in a meaningful way. Our finding indicates that the usefulness of a given representation is likely to depend on the task at hand, as they connect to complementary aspects of the phenomenon (Lampinen & McClelland, 2018). While it might be convenient to cite Faraday's law by writing down a generic statement, remembering an experiment or reproducing the formula, the field line representation was perceived as more useful for solving the selected task. Obviously, for a purely quantitative task, the mathematical representation would have been more helpful than other representations such as the field line representation. However, also the pre-instruction, where the magnetic flux was elaborated mainly using the intermediate representation, might have affected students' perceived usefulness. Finally, the perception of how helpful a representation is may also depend on the learners' prior knowledge (Lampinen & McClelland, 2018): learners who are more familiar with mathematical formalisms may be more adept in using it even in qualitative tasks. Note that this finding of the present study is exploratory; nevertheless, it indicates interesting directions for future research. That is, it could be worthwhile to put the intermediate representation in the focus of an instructional approach for Faraday's law. At least in science education, the intermediate representation often provides a potentially helpful bridge between concrete representations (e.g., hands-on experiments) and symbolic (mathematical) formulations of a concept.

As a last point, we refer to a limitation of the experimental materials, which provides interesting aspects for physics educators. For practical and didactical reasons, we used a simple experimental setup consisting of a bar magnet, a coil, and a voltmeter that does not require electrical supply or prior knowledge of electromagnets. In this setup, a voltage can only be induced when either the magnet or the coil is moved. However, magnetic induction can also occur without movement, for example, when a coil is put into a changing magnetic field. Hence, the limitation of our experimental material may have induced and solidified the misconception that induction requires movement (Jelicic et al., 2017). Fortunately, this was not the case, as the analysis of the open questions in the posttest showed: approximately 80% of the students in both conditions answered correctly when exposed to a new situation with a resting electromagnet. More interestingly, approximately 40% of the students who first said movement was necessary for induction were able to correctly explain the new situation. While the first item implies that the intervention may have produced the misconception that electromagnetic induction requires movement, the transfer item indicates that the misconception was not solidified. On the one hand, these inconsistent answer patterns of students might be a hint that the newly built knowledge was not yet well enough integrated but remained situated and thus dependent on the tasks and context. The first question made a possible overgeneralization from the experiments they conducted in the learning phase visible, while the second question exposed them to a new situation without movement; and students were able to correctly apply Faraday's law. On the other hand, it might also be possible that students could not spontaneously come up with an example of induction without movement and later, after answering the second question, they just did not go back to correct the first answer. In any case, the question about the resting electromagnet seems to have triggered a deeper examination of Faraday's law among students. For future use in the classroom, it would therefore make sense to include this case in the teaching materials, not just in the posttest. Considering all student results, our findings indicate that the relatively simple experimental setup together with the self-learning materials was suitable to help students build a solid basis of conceptual knowledge of Faraday's law. To make it accessible to a wide range of schools, we provide the self-learning materials used in the study (in German) in the Supporting Information.

8.1 | Limitations

We found no differences between the CF and SIM conditions immediately after the learning phase. While our measures indicated a high implementation fidelity with respect to the study design, our study has some limitations that should be targeted by future research. First, whether the benefits would materialize in a delayed posttest, as speculated by McNeil and Fyfe (2012), was not examined. Since we situated our study within normal physics classroom education, it was not possible to tell the teachers not to touch upon the content for one or even more weeks because they wanted to proceed with their instruction. Second, based on methodological considerations, we used different pre- and posttests (as described in the Pretest section above). However, this choice also comes with the disadvantage that we cannot exactly say how much learning actually occurred from pre- to posttest. For future studies, it will be helpful to further optimize our methodology so that pre-posttest gains can be determined. Third, there is a possible confounding variable in our study design. In the CF condition, the task types were presented simultaneously (i.e., each phase covered all four situations). In contrast, in the simultaneous condition, each step covered only one task type. This might have benefitted the CF condition since the tasks thus scaffolded generalizations about Faraday's law (therefore, concreteness fading may have yielded worse performance without this helpful learning material design feature). Fourth, considering the qualitative analysis of students' open answers, the interpretations must be treated with caution. When categorizing students' answers, we were restricted to their written texts. However, students might have thought about the other representations while mentioning only one of them or giving a generic statement in their responses. Finally, representations and content are always confounded. For example, in our study, Faraday's law was represented in different ways. In the experiments, the induced voltage produced by changing flux was illustrated with coils, magnets and voltmeters, the field line representations focused on the visualization of the change in flux that dynamically is linked to the induced voltage, and the graphs explicitly showed how the function of induced voltage over time can be derived from the flux. All three representations can illustrate the relation between magnetic flux and induced voltage, which is at the core of Faraday's law; that is, they describe the same content. However, they focus on different aspects and support different functions and therefore complement each other. This confounding of representations and content is not restricted to physics. For example, rational numbers can be represented as fractions, decimals, or percentages. All three representations can express the diverse interpretations of rational numbers (e.g., Tian & Siegler, 2018). Nevertheless, "decimals appear to be much more effective than fractions in conveying information about magnitude" (DeWolf et al., 2015, p. 128). Kokkonen and Schalk (2021) provide an extensive discussion on how representations and their relations to each other vary across mathematics, physics, chemistry, and biology. Accordingly, the confounding of representations and content is likely to vary between different topics within a subject, but also across subjects. It would be highly interesting to investigate in future research how the varying degree of confounding influences the effectiveness of different sequences of representations.

9 | CONCLUSION

We designed and applied two self-learning materials to introduce the challenging Faraday's law in physics classrooms at the upper secondary level. Both materials comprised multiple external representations (hands-on experiments, diagrammatic illustrations, mathematical formalisms) but presented them either sequentially following the concreteness fading approach or simultaneously. Our results from an experimental classroom study with a detailed posttest showed that the two presentation orders were equally effective. Future studies need to more closely examine the interaction between MERs and learning tasks regarding the triggering of effective learning processes and to scrutinize how students' selection and use of representations across various tasks affects the construction and development of conceptual knowledge. Moreover, the degree of confounding between representation and content needs to be carefully considered in future research since this may additionally impact the effectiveness of instructional sequences of representations. The lack of differences between the conditions in the present study implies that there are multiple ways for teachers to sequence MERs to teach Faraday's law in physics classrooms. Thus, supporting students in learning a challenging physics topic may not be only a matter of finding the best sequence of presenting MERs.

ACKNOWLEDGMENTS

We would like to thank Dr. Barbara Gränz for her help in coding the student answers. We would also like to express our appreciation to all the teachers and students who participated in our project. Open access funding provided by Eidgenossische Technische Hochschule Zurich.

ORCID

Andreas Lichtenberger https://orcid.org/0000-0001-7739-3904
Tommi Kokkonen https://orcid.org/0000-0003-3324-2690
Lennart Schalk https://orcid.org/0000-0002-7144-1431

REFERENCES

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2–3), 131–152. https://doi.org/10.1016/S0360-1315(99)00029-9

Ainsworth, S. (2006). Deft: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and practice in science education* (Vol. 3). Springer. https://doi.org/10.1007/978-1-4020-5267-5_9

Ainsworth, S. (2014). The multiple representation principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 464–486). Cambridge University Press. https://doi.org/10.1017/cbo9781139547369.024

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87–113. https://doi.org/10.1080/00461520.2013.775712

Baddeley, A. D., & Hitch, G. (1993). The recency effect: Implicit learning with explicit retrieval? *Memory & Cognition*, *21*, 146–155. https://doi.org/10.3758/BF03202726

Bagno, E., & Eylon, B.-S. (1997). From problem solving to a knowledge structure: An example from the domain of electromagnetism. *American Journal of Physics*, *65*(8), 726–736. https://doi.org/10.1119/1.18642

Becker, S., Klein, P., Gößling, A., & Kuhn, J. (2020). Using mobile devices to enhance inquiry-based learning processes. *Learning and Instruction*, *69*, 101350. https://doi.org/10.1016/j.learninstruc.2020.101350

Beichner, R. J. (1990). The effect of simultaneous motion presentation and graph generation in a kinematics lab. *Journal of Research in Science Teaching*, *27*(8), 803–815. https://doi.org/10.1002/tea.3660270809

Beichner, R. J. (1996). The impact of video motion analysis on kinematics graph interpretation skills. *American Journal of Physics*, *64*(10), 1272–1277. https://doi.org/10.1119/1.18390

Belenky, D. M., & Schalk, L. (2014). The effects of idealized and grounded materials on learning, transfer, and interest: An organizing framework for categorizing external knowledge representations. *Educational Psychology Review*, *26*(1), 27–50. https://doi.org/10.1007/s10648-014-9251-9

Börlin, J. (2012). *Das Experiment als Lerngelegenheit: Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität* [Doctoral Dissertation, Universität Basel]. Berlin, Germany.

Bouck, E. C., Satsangi, R., & Park, J. (2017). The concrete–representational–abstract approach for students with learning disabilities: An evidence-based practice synthesis. *Remedial and Special Education*, *39*(4), 211–228. https://doi.org/10.1177/0741932517721712

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching*, *24*(4), 385–395. https://doi.org/10.1002/tea.3660240409

Bruner, J. S. (1966). *Toward a theory of instruction*. Harvard University Press.

Brungardt, J. B., & Zollmann, D. (1995). Influence of interactive videodisc instruction using simultaneous-time analysis on kinematics graphing skills of high school physics students. *Journal of Research in Science Teaching*, *32*(8), 855–869. https://doi.org/10.1002/tea.3660320808

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, *105*(2), 380–400. https://doi.org/10.1037/a0031084

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing likert scales. *Medical Education*, *42*(12), 1150–1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Cattell, R. B. (1966). Scree test for number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477. https://doi.org/10.1016/0364-0213(94)90016-7

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47*(3), 177–188. https://doi.org/10.1080/00461520.2012.695709

Chi, M. T. H., & Wylie, R. (2014). The icap framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Ching, B. H.-H., & Wu, X. (2019). Concreteness fading fosters children's understanding of the inversion concept in addition and subtraction. *Learning and Instruction*, *61*, 148–159. https://doi.org/10.1016/j.learninstruc.2018.10.006

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal of Research in Mathematics Education*, *23*(1), 2–33. https://doi.org/10.2307/749161

Corradi, D., Elen, J., & Clarebout, G. (2012). Understanding and enhancing the use of multiple external representations in chemistry education. *Journal of Science Education and Technology*, *21*(6), 780–795. https://doi.org/10.1007/s10956-012-9366-z

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). Conceptual structure and the procedural affordances of rational numbers: Relational reasoning with fractions and decimals. *Journal of Experimental Psychology: General*, *144*(1), 127–150. https://doi.org/10.1037/xge0000034

Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics - Physics Education Research*, *5*(2), 020103. https://doi.org/10.1103/PhysRevSTPER.5.020103

Donovan, A. M., & Fyfe, E. R. (2022). Connecting concrete objects and abstract symbols promotes children's place value knowledge. *Educational Psychology*, *42*(8), 1008–1026. https://doi.org/10.1080/01443410.2022.2077915

Driver, R. (1983). *The pupil as scientist?* Open University Press.

Erfmann, C., & Berger, R. (2015). Ein elementarer zugang zur induktion. *Praxis der Naturwissenschaften - Physik in der Schule*, *64*(2), 13–25.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using g*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fredlund, T., Linder, C., Airey, J., & Linder, A. (2014). Unpacking physics representations: Towards an appreciation of disciplinary affordance. *Physical Review Special Topics - Physics Education Research*, *10*(2), 020129. https://doi.org/10.1103/PhysRevSTPER.10.020129

Fyfe, E. R., McNeil, N. M., & Borjas, S. (2015). Benefits of "concreteness fading" for children's mathematics understanding. *Learning and Instruction*, *35*, 104–120. https://doi.org/10.1016/j.learninstruc.2014.10.004

Fyfe, E. R., & Nathan, M. J. (2019). Making "concreteness fading" more concrete as a theory of instruction for promoting transfer. *Educational Review*, *71*(4), 403–422. https://doi.org/10.1080/00131911.2018.1424116

Gadgil, S., & Nokes, T. J. (2009). Analogical scaffolding in collaborative learning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *31*, 3115–3120. https://escholarship.org/uc/item/5b74x3jv

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34*(5), 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511–525. https://doi.org/10.1016/j.learninstruc.2006.10.001

Girwidz, R., Theyßen, H., & Widenhorn, R. (2021). Experiments in physics teaching. In H. E. Fischer & R. Girwidz (Eds.), *Physics education* (pp. 269–296). Springer. https://doi.org/10.1007/978-3-030-87391-2_10

Goldwater, M. B., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition*, *137*, 137–153. https://doi.org/10.1016/j.cognition.2014.12.001

Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition*, *118*(3), 359–376. https://doi.org/10.1016/j.cognition.2010.10.009

Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin*, *142*(7), 729–757. https://doi.org/10.1037/bul0000043

Gunstone, R. F. (1991). Reconstructing theory from practical experience. In B. Woolnough (Ed.), *Practical science* (pp. 67–77). Open University Press.

Hodson, D. (1993). Re-thinking old ways: Towards a more critical approach to practical work in school science. *Studies in Science Education*, *22*(1), 85–142. https://doi.org/10.1080/03057269308560022

Hong, J.-C., Hsiao, H.-S., Chen, P.-H., Lu, C.-C., Tai, K.-H., & Tsai, C.-R. (2021). Critical attitude and ability associated with students' self-confidence and attitude toward "predict-observe-explain" online science inquiry learning. *Computers & Education*, *166*, 104172. https://doi.org/10.1016/j.compedu.2021.104172

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185.

Jaakkola, T., & Veermans, K. (2018). Exploring the effects of concreteness fading across grades in elementary school science education. *Instructional Science*, *46*(2), 185–207. https://doi.org/10.1007/s11251-017-9428-y

Jelicic, K., Planinic, M., & Planinsic, G. (2017). Analyzing high school students' reasoning about electromagnetic induction. *Physical Review Physics Education Research*, *13*(1), 010112. https://doi.org/10.1103/PhysRevPhysEducRes.13.010112

Johnson, A. M., Reisslein, J., & Reisslein, M. (2014). Representation sequencing in computer-based engineering education. *Computers & Education*, *72*, 249–261. https://doi.org/10.1016/j.compedu.2013.11.010

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *18*(2), 178–191. https://doi.org/10.1037/a0026923

Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, *53*(6), 2576–2590. https://doi.org/10.3758/s13428-021-01587-5

Kohl, P. B., & Finkelstein, N. D. (2005). Student representational competence and self-assessment when solving physics problems. *Physical Review Special Topics - Physics Education Research*, *1*(1), 010104. https://doi.org/10.1103/PhysRevSTPER.1.010104

Kokkonen, T., Lichtenberger, A., & Schalk, L. (2022). Concreteness fading in learning secondary school physics concepts. *Learning and Instruction*, *77*, 101524. https://doi.org/10.1016/j.learninstruc.2021.101524

Kokkonen, T., & Schalk, L. (2021). One instructional sequence fits all? A conceptual analysis of the applicability of concreteness fading in mathematics, physics, chemistry, and biology education. *Educational Psychology Review*, *33*, 797–821. https://doi.org/10.1007/s10648-020-09581-7

Kozma, R. B., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Ed.), *Visualization in science education* (Vol. 1, pp. 121–145). Springer. https://doi.org/10.1007/1-4020-3613-2_8

Kozma, R. B., Russell, J., Jones, T., Marx, N., & Davis, J. (1996). The use of multiple, linked representations to faciiitate science understanding. In S. Vosniadou, E. de Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the design of technology-supported learning environments* (pp. 41–60). Routledge. https://doi.org/10.4324/9780203053386

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Sage.

Küchemann, S., Malone, S., Edelsbrunner, P., Lichtenberger, A., Stern, E., Schumacher, R., Brünken, R., Vaterlaus, A., & Kuhn, J. (2021). Inventory for the assessment of representational competence of vector fields. *Physical Review Physics Education Research*, *17*(2), 020126. https://doi.org/10.1103/PhysRevPhysEducRes.17.020126

Kurtz, K. J., Miao, C.-H., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences*, *10*(4), 417–446. https://www.jstor.org/stable/1466739

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Lampinen, A. K., & McClelland, J. L. (2018). Different presentations of a mathematical concept can support learning in complementary ways. *Journal of Educational Psychology*, *110*(5), 664–682. https://doi.org/10.1037/edu0000235

Lichtenberger, A., Wagner, C., Hofer, S. I., Stern, E., & Vaterlaus, A. (2017). Validation and structural analysis of the kinematics concept test. *Physical Review Physics Education Research*, *13*(1), 010115. https://doi.org/10.1103/PhysRevPhysEducRes.13.010115

Malone, S., Altmeyer, K., Vogel, M., & Brünken, R. (2020). Homogeneous and heterogeneous multiple representations in equation-solving problems: An eye-tracking study. *Journal of Computer Assisted Learning*, *36*(6), 781–798. https://doi.org/10.1111/jcal.12426

Maloney, D. P., O'Kuma, T. L., Hieggelke, C. J., & Van Heuvelen, A. (2001). Surveying students' conceptual knowledge of electricity and magnetism. *American Journal of Physics*, *69*, S12–S23. https://doi.org/10.1119/1.1371296

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science*, *29*(4), 587–625. https://doi.org/10.1207/s15516709cog0000_15

Mason, L. (2004). Fostering understanding by structural alignment as a route to analogical learning. *Instructional Science*, *32*, 293–318. https://doi.org/10.1023/B:TRUC.0000026512.88700.32

Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education*, *81*(3), 356–384. https://doi.org/10.1080/00220973.2012.727885

Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press. https://doi.org/10.1017/CBO9780511811678.011

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In *The cambridge handbook of multimedia learning* (pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). Cambridge University Press.

McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, *55*(6), 503–513. https://doi.org/10.1119/1.15104

McNeil, N. M., & Fyfe, E. R. (2012). "Concreteness fading" promotes transfer of mathematical knowledge. *Learning and Instruction*, *22*(6), 440–448. https://doi.org/10.1016/j.learninstruc.2012.05.001

Müller, A., Hettmannsperger, R., Scheid, J., & Schnotz, W. (2017). Representational competence, understanding of experiments, phenomena and basic concepts in geometrical optics: A representational approach. In D. F. Treagust, R. Duit, & H. E. Fischer (Eds.), *Multiple representations in physics education* (Vol. 10, pp. 209–229). Springer. https://doi.org/10.1007/978-3-319-58914-5_10

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*, 625–632. https://doi.org/10.1007/s10459-010-9222-y

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, *19*, 160940691989922. https://doi.org/10.1177/1609406919899220

Opfermann, M., Schmeck, A., & Fischer, H. E. (2017). Multiple representations in physics and science education—Why should we use them? In D. F. Treagust, R. Duit, & H. E. Fischer (Eds.), *Multiple representations in physics education* (Vol. 10, pp. 1–22). Springer. https://doi.org/10.1007/978-3-319-58914-5_1

Ottmar, E., & Landy, D. (2016). Concreteness fading of algebraic instruction: Effects on learning. *Journal of the Learning Sciences*, *26*(1), 51–78. https://doi.org/10.1080/10508406.2016.1250212

Özcan, G. E., & Uyanık, G. (2022). The effects of the "predict-observe-explain (poe)" strategy on academic achievement, attitude and retention in science learning. *Journal of Pedagogical Research*, *6*(3), 103–111. https://doi.org/10.33902/JPR.202215535

Planinic, M., Lana, I., & Ana, S. (2013). Comparison of university students' understanding of graphs in different contexts. *Physical Review Special Topics-Physics Education Research*, *9*, 020103. https://doi.org/10.1103/PHYSREVSTPER.9.020103

Rivera Gavidia, L. M., & Marrero Galván, J. J. (2023). Predictions and explanations about scientific situations in a high school context. *International Journal of Science Education*, *45*(2), 144–163. https://doi.org/10.1080/09500693.2022.2153095

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review*, *30*(3), 679–701. https://doi.org/10.1007/s10648-018-9435-9

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2011). Connecting instances to promote children's relational reasoning. *Jorunal of Experimental Child Psychology*, *108*(2), 260–277. https://doi.org/10.1016/j.jecp.2010.08.011

Suh, S., Lee, M., & Law, E. (2020, June 21-24). *How do we design for concreteness fading? Survey, general framework, and design dimensions*. IDC: Interaction Design and Children.

Thornton, R. K., & Sokoloff, D. R. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics*, *58*(9), 858–867. https://doi.org/10.1119/1.16350

Tian, J., & Siegler, R. S. (2018). Which type of rational numbers should students learn first? *Educational Psychology Review*, *30*(2), 351–372. https://doi.org/10.1007/s10648-017-9417-3

Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education*, *24*(4), 357–368. https://doi.org/10.1080/09500690110066485

Treagust, D. F., Duit, R., & Fischer, H. E. (Eds.). (2017). *Multiple representations in physics education* (Vol. 10). Springer. https://doi.org/10.1007/978-3-319-58914-5

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321–327.

White, R., & Gunstone, R. F. (1992). *Probing understanding* (1st ed.). Falmer Press.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179–217. https://doi.org/10.1207/s15516709cog2102_3

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lichtenberger, A., Kokkonen, T., & Schalk, L. (2024). Learning with multiple external representations in physics: Concreteness fading versus simultaneous presentation. *Journal of Research in Science Teaching*, 1–33. https://doi.org/10.1002/tea.21947