How to assess students' structure-property reasoning?

Otter, M.J. den; Juurlink, L.B.F.; Janssen, F.J.J.M.

Citation

Otter, M. J. den, Juurlink, L. B. F., & Janssen, F. J. J. M. (2022). How to assess students' structure-property reasoning? *Journal Of Chemical Education*, *99*(10), 3396−3405. doi:10.1021/acs.jchemed.2c00234

Version: Publisher's Version

JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc

Chemical Education Research

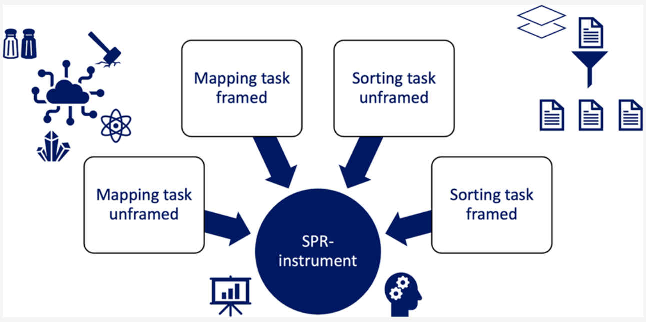
How to Assess Students' Structure–Property Reasoning?

Marie-Jetta den Otter,* Ludo B.F. Juurlink, and Fred J.J.M. Janssen

Cite This: *J. Chem. Educ.* 2022, 99, 3396–3405

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: This paper describes the design of an instrument to assess secondary school students' proficiency in structure–property reasoning (SPR). Design criteria for the instrument required that it should be based on a comprehensive model for structure–property reasoning, assess both reproductive and productive use of structure–property reasoning, be cost-effective, and be easy for teachers to adapt to their situation. An unframed and framed sorting task and an unframed and framed mapping task were included in the instrument. It was used to determine the proficiency in structure–property reasoning of two populations: 60 Dutch secondary school students on the preuniversity track and 108 Dutch first-year university chemistry students. Results were analyzed using established statistical techniques, and they confirmed that the SPR-instrument clearly discriminates between preuniversity and first-year chemistry students. The paper concludes by outlining the possibilities offered by the instrument and suggestions for further research.

KEYWORDS: *High School/Introductory Chemistry, Chemical Education Research, Testing/Assessment, Descriptive Chemistry, Structure–Property Reasoning, Chemical Thinking, Micro–Macro Thinking, Formative Assessment*

■ INTRODUCTION

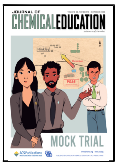
Learning to think like a chemist is considered an important goal of chemistry education.[1] A major part of chemical reasoning concerns structure–property reasoning.[2,3] Therefore, this type of reasoning is included in standards for chemistry education in many countries.[4,5]

Experts in chemistry use abstract models to explain and predict the properties of substances.[6,7] These models are based on invisible particles, e.g., atoms and molecules, and the interactions between such particles. These two levels of representation (the models of the particles and the properties of substances) have been referred to as the micro and macro levels.[8] Experts seamlessly switch between macroscopic properties of substances and their particulate structure. Whereas some properties, e.g. melting point or conductivity, are easily connected to the microscopic models of molecules or ions, other properties require a different scale.[2] For example, the ability of super absorbers to incorporate large quantities of water into their structure is explained using conglomerates of polymeric particles. It is this type of reasoning that is generally referred to as structure–property reasoning.[2,3]

Proficiency in structure–property reasoning is necessary for a good understanding of many chemical topics. For example, when working with problems concerning acids and bases or organic chemistry, students repetitively switch between the structure level and the property level. Lack of knowledge of the structure level or inability to apply the structure level leads to incorrect answers. The difficulties that students experience with structure–property reasoning have been widely discussed in the literature. Reasons for these difficulties range from the macroscopic propensity that students have due to previous experiences[3,8,9] to the lack of connection between the structure models and their own prior knowledge.[9] Furthermore, students find it very difficult to work with models; they struggle to learn how to recognize their limitations or how to apply them properly.[10]

In order to teach structure–property reasoning effectively, teachers require insight into students' proficiency in this specific skill.[1] Such insight enables teachers to adapt their teaching to enhance this method of reasoning. We aimed to design an instrument that teachers could use to assess the proficiency in structure–property reasoning of students in secondary education. The design criteria for such an instrument are the following. First, the instrument has to be based on a comprehensive model for structure–property reasoning. By "comprehensive" we meant that all facets of structure–property reasoning are made explicit and that the concepts associated with these facets cover a generic chemistry curriculum for secondary education. Previously, only specific

Received: March 16, 2022
Revised: August 19, 2022
Published: September 13, 2022



https://doi.org/10.1021/acs.jchemed.2c00234

aspects of structure−property reasoning have been emphasized.[10,11] In this paper, the focus was on the facets and the concepts needed for teaching years 11, 12, and 13 of the Dutch curriculum.[12] Second, as structure−property reasoning can be mastered on two levels, reproductive and productive use, the instrument has to assess structure−property reasoning at these two levels.[13] Third, the instrument had to be cost-effective, which means that the costs in terms of time, energy, and materials, e.g., should be as low as possible. Teachers have little time at their disposal;[14] therefore, preparation and administration of the instrument must be done within a limited time. The instrument also has to be applicable to large groups, such as the whole class. Finally, teachers have to be able to use the instrument repeatedly and adapt the tool to the grade, level, and content they teach to assess student development over multiple years.

After the design and development, the instrument was administered to two target groups (secondary school students on the preuniversity track and first-year university chemistry students) to determine whether it discriminates between these groups. Based on the results, we describe how the assessment instrument fulfilled the design criteria and what improvements are necessary.

DESIGN CRITERIA FOR STRUCTURE−PROPERTY REASONING ASSESSMENT INSTRUMENT

To fulfill the first criterion of the assessment instrument, i.e., that it is based on a comprehensive model for structure−property reasoning, the perspective for structure−property reasoning was used (Figure 1).[15,16] The following sections explain what a perspective is, how a perspective could facilitate structure−property reasoning, and how the perspective for structural−property reasoning was developed and validated.

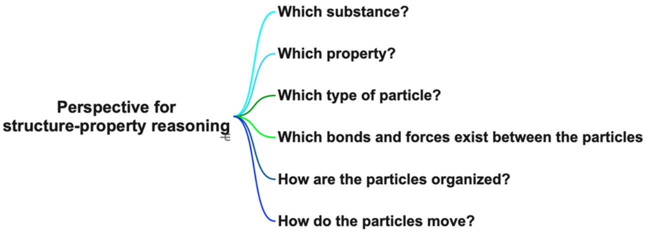


Figure 1. Illustration of the perspective for structure−property reasoning used as a comprehensive model for structure−property reasoning. The first two questions from the top address the macro aspect, and the last four questions address the micro aspect.

Complex domain-specific reasoning requires hierarchical problem solving.[17−19] In hierarchical problem solving, a concrete complex problem is first recognized as a problem of a certain kind. This problem type can be often divided in multiple abstract subproblems, each connected to multiple solution types. An abstract solution can be constructed by selection and recombination, and this abstract solution can be applied to the original concrete problem by refinement.

To demonstrate how structure−property reasoning works as hierarchical problem solving and how the perspective for structure−property reasoning can be used in this process, an example about the strength and elasticity of a spider's thread will be given. The questions of the perspective for structure−property reasoning (Figure 1) can be used to question and explain the structure−property relation between the elasticity and strength of a spider thread. When reasoning, a chemist may start with questions concerning the type of substance or material and the relevant properties. The chemist subsequently continues with questions concerning the type of particles, and their interactions (bonds and forces), organization, and movement. The answers combine and enable a chemist to formulate a possible answer. For example, the spider's thread consists of a copolymer (which particles?). In this polymer, β-sheets arise which are bounded tightly with hydrogen bonds (which bonds and forces exist between the particles?) and so on. The questions of the perspective help in structuring different aspects to the problem. The concepts represent the different answer options, and one develops a model by connecting several concepts.

In the example above, a domain-specific perspective, i.e., the perspective for structure−property reasoning (Figure 1), facilitated hierarchical problem solving. At its core, a perspective is an abstract schema that captures the core reasoning pattern in a domain and structures domain knowledge accordingly.[16,19−21] In an earlier study, Landa et al.[15] identified the perspective for structure−property reasoning as an important perspective that constitute an important part of the Dutch secondary school chemistry curriculum and formulate the following core reasoning schema for this perspective:

> *"The **properties** of **substances** can be explained by the nature of the **particles** of which it consists, the **bonds and forces** between them, and the **movement** and **organization** of those particles."*

The bolded words in this core reasoning are variables that can take different values. For instance, the word "particles" could refer to an atom, an ion of a molecule. A core reasoning schema serves as a template that concisely states how the multitude of concepts in a knowledge field meaningfully interact. To facilitate hierarchical problem solving required for domain-specific reasoning, the abstract schema underlying a scientific perspective is reformulated in a hierarchically connected set of questions and related concepts. Earlier studies suggest that perspectives could be used to scaffold and structure students' domain-specific reasoning.[15,16]

The perspective for structure−property reasoning, as one of the four theoretical chemical perspectives, was identified and validated in an earlier study in three steps.[15] First, a group of six chemistry experts were invited to reason about phenomena that we derived from the Dutch secondary school (preuniversity) chemistry curriculum syllabus to identify which distinct theoretical perspectives experts in fact applied in these concrete examples. Four chemical perspectives, namely, perspective for structure−property reasoning (particle perspective), kinetic perspective, thermodynamic perspective, and valence-shell perspective, were identified by the authors and validated by the experts. The following experts were involved in this study: a full professor who specializes in chemical drug design, two higher education teachers and researchers (with PhD in chemistry), and two secondary school chemistry teachers (with a PhD in chemistry). Next, the theoretical perspectives were refined by elaborating them as core reasonings and accompanying hierarchical question agendas. These were, again, validated by the experts and by mapping to the perspectives what various influential documents have identified as the big ideas in chemistry[22−25] in order to verify whether the chemical perspectives "covered" all the big ideas. Finally, the relevance of the four chemical perspectives for

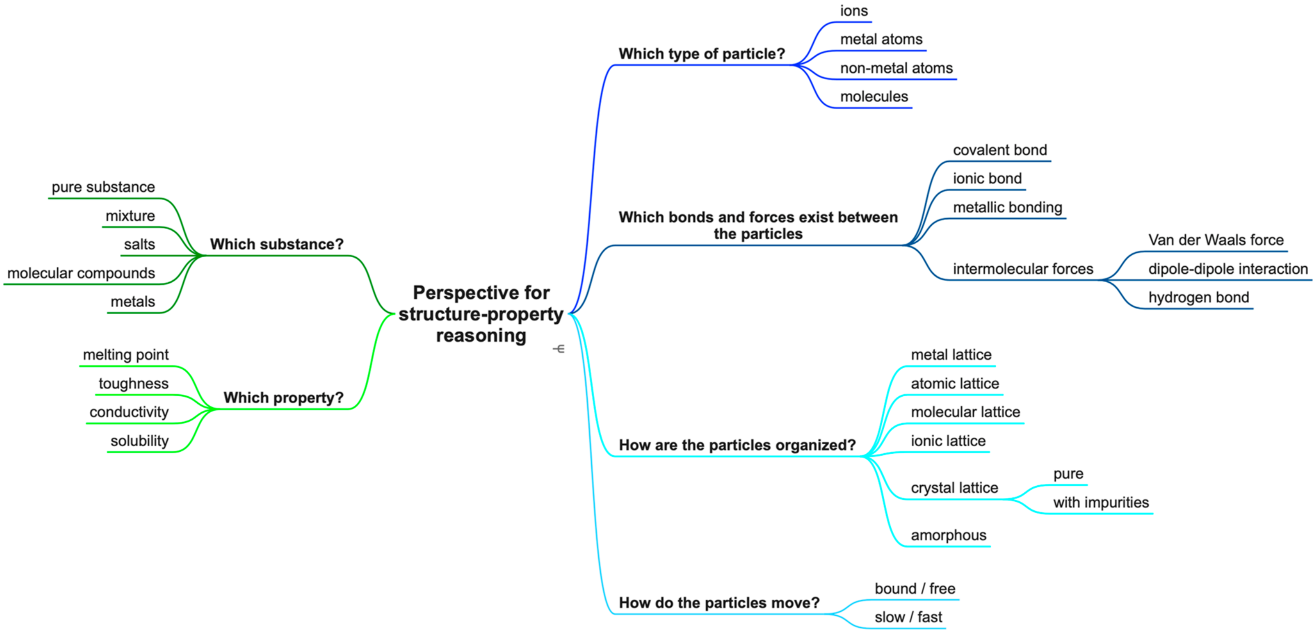


Figure 2. Model for structure−property reasoning with the corresponding answers to the questions in the branches.

secondary school chemistry education was validated by observing five secondary school chemistry teachers. It was established to which extent the four perspectives in their reasoning about why-questions were applied.[15]

The second design criterion required that the instrument assess structure−property reasoning at two levels, namely, reproductive and productive use. Reproductive use means that students can reproduce and understand the questions and chemical concepts including their mutual relations in the perspective for structure−property reasoning (Figures 1 and 2). Productive use means that students can use the questions and chemical concepts of the perspective for structure−property reasoning in problem situations in which they have to use structure−property reasoning for analyzing, explaining, predicting, creating, and/or evaluating.[13]

The third criterion for the instrument related to cost-effectiveness. Teachers have limited time and resources for preparing, teaching, and evaluating. Furthermore, they have to work with relatively large groups of students at the same time. Many innovations resulting from educational research require significant preparation and execution time or are difficult to implement in large groups. Lack of cost-effectiveness often results in instruments not being used by teachers.[14] This implies that an instrument should require as little time as possible for preparation and implementation. Furthermore, the instrument should be suitable for large groups.

The final criterion concerned the adaptability of the instrument. A ready-to-use instrument may save time. However, the scope of such an instrument may not match the teacher's requirements, for example, because of variations in curriculum or the employed textbook. In many cases, a ready-to-use instrument does not suit the teacher's class or needs exactly. The new instrument needed, therefore, to be very easily adapted by teachers to the year group and level they are teaching and/or to the curriculum or topics. The difficulty level and the choice of words needed to be adaptable as well as the subject the class is currently working on. Finally, it had to be possible to use the instrument repeatedly.

■ HOW TO ASSESS STRUCTURE−PROPERTY REASONING?

We devised a new instrument to assess structure−property reasoning with components of two existing instruments, namely, concept mapping techniques and sorting tasks. These techniques could be cost-effective and easy to adapt by users. Furthermore, they could be based on the perspective for structure−property reasoning. For these reasons, these two techniques were combined to fulfill the design criteria.

To estimate if a student understands and can reproduce answers to questions related to structure−property reasoning, a mapping task based on the concept mapping technique was designed. Concept mapping tasks can be used to measure the structure of a student's declarative knowledge (knowledge organization) in a certain domain.[26,27] The more common method of concept mapping was adapted to a mapping task in which the chemical concepts needed for structure−property reasoning had to be connected to the questions of the perspective (Figure 1) to produce a variation on Figure 2.

Based on the technique for producing a concept map,[26,28] two versions of the mapping task were designed, namely, a more difficult unframed and an easier framed version. In the unframed mapping task, students were not offered the concepts in advance, but they had to complete the questions from Figure 1 themselves with the appropriate chemical concepts. In this way, an indication about students' ability in reproducing and remembering the chemical concepts concerning structure−property reasoning could be obtained. In the framed mapping task, the students were offered the concepts they needed to place under the appropriate question of Figure 1 to estimate the understanding of the chemical concepts needed for structure−property reasoning.

Sorting tasks are designed to estimate the levels of cognitive processes applying, analyzing, and evaluating. Students are presented with a set of cards that each contain a problem, a statement, or a relation. Students sort the given cards into categories based on underlying commonality. Experts tend to sort the problems based on underlying conceptual features or

Table 1. General Overview of the Two Groups: Preuniversity Students (Secondary School) and First-Year Chemistry Students

	N	Male−Female	Mean Age (Year)	SD Age	Mean Final Grade[a] Preuniversity Chemistry	SD Final Grade Preuniversity Chemistry
Preuniversity students	60	33−27	16.6	0.78		
Group Y4 (age 15−16)	24	12−12	16.0	0.76		
Group Y5 (age 16−17)	36	21−15	16.9	0.52		
First-year chemistry students	108	75−33	18.4	1.09	7.67	0.75

[a] Students in The Netherlands take an exam at the end of their secondary school. This is combined with a part of the results they gained during the last two years of their education to produce their final grade. The range of this grade is from 1 (very bad) to 10 (no mistakes). The average final grade for chemistry of all preuniversity students in The Netherlands was 6.5 in 2016 and 6.6 in 2017.[36]

Table 2. Ideal Sorts for the Card-Sorting Task[a]

		Structure Aspects (Deep Features)			
		Molecular/Atomic Bonding	Molecular/Atomic Lattice	Ionic Bonding/Lattice	Metallic Bonding/Lattice
Property aspects (surface features)	Melting point	8	5	14	3
	Conductivity	11	7	2	12
	Toughness	16	13	4	10
	Solubility	1	6	9	15

[a] The columns are the structure aspects each card contains, and the rows represent the property aspects of each card. The numbers in the cells refer to the specific sample problem card. The design is based on the model of Krieter et al.[35].

"deep" features. Novices tend to sort the problems based on superficial features related to the presentation of the problem or the "surface" features.[29] Card-sorting tasks have been used in physics,[30] biology,[31,32] and chemistry.[33−35]

For the card-sorting task, deep and surface features were formulated and incorporated in 16 chemistry problems, all concerning structure−property reasoning. The deep features concerned the structure aspects of structure−property reasoning, i.e., the micro level. The surface features concerned the properties, i.e., the macro level. We would expect experts to sort the problems based on the micro level and novices to sort the problems based on the macro level.[29] The design of the card deck was based on the approach of Irby et al.[34] and Krieter et al.[35]

The sorting task was offered to participants framed (i.e., closed) in which they sorted a number of problem cards into predefined categories, and unframed (i.e., open) in which they sorted the problem cards into groups based on their own idea of the underlying chemical concepts.

The problems on the cards are all types of constructed response items, based on problems normally used in Dutch exams and books. However, the students were not expected to fully solve the problems. Both sorting tasks are used to indicate if the students can detect and describe underlying chemical concepts of the structure level to make decisions and evaluations. To achieve these tasks, the student had to be able to examine and break down information into parts to explore relationships and find generalizations.

■ METHOD

Participants and Context

To investigate whether the designed instrument was suitable for detecting differences in students' proficiency in structure−property reasoning, we selected two populations to apply the structure−property reasoning (SPR) instrument. The target group of the SPR-instrument was secondary school students on the preuniversity track. This target group was compared with a group of first-year university chemistry students, as they were expected to perform significantly better on the SPR-instrument.

The target group comprised 60 students attending a state secondary school in a Dutch city. These students were following the preuniversity track, which takes 6 years to complete. There were two groups: 24 students from year 4 and 36 students from year 5. They were taught chemistry by the first author or her teacher colleague.

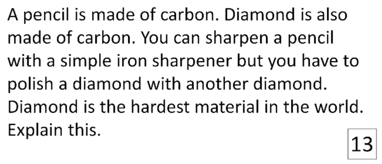
The preuniversity students' results were compared with those of 110 first-year chemistry students at Leiden University who were taking the General and Inorganic Chemistry course given by the second author. These students were almost at the end of this intensive course where the structure models learned in secondary education were repeated and elaborated. In total, 108 of the 110 (98%) students completed the assignments of the entire instrument.

Table 1 shows a general overview of the two groups of participants: the preuniversity students and the first-year chemistry students. At the time of completing the instrument, the preuniversity students had studied nearly all the chemical concepts needed for this instrument.

The Instrument

Card-Sorting Task. A deck of 16 cards with chemistry problems was designed as a basis for the card-sorting task. Four categories of physical properties (conductivity, melting point, toughness, and solubility) that are typically used in the Dutch curriculum in relation to structure−property reasoning were selected as surface features. The four structure aspects were ionic (bonding and lattice), metallic (bonding and lattice), molecular bonding, and molecular lattice, the deep features. Each problem card (the numbers in Table 2) contained a structure aspect (i.e., deep feature) and a property aspect (i.e., surface feature). The type of problems was to build an explanation of a structure−property relation. The choice was made because this type of problems is used by Dutch teachers in their lessons and exams to check students' structure−property reasoning.

The problems in the card deck were chosen and adapted from chemistry textbooks used in years 4–6 and from the Dutch national exams of the preuniversity track of secondary education. The most common properties that preuniversity students had to explain with the structure models were the four chosen property aspects (rows in Table 2). The four chosen structure aspects (columns in Table 2) were based on the structure models that students had to use to explain these properties. For example, card 13, shown in Figure 3, was designed with toughness as the property aspect and molecular/atomic and lattice as the structure aspect. When explaining the properties at the structure or micro level, the ionic lattice is inseparable from ionic bonding. Therefore, it was combined as one structure aspect when choosing the problems. The same applies for metal lattice and metal bonding.


A pencil is made of carbon. Diamond is also made of carbon. You can sharpen a pencil with a simple iron sharpener but you have to polish a diamond with another diamond. Diamond is the hardest material in the world. Explain this.

13

Figure 3. Sample problem card "13".

Three experts, the second author and two chemistry teachers, were consulted about the chosen problems. The second author is an associate professor in the field of catalysis and surface chemistry. In addition, he has extensive experience in the field of secondary and higher education and educational research. Both of the two chemistry teachers have a master's degree in chemistry and several years of experience with chemistry education in the preuniversity track of the secondary education in The Netherlands. They sorted the problems independently, and afterward, they gave feedback to the selected problems and the sorting task. After this consultation, the formulation of the situation on three cards of the original card set was adapted. The formulation of these original problems was suggested to be unclear.

Mapping Task. In the unframed mapping task, the participants were asked to complete the starting version of the perspective for structure–property reasoning (Figure 1) with all the chemical concepts that they could come up with and which seemed suitable as answers to the questions formulated in the model. In the framed mapping task, the participants were given 30 chemical concepts that needed to be placed at the appropriate question related to structure–property reasoning. The ideal outcome of this task is shown in Figure 2. This ideal outcome was based on the chemistry curriculum of the preuniversity track and on two experts: an experienced chemistry teacher with a master's degree in chemistry and the second author. The experts had no a priori knowledge of the task. From their unframed mapping task and the chemistry curriculum of the preuniversity track, the ideal outcome or the reference map was constructed. Next, this reference map was presented to these experts for feedback.

The participants were asked to complete the questions of the perspective unframed. The framed task, in which the questions of the model had to be completed using only the chemical concepts described as learning goals for preuniversity students, was only performed by the preuniversity students. To reduce the time needed for the experiment, the decision was made not to give this task to the first-year chemistry students. They were expected not to make any mistakes in this framed mapping task.

Procedure. The activities of the SPR-instrument were carried out in a classroom setting. Each participant noted his or her results on entry sheets. Adult participants were provided with an informed consent document approved by the Ethics Review Committee (IREC) of the university. For the underage participants, the parents received an informed consent letter approved by the IREC.

In Table 3, a short description and instruction are given for each task. The order in which the tasks were offered to the participants (unframed sorting task, framed sorting task, unframed mapping task, and framed mapping task) was chosen deliberately. Reversing the order of the tasks would have meant that the participants would be oriented in a certain direction.

The average time it took to take the test (including reading the instructions) was 50 min. The preuniversity students were tested in groups of 17–24 participants. The first-year chemistry students took the test in one group, all 108 students at once. The worksheets were designed to make it easy to collect the results and to assess large groups all at once.

The SPR-instrument and the corresponding worksheets to facilitate the administering of the SPR-instrument are provided in the Supporting Information.

Data Analysis. To determine the extent to which a student sorts on structure aspects, and therefore is more proficient in structure–property reasoning, the percentage of pairs (%*P*) made by a student was determined.[35] For each sort we compared the number of pairs that were common with one of the ideal sorts (i.e., the ideal sort on structure aspects and the ideal sort on property aspects (see Table 2)). The formed pairs and the total number of pairs were determined for each participant. When a single card was placed in a group, this was counted as a pair with a null card and considered as unexpected pairing. The pairs that the participant had in common with the ideal structure sort and the ideal property sort were counted to determine the number of pairs formed on the structure aspect and on the property aspect. The number of unexpected pairs consisted of the single cards and the pairs that were not in common with the ideal structure or property sort. The total number of pairs varied considerably between the participants. For this reason, the number of structure pairs, property pairs, and unexpected pairs was divided by the total number of pairs in a sort. The closer the similarity of a participant's sort to an ideal sort is, the higher the %*P* value is. A high %*P* for the structure aspect pairs (%*P*-structure) indicates that the participant sorted the cards more on structure aspects, i.e., deep features, meaning that the participant was thinking more like an expert.

In the unframed mapping task, the participants were asked to complete the questions related to structure–property reasoning (Figure 1) with the appropriate chemical concepts. The total number of answers or chemical concepts were counted as well as the total number of answers in accordance with the reference map (Figure 2). The number of extra chemical concepts given was also determined. The extra answers given by the participants were judged on correctness. In addition, the number of students who made one or more hierarchies in their answers were counted. The correctness of the hierarchy was also judged. For example, a student made a hierarchy in the question "which particles?". The first answer

Table 3. Description of the Procedure of the SPR-Instrument

Order of performance	Task	Performed By	Description	Instruction
1	Unframed sorting task	Preuniversity students and first-year chemistry students	16 problems, all containing a structure aspect and a property aspect, have to be sorted in groups. Each group should be given a name.	You received 16 cards with problems. Sort these cards in groups based on underlying common chemical concept. Give each group an appropriate name. Form at least 2 groups and maximal 15 groups.
2	Framed sorting task	Preuniversity students and first-year chemistry students	16 problems, same as unframed, have to be sorted in four groups, namely, molecular/bonding, molecular/lattice, ionic, metallic	Shuffle your 16 cards and sort them in the four groups as stated on your worksheet: molecular/bonding, molecular/lattice, ionic, metallic. Every group should contain at least one card.
3	Unframed mapping task	Preuniversity students and first-year chemistry students	Participants receive questions of perspective for structure−property reasoning (see Figure 1). The questions should be complete by answers in form of chemical concepts. Creating hierarchy is allowed.	In front of you, you see the questions of the perspective for structure−property reasoning. A perspective is a way of questioning your topic or problem. Complete the questions with the appropriate chemical concepts. You are allowed to form a hierarchy. (An example is added, see the Supporting Information.)
4	Framed mapping task	Preuniversity students	Participants receive questions of perspective for structure−property reasoning (see Figure 1) and 30 chemical concepts. The concepts should be placed at the appropriate question. Creating hierarchy is allowed.	Again, you are given the questions of the perspective for structure−property reasoning. Complete the questions with the given 30 chemical concepts. You are allowed to form a hierarchy.

"nucleus particles" was divided further into "protons" and "neutrons".

A second researcher independently counted and judged the total number of answers, the number of corresponding answers, and the number of extra answers in the unframed maps of 38 participants (23%). The level of consistency among the researchers was then determined by comparing the determined numbers with those of the researcher. This resulted in 691 agreements out of a total of 808 answers. The percentage of agreements is therefore 86%. The differences between the two researchers were discussed until agreement was reached. An example of discussion was whether an extra answer was correct, like soluble in water and through the air given with the question about the movement of the particles. It was agreed that these answers were not counted as a correct extra answer, because it refers to the macro level.

The framed mapping task was analyzed on the number of correctly placed chemical concepts compared to the reference map in Figure 2.

To determine whether the results of the preuniversity students and the first-year chemistry students were different, *t* tests assuming unequal variances were performed. In addition, effect sizes (Cohen's *d*) were calculated by dividing the absolute difference between the means by the $\mathrm{SD}_{\mathrm{within}}$.

The names given to the groups in the unframed sort were analyzed by systematic coding. First the names were coded by type of category name with the codes "referring to structure", "referring to property", and "other". The category names were then analyzed at a deeper level, i.e., the given group names were subdivided into the categories of ideal sort as shown in Table 4. Interrater reliability was estimated by double coding a pseudorandom sample of 48 participants: 32 first-year students and 16 preuniversity students. 29% of the category names were double coded. Two raters assigned identical codes to 89% of the category names. The Cohen's κ was 0.83.

Table 4. Main Codes Used for Analyzing the Group Names on a Deeper Level

Aspects of Ideal Sort	Examples of Group Names
Property—Melting point	Melting point, boiling point, or both
Property—Solubility	Solubility, hydrophobic, hydrophilic
Property—Conductivity	Conductivity
Property—Hardness	Hardness, firmness
Structure—Molecular bonding	Hydrogen bond, van der Waals bond
Structure—Molecular lattice	Lattice, atomic lattice
Structure—Ionic	Ions, ionic bond
Structure—Metallic	Metal lattice

■ RESULTS

The analysis was focused on the comparison between preuniversity students and first-year chemistry students. The results are presented in reverse order, starting with the framed mapping task.

Framed Mapping Task

Sixty preuniversity students finished the framed mapping task in which they had to complete the questions from the perspective for structure−property reasoning (Figure 1) with 30 given concepts (ideal outcome shown in Figure 2).

Table 5 shows that the preuniversity students barely made any mistakes in this task. Out of 30 chemical concepts that had

Table 5. Results of the Framed Mapping Task (Preuniversity Students Only)

Group	Average Correct Answers	SD
Y-12	26.4 (88%)	3.18
Y-11	27.7 (92%)	1.88
Total	26.9 (90%)	2.79

to be placed, they obtained an average score of 27 correct answers (90%). This suggests that all these students recognized the questions of the perspective for structure−property reasoning. Furthermore, they were able to place the appropriate chemical concepts at the questions, both a precondition for proficient structure−property reasoning.

Unframed Mapping Task

Table 6 shows the results of the unframed mapping task. The average number of answers given in accordance with the reference map (Figure 2) was approximately the same for the preuniversity students (9.3 answers) as the first-year chemistry students (9.7 answers). However, the average number of answers (in total) and the average number of additional answers were significantly different. The first-year chemistry students gave more correct answers (23.5 answers) than the preuniversity students (17.7 answers) and, consequently, more additional answers (10.7 answers versus 5.9 answers). Moreover, 50% of the first-year chemistry students created a correct hierarchy in their answers compared to only 15% of the preuniversity students. These results indicate that the first-year chemistry students gave more elaborate answers to the questions related to structure−property reasoning (Table 6).

The additional answers given by the preuniversity students were mainly on the questions: "Which type of particle?", "How does the particle move?", and "How are the particles organized?". When answering the question "Which type of particle?", they often referred to the particles of the atomic model, neutrons, protons, and electrons.

A good proportion of the additional answers given by the first-year chemistry students were the same as those provided by preuniversity students. They also showed more repertoire for the questions "Which substance?" using terms such as acids, bases, and alloys. For the question "Which type of bond or force?", they used, for example, London dispersion forces, Coulomb interactions, and ion−dipole interactions. For the question "Which organization?", they used, for example, types of crystal lattices, and for the question "Which movement?", terms such as vibration, rotation, and translation. These additional answers were directly related to the topics and terms discussed in the General and Inorganic Chemistry course.

Recognition of the chemical concepts affiliated with the questions from Figure 1 was not very high in the case of both the preuniversity students and the first-year chemistry students. With six questions to be answered, the first-year chemistry students gave an average of four answers per question; the preuniversity students only three answers per question. However, the results show that first-year chemistry students gave more additional corresponding chemical concepts compared to preuniversity students. On average, the first-year chemistry students provided 4.8 additional answers. Furthermore, they incorporated more hierarchy into their answers. As data in Table 6 indicate that first-year chemistry students remembered, related, and identified more chemical concepts with the appropriate question, they are better equipped for structure−property reasoning.

Framed Sorting Task

Sixty preuniversity students and 106 first-year chemistry students completed the framed sorting task. Two first-year students did not complete the task as intended, and their framed sorts were excluded from the results.

The percentage of pairs (%P) for each participant was determined (Table 7 and Figure 4). The %P-structure was

Table 7. Results of the Framed Card-Sorting Task

		%P-Structure		%P-Property	
Framed Sort	N	Mean	SD	Mean	SD
Preuniversity students	60	35%	13%	24%	13%
First-year chemistry students	106	58%	15%	13%	6%

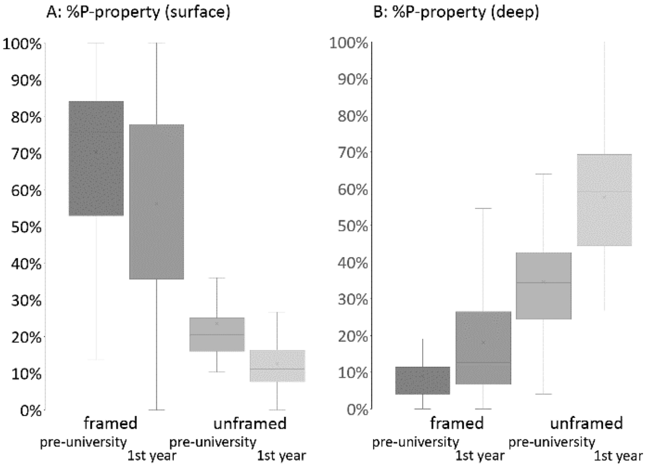


Figure 4. Percentage pairs (%P)-comparison of the preuniversity students' and the first-year chemistry students' sorts, both unframed and framed, relative to the ideal sort on property (A) and structure (B) aspects.

Table 6. Results of the Unframed Mapping Task

	Average Number of Answers Given (SD)	Average Number of Answers Given in Accordance with the Reference Map (SD)	Average Number of Additional Answers Given (SD)	Correct Hierarchy Present
Preuniversity students (N = 60)	17.7 (4.6)	9.3 (2.9)	5.9 (3.8)	9 students[a] (15%)
First-year chemistry students (N = 107)	23.5 (9.6)	9.7 (3.2)	10.7 (6.2)	53 students (50%)[b]
Significant?	Yes, p < 0.0001	No, p = 0.19	Yes, p < 0.0001	

[a]One preuniversity student made a hierarchy who was not judged as correct. [b]Two first-year chemistry students made a hierarchy who was not judged as correct.

Table 8. Results of the Unframed Card-Sorting Task

		Number of Groups		%*P*-Structure		%*P*-Property		Group Names: Words Referring to		
Unframed Sort	*N*	Mean	SD	Mean	SD	Mean	SD	Structure	Property	Other
Preuniversity students	60	4.8	1.28	9%	11%	70%	20%	8%	74%	18%
First-year chemistry students	107	5.5	1.65	18%	16%	56%	24%	40%	39%	21%

significantly lower for the preuniversity students, and the %*P*-property was significantly higher, compared to those for the first-year chemistry students (both $p < 0.0001$). The effect sizes (*E*) of both the %*P*-property as %*P*-structure were large (*E*-structure = 1.6; *E*-property = 1.2). These scores indicate that the first-year chemistry students were better at identifying and classifying the problems and categorizing and sorting them accordingly. This suggests that proficiency in structure−property reasoning may be higher in the first-year chemistry students.

Unframed Sorting Task

Sixty preuniversity students and 107 first-year students completed the unframed sorting task. One first-year student did not complete the task as intended and was excluded from the results. Table 8 gives a summary of the results of this unframed card-sorting task. In the unframed sort, the preuniversity students made 288 groups in total (mean 4.8 groups, SD = 1.28). The first-year chemistry students made 590 groups in total (mean 5.5 groups, SD = 1.65). There was no significant difference in the average between the two groups.

Analysis of the category names of the groups formulated by the participants (Table 8) showed that the first-year chemistry students more often used words relating to the structure aspect of the problem than the preuniversity students (43% versus 8%). For preuniversity students, words that referred to the property aspects of the problem dominated (74% versus 40%). Category names commonly given by the preuniversity students were conductivity (10%), hardness (15%), solubility (20%), and melting point (19%). "Density" as a category name also appeared (5%). Preuniversity students used "density" in the context of packing of particles, probably due to the Dutch word for density which also means packing or tightness. For categories specifying a structure aspect, preuniversity students mainly used a category name referring to molecular bonding (62%) or molecular lattice (17%).

The category names given by the first-year chemistry students also referred frequently to the four groups on the property level: conductivity (22%), hardness (7%), solubility (24%), and melting point (14%). The category name "density" was given by 3% of them. For categories specifying a structure aspect, first-year chemistry students used category names referring to molecular bonding (24%) and molecular lattice (24%).

The average %*P*-structure (Table 8 and Figure 4) of the first-year chemistry students (18%) was significantly higher ($p < 0.0001$) than that of the preuniversity students (9%). As anticipated, the average %*P*-property of the first-year chemistry students (56%) was significantly lower ($p < 0.0001$) than that of the preuniversity students (70%). The effect size (*E*) for both %*P*-structure and %*P*-property were considered as medium (*E*-structure = 0.63; *E*-property = 0.61). First-year chemistry students sorted the problems more on structure aspects. Preuniversity students sorted more on the property aspects of the problems. These findings corroborate previous research indicating that novices sort more on surface features, in this case the property aspects, and experts sort more on deep features, in this case the structure aspects.[29]

Despite first-year chemistry students using many more category names referring to a structure aspect, the value of %*P*-structure was not as high as expected. Evidently, first-year chemistry students still sorted the problems on property aspects rather than the structure aspect. For example, one student made a group with all the problems with the property aspect conductivity and named the group "Electrons". This could indicate that the student had misunderstood the concept. On the other hand, there may have been other levels of sorting possible which would indicate a certain proficiency in structure−property reasoning. To obtain insight into the proficiency of each student, it is important therefore to look at the individual scores.

■ DISCUSSION AND CONCLUSIONS

This article focuses on a tool for chemistry teachers to assess students' ability in structure−property reasoning: the SPR-instrument. The tool uniquely combines a sorting task and a concept mapping task, both framed and unframed. Results show that the SPR-instrument clearly discriminates between secondary school students on the preuniversity track and first-year chemistry students at university. The first-year chemistry students performed better on the instrument than the preuniversity students. In the unframed mapping task, the first-year chemistry students gave more elaborate and richer answers to the questions related to structure−property reasoning than the preuniversity students. In the framed sorting task, the %*P*-structure was higher for the first-year chemistry students, meaning that they sorted the problems more in line with the ideal structure sort compared to the preuniversity students. In the unframed sorting task, the first-year chemistry students used more category names referring to structure aspects for their formed categories, and their %*P*-structure was higher than that for the preuniversity students. We conclude that the aim to develop an instrument to discriminate in students' structure−property reasoning abilities was reached.

We also intended to critically review the instrument against the four design criteria. With respect to the first criterion, the SPR-instrument was indeed based on a rather comprehensive model for structure−property reasoning, namely, the perspective for structure−property reasoning (Figure 1). This model covers the aspects of structure−property reasoning needed in the chemistry curriculum for secondary education. In designing the SPR-instrument, we focused on the concepts needed for the last three years of the Dutch secondary education.[12] This perspective for structure−property reasoning was used as the base for the design of a sorting task and a mapping task. Although the SPR-instrument seems to work to discriminate between the educational levels included in this study, it is uncertain whether it could be extended for use in lower school years and for more experienced chemists, such as undergraduates and postgraduates. This could be done by adapting the situations in the sorting task and extending the questions and corresponding concepts in the perspective for structure–property reasoning and thus in the mapping task.

The used perspective in this study is one of four chemical perspectives identified by Landa et al.[15] The valence-shell perspective can be well-integrated in the perspective for structure–property reasoning. Further research should reveal to what extent the thermodynamic perspective and the kinetic perspective could be integrated in the perspective for structure–property reasoning.

Second, the SPR-instrument was designed, as intended, to assess structure–property reasoning at two levels of use. It consists of a framed and unframed mapping task to assess the level of reproductive use of structure–property reasoning, and a framed and unframed sorting task to assess the level of productive use of structure–property reasoning as analyzing, evaluating, creating, and predicting. Structure–property reasoning was described as hierarchical problem solving, and the perspective for structure–property reasoning could be used as a scaffold for this. The card-sorting task can be used to check whether students recognize a concrete problem as a certain type of abstract problem and its associated concepts. The mapping task can be used to check whether students have an adequate abstract structure of problems and subproblems and associated concepts.

One limitation of the study concerns the type of problems that were used in the sorting task. The problems were of the type to build an explanation for a structure–property relation. To address other aspects of structure–property reasoning, like making predictions and constructing models, other types of problems could be used. For example, a problem could be of the type of synthesizing: “Iron is a good material to make various utensils. How can the toughness of iron be increased?”.

Another limitation of the SPR-instrument concerns the level of abstraction at which students construct explanations. The SPR-instrument provides relatively quick insight into the types of models students possibly consider for explaining chemical phenomena. For example, the results of the (framed) sorting task showed that students consider the atomic lattice to explain the conductivity of graphite. However, the SPR-instrument does not invite students to actually construct this mechanistic explanation using this model. Therefore, teachers should regularly ask students to construct specific mechanistic explanations to verify that they can also adequately specify and use the chosen model.

The instrument also complied with the third criterion: cost-effectiveness. Preparation and administration of the instrument was not time-consuming. The average time a participant needed to complete the four tasks was 50 min. The instrument is suitable for large groups. In this study, the test was used in groups of 20 up to 110 participants. A disadvantage of the SPR-instrument is that analyzing the test results is complex and time-consuming. By using a computer applet for the test,[37] the analysis time for teachers could be reduced significantly.

Finally, the SPR-instrument is adaptable to a teacher’s own teaching goals. It can be adapted to the year group (for example to cater for the specific learning goals of year 6) and the proficiency level of the students by, for example, adding more or fewer concepts to the framed mapping task. The set of problems used for the sorting task can also be easily adapted to school year, proficiency level, or learning goals of participants, e.g., by taking problems from the textbook used in the chemistry class. In the framed sorting task, the categories could be chosen otherwise, for example, the questions of the perspective for structure–property reasoning could be used.

Furthermore, the tool could be used repeatedly to estimate the development of students’ structure–property reasoning. The bias which could occur because the student gets acquainted with the used problems and chemical concepts when performing the tasks on a regular basis can be reduced by adapting the chemical problems to more difficult ones. Furthermore, the number of chemical concepts or problems used could be expanded. When a student progresses in the curriculum and extends its structure–property reasoning, a growth in the number of chemical concepts could be expected. A disadvantage of this adapting and/or expanding of the items in the SPR-instrument could be a decrease in cost-effectiveness. Nevertheless, we think that repeatedly offering the (slightly adapted) SPR-instrument to students could give the teacher and the students insight into their progression in structure–property reasoning. To increase the cost-effectiveness, the framed mapping task could be dropped when using the SPR-instrument repeatedly.

Many curricula are topic centered. Using the perspective for structure–property reasoning ensures that concepts that are typically offered as fragmented are integrated into a perspective that facilitates structure–property reasoning. Using this perspective, teachers can build the curriculum to be more structure–property reasoning focused by making small adjustments. In another study, this was done for the introduction of the structure models for metals, salts, and molecular compounds by using POE (Predict-Observe-Explain) demonstrations and the perspective for structure–property reasoning as a scaffold.[38]

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.2c00234.

Cards for the sorting task, worksheets, and instruction (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

Marie-Jetta den Otter – *Leiden University Graduate School of Teaching (ICLON), Leiden University, 2300 AX Leiden, The Netherlands;* orcid.org/0000-0003-1886-7246; Email: m.den.otter@iclon.leidenuniv.nl

Authors

Ludo B.F. Juurlink – *Leiden Institute of Chemistry (LIC), Leiden University, 2300 RA Leiden, The Netherlands;* orcid.org/0000-0002-5373-9859

Fred J.J.M. Janssen – *Leiden University Graduate School of Teaching (ICLON), Leiden University, 2300 AX Leiden, The Netherlands*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.2c00234

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The research reported in this article was carried out within the Dudoc-Bèta program with financial support of the Dutch Ministry of Education, Culture and Science.

■ REFERENCES

(1) Sevian, H.; Talanquer, V. Rethinking Chemistry: A Learning Progression on Chemical Thinking. *Chem. Educ. Res. Pract.* **2014**, *15* (1), 10−23.

(2) Meijer, M. R. Macro-Meso-Micro Thinking with Structure-Property Relations for Chemistry. *An Explorative Design-Based Study*; Utrecht University, 2011.

(3) Talanquer, V. Progressions in Reasoning about Structure−Property Relationships. *Chem. Educ. Res. Pr* **2017**, 998.

(4) *Taking the Lead in Science Education: Forging Next-Generation Science Standards*; Achieve, 2010; pp 1−66.

(5) *The Next Generation Science Standards For States, By States*; National Academies, 2014. DOI: 10.17226/18290

(6) Gilbert, J. K.; Treagust, D. F. Introduction: Macro, Submicro and Symbolic Representations and the Relationship Between Them: Key Models in Chemical Education. In *Multiple Representations in Chemical Education*; Gilbert, J. K., Treagust, D., Eds.; Springer Netherlands: Dordrecht, 2009; pp 1−8. DOI: 10.1007/978-1-4020-8872-8_1

(7) Harrison, A. G.; Treagust, D. F. The Particulate Nature of Matter: Challenges in Understanding the Submicroscopic World. In *Chemical Education: Towards Research-based Practice*; Gilbert, J. K., De Jong, O., Justi, R., Treagust, D. F., Van Driel, J. H., Eds.; Springer Netherlands: Dordrecht, 2003; pp 189−212. DOI: 10.1007/0-306-47977-X_9

(8) Johnstone, A. H. Why Is Science Difficult to Learn? Things Are Seldom What They Seem. *J. Comput. Assist. Learn.* **1991**, *7* (2), 75−83.

(9) Gabel, D. L. Improving Teaching and Learning through Chemistry Education Research: A Look to the Future. *J. Chem. Educ.* **1999**, *76* (4), 548.

(10) de Jong, O.; Taber, K. S. The Many Faces of High School Chemistry. *Handb. Res. Sci. Educ. Vol. II* **2015**, 457−480.

(11) Talanquer, V.; Pollard, J. Let's Teach How We Think Instead of What We Know. *Chem. Educ. Res. Pr.* **2010**, *11* (2), 74−83.

(12) *Syllabus Centraal Examen, Scheikunde Vwo*; College voor Toetsen en Examens, 2016; Vol. *2016*.

(13) Tienken, C. H.; Goldberg, S.; Dirocco, D. Questioning the Questions. *Kappa Delta Pi Rec.* **2009**, *46* (1), 39−43.

(14) Janssen, F. J. J. M.; Westbroek, H. B.; Doyle, W.; van Driel, J. H. How to Make Innovations Practical. *Teach. Coll. Rec.* **2013**, *115* (7), 1−42.

(15) Landa, I.; Westbroek, H. B.; Janssen, F. J. J. M.; Muijlwijk-Koezen, J.; van Meeter, M. M. Scientific Perspectivism in Secondary-School Chemistry Education. *Sci. Educ.* **2020**, *29* (5), 1361−1388.

(16) Janssen, F. J. J. M.; Westbroek, H. B.; Landa, I.; Ploeg, B.; van der Muijlwijk-Koezen, J. Perspectives for Teaching About How Science Works. In *Nature of Science in Science Instruction*; McComas, W., Ed.; Springer International Publishing: Cham, 2020; pp 253−269.

(17) Ohlsson, S. *Deep Learning: How the Mind Overrides Experience*; Cambridge University Press, 2011.

(18) Nokes, T. J.; Schunn, C. D.; Chi, M. T. H. Problem Solving and Human Expertise. In *International Encyclopedia of Education*; 2010. DOI: 10.1016/B978-0-08-044894-7.00486-3

(19) Wimsatt, W. C. *Re-Engineering Philosophy for Limited Beings*; Cambridge: Havard University Press, 2007.

(20) Giere, R. N. *Scientific Perspectivism*; University of Chicago Press, 2010.

(21) Thagard, P. *The Cognitive Science of Science: Explanation, Discovery, and Conceptual Change*; MIT Press, 2012.

(22) Atkins, P. Chemistry's Core Ideas. *Chem. Educ. New Zeal.* **2010**, *2* (3), 8−12.

(23) Gillespie, R. J. The Great Ideas of Chemistry. *J. Chem. Educ.* **1997**, *74* (7), 862.

(24) College Board. AP Chemistry Course and Exam Description. https://apcentral.collegeboard.org/pdf/ap-chemistry-course-and-exam-description.pdf.

(25) Claesgens, J.; Scalise, K.; Wilson, M.; Stacy, A. Mapping Student Understanding in Chemistry: The Perspectives of Chemists. *Sci. Educ.* **2009**, *93* (1), 56−85.

(26) Novak, J. D.; Gowin, D. B. *Learning How to Learn*; Cambridge University Press, 1984; p 199. DOI: 10.1037/e362742004-020

(27) Ruiz-Primo, M. A.; Shavelson, R. J. Problems and Issues in the Use of Concept Maps in Science Assessment. *J. Res. Sci. Teach.* **1996**, *33* (6), 569−600.

(28) Lambiotte, J. G.; Dansereau, D. F.; Cross, D. R.; Reynolds, S. B. Multirelational Semantic Maps. *Educ. Psychol. Rev.* **1989**, *1* (4), 331−367.

(29) Chi, M. T. H.; Feltovich, P. J.; Glaser, R. Categorization and Representation of Physics Problems by Experts and Novices. *Cogn. Sci.* **1981**, *5* (2), 121−152.

(30) Taconis, R. Understanding Based Problem Solving: Towards Qualification Oriented Teaching and Learning in Physics Education = (Probleemoplossen Op Basis van Begrip). Thesis, Eindhoven University of Technology. 1995. DOI: 10.6100/IR439275

(31) Smith, J. I.; Combs, E. D.; Nagami, P. H.; Alto, V. M.; Goh, H. G.; Gourdet, M. A. A.; Hough, C. M.; Nickell, A. E.; Peer, A. G.; Coley, J. D.; Tanner, K. D. Development of the Biology Card Sorting Task to Measure Conceptual Expertise in Biology. *CBE Life Sci. Educ.* **2013**, *12* (4), 628−644.

(32) Bissonnette, S. A.; Combs, E. D.; Nagami, P. H.; Byers, V.; Fernandez, J.; Le, D.; Realin, J.; Woodham, S.; Smith, J. I.; Tanner, K. D. Using the Biology Card Sorting Task to Measure Changes in Conceptual Expertise during Postsecondary Biology Education. *CBE Life Sci. Educ.* **2017**, *16* (1), ar14.

(33) Kozma, R. B.; Russell, J. Multimedia and Understanding: Expert and Novice Responses to Different Representations of Chemical Phenomena. *J. Res. Sci. Teach.* **1997**, *34* (9), 949−968.

(34) Irby, S. M.; Phu, A. L.; Borda, E. J.; Haskell, T. R.; Steed, N.; Meyer, Z. Use of a Card Sort Task to Assess Students' Ability to Coordinate Three Levels of Representation in Chemistry. *Educ. Chem.* **2016**, *17* (2), 337−352.

(35) Krieter, F. E.; Julius, R. W.; Tanner, K. D.; Bush, S. D.; Scott, G. E. Thinking Like a Chemist: Development of a Chemistry Card-Sorting Task to Probe Conceptual Expertise. *J. Chem. Educ.* **2016**, *93* (5), 811−820.

(36) Cito. Central Exam Reports. https://www.cito.nl/onderwijs/voortgezet-onderwijs/centrale-examens-voortgezet-onderwijs/tools-en-informatie-voor-docenten/examenverslagen/examenverslag-2018.

(37) Chen, D. A.; Hoople, G. D.; Ledwith, N.; Burlingame, E.; Bush, S. D.; Scott, G. E. Exploring Faculty and Student Frameworks for Engineering Knowledge Using an Online Card Sorting Platform. *Int. J. Eng. Pedagog.* **2020**, *10* (1), 62−81.

(38) den Otter, M.-J.; Dam, M.; Juurlink, L. B. F.; Janssen, F. Two Design Principles for the Design of Demonstrations to Enhance Structure−Property Reasoning. *Education Sciences* **2021**, *11*, 504.