PHYSICAL REVIEW PHYSICS EDUCATION RESEARCH **20,** 020104 (2024)

Evaluation of a new teaching-learning sequence on the particulate nature of matter using crystal structures

Florian Budimaier[1,2] and Martin Hopf[1]

[1]University of Vienna, Physics Education Research Group, 1090 Vienna, Austria
[2]University College of Teacher Education, Grenzackerstraße 18, 1100 Vienna, Austria

(Received 26 April 2024; accepted 9 July 2024; published 29 July 2024)

Although there has been extensive research on students' understanding of the particulate nature of matter (PNM), there is still a lack of research on contexts that can be used to teach this challenging topic. In a previous design-based research study, the authors developed a teaching-learning sequence (TLS) on the PNM in the context of crystal structures based on 40 student interviews using the method of probing acceptance. Data suggested that salt and snow crystals form an effective context for learning the concept of emergence and therefore gaining a better understanding of the PNM. To test whether the TLS also promotes students' use of the PNM in a realistic classroom setting, a proof of principle study was conducted. In six eighth-grade classes in Vienna, students' use of the PNM was assessed with a pretest before they were taught the TLS during four lessons. After the intervention, students were given a post-test. Open-ended questions were coded using evaluative qualitative content analysis so that quantitative analysis could be applied. T-tests comparing the means of students' scores on both tests show significant improvements in students' use of the PNM in the post-test. The context of crystal structures seems to be helpful to students, as most of them use the PNM when asked about crystal formation. In addition, in the post-test, students more often accepted the idea of empty space between particles and associated particle motion with temperature. However, when asked about phase changes, most students remained in a continuous conception of matter.

DOI: 10.1103/PhysRevPhysEducRes.20.020104

I. INTRODUCTION

The particulate nature of matter (PNM) is a key concept in science education in many countries. Consequently, it has already been extensively researched. Some of this research has focused on documenting students' difficulties and misconceptions in this field [1–3], while others have developed methods to overcome those difficulties [4,5]. Further research has described students' progressions when learning about the PNM throughout their school careers [6–8]. Nevertheless, there are still aspects of teaching and learning the PNM that require further investigation. Following a meta-analysis on the effectiveness of PNM-based intervention studies, Çalik *et al.* [9] argued that research in this field should provide better contexts to teach the PNM. By better understanding the relationships between the context and the PNM, it might be possible to improve students learning.

In a previous study [10], the authors utilized the context of crystal structures to develop a teaching-learning sequence (TLS) on the PNM. A TLS is designed to enhance teaching and learning by integrating findings from educational research into practice [11]. Research indicates that the implementation of a TLS can be a significant factor in influencing teaching and learning in classrooms [12–14]. There is an ongoing discussion on the development of methodological frameworks to serve as guidelines for the development and evaluation of TLSs [12,15–17]. One such framework, namely design-based research (DBR), was used to develop our TLS on the PNM using crystal structures.

The objective of this study was to ascertain whether the TLS is conducive to students' abilities to apply the PNM. Additionally, we sought to ascertain teachers' perceptions of the TLS, with a view to identifying any potential difficulties they may have encountered in implementing the TLS in their classrooms.

II. THEORETICAL FRAMEWORK

A. Teaching and learning the PNM

Given the significance of the PNM in science education, there is a substantial corpus of theoretical knowledge on

---

teaching and learning the PNM in science education research. Consequently, we will limit our discussion to a few theoretical aspects that are particularly relevant to our work.

Research on students' conceptions of the PNM forms an important basis for understanding how students process new information on that topic. Although most students are familiar with terms like "atom" or "molecule," they tend to view matter as continuous [18–20], as this is more in line with their everyday experiences. When trying to make use of the PNM, students often apply properties of macroscopic matter to atoms and molecules [1,21,22]. For instance, students believe that the volume of atoms increases when an object is heated because the size of the object itself increases [23,24]. This kind of conception is referred to as *hybrid*, as students integrate newly learned information while still applying nonscientific concepts [25]. This also results in students experiencing difficulties with the idea that there is nothing between particles [26]. Although students accept the existence of particles, they tend to believe that these particles are embedded in a continuous substance. For example, they believe that there is always air between the particles because they perceive that everything is surrounded by air in their everyday experience [27].

Another significant area of research concerns different methodological approaches and their suitability for learning the PNM. Computer simulations [28–30], computer animations [5,31,32], or other digital learning environments [4] are frequently employed to illustrate interaction at the particle level. Regarding the connections between macroscopic and submicroscopic levels, approaches based on multiple representations have been investigated in particular [33,34]. Moreover, students should be persuaded of the veracity of the PNM through the observation of experiments [35–37].

In addition to different methodological approaches, various contexts for the introduction of the PNM have already been tested. Commonly encountered are phenomena from thermodynamics such as diffusion [30,38], the behavior of gases [39], and fire and ice [40]. Since a fundamental understanding of the behavior of particles is necessary for (statistical) thermodynamics, this context is particularly closely related to the learning objectives of the PNM. In contrast, other approaches attempt to demonstrate the importance of the PNM by presenting phenomena that cannot be explained without the assumption of atoms. For example, observations of images produced by scanning tunneling microscopes suggest the existence of atoms [41,42]. Natural phenomena, such as the regular geometric shape of salt crystals, can only be explained by assuming the existence of atoms [43–45].

Another relevant aspect of teaching the PNM is the consistent use of terms that connect to scientific language but avoid its complexity and partly historically grown contradictions [46]. The term "particle" is not a clearly defined term in everyday language and can be associated with dust particles or grains of sand in addition to atoms, molecules, and subatomic particles [47]. Another argument against the use of the term "particle" in teaching has been made by Pfundt [48]. According to her, the frequently used approach of mentally dividing matter further and further until indivisible particles remain at a certain point rather leads to the transfer of macroscopic properties to atoms. If, for example, a piece of yellow sulfur is divided into smaller and smaller particles, it is not surprising if students assume that these particles are also yellow [49].

Representations also play an important role in teaching the PNM. Since atoms and molecules are inherently invisible, visualizations are necessary in the classroom to illustrate their fundamental properties. However, the prevalent visualizations of atoms and molecules in the form of small spheres [50] can potentially lead to misconceptions [51]. Wiener *et al.* have therefore proposed the use of typographic representations as an alternative to spherical representations [46]. Their acceptance has already been investigated in the context of the atomic model and the PNM and was received positively by students [52,53].

Additionally, insights into the design of a TLS on the PNM can be found in the field of conceptual change theories. For instance, Chi explains the challenges students face in comprehending the PNM due to misconceptions at the ontological level [54]. A multitude of phenomena associated with the particle model adheres to an emergent ontology, whereby the properties of the system emerge from the interaction of its components. The properties of the system are distinct from the properties of the components. Emergent phenomena occur in physics whenever the number of submicroscopic constituents tends to infinity [55]. This is exemplified by phenomena such as temperature, pressure, heat transport, or diffusion. However, students are usually unfamiliar with emergence, as it does not play a role in their everyday experiences. Therefore, they interpret emergent phenomena with a *direct-causal* schema. For example, students tend to explain the cooling of an object by the release of "hot particles" because this "narrative-like" explanation is more familiar to them [56]. Consequently, Chi *et al.* [38] emphasize that students must first be familiarized with an *emergent ontology* to understand the PNM.

B. Development of a TLS within the framework of design-based research

How can these implications from theory be used in classroom practice? One answer to this question can be found in the development of teaching-learning sequences (TLS). "A TLS is both an interventional research activity and a product, like a traditional curriculum unit package, which includes well-researched teaching–learning activities empirically adapted to student reasoning" [15]. Over the past

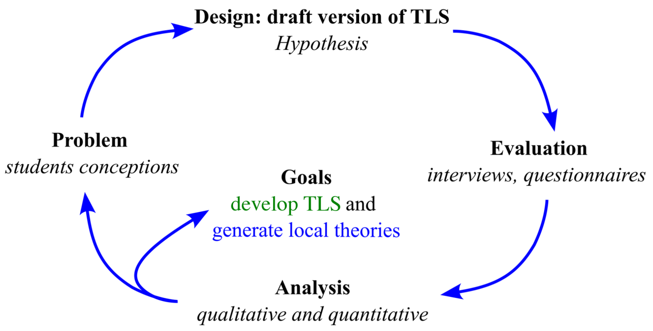


FIG. 1. Process of design-based research (source: own illustration adapted from Ref. [60]).

decades, several frameworks for developing TLSs have been developed [14].

One branch of the studies on TLS design and implementation uses design-based research (DBR) [57]. This methodological framework takes up problems from classroom practice and seeks theory-driven solutions through cyclical testing, evaluation, and further development of learning opportunities [58]. Given its focus on theory and practice [59], the goal of DBR is to develop instructional designs, as well as to generate local theories about teaching and learning the subject matter at hand [60]. Figure 1 shows the cyclical progression of a DBR project.

The concrete translation of theoretical findings into guidelines for classroom practice occurs through design principles that guide the research process [13]. Haagen-Schützenhöfer and Hopf [60] distinguish between general and domain-specific design principles. While the former primarily influences fundamental decisions about teaching and learning, the latter relates specifically to the learning of the subject matter (here: the PNM). Design principles are theory-driven and are continuously developed during the cyclical course of the DBR process.

C. Research question

Despite great efforts in science education research, the PNM is still a difficult topic for students to learn and there is a lack of different contexts for teaching the PNM. Therefore, the authors aimed to develop and evaluate a TLS within the methodological framework of DBR. Several cycles in DBR have already been carried out to develop the instructional design. The present study is dedicated to the evaluation of the instructional design, which leads to the following research question:

Is the developed TLS, based on empirically evaluated design principles, effective in promoting students' use of the PNM?

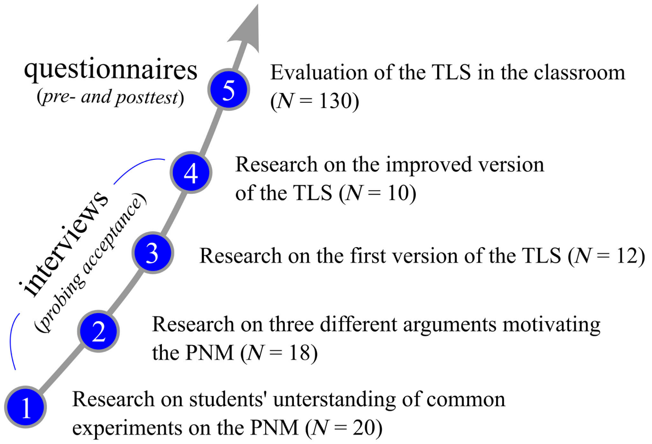


FIG. 2. Sequence of iterations in the DBR project. Results from cycle 1 are published in Ref. [36]. Results from cycles 2 to 4 in Ref. [10]. Cycle 5 is the present study (source: own depiction adapted from Ref. [60]).

III. TLS DESIGN AND EVALUATION

Following the framework of DBR, a TLS on the PNM was developed through multiple iterations of design, evaluation, and analysis. Figure 2 shows the sequence of iterations in the DBR project.

A. Development of the design principles and key ideas

As a result of the first four iterations of the DBR project (see Fig. 2), four domain-specific design principles and seven key ideas were found to improve students' use of the PNM (see Tables I and II). We here describe these only briefly, for a detailed report, see Budimaier and Hopf [10].

The choice of crystal structures as the context for introducing the PNM was based on the idea of Franzbecker and Quast [43]. Results from the DBR project suggested that crystal structures allow students to connect macroscopic and submicroscopic levels of matter [10]. Teaching them the emergent ontology [38] and emphasizing the emergent aspects of the PNM further helped

TABLE I. Domain-specific design principles developed throughout the first four cycles of the DBR project.

Domain-specific design principles
DP 1: Using crystal structures as a starting point for the introduction of the PNM.
DP 2: Use of “building blocks” instead of “particles.”
DP 3: Stressing the emergent aspects of the PNM.
DP 4: Using typographic representations instead of ball-shaped drawings of atoms and molecules.

TABLE II. Key ideas developed throughout the first four cycles of the DBR project.

Key ideas
KI 1: Everything that can be touched is imagined to be composed of very many, small, nonvisible building blocks.
KI 2: The properties of a building block are not the properties of the object.
KI 3: The bonding of the building blocks determines the properties of the object.
KI 4: The movement of the building blocks determines the temperature of an object.
KI 5: At the melting point, the bond becomes flexible, and the rigid structure disappears.
KI 6: There is nothing between the building blocks.
KI 7: At the boiling point, the bonding of the building blocks ceases and they move away from each other.

students to make that connection. To avoid misconceptions based on flawed representations of atoms and molecules as small spheres [51], we decided to use typographic representations, based on the idea of Wiener *et al.* [46]. Based on previous research [48], we decided to use the term "building blocks" instead of "particles," to also avoid misconceptions based on the fact that "particle" is not well defined in everyday language [47].

The seven key ideas (see Table II) were also influenced by the domain-specific design principles. For example, atoms and molecules are always referred to as "building blocks," and the emergent aspects of the PNM are explained in key idea 2. The key ideas were not formulated all at once but were developed over the first four cycles of the DBR project. New key ideas were added with each cycle to increase the explanatory power of the TLS and to respond to students' needs. For example, key idea number 6 was added after some students mentioned that they could not imagine that there is nothing between the building blocks. In addition, existing key ideas were adjusted when the results from the previous cycle showed that students were having difficulties learning that key idea. The order of the key ideas was also changed several times.

B. Development of the teaching materials

Given the overall satisfactory results in cycle 4, the TLS was considered ready to be tested on a larger scale in classrooms. To increase the likelihood that teachers would volunteer to participate in the study with one of their classes, the researchers decided to keep the number of lessons to a minimum. The goal was to incorporate all seven key ideas into 4 h of instruction.

The domain-specific design principles and key ideas guided the creation of the instructional materials. Snow and salt crystals serve as the context for teaching the PNM because the formation of crystals can only be explained in terms of their atomic structure. Since the shape of the crystal results from the lattice structure and not from the shape of the individual atoms, students need to distinguish between the properties of substances and those of particles. To help students understand the underlying concept of emergence, several examples from everyday life, such as jigsaw puzzles or pixels on a screen, are used. When students see that the building blocks of a puzzle have a different shape than the puzzle itself, they can apply this idea to the building blocks of a salt crystal. To visualize this connection, 3D-printed models of crystal structures are used. To emphasize the different properties of particles and substances, atoms and molecules are represented by their chemical symbol. In this way, students are encouraged to understand the relationship between building blocks and matter and to build their own mental model of matter. By the end of the TLS, they should be able to apply this model in small hands-on experiments, such as compressing air in a syringe and making assumptions about why this is possible.

To get students thinking about the PNM, several worksheets were created with tasks for students to complete alone, in pairs, or in groups. For example, in the first lesson, students should watch a video about the formation of salt or snow crystals and then explain what they saw in the video to a classmate. Students should use 3D-printed crystal structures to investigate how the crystals are formed. Over the course of the four lessons, students were given a total of seven worksheets, which covered everything they would later be asked to do in the post-test on a conceptual level.

Teachers who participated in the study received a manual that included a description of the key ideas, design principles, and detailed planning for the four lessons, including learning objectives and materials to be used. Teachers were free to use the lesson plans as they were or to adapt them according to their preferences or the needs of their classes. Teachers were also given a presentation file with several slides for each lesson. The slides included key ideas, illustrative graphics, and videos of simple experiments. There were also materials for teachers to demonstrate, such as small experiments or 3D-printed models to support the teacher's explanations.[1]

C. TLS evaluation

A pre-post test format was chosen to examine students' performance. For the pretest and post-test, mostly the items from the interviews in cycles 3 and 4 were used, supplemented by some items from a German translation of the

---

[1] All teaching materials can be accessed via the following link: https://tinyurl.com/33dp2424.

Chemistry Concept Inventory (CCI) [61]. There were both single-choice and open-ended items. The same items were used for the pretest and the post-test.

To evaluate the pretest and post-test items, a small pilot study was conducted. First, a test containing all items was administered to six students from a school where the first author teaches. All students had previously been taught the PNM. The students not only answered the questions but were also asked to explicitly tell the researcher if they had any problems or did not understand the questions. As the main problem, most of the six students thought that the test was too long. In addition, they did not use the PNM to answer the open-ended questions. To address these issues, the authors then met with other members of their working group to discuss possible revisions to the test. The discussion resulted in several changes being made to the pretest and post-tests. One item required students to do a lot of reading because it consisted of 13 statements that students had to decide were true or false. The number of statements was reduced by more than half. To make it clearer to students how to answer the open-ended questions, each item explicitly stated that they should use the PNM for their answer. In addition, several minor changes were made to most of the items. The refined version of the pretest was then administered to a class of 23 students. Because some students still took a long time to answer the questions, one more item was removed from the pretest.

The final version of the pretest consisted of seven general questions about the PNM, which the students should already have known from previous instruction according to the national curriculum in Austria. The post-test consisted of 11 questions: 7 questions from the pretest and 4 questions more specific to the TLS.

Four teachers from three different schools in Vienna agreed to teach the TLS to an eighth-grade class. The teachers' experience in teaching physics ranged from 2 to 6 years. All principals of the schools agreed to the implementation of the study. Informed consent was obtained from all parents of participating students. Teachers were also instructed on how to administer the pretest and post-tests in a way that would not reveal the identities of the students. For 130 students, a dataset of pretest and post-tests was gathered. An overview of the sample is shown in Table III.

TABLE III. Sample of students participating in the study.

School	Teacher	Class	Number of students
I	A	1	23
		2	22
	B	3	25
		4	22
II	C	5	20
III	D	6	18

Our aim in this study was to test the feasibility of the TLS in some classrooms. Rather than trying to create generalized findings, the goal was to find what is commonly referred to as existence proof or proof of principle [62,63]. As stated in the research question, we wanted to see if the developed TLS, which already worked with students in one-on-one interviews, can also enhance students' use of the PNM in a realistic classroom setting. This approach, looking for "what works" and not for generalized statements, is characteristic of DBR as it aims to work on problems from classroom practice, which are influenced by a huge number of factors [58,63].

Given the aim and purpose of the study, students were not divided into experimental and control groups, i.e., this was not a randomized controlled trial (RCT). RCTSs stem from research in agriculture, where different treatments and their effect on crops should be tested and are often applied in the natural sciences. In contrast to natural sciences, teaching and learning in a classroom cannot be described by causal laws. Therefore, using RCTs in education research is often criticized. [63]. Rather, the purpose of our study was to investigate whether the TLS would encourage students in the sample to use the PNM at all. Using a control group would have been difficult as, e.g., the context of crystal structures would not have been taught in a more traditional approach to the PNM. Likewise, a relevant number of other studies engaging in the evaluation of a TLS [64–71] also did not use a control group.

Furthermore, it was of interest whether teachers were able to implement the TLS in a meaningful way. Therefore, each teacher was interviewed by the first author to get feedback on how the implementation of the TLS worked out. The interviews were semistructured, each teacher was asked a minimum of five basic questions, but depending on the course of the conversation, the interviewer also asked additional questions to gain more insight. All interviews were recorded with the permission of the teachers and later transcribed for qualitative content analysis.

For data analysis, student responses from the paper and pencil tests were transferred to a spreadsheet for further processing. The open-ended responses were coded on the item level using evaluative qualitative content analysis [72]. In this type of qualitative content analysis, a set of preexisting categories is used to code the material. Categories for coding were based on students' different mental models of matter [18]. If students correctly applied the PNM, their response to an item was coded as "particulate model." If the students used the same properties for particles as for macroscopic materials in their response to an item, it was coded as "hybrid model." If students did not use particles at all in their explanation, their answer to that item was coded as a "continuous model." Thus, the coding focused specifically on students' use of the PNM. Students could answer an item correctly on a macroscopic level, but without using particles in their answer, the coding, in that case, would have been a continuous model. Interrater reliability was examined by two researchers from the authors' working group using 10% of the dataset. Cohen's kappa was calculated to be $\kappa = 0.73$. This represents a good interrater agreement [73].

Responses to the open-ended and closed-ended questions were then converted into scores for use in a quantitative analysis. Wrong answers to the closed-ended questions were coded as 0 points and correct answers were coded as 1 point. For the open-ended questions, students received 0 points if the answers were based on the continuous model, 1 point if it was based on the hybrid model, and 2 points if it was based on the particulate model. Therefore, all items are ordinally scaled. The analysis was performed in jasp, a free statistics program [74]. Because the data are ordinal scaled, only nonparametric tests can be used. The Wilcoxon singed-rank test was used to test whether the differences between students' responses in the pretest and post-test were statistically significant. To gain an insight into the overall performance of the students in the pretest and post-test, the sum of their scores on the seven identical questions from the two tests was compared. Therefore, a paired samples t test was used to calculate *Cohen's d* for each class individually as well as for the sum of all participants.

The transcripts of the interviews with teachers were analyzed using structuring qualitative content analysis [72]. In contrast to the evaluative content analysis used with the open-ended questions from the pretest and post-test, the goal was not to find preexisting categories in the data. Rather, teachers' ideas about the TLS were to be identified through an inductive analysis of the transcribed interview data. Coding was based on the five questions in the interview guidelines. First, sections of the transcript containing answers to one of the five questions were marked. The content of the marked sections was then examined for relevant statements made by the teachers. Each statement was given a code in order to make it possible to compare the statements of the individual teachers. If the same code was assigned to the statements of several teachers, this statement was considered relevant for the presentation of the results.

IV. RESULTS

The results are based on the scoring of the pretest and post-test as well as the interviews. For comparison between the pretest and the post-test, the seven items that were part of both tests were used. Items that were only part of the post-test were analyzed separately.

A. Pre-post comparison

Overall, students used the PNM more often in the post-test than in the pretest. The Sankey diagram (Fig. 3) shows, how many students answered a question within the same mental model in both tests and changed their mental model from pretest to post-test. In the pretest, most of the students' answers can be described as a continuous model of matter, followed by a hybrid model and a particulate model. In the post-test, the continuous model is still the most prevalent, although fewer students' responses fall within that model. Fewer students use a hybrid model while students' use of the particulate model more than doubles. The most noticeable changes are students switching their answers from a continuous to a particulate model and students switching from a hybrid to a particulate model. However, most of the students' answers that were in the continuous model in the pretest remained in that model in the post-test.

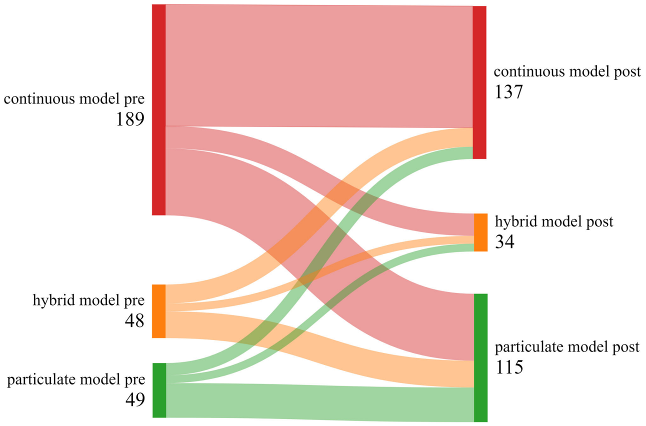


FIG. 3. Sankey diagram of students' mental models of matter before and after the intervention. Note that the numbers are pooled responses to the three open-ended questions that were the same on pretest and post-test. Students' answers to these questions were excluded from the analysis if they did not answer both the pretest and the post-test (source: own illustration using sankeymatic.com).

A paired-sample t test was conducted to compare the means of students' scores in the pretest and post-test. As the pretest and post-test consisted of a different number of items, only the seven items that were the same in both tests were used for comparison. If a student had answered all of these seven items correctly, meaning that they always applied the PNM, they would have scored 14 points (how the points were awarded is explained in Sec. III C). Students' mean scores were 5.38 out of 14 points ($\mathrm{SD} = 2.54$) in the pretest and 7.34 out of 14 points ($\mathrm{SD} = 3.01$) in the post-test. The t test showed a significant difference between those means as $t(130) = -7.03$, $p < 0.001$. The effect size, measured by Cohen's d, was $d = -0.62$, indicating a medium effect [75]. The negative effect size indicates a higher score on the post-test. Table IV shows different effect sizes for the six classes of students who participated in the study. Five out of the six classes scored significantly higher on the post-test than on the pretest. Cohen's d varies from small ($d_1 = -0.428$) to very large effects ($d_4 = -1.451$). Class 5 scored lower on the post-test than on the pretest ($d_5 = 0.033$). Possible reasons

TABLE IV. Comparison between pretest and post-test in different classes (paired samples *t* test). For all tests, the alternative hypothesis specifies that the pretest is less than the post-test.

Class	Mean pretest	Mean post-test	t	d.o.f.	p	Cohen's d
1	5.826	7.043	−2.053	22	0.026	−0.428
2	4.182	6.773	−3.694	21	<0.001	−0.787
3	6.240	8.520	−4.174	24	<0.001	−0.835
4	4.136	8.273	−6.807	21	<0.001	−1.451
5	5.250	5.150	0.149	19	0.558	0.033
6	6.722	8.056	−1.856	17	0.040	−0.437
All	5.377	7.338	−7.032	129	<0.001	−0.617

why students in class 5 could not benefit from the TLS are discussed in Sec. V A.

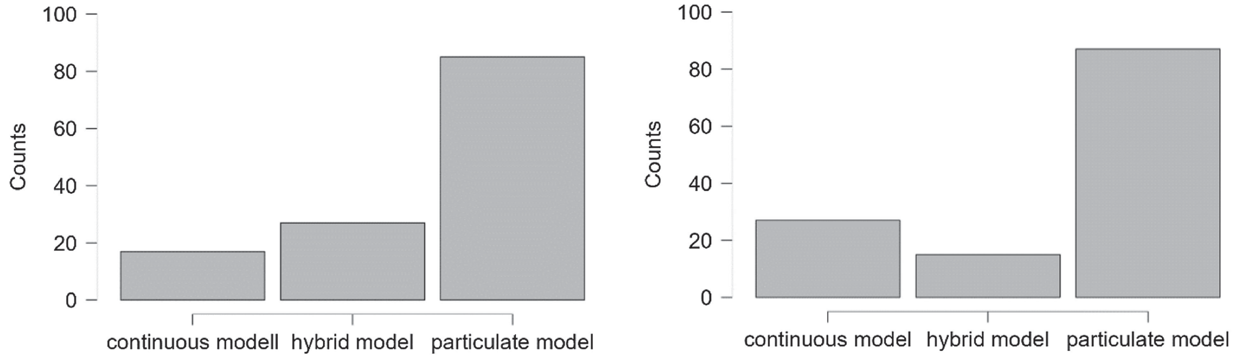
Table V shows the results of the Wilcoxon signed-rank test for the seven identical items in the pretest and post-test. Since item 1 consisted of five questions asking students to rate whether a statement is true or false, the total number of compared measures was 11. Table V shows that for 6 of these 11 items, the difference between the pretest and post-test is statistically significant as $p < 0.05$.

Three out of these six items are related to the idea that there is nothing between the particles (1b, 2, and 7). In item 1b, students were asked if they agreed with the statement "There is empty space between the building blocks." There was a significant difference in students' mean acceptance of this statement between the pretest ($M = 0.51$, $\mathrm{SD} = 0.50$) and the post-test ($M = 0.70$, $\mathrm{SD} = 0.46$); $z = -2.69$, $p < 0.001$. The effect size measured by the rank-biserial correlation is $r_\mathrm{B} = -0.40$, which represents a medium effect [75]. In item 2, students were asked what is between two adjacent water building blocks when the water has evaporated. They could choose from three options: "air," "water vapor," or "nothing." In the post-test, significantly more students chose the correct answer "nothing" ($M = 0.43$, $\mathrm{SD} = 0.50$) than in the pretest ($M = 0.20$, $\mathrm{SD} = 0.40$); $z = -3.45$, $p < 0.001$. The effect size measured by the rank-biserial correlation is $r_\mathrm{B} = -0.56$, indicating a large effect. In item 7 (11 in the post-test), students were asked why a sealed syringe filled with air can be compressed. They were also asked to use the PNM in their explanation. If they used the PNM correctly, their answer would receive 2 points. If they used the PNM but also applied misconceptions, their answer would receive 1 point. If they did not use the PNM, they would always receive 0 points. On the post-test, significantly more students gave answers based on the PNM ($M = 1.53$, $\mathrm{SD} = 0.76$) than in the pretest ($M = 0.66$, $\mathrm{SD} = 0.87$); $z = -5.36$, $p < 0.001$. The effect size measured by the rank-biserial correlation is $r_\mathrm{B} = -0.89$, indicating a large effect. The data from these three items suggest that most of the students accepted the idea, that there is nothing between the building blocks in the post-test.

Two items with significant differences between pretest and post-test (3 and 5) are related to the idea that the temperature of an object is determined by the motion of the particles. In item 3, students were asked: "What happens to a single building block of water when water vapor is heated?" The students were presented with four options, which they could select according to their understanding of the phenomenon in question. These were as follows: (a) "The building block's density decreases," (b) "The building block

TABLE V. Wilcoxon signed-rank test. Note that for all tests, the alternative hypothesis specifies that pretest is less than post-test. For example, 1a_Pre is less than 1a_Post. For the last three items, the numbers do not align because the post-test had more items than the pretest.

Pretest		Post-test	W	z	p	Rank-biserial correlation	SE rank-biserial correlation
1a_Pre	–	1a_Post	574.000	2.204	0.994	0.400	0.179
1b_Pre	–	1b_Post	549.000	−2.694	<0.001	−0.400	0.147
1c_Pre	–	1c_Post	632.500	−0.947	0.139	−0.148	0.155
1d_Pre	–	1d_Post	627.000	−1.395	0.055	−0.214	0.152
1e_Pre	–	1e_Post	231.000	−0.617	0.243	−0.125	0.200
2_Pre	–	2_Post	280.500	−3.446	<0.001	−0.560	0.161
3_Pre	–	3_Post	375.000	−2.362	0.003	−0.388	0.163
4_Pre	–	4_Post	400.000	−2.114	0.008	−0.347	0.163
5_Pre	–	8_Post	411.000	−3.301	<0.001	−0.503	0.151
6_Pre	–	10_Post	242.500	−0.940	0.168	−0.185	0.194
7_Pre	–	11_Post	65.000	−5.364	<0.001	−0.889	0.164



(a) Why do all snow crystals have six prongs?

(b) How can diamond and graphite have different properties when they both solely consist of carbon?

FIG. 4. Coded answers of students to questions in the context of crystal structures.

expands," (c) the building block slows down, and (d) "The building block speeds up." In the post-test, a significantly greater number of students chose the correct answer, option d ($M = 0.61$, $\mathrm{SD} = 0.49$) in comparison to the pretest ($M = 0.46$, $\mathrm{SD} = 0.50$); $z = -2.36$, $p = 0.003$. The effect size measured by the rank-biserial correlation is $r_\mathrm{B} = -0.38$, which represents a medium effect. In item 5 (item 8 in the post-test), students were asked: "An iron ball does not fit through an iron ring at room temperature. Describe what could be done to make it fit through. Justify your answer with the building blocks of matter!" Students' answers were coded in the same manner as with item 7 above. In the post-test, a significantly greater proportion of students provided responses based on the PNM ($M = 0.79$ and $\mathrm{SD} = 0.92$) than in the pretest ($M = 0.46$, $\mathrm{SD} = 0.72$); $z = -3.17$, $p < 0.001$. The effect size measured by the rank-biserial correlation is $r_\mathrm{B} = -0.49$, representing a medium effect. The data from these two items indicate that the majority of students accepted the notion that the movement of the building blocks determines the temperature of an object.

The final item with statistically significant differences between pretest and post-test is item 4. In this item, students were presented with the following scenario: "A balloon filled with helium is pushed to the bottom of the swimming pool. It immediately becomes smaller. How can this decrease in volume be explained?". Students were then asked to choose from four different answers.[2] In the post-test, a significantly greater proportion of students provided responses based on the PNM ($M = 0.75$, $\mathrm{SD} = 0.44$) than in the pretest ($M = 0.61$, $\mathrm{SD} = 0.49$); $z = -2.11$, $p = 0.008$. The effect size measured by the rank-biserial correlation is $r_\mathrm{B} = -0.35$, representing a medium effect. Students were less likely to choose option b, indicating that they less often used the same properties for particles as for substances.

B. Post-test analysis

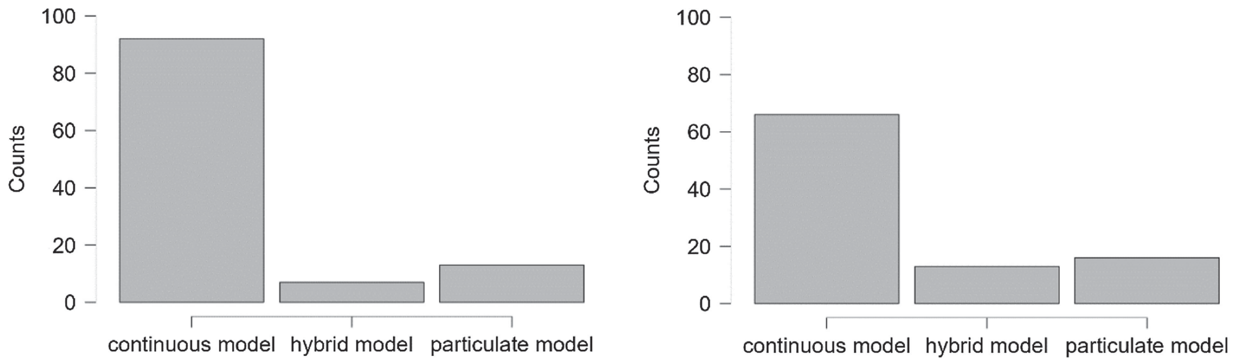
In the TLS, the PNM was introduced via crystal structures. In the post-test, students were required to answer two single-choice questions. They were asked to select the statement in a concept cartoon that most closely aligned with their beliefs. The first question pertained to the formation of snow crystals. Three statements were presented to explain why all snow crystals possess six prongs (see the Supplemental Material [76]). Each statement represented a person having a continuous, hybrid, or particulate mental model of matter. About 66%[3] of the students selected the statement representing a particulate model of matter [see Fig. 4(a)]. The second question concerned the reason for the differing properties of diamond and graphite, despite both being composed solely of carbon atoms. Once again, three different statements were presented to explain this phenomenon, representing a continuous, hybrid, and particulate mental model of matter. In this item, 67% of the students selected the statement representing a particulate model of matter [see Fig. 4(b)].

One of the most relevant challenges in teaching the PNM is that students tend to apply the same properties to particles as they do to macroscopic objects. Despite students demonstrating higher performance on the post-test, this misconception persisted in some cases. When asked about the statement: "A railroad track expands when heated because the iron building blocks expand," 55% of the students agreed. However, when asked: "What happens to a single water building block when water vapor is heated?" 61% of the students selected the correct response: "The building block speeds up." Only 14% selected the incorrect response: "The building block expands." It seems that

---

[2] (a) The cold water causes helium to become liquid; (b) The helium building blocks are getting smaller; (c) Helium building blocks migrate through the rubber skin of the balloon into the water; (d) The distances between the helium building blocks become smaller.

[3] These are the valid percentages; missing answers were omitted from the analysis.


(a) Why does an ice cube melt over a flame but not an aluminum can?

(b) When heating salt water: why does the water evaporate but not the salt?

FIG. 5. Coded answers of students to questions in the context of melting point (a) and boiling point (b).

students are more likely to transfer the properties of macroscopic objects to submicroscopic particles when they think about a solid object (55% use a hybrid model) than when they think about a gas (14 use percent hybrid model).

When it comes to explaining melting and boiling points, students rarely utilize the PNM. In the post-test, only 17% of students in the post-test utilized the PNM to explain the differing boiling points of various substances. Of the valid responses, 70% were based on a continuous model of matter [see Fig. 5(b)]. Furthermore, a smaller proportion of students employed the PNM when asked to explain why ice immediately melts over a flame while aluminum does not. A mere 12% of students employed a particulate model of matter to answer the question, while 82% applied a continuous model [see Fig. 5(a)].

C. Teachers' perspective

In general, the feedback from the four teachers on the TLS was positive. They found the materials to be "good" and "helpful," and they described the implementation of the TLS in the classroom as "easy." Teachers generally liked introducing the PNM with crystal structures. However, two of them said that they would not discuss crystal structures in as much detail as suggested by the TLS, while teacher D said it would be beneficial to go even more into detail. All four teachers indicated that they would utilize the TLS again, although one teacher expressed reservations about doing so under certain conditions. Teacher A expressed a preference for not using the model experiment demonstrating the spacing between particles in the gaseous state, stating concerns that it lacked sufficient descriptive detail for students.

Some of the domain-specific design principles of the TLS were novel for the teachers. For instance, the use of "building blocks" instead of "particles" was unfamiliar to all of them. The teachers indicated that they were accustomed to the term "particles" and that the textbooks also employed this terminology. This may present a challenge, as teachers may encounter difficulties integrating their textbooks with the TLS. Additionally, the concept of utilizing crystal structures and 3D-printed models to teach the PNM was novel to some teachers. They noted that the handling of the 3D-printed model was challenging, especially the need for substantial storage space and the potential for breakage. Nevertheless, all of the teachers indicated that the models were well received by the students. The teachers' descriptions of the students' prior knowledge of the PNM differed. Teachers A and B expressed surprise at the limited recollection of the PNM by the students, whereas Teacher D observed that the students demonstrated a high level of familiarity with the topic due to their previous studies in physics and chemistry. This finding aligns with the observation that class 6 achieved the highest average score in the pretest. Teacher D additionally proposed that supplementary materials be included in the teachers' manual for students who demonstrated a high level of interest. These materials could include a more detailed description of how crystal structures are formed.

All teachers felt that most of their students understood what they were taught during the TLS. Teachers C and D mentioned that the learning materials and the hands-on nature of the TLS were beneficial to their students' learning. Teacher B felt that when talking to the students, the majority of them seemed to understand the content. However, they may soon forget if they do not review and reinforce this knowledge.

None of the four teachers reported making any major changes to the TLS. They used all of the student worksheets and presentation slides provided by the researchers. Only the experiments demonstrating the expansion of an evaporating liquid in a food bag did not work well for the teachers. Teachers A and B reported that it did not work at all, no matter what they tried. They showed the video of the experiment instead. Teacher C also had some difficulties at first but managed to make the experiment work for the students. Teacher D did not let the students do the experiment themselves because there was not enough time and they had to try more than once to get it to work. It seems that this hands-on activity is difficult for teachers to use in the classroom, although they said that it would be helpful for students learning the PNM if it worked.

V. DISCUSSION

The purpose of the study was to evaluate whether a TLS developed based on domain-specific design principles would promote students' use of the PNM. Domain-specific design principles are "local theories" [60], meaning they are not generalized theories, but rather practical guidelines within the context in which they operate [13]. They inform instructional activities in a particular context to promote learning [77]. In this study, the four domain-specific design principles in Table I were intended to guide students' application of the PNM.

The nongeneralized nature of the design principles makes it challenging to determine which of these principles is responsible for which outcome. The higher performance of students in the post-test can be attributed to the collective influence of all domain-specific design principles. However, it is not possible to determine which one had the greatest impact on students' use of the PNM. Some of the key ideas, in particular, make use of a design principle. For instance, the concept of empty space between particles is elucidated to students through the use of various typographic representations. Consequently, it can be posited that typographic representations influenced students' acceptance of this idea more frequently in the post-test. Similarly, the emphasis on the emergent properties is of paramount importance for students to comprehend temperature. As evidenced by the fact that the majority of students in the post-test connected temperature with the movement of particles, an understanding of the concept of emergence might have facilitated their ability to apply the PNM correctly. Comparable to the study by Henderson *et al.* [56], explaining to students the concept of emergence seems to have helped them to argue within the appropriate ontology.

Only some of the results from the previous study could be replicated in the classroom setting. Crystal structures appeared to be a fruitful context for the introduction of the PNM, as students in the interviews as well as in the post-test were able to apply the PNM within that context. Even the students in class 5, who scored poorly on the post-test, predominantly applied the PNM in connection with crystal structures.

However, regarding phase transitions, there are differences between the previous and the current study. In the previous study, students were able to explain what happens to the building blocks during phase transitions when specifically asked to do so in the interviews. In this study, students were also asked to use the building blocks for their explanation of phase transitions, but most of them did not do so. It remains unclear why students did not use the PNM for their explanation of melting and boiling points. It is possible that the examples provided for the items were not optimal, as students may perceive the melting of a substance and the heating of another substance without undergoing a phase transition as two distinct phenomena. Consequently, they may adhere to the continuous model, as it is more familiar to them, and they can justify the observed differences in melting and boiling points. However, the same examples were also utilized in interviews with students in the previous study. In this instance, students encountered only minor difficulties in utilizing the PNM to elucidate phase transitions.

The students in this study showed what has been termed "level confusion" [22] or "hybrid conceptions" [25]. They employed the same properties for submicroscopic particles as they did for macroscopic objects. This conception appeared to be most prevalent in connection with solids, as more than half of the students imagined expanding iron particles causing the expansion of a railroad track. This phenomenon might be attributed to students' belief that particles cannot move in a solid. Conversely, students in general accepted the idea of moving particles in gases. Only a minority of students in this study argued that particles expand in a gas. This finding is consistent with the observation that students have difficulties explaining phase transitions with the PNM. They do not perceive the different phases of matter as the same particles having different velocities and bonding, but rather as three distinct entities. This kind of students' conception has already been documented by Johnson and Papageorgiou [78].

A. Limitations of the study

The objective of this study was to assess the feasibility of utilizing the developed TLS in instructing the PNM to eighth-grade students from Austria. As such, our aim was to find a proof of principle, rather than conducting a comprehensive quasiexperimental study. The study involved four teachers and six classes comprising a total of 130 students from three distinct schools in Vienna. The participation of teachers and their respective classes was entirely voluntary. All students received the same instructional approach, and there were no control groups. Conducting an educational quasiexperiment with a larger sample would have enhanced the reliability and objectivity of the study. However, many published studies investigating the effectiveness of a TLS did also not engage in educational quasiexperiments.

Teachers were provided with guidelines outlining the most crucial aspects of the TLS. These guidelines also included detailed lesson plans. Additionally, they received all necessary materials (students' worksheets, presentation slides, experimental materials) for implementing the intervention. However, the researchers did not observe teaching. It is therefore not possible to conclude that all students were taught exactly in the same way. However, it can be argued that during the follow-up interviews, none of the teachers made any remarks about teaching in a way that differed from what was described in the guidelines.

All of the teachers reported that using the term "building blocks" instead of "particles" was new to them and not in line with the textbooks they normally used. Although the teachers also mentioned that they would use the TLS again in their future teaching, this might limit the acceptance of the TLS by other teachers who are very focused on the textbook in their teaching. Nevertheless, teachers mentioning difficulties in using new terms when implementing a TLS is also reported by other studies [79]. Further studies have revealed that teachers in general struggle to accept new features of empirically developed curriculum materials as they do not see their value in fostering student learning [80,81]. Supporting teachers in using the TLS therefore would be important to promote its implementation in classrooms.

The results of the six classes differed from each other. Five of the classes showed a significant improvement in the post-test, while class 5 scored worse in the post-test than in the pretest. Class 4 exhibited the greatest improvement in mean test scores between the pretest and post-test. The use of a single worksheet by teacher B for grading students' performance may have contributed to the high learning gain observed in class 4. However, it remains unclear whether this alone can explain the significantly higher learning gain observed in class 4 compared to the other classes.

If we assume that all teachers adhered to the guidelines as they stated, another source of variation in student outcomes may be due to differences in the time required to teach the intervention. As the intervention was planned for four consecutive lessons, teachers would need between 2 and 4 weeks, depending on the number of lessons in a week. However, in the interview, teacher C stated that they did not teach the TLS in four consecutive school lessons as planned but had to interrupt the TLS to teach something else. The intermission lasted for more than a month. Only then the intervention was completed, and the post-test was administered. This most probably was the reason for class 5 scoring worse on the post-test than on the pretest.

Another limitation is students' attitudes toward the study. Teachers reported that students did not take it as seriously as their regular teaching, as they were aware that the pretest and post-test would not be relevant to their grades. Consequently, students frequently failed to consider their responses in depth or even to respond to open-ended questions at all. It can be posited that had their results on the post-test been relevant for their grade, students would have invested more effort in providing accurate responses.

VI. CONCLUSION

In general, teachers were satisfied with the materials of the TLS and found them useful for teaching. Students overall used the PNM more after the intervention. Therefore, it can be concluded that the domain-specific design principles promoted students' use of the PNM. These results suggest that the TLS can be used for teaching the PNM in eighth-grade classes. In the case of the Austrian physics curriculum, the TLS can serve as a basic introduction to the PNM in the context of thermodynamics. However, the TLS might also be used in other physics or chemistry contexts. As the TLS focuses on the most fundamental ideas of the PNM, it might also be expanded to cover a broader range of physics concepts, such as heat transfer and diffusion. As the term "building blocks" encompasses atoms, molecules, and ions, it is possible to delve deeper into the distinctions between these terms and their underlying concepts.

As previously mentioned in the limitations section, the objective of this study was to find a proof of principle for implementing the TLS in eighth-grade physics classes. The presented results cannot be generalized, as the sample size was limited. Future research should therefore conduct an educational quasiexperiment to test the TLS with a larger number of students. In reverse, a more qualitative follow-up study could investigate the effects of the different design principles used in the TLS in more detail. As for the results of this study, it is often not possible to determine which of the domain-specific design principles proves most effective in fostering students' use of the PNM. Moreover, students' reasoning with the PNM in the different phases of matter should be investigated more closely. Science education research might also focus more on investigating the teaching of the concept of emergence, as it is key for explaining many phenomena in science education.

ACKNOWLEDGMENTS

Open-access funding is provided by the University of Vienna.

---

[1] A. K. Griffiths and K. R. Preston, Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules, J. Res. Sci. Teach. **29**, 611 (1992).

[2] A. Albanese and M. Vicentini, Why do we believe that an atom is colourless? Reflections about the teaching of the particle model, Sci. Educ. (Dordrecht) **6**, 251 (1997).

[3] H. Ö. Kapıcı and H. Akcay, Students' conceptions of the particulate nature of matter at secondary and tertiary level, Int. J. Sci. Math. Educ. **8**, 165 (2009).

[4] J. Snir, C. L. Smith, and G. Raz, Linking phenomena with competing underlying models: A software tool for introducing students to the particulate model of matter, Sci. Educ. **87**, 794 (2003).

[5] V. M. Williamson and M. R. Abraham, The effects of computer animation on the particulate mental models of college chemistry students, J. Res. Sci. Teach. **32**, 521 (1995).

[6] J. C. Hadenfeldt *et al.*, Students' progression in understanding the matter concept, J. Res Sci. Teach. **53**, 683 (2016).

[7] L. Morell, T. Collier, P. Black, and M. Wilson, A construct-modeling approach to develop a learning progression of how students understand the structure of matter, J. Res. Sci. Teach. **54**, 1024 (2017).

[8] S. Y. Stevens, C. Delgado, and J. S. Krajcik, Developing a hypothetical multi-dimensional learning progression for the nature of matter, J. Res. Sci. Teach. **47**, 687 (2010).

[9] M. Çalik, N. Ültay, H. Bağ, and A. Ayas, Effectiveness of particulate nature of matter (PNM)-based intervention studies in improving academic performance: A meta - analysis study, Chem. Educ. Res. Pract. **24**, 938 (2023).

[10] F. Budimaier and M. Hopf, Development of a new teaching-learning sequence on the particulate nature of matter using crystal structures, Phys. Rev. Phys. Educ. Res. **19**, 020169 (2023).

[11] J. Guisasola, K. Zuza, J. Ametller, and J. Gutierrez-Berraondo, Evaluating and redesigning teaching learning sequences at the introductory physics level, Phys. Rev. Phys. Educ. Res. **13**, 020139 (2017).

[12] Designing theory-based teaching-learning sequences for science education, edited by K. Kortland and K. Klaasen, in *Proceedings of the symposium in honour of Piet Lijnse at the time of his retirement as Professor of Physics Didactics at Utrecht University* (CDBeta Press, Utrecht, 2010).

[13] T. Anderson and J. Shattuck, Design-based research: A decade of progress in education research?, Educ. Res. **41**, 16 (2012).

[14] D. Psillos and P. Kariotoglou, Theoretical issues related to designing and developing teaching-learning sequences, in *Iterative Design of Teaching-Learning Sequences. Introducing the Science of Materials in European Schools*, edited by D. Psillos and P. Kariotoglou (Springer Netherlands, Dordrecht, 2016), p. 11.

[15] M. Méheut and D. Psillos, Teaching–learning sequences: Aims and tools for science education research, Int. J. Sci. Educ. **26**, 515 (2004).

[16] P. Lijnse, Didactical structures as an outcome of research on teaching–learning sequences?, Int. J. Sci. Educ. **26**, 537 (2004).

[17] V. Muñoz-Campos, A.-J. Franco-Mariscal, and Á. Blanco-López, Integration of scientific practices into daily living contexts: A framework for the design of teaching-learning sequences, Int. J. Sci. Educ. **42**, 2574 (2020).

[18] P. Johnson, Progression in children's understanding of a 'basic' particle theory: A longitudinal study, Int. J. Sci. Educ. **20**, 393 (1998).

[19] O. Lee, D. C. Eichinger, C. W. Anderson, G. D. Berkheimer, and T. D. Blakeslee, Changing middle school students' conceptions of matter and molecules, J. Res. Sci. Teach. **30**, 249 (1993).

[20] A. G. Harrison and D. F. Treagust, Secondary students' mental models of atoms and molecules: Implications for teaching chemistry, Sci. Ed. **80**, 509 (1996).

[21] A. Derman, N. Koçak, and I. Eilks, Insights into components of prospective science teachers' mental models and their preferred visual representations of atoms, Educ. Sci. **9**, 154 (2019).

[22] U. Wilensky and M. Resnick, Thinking in levels: A dynamic systems approach to making sense of the world, J. Sci. Educ. Technol. **8**, 3 (1999).

[23] M. B. Nakhleh, Why some students don't learn chemistry: Chemical misconceptions, J. Chem. Educ. **69**, 191 (1992).

[24] Y. Boz, Turkish pupils' conceptions of the particulate nature of matter, J. Sci. Educ. Technol. **15**, 203 (2006).

[25] S. Vosniadou, The development of students' understanding of science, Front. Educ. **4**, 1 (2019).

[26] S. Novick and J. Nussbaum, Pupils' understanding of the particulate nature of matter: A cross-age study, Sci. Educ. **65**, 187 (1981).

[27] J. Peuckert, Stabilität und Ausprägung von Teilchenvorstellungen, in *Die Teilchenstruktur der Materie im Physik- und Chemieunterricht*, edited by H. Fischler (Logos-Verl., Berlin, 2006), p. 77.

[28] S. Levy and U. Wilensky, Students' learning with the Connected Chemistry (CC1) Curriculum: Navigating the complexities of the particulate world, J. Sci. Educ. Technol. **18**, 243 (2009).

[29] H. Tang and M. R. Abraham, Effect of computer simulations at the particulate and macroscopic levels on students' understanding of the particulate nature of matter, J. Chem. Educ. **93**, 31 (2016).

[30] J. Kelter, J. D. Emery, and U. Wilensky, Learning about diffusion at two levels: Agent-based micro-scale and equation-based macro scale, in *Proceedings of ASEE's Virtual Conference, edited by American Society for Engineering Education, Virtual* (2020).

[31] H.-Y. Chang, C. Quintana, and J. S. Krajcik, The impact of designing and evaluating molecular animations on how well middle school students understand the particulate nature of matter, Sci. Educ. **94**, 73 (2009).

[32] Y. Chophel, Remediating misconceptions related to particulate nature of matter using video animation: An action research, Int. Res. J. Sci. Technol. Educ. Manage. **2**, 75 (2022).

[33] E. Adadan and M. M. Ataman, Promoting senior primary school students' understanding of particulate nature of matter through inquiry instruction with multiple representations, Education 3–13 **49**, 317 (2021).

[34] E. Adadan, K. E. Irving, and K. C. Trundle, Impacts of multi-representational instruction on high school students' conceptual understandings of the particulate nature of matter, Int. J. Sci. Educ. **31**, 1743 (2009).

[35] A. M. Hitt and J. S. Townsend, The heat is on! Using particle models to change students' conceptions of heat and temperature, Sci. Act. **52**, 45 (2015).

[36] F. Budimaier and M. Hopf, Students' ideas on common experiments about the particulate nature of matter, J. Baltic Sci. Educ. **21**, 381 (2022).

[37] L. Glatz, A. Teichrew, and R. Erb, Assessing the persuasiveness of experiments on the particle model of matter, in *Fostering Scientific Citizenship in an Uncertain World. Proceedings of ESERA 2021 online conference*, edited by G. S. Carvalho, A. S. Afonso, and Z. Anastácio (Braga, 2022), p. 37.

[38] M. T. H. Chi, R. D. Roscoe, J. D. Slotta, M. Roy, and C. C. Chase, Misconceived causal explanations for emergent processes, Cogn. Sci. **36**, 1 (2012).

[39] A. K. Bílgín, F. N. D. Yürükel, and N. Yígít, The effect of a developed REACT strategy on the conceptual understanding of students: "Particulate nature of matter", J. Turk. Sci. Educ. **14**, 65 (2017).

[40] C. F. Bauer and J. Y. K. Chan, Non-science majors learn about heat, temperature, and thermodynamics using the particulate nature of matter and guided-inquiry instruction, Am. J. Phys. **87**, 550 (2019).

[41] P. Y. Furlan, Engaging students in early exploration of nanoscience topics using hands-on activities and scanning tunneling microscopy, J. Chem. Educ. **86**, 705 (2009).

[42] H. Margel, B.-S. Eylon, and Z. Scherz, "We actually saw atoms with our own eyes". Conceptions and convictions in using the scanning tunneling microscope in junior high school, J. Chem. Educ. **81**, 558 (2004).

[43] W. Franzbecker and U. Quast, Entwicklung von Hypothesen und Modellen zum Aufbau der Materie im Physikunterricht der Sekundarstufe I, Der Phys. **9**, 43 (1975).

[44] J. V. Ebenezer, A hypermedia environment to explore and negotiate students' conceptions: Animation of the solution process of table salt, J. Sci. Educ. Technol. **10**, 73 (2001).

[45] D. W. Smithenry, Teaching with crystal structures, Sci. Teach. **76**, 52 (2009).

[46] G. J. Wiener, S. Schmeling, and M. Hopf, Can grade-6 students understand quarks? Probing acceptance of the subatomic structure of matter with 12-year-olds, Eur. J. Sci. Math. Educ. **3**, 313 (2015).

[47] K. S. Taber and A. García-Franco, Learning processes in chemistry: Drawing upon cognitive resources to Learn about the particulate structure of matter, J. Learn. Sci. **19**, 99 (2010).

[48] H. Pfundt, Das Atom—letztes Teilstück oder erster Aufbaustein? Zu den Vorstellungen, die sich Schüler vom Aufbau der Stoffe machen, Chim. Didact **7**, 75 (1981).

[49] W. Dierks, An approach to the educational problem of introducing the discontinuum concept in secondary chemistry teaching and an attempted solution, in *Relating macroscopic phenomena to microscopic particles. A central problem in secondary science education; proceedings of a seminar [from ober 23–26, 1989 near Utrecht*, edited by P. L. Lijnse (CD-ß Press, Utrecht, 1990), p. 177.

[50] D. M. Bunce and D. Gabel, Differential effects on the achievement of males and females of teaching the particulate nature of chemistry, J. Res. Sci. Teach. **39**, 911 (2002).

[51] A. G. Harrison and D. F. Treagust, Particles and matter: Problems in learning about the submicroscopic world, in *Die Teilchenstruktur der Materie im Physik- und Chemieunterricht*, edited by H. Fischler (Logos-Verl., Berlin, 2006), p. 53.

[52] G. J. Wiener, S. M. Schmeling, and M. Hopf, Introducing 12 year-olds to elementary particles, Phys. Educ. **52**, 044001 (2017).

[53] F. Budimaier and M. Hopf, Alternative Darstellungsformen des Teilchenmodells, in *Forschungsperspektiven 15*, edited by C. Fridrich *et al.* (LIT Verlag, Berlin, Münster, London, Wien-Zürich, 2023), p. 97.

[54] M. T. H. Chi, Commonsense conceptions of emergent processes: Why some misconceptions are robust, J. Learn. Sci. **14**, 161 (2005).

[55] S. Kivelson and S. A. Kivelson, Defining emergence in physics, npj Quantum Mater. **1**, 16024 (2016).

[56] B. Henderson, E. Langbeheim, and M. Chi, Addressing robust misconceptions through the ontological distinction between sequential and emergent processes, in *Converging Perspectives on Conceptual Change. Mapping an Emerging Paradigm in the Learning Sciences*, edited by T. G. Amin and O. Levrini (Routledge, London, 2017).

[57] S. A. Barab and K. Squire, Design-based research: Putting a stake in the ground, J. Learn. Sci. **13**, 1 (2004).

[58] Design-Based Research Collective, Design-based research: An emerging paradigm for educational inquiry, Educ. Res. **32**, 5 (2003).

[59] A. L. Brown, Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings, J. Learn. Sci. **2**, 141 (1992).

[60] C. Haagen-Schützenhöfer and M. Hopf, Design-based research as a model for systematic curriculum development: The example of a curriculum for introductory optics, Phys. Rev. Phys. Educ. Res. **16**, 020152 (2020).

[61] J. Barbera, A psychometric analysis of the chemical concepts inventory, J. Chem. Educ. **90**, 546 (2013).

[62] H. Borko, Professional development and teacher learning: Mapping the terrain, Educ. Res. **33**, 3 (2004).

[63] A. Bakker, Design research in education, in *A practical guide for early career researchers* (Routledge Taylor & Francis Group, London, New York, 2019).

[64] M. Méheut, Designing, and validating two teaching–learning sequences about particle models, Int. J. Sci. Educ. **26**, 605 (2004).

[65] S. Toffaletti *et al.*, Guiding Students towards an understanding of climate change through a teaching–Learning sequence, Educ. Sci. **12**, 759 (2022).

[66] O. R. Battaglia, A. Agliolo Gallitto, G. Termini, and C. Fazio, Outcomes of a teaching learning sequence on modelling surface phenomena in liquids, Educ. Sci. **13**, 425 (2023).

[67] I. Arriassecq and I. M. Greca, A teaching–learning sequence for the special relativity theory at high school level historically and epistemologically contextualized, Sci. Educ. (Dordr.) **21**, 827 (2012).

[68] P. Onorato, M. Malgieri, and A. de Ambrosis, Three years evaluation of a teaching learning sequence on rolling motion based on a blended learning environment, J. Phys. Conf. Ser. **1286**, 012053 (2019).

[69] P. A. Archila, Using history and philosophy of science to promote students' argumentation, Sci. Educ. (Dordr.) **24**, 1201 (2015).

[70] F. Kabapinar, J. Leach, and P. Scott, The design and evaluation of a teaching–learning sequence addressing the solubility concept with Turkish secondary school students, Int. J. Sci. Educ. **26**, 635 (2004).

[71] E. West and A. Wallin, Students' learning of a generalized theory of sound transmission from a teaching–learning sequence about sound, hearing and health, Int. J. Sci. Educ. **35**, 980 (2013).

[72] U. Kuckartz and S. Rädiker, Qualitative content analysis, in *Methods, Practice and Software* (SAGE, London, Thousand Oaks, New Delhi, Singapore, 2023).

[73] J. R. Dettori and D. C. Norvell, Kappa and beyond: Is there agreement?, Global Spine J. **10**, 499 (2020).

[74] J. Love *et al.*, JASP: Graphical statistical software for common statistical designs, J. Stat. Softw. **88** (2019).

[75] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Taylor and Francis, Hoboken, 1988).

[76] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.20.020104 for the post-test (including also the questions from the pretest).

[77] D. Euler, Design principles as bridge between scientific knowledge production and practice design, EDeR **1** (2017).

[78] P. Johnson and G. Papageorgiou, Rethinking the introduction of particle theory: A substance-based framework, J. Res. Sci. Teach. **47**, 130 (2009).

[79] V. Tobias, M. Hopf, C. Waltner, Thomas Wilhelm, and H. Wiesner, Der Einfluss der Sachstruktur im Mechanikunterricht. Qualitatives Forschungsvorhaben im Rahmen einer integrativen Studie, in *Didaktik der Physik. Bochum 2009*, edited by V. Nordmeier and H. Grötzebauch (Lehmanns Media, Berlin, 2009).

[80] J. Breuer, Implementierung fachdidaktischer Innovationen durch das Angebot materialgestützter Unterrichtskonzeptionen, *Fallanalysen zum Nutzungsverhalten von Lehrkräften am Beispiel des Münchener Lehrgangs zur Quantenmechanik* (Logos Verlag Berlin, Berlin/Germany, 2021), Vol. 314.

[81] M. Obczovsky, T. Schubatzky, and C. Haagen-Schützenhöfer, Supporting preservice teachers in analyzing curriculum materials, Educ. Sci. **13**, 518 (2023).