



Cite this: DOI: 10.1039/d0rp00216j

Investigating the impact of three-dimensional learning interventions on student understanding of structure–property relationships

Sonia M. Underwood, * Alex T. Kararo and Gabriela Gadia

The ability to predict macroscopic properties using a compound's chemical structure is an essential idea for chemistry as well as other disciplines such as biology. In this study we investigate how different levels of interventions impact the components of students' explanations (claims, evidence, and reasoning) of structure–property relationships, particularly related to boiling point trends. These interventions, aligned with Three-Dimensional Learning (3DL), were investigated with four different cohorts of students: Cohort 1 – a control group of students enrolled in an active learning general chemistry course; Cohort 2 – students enrolled in the same active learning general chemistry course but given Intervention 1 (a 3DL worksheet administered during class time); Cohort 3 – students enrolled in the same active learning general chemistry course but given Intervention 1 and Intervention 2 (a 3DL course exam question administered after instruction); and Cohort 4 – a reference group of students enrolled in a transformed active learning general chemistry curriculum in which 3DL is an essential feature and includes Intervention 1 and Intervention 2 as part of the curriculum. We found that Cohort 2 students (with the 3DL worksheet intervention) were more likely than the control group (Cohort 1) to correctly predict the compound with a higher boiling point as well as incorporate ideas of strength of intermolecular forces into their explanations of boiling point differences. When a 3DL exam question was given as a follow up to the 3DL worksheet, students in Cohort 3 were more likely than Cohorts 1 and 2 to correctly identify the claim. Further comparison showed that Cohort 4 (transformed general chemistry curriculum) were more likely than Cohorts 1–3 to also include the ideas of energy needed to overcome stronger forces for a more sophisticated explanation (50% of Cohort 4 students compared to 17–33% for Cohorts 1–3). In addition, 80% of Cohort 4 students were able to construct a correct representation of hydrogen bonding as a non-covalent interaction compared to 13–57% for the other three cohorts.

Received 21st July 2020,
 Accepted 25th November 2020

DOI: 10.1039/d0rp00216j

rsc.li/cerp

Introduction

Being able to use a compound's chemical structure to predict physical and chemical properties is fundamental not only in chemistry, (Cooper *et al.*, 2017a; Murphy *et al.*, 2012; The College Board, 2009) but also in other disciplines such as biology in which students need to use structures to predict function (AAAS, 2011; Tansey *et al.*, 2013). However, students frequently struggle to identify proper structure–property relationships (DeFever *et al.*, 2015; Shane and Bodner, 2006; Talanquer, 2018) or broader structure–function relationships (Kohn *et al.*, 2018). The study presented here is part of a larger project in which we are investigating how students develop an understanding of structure–property relationships (Cooper *et al.*, 2010, 2013, 2016;

Cooper *et al.*, 2012a; Cooper *et al.*, 2012b; Cooper *et al.*, 2015b; Crandell *et al.*, 2019; Kararo *et al.*, 2019; Underwood *et al.*, 2015, 2016; Williams *et al.*, 2015), particularly those related to boiling point trends. For students to correctly predict and explain relative boiling points for covalent compounds, many skills (*e.g.*, drawing a Lewis structure, determining its three-dimensional shape, and the intermolecular forces present within a substance) must be used and connected. Often students have difficulties with each of these steps and lack an understanding of why this topic is central to chemistry (Cooper *et al.*, 2010; Cooper *et al.*, 2012a). Previous research has shown that students often rely on instructor-driven or self-created heuristics (Cooper *et al.*, 2013), or “cognitive shortcuts” to answer these tasks (Cooper *et al.*, 2013; Maeyer and Talanquer, 2010). Though heuristics may get students to a correct answer, it is frequently the case that students lack the underlying reasoning to explain the causal mechanism for why a substance would have a higher boiling point (Cooper *et al.*, 2013; Maeyer and Talanquer, 2010; Schmidt *et al.*, 2009). Here we go beyond the

Department of Chemistry & Biochemistry and STEM Transformation Institute,
 Florida International University, Miami, FL 33199, USA.
 E-mail: sonia.underwood@fiu.edu

correctness of a students' boiling point ranking abilities and investigate the impact of different levels of interventions on general chemistry students' explanations of a structure–property relationship.

For the development of these interventions, we build upon prior research which has highlighted the need to assist students with developing a robust framework of knowledge that can be drawn upon appropriately when necessary. According to the *Education for Life and Work* report, deeper learning occurs when a person is able to not only apply their knowledge to new situations but also understand when, how, and why to appropriately apply that knowledge (National Research Council, 2012b). The *Framework for K-12 Science Education* (National Research Council, 2012a) (referred to as Framework in this paper) and *Next Generation Science Standards* (NGSS, 2013) call for this type of deeper learning of core ideas through the process of active participation in authentic disciplinary practices and crosscutting concepts (*i.e.*, Three-Dimensional Learning – 3DL). It has also been noted that the modified teaching practices should align with course evaluation, as assessments have been shown to influence how students approach studying and what they perceive as important (Crooks, 1988; Hora and Oleson, 2017; Momsen *et al.*, 2013; Scouller, 1998; Scouller and Prosser, 1994; Snyder, 1973; Struyven *et al.*, 2005). Therefore, in this study we implement different levels of 3DL interventions, which incorporate these ideas, to investigate their impact on students' explanations regarding boiling point trends.

The first intervention consists of a worksheet aimed at having students integrate together individual structure–property skills to predict and explain boiling point trends within a single activity. As further discussed below, these skills are not always used to explain phenomena and it was intended that this activity would assist students with constructing the larger picture of structure–property relationships. The second intervention consists of a constructed response (open-ended) question on the students' unit exam in their general chemistry course to determine how emphasizing this material on an exam impacts student explanations of boiling point trends. Lastly, the impact of these interventions was compared to a transformed curriculum (*Chemistry, Life, the Universe and Everything* – CLUE (Cooper and Klymkowsky, 2013)), discussed below, in which students are consistently asked to incorporate multiple core ideas to predict and explain phenomena. Therefore, this study investigates the effectiveness of different levels of these interventions using student responses to an end-of-course assessment task that requires them to use and integrate knowledge.

Theoretical frameworks – three-dimensional learning (3DL) and explanations

As part of the Framework (National Research Council, 2012a) and NGSS (2013), educators are urged to question (1) what do we really want students to know from a course/program (core ideas), (2) what do we want students to do with that knowledge

(scientific practices), and (3) what productive lenses or tools can be used to explore phenomenon within and across disciplines (crosscutting concepts (Cooper, 2020; Rivet *et al.*, 2016)). These three dimensions (core ideas, scientific practices, and crosscutting concepts) blended together are referred to as Three-Dimensional Learning (3DL) and when integrated throughout a curriculum, students should develop a deeper understanding of science (National Research Council, 2012a). With the 3DL framework, students are expected to anchor knowledge to larger core ideas within the discipline in order to deepen their understanding through the use and application of knowledge (Harris *et al.*, 2019; National Research Council, 2012a; NGSS, 2013). This differs from prior reform efforts which often separated content and “inquiry” whereas 3DL explicitly links these two aspects together. The 3DL framework can therefore be used to identify what should be learned and assessed and has shown to be an effective framework for larger transformational efforts at the K-12 level (Anderson *et al.*, 2018) and college level (Matz *et al.*, 2018).

In this study, we present one type of structure–property relationship, boiling point trends. In order to analyze students' knowledge and integration of their knowledge for that trend, we analyzed the data using the scientific practices within 3DL, engaging in argumentation/constructing explanations. While Cooper *et al.* (2012b) provided a learning progression for structure–property relationships and found improvements in students' abilities to construct structures and decode information, there is limited research on reasoning as it relates to structure–property relationships. Talanquer (2018) suggested progressions related to reasoning about structure–property appear slowly and gradually as students interact with knowledge and models. Students, even advanced ones, have difficulty connecting all the different scales to mechanistically explain the structure–property relationship.

In general, explanations are used to describe the occurrence of phenomena (Chin and Brown, 2000) and requires a deep understanding of scientific principles in order to integrate scientific knowledge with evidence and provide a coherent reasoning (Driver *et al.*, 2000; McNeill and Krajcik, 2008). An explanation typically consists of three components: claim, evidence and reasoning. Claim is the conclusion related to how or why a phenomenon occurs, the evidence is the scientific principles or data used to support the claim and the reasoning is used to link the claim and evidence through further explanation. It has been shown that while students may be able to easily make a claim, they can struggle with identifying appropriate evidence or being able to clearly link evidence to the claim through reasoning (Kuhn, 1991; Sadler, 2004; Sandoval and Millwood, 2005). Therefore, students should integrate these three components together to create a richer explanation. These components needed to construct an explanation are similar to those needed for an argument and while Toulmin's framework consists of more components (*e.g.*, warrants and backing (1958)), simplified versions of argumentation also have been found to consist of claim, evidence, and reasoning. For the purposes of this study, we are using explanation as the overarching term for both constructing an argument and explanation.

In this project, we used 3DL to investigate the impact of implementing different levels of interventions to promote and emphasize knowledge-in-use for structure–property relationships. Particularly, we wanted students to integrate their knowledge of intermolecular forces and energy to construct an explanation about why one substance would have a higher boiling point within a single activity. Lavery *et al.* have built upon the descriptions provided in the Framework, as part of a collaborative project with chemistry, biology, and physics faculty, to develop the *Three-Dimensional Learning Assessment Protocol* (3D-LAP) (2016), which can be used to characterize whether an assessment task has the potential to elicit a core idea, scientific practice, and/or crosscutting concept. This protocol has shown useful in creating and modifying assessment questions (Underwood *et al.*, 2018) and more recently in determining the impact of large-scale transformational efforts by characterizing course-level changes in chemistry, biology, and physics over a five-year period of time (Matz *et al.*, 2018). The 3DL framework was used to guide the development of this study to promote and assess knowledge-in-use and the 3D-LAP was used to align the interventions and end-of-course assessment to the 3DL framework. More information about how the 3D-LAP was used in this study is described in Appendix 1.

Research Question: How do different levels of 3DL interventions impact the components of students' explanations (claims, evidence, and reasoning) of a structure–property relationship?

Methods

Overview of study design

This study (Fig. 1) includes four different cohorts of students to investigate the impact of the levels of 3DL interventions on student components of an explanation related to a structure–property relationship as compared to a transformed curriculum in which 3DL is consistently incorporated throughout (referred to as CLUE in this study).

The students included in Cohorts 1–3 were enrolled in a general chemistry curriculum with a commercially available atoms-first approach textbook (Burdge and Overby, 2012) and online pre-instruction assignment system and online post-instruction homework system. This curriculum has about 200 students per section working through guided worksheets every class period in small groups of four. Multiple undergraduate

learning assistants (LAs) provide support by walking around the class to guide students through these worksheets. In addition, in-class response systems were used to ask students questions during class time (*i.e.*, Top Hat or clickers). While the content delivered in this course is traditional in nature (*i.e.*, follows a well preceded sequence of topics (Sienko and Plane, 1966)), how the material was presented during class time was reformed. That is, the pedagogy setting consisted of a flipped, active learning classroom approach in which students had access to lecture videos and preparation materials outside of class time and worked in groups of four during class time to complete guided worksheets designed by the institution itself. These worksheets were guided in that students were given learning objectives followed by information about the concepts and free-response questions.

Description of Cohort 1 students' structure–property unit

Cohort 1 students (referred to as GC), 2015–2016, serve as a control group in this study to characterize the original instructional impact on student understanding of structure–property relationships at the institution of interest. The original curriculum materials related to structure–property relationships described here is relevant for Cohorts 1–3, which consisted of about two weeks. The unit started with the construction of Lewis structures followed by determining electronic and molecular geometry, polarity, and the effect of intermolecular forces. Students were even asked to engage in activities such as constructing molecules in class using toothpicks and candies in order to help with three-dimensional visualization of molecules. At the end of this unit, Cohort 1 students (control group) devoted 1.5 class periods, 75 minutes total, to a lesson which involved the use of Lewis structures as models to predict chemical and physical properties such as boiling point trends. At the beginning of this lesson students completed a pre-instruction assignment, which included a reading assignment with closed-ended questions (*e.g.*, True/False and multiple-choice questions) followed by a post-instruction online homework assignment related to the content upon completion of the lesson. The post-instruction assignment also included closed-ended questions. This lesson on connecting structure and property relationships had students work through a guided worksheet during class time in small groups that included information about types of intermolecular forces (London dispersion forces, dipole–dipole interactions, hydrogen bonding, and ion–dipole interactions), with interspersed questions asking students to make claims such as ranking boiling points (Fig. 2, Example A), draw IMFs such as hydrogen bonding between ammonia molecules (Fig. 2, Example B), and provide explanations such as why dimethyl ether is a gas while ethanol is a liquid at room temperature (Fig. 2, Example C). This original worksheet was 14 pages in length with an introduction to the content providing representations along with questions for the students to answer. Students were asked to complete this original guided worksheet on structure–property relationships in small groups in which the instructor reviewed the worksheet as part of a whole-class discussion. With Cohort 1, the scientific practice of constructing

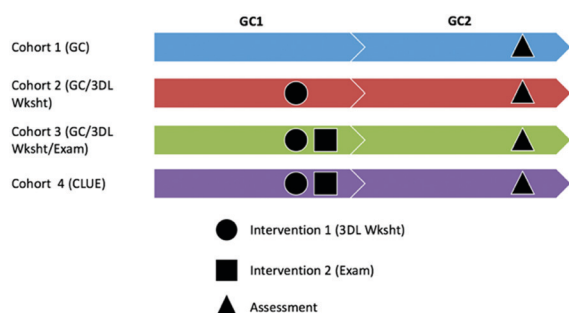
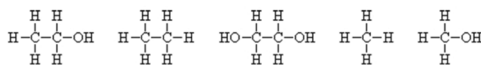


Fig. 1 Outline of study design.

Example A. Question asking students to make a claim on the guided worksheet.

6) Arrange the following in order of increasing boiling point



Example B. Question asking students to draw on the guided worksheet.

Example C. Question asking students to provide an explanation on the guided worksheet after given the name, formula, space filling model, molar mass, structure, boiling point and melting points of ethanol and dimethyl ether.

Explain why dimethyl ether is a gas at room temperature while ethanol, which has the same mass is a liquid?

Fig. 2 Example questions from the original general chemistry guided worksheet used in class for Cohort 1.

explanations was not emphasized in this original curriculum, meaning that students were not expected to include claim, evidence, and reasoning in their responses to any questions asked. Furthermore, these students were not explicitly assessed on representations of intermolecular forces or connecting structure–property relationships. Thus, it was important to determine the impact of bringing all of these skills together to predict and explain properties in addition to how emphasis on an assessment question influences explanations.

Description of Cohort 2 students' modified structure–property unit

Cohort 2 students (referred to as GC/3DLWksht), 2017–2018, had the same active learning curriculum described for Cohort 1 but replaced the original guided worksheet with Intervention 1 – a 3DL worksheet (Fig. 3). When adding in Intervention 1 to the active learning general chemistry curriculum for this study, the original guided worksheet was assigned as individual

- Draw out 3 molecules of CH_3OH (methanol) showing the strongest interactions that are present between the molecules.
 - What type of interaction did you show in your picture? Identify the important features of the interaction and how they were depicted in your picture. Revise your picture if necessary.
 - What intermolecular forces are present in liquid CH_3OH ?
- Draw out 3 molecules of CH_3OCH_3 (dimethyl ether) showing the strongest interactions that are present between the molecules.
 - What type of interaction did you show in your picture? Identify the important features of the interaction and how they were depicted in your picture. Revise your picture if necessary.
 - What intermolecular forces are present in liquid CH_3OCH_3 ?
- Draw out 3 molecules of CH_3CH_3 (ethane) showing the strongest interactions that are present between the molecules.
 - What type of interaction did you show in your picture? Identify the important features of the interaction and how they were depicted in your picture. Revise your picture if necessary.
 - What intermolecular forces are present in liquid CH_3CH_3 ?
- What would you predict would be the relative boiling points of methanol (CH_3OH), dimethyl ether (CH_3OCH_3) and ethane (CH_3CH_3)? Explain your answer, being sure to use the ideas of forces and energy.

Fig. 3 Intervention 1–3DL worksheet on boiling point ranking.

homework pre-instruction and followed up with a few clicker questions at the beginning of class.

The students worked in small groups during the class period on the 3DL worksheet in which the instructor discussed it at the end of the class period. Similar to Cohort 1, the LAs and course instructor walked around the classroom to guide students, while not providing them the answer directly. Before the discussion, LAs walked around and checked the worksheets for participation, but did not collect them so the students could reference them while the instructor discussed the worksheet with the whole class. Cohorts 1–3 spent similar amounts of class time connecting structure–property relationships; Cohort 1 spent 1.5 class periods or 75 minutes and Cohorts 2–3 spent one class period or 50 minutes. Essentially the same time was spent on this task when compiling all assignments and classroom activity for Cohorts 1–2 (Table 1). That is, what the students were asked to do in the classroom was what changed.

The purpose of this in-class 3DL activity worksheet (Intervention 1) was to have students apply and integrate their

Table 1 Summary of lesson implementation for Cohorts 1–3 where an X indicates that all three cohorts did the same assignment

	Structure–property unit for Cohorts 1–3 (GC)		
Cohort	Cohort 1	Cohort 2	Cohort 3
Length of unit	2 weeks	2 weeks	2 weeks
Pre-instruction textbook reading	X	X	X
Pre-instruction assignment	X	X	X
Original guided active learning worksheet	1.5 class periods (50 min each)	Given as pre-instruction homework	Given as pre-instruction homework
Intervention 1–3DL worksheet		1 class period (50 min)	1 class period (50 min)
Intervention 2–3DL exam question			5 pt question (5% of exam)
Post-instruction assignment	X	X	X

knowledge by providing components of an explanation (*i.e.*, claim, evidence, reasoning) for a boiling point trend. While the original guided worksheet (given in class to Cohort 1 and as homework to Cohorts 2–3) also provided students the opportunity to make claims (Fig. 2, Example A), draw hydrogen bonding (Fig. 2, Example B), and provide explanations (Fig. 2, Example C), these types of questions were split up throughout the worksheet and the specific questions were tied to different phenomena. Therefore, the difference in these two worksheets (original guided worksheet *versus* 3DL worksheet) lies in the organization and integration of the questions that students were asked. The 3DL worksheet was organized in a way that encouraged students to build visual models and use that model as well as their knowledge to explain the boiling point trend of one phenomenon. In contrast, the original guided worksheet asked students to apply individual skills (drawing IMFs and predicting boiling points), but were not organized in a way that required student to integrate these skills to explain the phenomenon nor was it scaffolded in way to explicitly show the integration between concepts.

The 3D-LAP coding of this worksheet (Appendix 1 – Table 6) has the potential to elicit core ideas of *atomic/molecular structure and properties*, *energy*, and *electrostatic and bonding interactions* through scientific practices of *constructing explanations/engaging in argument from evidence* and *developing and using models* and crosscutting concepts *structure and function* and *cause and effect*.

Description of Cohort 3 students' assessment emphasis

Cohort 3 students (referred to as GC/3DLWksht/Exam), 2016–2017, experienced the same curriculum and teaching methods as Cohort 2 who completed Intervention 1 (3DL worksheet). Cohort 3 students, however, were explicitly assessed on the connection of structure–property relationships, unlike Cohort 1 or 2, since material on exams can reinforce its importance to students (Hora and Oleson, 2017). The exam structure for Cohorts 1–3 typically consisted of 20 multiple-choice questions and one open-ended question worth a total of 105 points for each exam. For Cohorts 1–2 the open-ended question for the students' third exam in general chemistry 1 (GC1) consisted of drawing Lewis structures or a mathematical question unrelated to this topic. For Cohort 3 the open-ended question consisted of Intervention 2 (3DL exam question – Fig. 4), which was 5% of their total exam score. Again, Cohort 1–3 students essentially spent the same time on task for this unit (Table 1), which allows for comparison on the impact of the interventions since the classroom structure, other assignments, and instructors were consistent.

Description of Cohort 4 students' CLUE curriculum

Cohort 4 students (referred to as CLUE), 2016–2017, were enrolled in a transformed two-semester general chemistry curriculum called *Chemistry, Life, the Universe and Everything* (CLUE (Cooper and Klymkowsky, 2013)). Cohort 4 students serve as a comparison group to determine how the use of a single 3DL activity and exam emphasis compare to the impact of consistent use of 3DL throughout a curriculum. CLUE is

Fluoromethane (CH_3F) and propanol ($\text{CH}_3\text{CH}_2\text{CH}_2\text{OH}$) are both common substances, however propanol has a higher boiling point than fluoromethane.

- Draw the Lewis structure for fluoromethane and propanol.
- For a container of fluoromethane, what intermolecular forces would be present?
- For a container of propanol, what intermolecular forces would be present?
- Draw three molecules of CH_3F and then $\text{CH}_3\text{CH}_2\text{CH}_2\text{OH}$ showing how the strongest type of intermolecular forces act between the molecules.
- Using your responses to Parts A–D, explain why fluoromethane has a lower boiling point than propanol.

Fig. 4 Intervention 2–3DL course exam open-ended question.

an evidence-based general chemistry curriculum developed with consideration of how people learn (National Research Council, 1999) and the theory of learning progressions (Corcoran *et al.*, 2009; Duschl *et al.*, 2011).

With respect to structure–property relationships, CLUE starts in Chapter 1 examining boiling point trends using simple atoms (helium vs xenon) and molecules (hydrogen vs oxygen), introducing students to the concept of London dispersion forces, asking students to draw and interpret particulate level representations of the different phases, and requiring students to construct explanations with claim, evidence, and reasoning regarding the relative amounts of energy needed to overcome stronger London dispersion forces. For example, students are expected to explain that xenon has a higher boiling point than helium since it has a larger electron cloud which can be more polarizable creating a stronger temporary dipole interaction (London dispersion forces) between the two xenon atoms compared to the two helium atoms. Thus, more energy would be needed to overcome the stronger interactions between xenon atoms compared to helium atoms which would result in a higher boiling point for xenon. Additional complexity is added in Chapter 4 as students are introduced to more types of intermolecular forces such as dipole–dipole interactions and hydrogen bonding and expected to apply this knowledge to more complex molecules such as propane vs methane or water vs ammonia. This concept is further extended in Chapter 6 with the introduction of solutions and Chapter 7 with acid–base reactions. As described, the CLUE curriculum places significant and explicit emphasis on students using their understanding of forces between and within molecules to construct predictions, explanations, and models that link atomic–molecular structures to observable properties (*i.e.*, scientific practices).

Prior research has shown that students in the CLUE curriculum show significant improvements with their understanding of structure–property relationships: students were better able to self-report the connection of Lewis structures in predicting physical/chemical properties (Cooper *et al.*, 2012b; Underwood *et al.*, 2016), showed success in constructing Lewis structures (Cooper *et al.*, 2012b), successfully drew how multiple molecules can interact through intermolecular forces (Becker *et al.*, 2016; Williams *et al.*, 2015), and used these structures to provide causal

mechanistic understanding about why reactions happen (Crandell *et al.*, 2019). At the institution of interest, the CLUE curriculum was taught with an active learning approach where students worked in groups of six or nine depending on class size (either 174 or up to 300). Most of class time was spent with students working on worksheets in groups accompanied by mini-lectures and homework review.

The 3DL worksheet (Intervention 1) and 3DL exam question (Intervention 2) were developed as part of the CLUE curriculum structure–property learning progression (Cooper *et al.*, 2012b) and therefore, both of these interventions were given to the CLUE students as part of a larger progression of ideas. Cohort 4 exams typically consisted of 20 multiple-choice questions and three multi-step open-ended questions for a total of 110 points per exam in which Intervention 2 consisted of about 10% of the exam total.

Post-instruction 3DL structure–property assessment

The *structure–property assessment* administered to Cohorts 1–4 (Table 2) was adapted from a previous study which investigated how the question prompt can impact student responses (Kararo *et al.*, 2019) and has since been used to investigate student understanding in high school (Stowe *et al.*, 2019). The assessment was developed by combining several previously published assessment tasks. Question 1, *Implicit Information from Lewis Structures Instrument* (IILSI) (Cooper *et al.*, 2012a), asked students to self-report what type of information they can predict using a Lewis structure. Question 2 asked students to

select which structure has a higher boiling point (*i.e.*, claim) and make an argument as to why, a common task for general chemistry students. Students were then presented with macroscopic property information about ethanol and dimethyl ether as well as the correct claim with evidence as to why. Next students were asked to construct a representation of hydrogen bonding (Question 3), as previous research has shown that students have difficulty differentiating between bonding and intermolecular forces (Cooper *et al.*, 2015b; Peterson and Treagust, 1989; Schmidt *et al.*, 2009; Villafañe *et al.*, 2011) along with the misunderstanding that covalent bonds are broken during the boiling process (Bodner, 1991; Henderleiter *et al.*, 2001; Osborne and Cosgrove, 1983). Therefore, this question was important to determine whether students had the correct idea of hydrogen bonding as a non-covalent interaction between different ethanol molecules. Questions 4–5 asked students to represent dipole–dipole interactions and London dispersion forces (respectively). Questions 3–5 were taken from the *intermolecular forces assessment* (Cooper *et al.*, 2015b). Lastly, Question 6 provided students with their hydrogen bonding representation from Question 3 to use for their explanation of the phenomenon. The 3DL alignment of the *structure–property assessment* (Appendix 1 – Table 8) was coded as a single entity as recommended by the 3D-LAP (Lavery *et al.*, 2016) and is similar to Intervention 2 coding (Appendix 1 – Table 7). It is important to note that the explicit scaffolding/prompting to use the core idea of energy was removed from Question 6 to determine how students integrated these multiple core ideas into their

Table 2 The structure–property assessment (adapted from Kararo *et al.*, 2019) administered at the end of General Chemistry 2 (GC2)

Question number	Question details
Question 1: IILSI (Cooper <i>et al.</i> , 2012a)	“What information could you determine using a Lewis structure and any other chemistry knowledge you have? (Mark all that apply)” 20 answer choices including information such as “hybridization”, “element(s) present”, “Relative melting point”, etc.
Question 2: boiling point ranking task and reasoning	“If given the two compounds below [ethanol and dimethyl ether Lewis structures provided], which compound would you predict has the higher boiling point? Please explain your reasoning.”
Informational slide 1	Same chemical formula, same molecular mass; dimethyl ether is a gas and ethanol is a liquid at room temperature; given the boiling point of each substance
Informational slide 2	“The difference in properties between the two compounds is because ethanol can participate in hydrogen bonding.”
Question 3: draw hydrogen bonding for three molecules of ethanol (Cooper <i>et al.</i> , 2015b)	“Please draw and label a representation below in the box provided that clearly indicates where hydrogen bonding is present for three molecules of ethanol (CH ₃ CH ₂ OH). Please explain in words in the box provided, what you were trying to show in your drawing. Note: If you do not think this interaction is present, please write ‘not present’”
Question 4: draw dipole–dipole interactions for three molecules of ethanol (Cooper <i>et al.</i> , 2015b)	“Please draw and label a representation below in the box provided that clearly indicates where dipole–dipole interactions are present for three molecules of ethanol (CH ₃ CH ₂ OH). Please explain in words in the box provided, what you were trying to show in your drawing. Note: If you do not think this interaction is present, please write ‘not present’”
Question 5: draw London dispersion forces for three molecules of ethanol (Cooper <i>et al.</i> , 2015b)	“Please draw and label a representation below in the box provided that clearly indicates where London dispersion forces are present for three molecules of ethanol (CH ₃ CH ₂ OH). Please explain in words in the box provided, what you were trying to show in your drawing. Note: If you do not think this interaction is present, please write ‘not present’”
Question 6: explain why ethanol has a higher boiling point due to hydrogen bonding	“Now let’s go back to the comparison between the boiling points for ethanol and dimethyl ether. Using your representation of hydrogen bonding shown in the box below, explain in the black box why the ability of ethanol to form hydrogen bonds results in ethanol having a higher boiling point than dimethyl ether.”

explanation regarding the boiling point trend. The assessment was administered in *beSocratic* – an online program that allows students to draw representations and graphs as well as predict and explain phenomena (Bryfczynski, 2012; Cooper *et al.*, 2014) – at the end of GC2 as an extra credit assignment. The instructors notified students of the extra credit opportunity through their course management system and the activity typically took about 20–30 minutes to complete. The questions were asked sequentially one at a time and students were informed that they could not move backwards through the activity.

Student participants

This study includes university students enrolled in a two-semester general chemistry course at a large southeastern Hispanic-serving university with Carnegie classification of very high research activity. At the institution of interest there are about 1400 students who enroll in the first semester of general chemistry for the fall semester. Of these students, about 40% of them continue on to the second semester of general chemistry with about 500 additional students entering the second semester without taking the first semester immediately before at the same institution; primarily with reasons due to being a transfer student into the institution from the nearby two-year college, repeating a course, or taking a gap semester(s) between the two semesters of general chemistry. Therefore, only students enrolled in both general chemistry semesters consecutively at the institution were included to determine the impact of 3DL on structure–property relationships since this topic is introduced in the first semester of general chemistry and recapped in second semester. In addition, for Cohorts 1–3 only students from two instructors who consistently taught general chemistry at the institution were included to minimize the variation of non-permanent instructors on the findings of this study (*i.e.*, instructor effect).

Given the criteria listed above, the number of students eligible for this study consisted of about 250 students per group. However, since the assessment task was optional, about 40% of the students completed the task. Appendix 2 Table 9 highlights that the students who completed the assessment task performed similarly on their course grade for GC1 and/or GC2 to students who did not complete the assessment tasks. Therefore, the total number of students who participated in this study are: Cohort 1 ($N = 119$), Cohort 2 ($N = 106$), Cohort 3 ($N = 86$), Cohort 4 ($N = 70$). Pre-instruction assessments with regards to the students' understanding at the beginning of the course are often minimal with respect to intermolecular forces and structure–property relationships (Williams, 2015). In fact, students often get frustrated with not understanding what was asked of them prior to instruction, as these are open-ended questions, so all pre-test administrations of these types of assessment tasks were not conducted. However, students' ACT composite scores (Appendix 2 Table 10) were used to identify matched cohorts of students. In addition, if there were multiple scores for a single student only the highest ACT composite score was included. All students in this study were provided an informational page as part of the IRB approved protocol.

Analysis of the structure–property assessment at the end of GC2

The data analyses from the 3DL worksheet and 3DL exam question are not included in this study since the goal was to investigate how the different levels of 3DL interventions could impact students' construction of explanations at the end of GC2 as well as remove any variation related to grading by different instructors. In addition, the exam was not administered to all cohorts of students (only Cohorts 3 and 4 were administered the exam question) and thus cannot be used as a comparison across all cohorts. Instead, the 3DL structure–property assessment was administered at the end of GC2 to determine the impact of the different levels of interventions.

Within this assessment, students should ideally identify that ethanol has a higher boiling point than dimethyl ether (claim – Question 2, Table 2). This is because ethanol molecules can interact through London dispersion forces, dipole–dipole interactions, and hydrogen bonding whereas dimethyl ether molecules can interact only through London dispersion forces and dipole–dipole interactions (evidence or scientific principles). Therefore, (1) the intermolecular forces between ethanol molecules are stronger resulting in the ethanol molecules being more strongly attracted to each other as compared to the dimethyl ether molecules, (2) molecules that are more strongly attracted, or have stronger intermolecular forces, require more energy to overcome these interactions, and (3) due to the stronger intermolecular forces between the ethanol molecules compared to the intermolecular forces between the dimethyl ether molecules, ethanol requires more energy to overcome the stronger attraction resulting in a higher boiling point (reasoning – Question 6 (Cooper, 2015)).

Prior research has shown that students' explanations (Question 6), after the informational slides (Table 2) are presented, are more sophisticated than students' original arguments (Question 2) (Kararo *et al.*, 2019); therefore, Question 2 was only analyzed for the correctness of the students' original claim as to which substance has a higher boiling point. It was also important to determine students' views of hydrogen bonding (Question 3) since their explanations would center around why hydrogen bonding (evidence or scientific principle) resulted in ethanol having a higher boiling point. A previously developed coding scheme (Cooper *et al.*, 2015b) was used in which student drawings were coded as representing hydrogen bonding *within* the molecule (indicated as a covalent bond between hydrogen and oxygen), *between* molecules (a non-covalent interaction between different ethanol molecules), or *ambiguous* (there is not enough information provided to classify the student drawing as one of the other two categories) as shown in Appendix 3 – Fig. 8. The second and third authors of this paper completed inter-rater reliability and produced Cohen's kappa (κ) from 0.9 to 1.0. Any disagreements were discussed until consensus was reached.

Lastly, students' explanations (Question 6) were analyzed for students' reasoning to the boiling point ranking question using a previously established coding scheme as detailed in Appendix 3 – Table 11 (Kararo *et al.*, 2019). The codes increase

in completeness of a possible explanation that students could make: *hydrogen bonding*; *hydrogen bonding and strength of bonds/interactions*; and *hydrogen bonding, strength of bonds/interactions, and energy*. It should be noted that since students find the term “hydrogen bond” confusing, we combined student responses that used the terms “bond” and “interaction” since we are unable to decipher between the two within students’ written responses. In addition, the codes *does not know/no response* and *non-normative* (scientifically inaccurate) were included. Inter-rater reliability was assessed between the second and third authors of this paper which produced Cohen’s kappa (κ) from 0.8 to 1.0. Discussion occurred between any disagreements until consensus was reached. Chi-square test was used in the comparison of the data related to claim, evidence, drawing, and explanation to look for differences with each additional intervention due to the categorical and independent nature of the data.

Student responses to the ILSI (Question 1 – Table 2) were exported from *beSocratic* as dichotomous data (either students did not select the type of information – coded with a value of 0 or did select the type of information – coded with a value of 1) in the format of a csv file. These results are shown only in Appendix 4 as they were used to probe students’ abilities to self-report on the relationship between a Lewis structure and physical and chemical properties and are not directly tied to the specific boiling point phenomenon presented here. Although Questions 4 and 5 do apply to this phenomenon, they are not essential to the explanation being made for why hydrogen bonding results in ethanol having a higher boiling point and were therefore not examined as part of this study.

Results

RQ: How do different levels of 3DL interventions impact the components of students’ explanations (claims, evidence, and reasoning) of a structure–property relationship?

The goal of this study is to investigate how different 3DL interventions impact students’ construction of an explanation related to a specific structure–property relationship, boiling point trends. Therefore, we focus our results and discussion on Questions 2, 3, and 6 and present the results in terms of claim, evidence, and reasoning for the four cohorts of students: flipped, active learning original general chemistry curriculum (Cohort 1 – GC), same curriculum with Intervention 1 (Cohort 2 – GC/3DLWksht), same curriculum with Interventions 1 and 2 (Cohort 3 – GC/3DLWksht/Exam), and a different curriculum where 3DL is an essential feature of the curriculum (Cohort 4 – CLUE). Each of the findings below represent the percentage of student responses (frequency value of how many students) for the claim, evidence, and reasoning. Chi-squared analyses were conducted for each of these measures given the nature of the data and research question.

Claim – ethanol has a higher boiling point (Question 2)

As a first step it is essential for students to identify the correct claim. Fig. 5 shows that with each additional level of intervention,

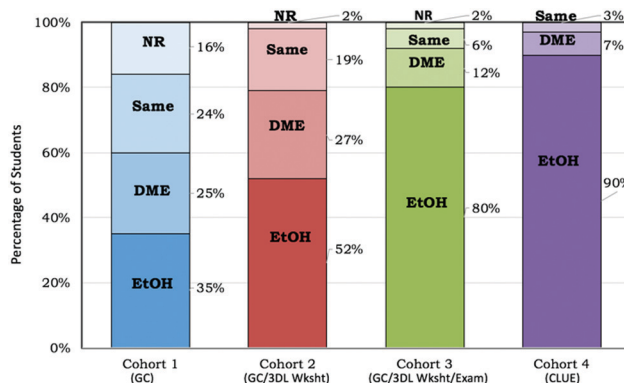


Fig. 5 Proportions of student responses to Question 2 on the 3DL structure–property assessment for all four cohorts of students. The claim options in figure from bottom to top for each cohort are: ethanol (denoted EtOH), dimethyl ether (denoted DME), same boiling point (denoted Same), or no response/cannot predict (denoted NR).

Table 3 Chi-squared analyses for each cohort comparison for Question 2 on students’ selection of ethanol as the correct claim

Cohort comparison	Chi-square	p-Value ^a	Effect size ^b (ϕ)
Cohort 1 and 2	15.5	<0.001	0.3
Cohort 2 and 3	6.3	0.012	0.2
Cohort 3 and 4	2.1	0.15	—

^a A Bonferroni correction was determined as three comparisons were performed so a p -value < 0.017 was considered significant instead of < 0.05. ^b Effect size interpretation of 0.1 for small, 0.3 for medium, and 0.5 for large (Cohen, 1988).

students became progressively more likely to correctly identify that ethanol would have the higher boiling point. That is, merely having students complete the 3DL worksheet (Intervention 1) resulted in a higher proportion of students’ selecting ethanol as the correct compound from 35% to 52%, which was further increased to 80% for Cohort 3 when these ideas were further emphasized on the course exam. Table 3 shows the statistical results in which a medium effect size is reported between Cohorts 1 and 2, a small to medium effect size between Cohorts 2 and 3, and no significant differences found between Cohorts 3 and 4. Therefore, overall a small change in instruction (*i.e.*, 3DL worksheet) and one 3DL open-ended question on an exam appear to have an impact on students’ identification of the compound with the highest boiling point (correct claim). This small difference between Cohort 3 and 4 could also be due to a ceiling effect in which most students correctly identified the claim in Cohort 3.

Evidence – correct representation of hydrogen bonding (Question 3)

Once students made an initial claim for this phenomenon, students were presented with the correct claim that ethanol has a higher boiling point and the evidence that it is due to hydrogen bonding. Research shows that students need to draw and write for instructors and researchers to develop a more complete understanding of the students’ knowledge (Becker *et al.*,

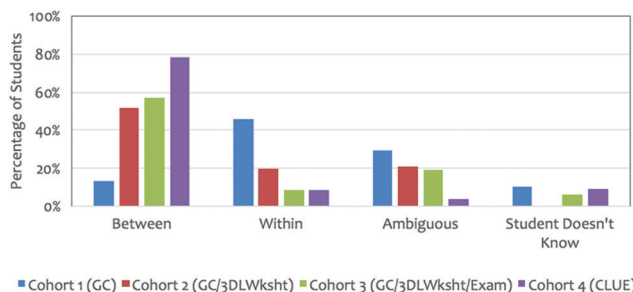


Fig. 6 Student representations of hydrogen bonding for three ethanol molecules *between* molecules represents a student response of a non-covalent interaction between different ethanol molecules (correct answer), a *within* the molecule indicates a student response of a covalent bond between hydrogen and oxygen atoms, *ambiguous* where there is not enough information provided to classify the student drawing as one of the other two categories.

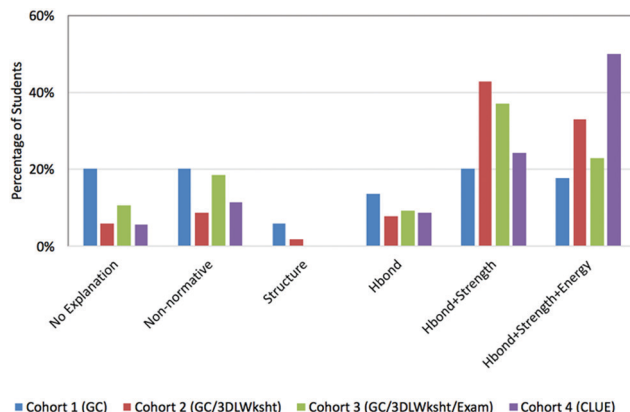


Fig. 7 Student responses for Question 6 boiling point ranking post-prompting explanation.

2016; Cooper *et al.*, 2017b; Cooper *et al.*, 2015b). Therefore, in this study we asked students to construct a representation of hydrogen bonding between three ethanol molecules (Question 3, Table 2). As shown in Fig. 6, less than 20% of the students in Cohort 1 could accurately represent hydrogen bonding as a non-covalent interaction between molecules. The students in Cohorts 2–4, however, appeared to have a better ability to draw this representation. That is, Intervention 1 (Cohort 2) increased the correctness of students' representations of hydrogen bonding (52%) with no additional gains for combined Interventions 1 and 2 (Cohort 3 – 57%) when compared to Cohort 2. Cohort 4 showed additional improvement in students' correctness of representing hydrogen bonding as a non-covalent interaction (79%). Table 4 reflects the significant difference found between Cohorts 1 and 2 with medium effect size, no significant difference between Cohorts 2 and 3, and significant difference between Cohorts 3 and 4 with small effect size.

Reasoning – why hydrogen bonding results in ethanol having a higher boiling point (Question 6)

Students were then asked to use their representation of hydrogen bonding to explain why ethanol had a higher boiling point due to hydrogen bonding. The level of sophistication in students' explanations were determined using the previously developed coding scheme as shown in Appendix 3 – Table 11 (Kararo *et al.*, 2019). In Fig. 7, 40% of Cohort 1 provided no

explanation or non-normative (*i.e.*, inaccurate scientific reasoning) responses for this phenomenon. The sophistication of student explanations through the incorporation of a comparison of relative strength of intermolecular forces increased with Intervention 1, the 3DL worksheet (76% for Cohort 2), with similar results for combined Interventions 1 and 2 (60% for Cohort 3). For the most sophisticated reasoning (*i.e.*, hydrogen bonding, comparison of strength of interactions, and more energy needed for stronger interactions), there was little change for Cohorts 1–3. There were, however, almost twice as many students in Cohort 4 (CLUE – 50%) incorporating all of these core ideas, including energy, without being prompted. Table 5 presents the most sophisticated type of reasoning (Hbond + Strength + Energy), in which a significant difference was found between Cohorts 1 and 2 with a small effect size, no significant difference between Cohorts 2 and 3, and significant difference between Cohorts 3 and 4 with a medium effect size.

Discussion

This study aimed to determine how levels of Three-Dimensional Learning interventions, particularly the emphasis on integrating student knowledge together for a complete picture of the phenomenon, affected their construction of an explanation (claim, evidence, and reasoning) related to a specific structure–property relationship. Thus, the discussion presented here is organized by cohort to discuss how the levels of 3DL interventions and emphasis on integrating student knowledge of these concepts

Table 4 Chi-squared analyses for code “between” which signifies a non-covalent interaction for the students' drawing of hydrogen bonding for three molecules of ethanol

Cohort comparison	Chi-square	<i>p</i> -Value ^a	Effect size ^b (Φ)
Cohort 1 and 2	37.4	<0.001	0.4
Cohort 2 and 3	0.3	0.58	—
Cohort 3 and 4	7.2	0.007	0.2

^a Note a Bonferroni correction was determined as three comparisons were performed so a *p*-value <0.017 was considered significant instead of <0.05. ^b Effect size interpretation of 0.1 for small, 0.3 for medium, and 0.5 for large.

Table 5 Chi-squared analyses for code “hydrogen bonding, strength of interactions, and energy”

Cohort comparison	Chi-square	<i>p</i> -value ^a	Effect size ^b (Φ)
Cohort 1 and 2	7.0	0.008	0.2
Cohort 2 and 3	2.2	0.15	—
Cohort 3 and 4	11.0	0.001	0.3

^a Note a Bonferroni correction was determined as three comparisons were performed so a *p*-value <0.017 was considered significant instead of <0.05. ^b Effect size interpretation of 0.1 for small, 0.3 for medium, and 0.5 for large.

impacted students' constructions of the components of an explanation related to a boiling point trend.

Cohort 1 students represent the control group for this study in that they were enrolled in the general chemistry curriculum (GC) at the institution of interest. These students were part of a general chemistry curriculum that was traditional in content but had a flipped active learning classroom approach in which students worked on guided worksheets in groups during class time. The majority of students in the control group incorrectly identified the claim (Fig. 5), incorrectly drew hydrogen bonding for liquid ethanol (Fig. 6), and provided no response/non-normative or less sophisticated reasoning for ethanol having a higher boiling (Fig. 7). With this curriculum, the instruction included a reading assignment and homework assignment given before class and 1.5 class periods spent working through a guided worksheet on structure–property relationships in small groups. While the worksheet did prompt students to make claims, draw, and provide explanations, these were isolated tasks throughout the worksheet and not the primary focus. That is, the original worksheet treated this relationship as individual tasks and were not scaffolded in a way that prompted students to integrate these skills for a single phenomenon as the 3DL worksheet did. In addition, this information was not emphasized on the students' exams.

Cohort 2 students experienced the same curriculum as Cohort 1 students with the movement of the original guided worksheet to individual homework and the completion of Intervention 1 (3DL Worksheet) during one class period in small groups (Fig. 3). Just this one 3DL worksheet had a notable impact on students' correctness of their claim (Fig. 5), correctly draw hydrogen bonding as between molecule interactions (Fig. 6), and provide an explanation that at least included strength of interactions between molecules (Fig. 7).

The inclusion of a 3DL exam question in addition to the 3DL worksheet appeared to improve Cohort 3 students' prediction of the correct substance with the higher boiling point at the end of the two-semester sequence (Fig. 5), but had little additional impact on the correctness of their drawings of hydrogen bonding interactions as non-covalent interactions (Fig. 6) or provide more sophisticated explanations beyond that of the 3DL worksheet alone (Fig. 7).

The consistent and explicit integration of skills and scientific practices, building visual models and constructing explanations, within the curriculum impacted Cohort 4 students' prediction of which substance has a higher boiling point similarly to the administered interventions of a 3DL worksheet and exam question (Fig. 5). In addition, it did significantly impact students' correctness in drawing non-covalent hydrogen bonding interactions (Fig. 6) and provide more sophisticated explanations to include the need for energy to overcome interactions as compared to the worksheet and exam question alone (Fig. 7). These students ultimately constructed more accurate components of an explanation (claim, evidence and reasoning) and integration of core ideas related to this phenomenon compared to the other cohorts. It is within the design of the CLUE curriculum that students are asked to use their understanding of multiple core ideas to

predict, model, and explain phenomena throughout the curriculum as a means for promoting knowledge-in-use (Cooper and Klymkowsky, 2013). Prior research has shown that students in CLUE are not only better able to identify a relationship between structure and physical and chemical properties (Underwood *et al.*, 2016) but also better able to draw molecular structures (Cooper *et al.*, 2012b) and represent intermolecular forces as non-covalent interactions (Williams *et al.*, 2015). Together, prior research and this study, provide evidence that individual interventions (*i.e.*, 3DL worksheet and exam question) do assist students in integrating knowledge for specific phenomena, as shown by the performance of Cohorts 2 and 3 with identifying a claim and drawing hydrogen bonding. However, continued student engagement with well-structured tasks that encourages students to apply and integrate knowledge in creating visual models and building causal explanations lead to a group of students being able to more accurately draw hydrogen bonding and provide more complete, in-depth explanation.

Conclusions

In summary, one well-structured task that required students to draw and use their knowledge to build causal explanations (3DL worksheet – Intervention 1) resulted in improvements of students' construction of an explanation. That is, Cohort 2 students were more likely to identify the correct claim for the boiling point ranking task, draw hydrogen bonding as non-covalent interactions, and use the idea of strength of interactions to explain why ethanol has a higher boiling point than dimethyl ether. When further emphasizing the integration of knowledge for one phenomenon on a course exam (combination of Interventions 1 and 2), the students only improved upon identifying the correct claim. While this improvement supports prior research in which including content on exams reinforces its importance to students (Hora and Oleson, 2017), the exam emphasis did not improve student accuracy in drawing hydrogen bonding or provide a more complete reasoning for the boiling point trend. The findings from the last group of students (Cohort 4) support that continued well-structured tasks focusing on building visual models and using that to build causal models is needed to improve students' construction of explanations of phenomena. Cohort 4 students were significantly more likely than the other cohorts to represent hydrogen bonding as a non-covalent interaction between ethanol molecules as well as incorporate multiple core ideas of bonding and interactions, structure and property, and energy into their explanation of boiling point trends to present a more coherent reasoning.

Implications for teaching and future work

There has been a strong push in recent years to reconsider the general chemistry curriculum (what is taught and how it is taught) to better serve students. With extensive research showing

students' difficulties with understanding core concepts in chemistry such as structure–property relationships, bonding and interactions, and energy, it would make sense that the integration of these core ideas would be given great significance and taught in a way that stresses long-term understanding for future chemistry and biology courses. However, this is often not the case. For example, in some general chemistry textbooks, the construction of Lewis structures is taught separate from the rationale for why bonds form as well as how molecules interact through intermolecular forces. Further these concepts may be isolated within the textbook itself by multiple chapters of unrelated content interspersed. Because topics are typically taught and assessed without being grounded to core ideas, typically as individual tasks rather than part of a working whole, it is not surprising that student understanding of structure–property relationships can be fragmented. On the other hand, curricula like CLUE, built around core ideas, starting with simpler systems and building to more complex, as well as three-dimensional learning can help students build a more cohesive understanding of chemistry.

The results of this study provide insights into the impact of well-structured tasks using the 3DL as a guiding framework to impact students' causal explanations and visual representations for a boiling point trend. This study also shows the use of 3DL integrated instruction and assessments consistently throughout a course supports students in constructing more complete and sophisticated explanations of these relationships. Thus, it is essential that we provide students with multiple opportunities to use their knowledge within the classroom and on exams. While multiple-choice questions may be ideal for large courses due to limitations of time available for grading and personalized feedback to students, it is necessary to ask students to draw and explain their understanding in the classroom. Otherwise students can successfully move through courses without ever being asked to draw or explain their understanding which could allow them to exit a course with a false or incomplete notion of the content. Future work can involve the investigation of how the incorporation of 3DL activities compared to the CLUE curriculum in different environments (active vs lecture setting) impact students' understanding of core ideas like structure–property relationships. Future work can also include the investigation of the coherence of student explanations as a whole.

Limitations

As with every study there are limitations that should be mentioned. First, the data from this study only came from one institution so we are unable to generalize how these results would appear beyond this institution. Meaning that other institutions initial data (control group) may differ from the results presented here; however, this is part of our future work to investigate the efficacy of the CLUE curriculum at multiple institutions. Second, while the first author did go to the LA meetings to talk the professors of Cohorts 2 and 3 as well as the LAs through the 3DL worksheet, the actual class period in

which the worksheet was administered was not observed. Because of this, we are unable to make any claims about the way in which the instructors administered the worksheet to their class. Third, the structure–property assessment administered at the end of the second semester of general chemistry was given as an extra credit opportunity; therefore, it is possible that students did not put maximum effort into that assignment or provided their complete understanding of boiling point trends. In our experience, however, students do typically take these assignments seriously as noted by genuine responses to open ended feedback on assignments (Noyes and Cooper, 2019).

Conflicts of interest

There are no conflicts of interest to declare.

Appendix 1

Three-dimensional learning assessment protocol (3D-LAP) coding of interventions and assessment

Laverty *et al.* have built upon the descriptions provided in the *Framework* (National Research Council, 2012a), as part of a collaborative project with chemistry, biology, and physics faculty, to develop the *Three-Dimensional Learning Assessment Protocol* (3D-LAP) (Laverty *et al.*, 2016) to characterize whether an assessment task has the potential to elicit a core idea, scientific practice, and/or crosscutting concept. Through faculty input, a list of core ideas (Cooper *et al.*, 2017a) was created based upon other national initiatives at the university level since the core ideas listed within the *Framework* were intended for K-12 curriculum. The four core ideas for chemistry identified in the 3D-LAP were *electrostatic and bonding interactions*, *atomic/molecular structure and properties*, *change and stability in chemical systems*, and *energy* (at the macroscopic, atomic/molecular, and quantum levels) (Laverty *et al.*, 2016). As part of the 3D-LAP a description of each core idea is provided, and the assessment is coded based on its alignment with that description. The scientific practices (*e.g.*, *developing and using models* or *engaging in argumentation*) and crosscutting concepts (*e.g.*, *structure and function* and *energy*) were used with minimal modification from the *Framework* (National Research Council, 2012a) to develop the criteria for the 3D-LAP, as they are applicable to the college level (Cooper *et al.*, 2015a). While assessments for crosscutting concepts were coded similarly to core ideas in which the task was coded based on its alignment with the description of the crosscutting concept, the dimension of scientific practices was coded differently. For each scientific practice, a list of criteria for that practice was provided and all criteria had to be met in order to be coded as that practice. For this study, the 3D-LAP was used to verify the alignment of our interventions to the framework of 3DL.

As part of the recommendation within the 3D-LAP, clusters of questions or questions with various parts should be coded as a single item. Therefore, Table 6 shows how Intervention 1 (3DL worksheet) as a single unit aligns with the 3D-LAP. When

Table 6 The 3D-LAP coding of Intervention 1 – the 3DL worksheet

Dimension	Type
Core idea	Atomic/molecular structure and properties Electrostatic and bonding interactions Energy – atomic/molecular & macroscopic
Scientific practices	Constructing explanations/engaging in argument from evidence Developing and using models
Crosscutting concepts	Cause and effect: mechanism and explanation Structure and function

Table 7 The 3D-LAP coding of Intervention 2 – the 3DL short-answer exam question

Dimension	Type
Core idea	Atomic/molecular structure and properties Electrostatic and bonding interactions
Scientific practices	Constructing explanations/engaging in argument from evidence Developing and using models
Crosscutting concepts	Cause and effect: mechanism and explanation Structure and function

Table 8 The 3D-LAP coding of the 3DL structure–property assessment administered at the end of GC2

Dimension	Type
Core idea	Atomic/molecular structure and properties Electrostatic and bonding interactions
Scientific practices	Constructing explanations/engaging in argument from evidence Developing and using models
Crosscutting concepts	Cause and effect: mechanism and explanation Structure and function

compared to Intervention 2 (Table 7) and the 3DL assessment (Table 8) at the end of the second semester of general chemistry the 3D-LAP coding is similar, except for one main difference regarding the core idea of energy. It is important to note that after the initial prompting for energy within Intervention 1 activity, the scaffolding for energy was removed to determine how students integrated multiple core ideas without prompting.

Appendix 2

Grade comparison for students who participated and students who did not participate

As stated above, specific criteria were considered when identifying students to participate in this study: (1) needed to be enrolled in consecutive semesters of GC1 and GC2 and (2) needed to be enrolled in a course taught by one of two main instructors who consistently taught GC1 and GC2. Therefore, only about 250 students per year were eligible to participate in this study. Of those students about 40% completed the extra credit assessment tasks. A non-parametric one sample Kolmogorov–Smirnov test was performed to determine if the data was normalized (p -values for each cohort were less than

Table 9 Mann–Whitney U analyses for students who did and did not participate within Cohorts 1–3

Cohort	GC1 results	GC2 results
Cohort 1	$U = 11\,404, z = -0.4, p = 0.7$	$U = 11\,403, z = -0.4, p = 0.7$
Cohort 2	$U = 4212, z = -0.7, p = 0.5$	$U = 3691, z = -2.1, p = 0.04$
Cohort 3	$U = 5361, z = -1.9, p = 0.05$	$U = 5595, z = -1.5, p = 0.1$

0.001); therefore, the null hypothesis that the student grades were normally distributed was rejected and instead Mann–Whitney U analyses were performed. Table 9 highlights how the students who took the assessment tasks received a similar course grade to the students who did not take the assessment tasks. For Cohort 1, 314 students were eligible (119 completed); for Cohort 2, 190 students were eligible (86 completed); for Cohort 3, 241 students were eligible (106 completed); for Cohort 4 89 students were eligible (70 completed). It should be noted that not all GC1 grades were accessible for Cohort 1 and therefore only students with data were included in the test below (about 83% of the students). For all of these comparisons the p -value for the Mann–Whitney U test results were above 0.05 except for Cohort 2 GC2 (these results presented a small effect size 0.15). Therefore, for this study we considered the students who participated in this study not statistically different from the students who did not participate in this study.

ACT comparison analysis for Cohorts 1–4

Before being able to determine the impact of the levels of 3DL interventions on students' construction of an explanation related to a boiling point trend, we needed to identify four matched cohorts of students. At this institution the student population is divided among whether the students have a score for the SAT or ACT; therefore, all students' SAT scores were converted to an ACT composite score to compare the students' scores. That is, when converting between the two standardized tests a range of SAT scores corresponds to a single ACT value. If the reverse would have been performed it would have been difficult to pinpoint a value within the SAT range that the ACT score would represent. Once the ACT composite scores were determined, a non-parametric one sample Kolmogorov–Smirnov test was performed to determine if the data was normalized (p -value was 0.042); therefore, rejecting the null hypothesis that the data was normally distributed. A Kruskal–Wallis test was then performed to determine if the students in Cohorts 1–4 were significantly different in their ACT composite scores [$\chi^2 = 9.688, p = 0.021, df = 3$]. Since there was a significant difference among the groups, further paired analyses of Mann–Whitney U were

Table 10 Mann–Whitney U analyses for Cohorts 1–4

Cohort comparison	Results
Cohorts 1&2	$U = 2092, z = -1.7, p = 0.08$
Cohorts 1&3	$U = 2691, z = -2.3, p = 0.02$
Cohorts 1&4	$U = 2835, z = -2.9, p = 0.004$
Cohorts 2&3	$U = 2371, z = -0.3, p = 0.8$
Cohorts 2&4	$U = 3397, z = -0.8, p = 0.4$
Cohorts 3&4	$U = 2414, z = -1.1, p = 0.3$

conducted to determine which groups were different (Table 10). The medians for Cohort 1 (GC) is 24, Cohort 2 (GC/3DL Wksht) is 23, Cohort 3 (GC/3DL Wksht/Exam) is 23, and Cohort 4 (CLUE) is 23. The only two significant differences found were between Cohorts 1&3 and Cohorts 1&4, with Cohort 1 having a higher ACT scores than Cohorts 3 and 4. Therefore, for the purpose of our study since Cohort 1 was the control group, the four groups were compared to determine the impact of the interventions.

Appendix 3

Analysis coding scheme for 3DL structure–property assessment

Here we present the previously established coding scheme for Question 3 for the representation of hydrogen bonding and Question 6 for the explanation of why ethanol has a higher

boiling point than dimethyl ether. Fig. 8 shows the coding scheme for Question 3 in Table 2. Table 11 shows the coding scheme for Question 6 in Table 2.

Appendix 4

IILSI responses for the chemical and physical properties for the 3DL structure–property assessment

Question 1 in Table 2 – Implicit Information from Lewis Structures Instrument (IILSI) (Cooper et al., 2012a)

The IILSI was administered to determine whether students self-report that Lewis structures can be used to predict various types of information. We present here only the five types of chemical and physical properties (Fig. 9) which have been previously found to be more difficult types of information for

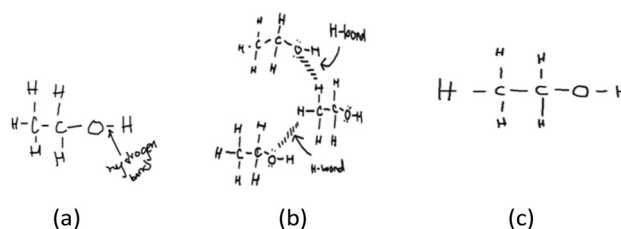


Fig. 8 Coding example for Question 3 student drawings of hydrogen bonding based on original coding scheme (Cooper et al., 2015b): (a) within the molecule, (b) between molecules, and (c) ambiguous.

Table 11 Previously established coding scheme for why ethanol has a higher boiling point than dimethyl ether (Questions 2&6) (Kararo et al., 2019)

Code	Definition	Examples of student responses
Student does not know/no response/can't be predicted	Student expresses that they do not know the answer, they do not provide any reasoning, or that boiling point cannot be predicted from a Lewis structure	"There needs to be more information for me to predict the boiling point." – Derek
Non-normative	Student uses scientifically inaccurate or unrelated reasoning	"Because ethanol will have more bonds than dimethyl ether, therefore have more bonds to break to reach a boiling point" – Lindsay
Hydrogen bonding	Student explicitly mentions hydrogen bonding in their reasoning	"Ethanol can form hydrogen bonds because they form between a H bonded to an O, N, F and the electron pair of another element (only N, O, F)" – Bryan
Hydrogen bonding and strength of bonds/interactions	Student explicitly mentions hydrogen bonding and compares strength of bonds/interactions	"Ethanol has a hydrogen bond which is much stronger and harder to break." – Destiny
Hydrogen bonding and strength of bonds/interactions and energy	Student explicitly mentions hydrogen bonding and strength of bonds/interactions; Student mentions energy in terms of it being higher/lower or more/less than another entity.	"The ethanol can form hydrogen bonds which are the strongest IMF and would require a lot of energy to break leading to a higher BP" – Tiffany

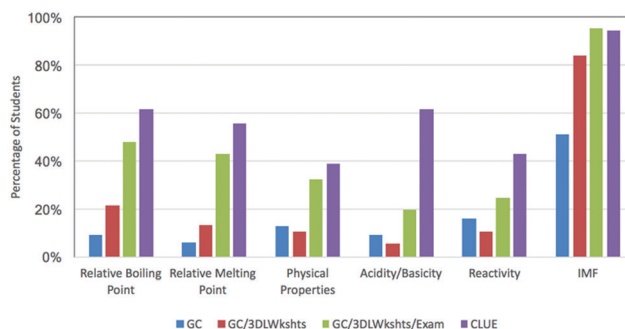


Fig. 9 Student responses to the IILSI (Question 1) for all four cohorts on the 3DL structure–property assessment.

students to predict (Underwood *et al.*, 2015). In addition, the item *intermolecular forces* (IMF) is included since it is needed for the explanation of why a liquid substance of ethanol has a higher boiling point than dimethyl ether.

Acknowledgements

This work is supported by startup funds from Florida International University, State of Florida, for support of the UP:LIFT project, NSF IUSE 1725609, the Howard Hughes Medical Institute Grant No. HHMI 52008097. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors and do not necessarily reflect the views of the funding agencies. The authors also want to thank Zahilyn D. Roche Allred for her assistance with the figures in this paper.

References

- American Association for the Advancement of Science (AAAS), (2011), in Brewer C. A. and Smith D. (ed.), *Vision and change in undergraduate biology education: A call to action*, American Association for the Advancement of Science.
- Anderson C. W., de los Santos E. X., Bodbyl S., Covitt B. A., Edwards K. D., Hancock J. B., Lin Q., Thomas C. M., Penuel W. R. and Welch M. M., (2018), Designing educational systems to support enactment of the Next Generation Science Standards. *J. Res. Sci. Teach.*, **55**(7), 1026–1052.
- Becker N. M., Noyes K. and Cooper M. M., (2016), Characterizing students' mechanistic reasoning about London dispersion forces. *J. Chem. Educ.*, **93**(10), 1713–1724.
- Bodner G. M., (1991), I have found you an argument: The conceptual knowledge of beginning chemistry graduate students. *J. Chem. Educ.*, **68**(5), 385–388.
- Bryfczynski S. P., (2012), *BeSocratic: An intelligent tutoring system for the recognition, evaluation, and analysis of free-form student input*, (UMI No. 3550201) Doctoral Dissertation, Clemson University.
- Burdge J. and Overby J., (2012), *Chemistry: Atoms First*, McGraw Hill.
- Chin C. and Brown D. E., (2000), Learning in science: A comparison of deep and surface approaches. *J. Res. Sci. Teach.*, **37**, 109–138.
- Cohen J., (1988), *Statistical power analysis for the behavioral sciences (Second)*, Lawrence Erlbaum Associates.
- Cooper M. M., (2015), Why ask why? *J. Chem. Educ.*, **92**, 1273–1279.
- Cooper M. M., (2020), The Crosscutting Concepts: Critical Component or "Third Wheel" of Three-Dimensional Learning? *J. Chem. Educ.*, **97**(4), 903–909.
- Cooper M. M. and Klymkowsky M. W., (2013), Chemistry, Life, the Universe and Everything: A new approach to general chemistry, and a model for curriculum reform. *J. Chem. Educ.*, **90**, 1116–1122.
- Cooper M. M., Grove N., Underwood S. M. and Klymkowsky M. W., (2010), Lost in Lewis structures: An investigation of student difficulties in developing representational competence. *J. Chem. Educ.*, **87**(8), 869–874.
- Cooper M. M., Underwood S. M. and Hilley C. Z., (2012a), Development and validation of the implicit information from Lewis structures instrument (IILSI): Do students connect structures with properties? *Chem. Educ. Res. Pract.*, **13**(3), 195–200.
- Cooper M. M., Underwood S. M., Hilley C. Z. and Klymkowsky M. W., (2012b), Development and assessment of a molecular structure and properties learning progression. *J. Chem. Educ.*, **89**(11), 1351–1357.
- Cooper M. M., Corley L. M. and Underwood S. M., (2013), An investigation of college chemistry students' understanding of structure–property relationships. *J. Res. Sci. Teach.*, **50**(6), 699–721.
- Cooper M. M., Underwood S. M., Bryfczynski S. P. and Klymkowsky M. W., (2014), A short history of the use of technology to model and analyze student data for teaching and research, in Cole R. S. and Bunce D. (ed.), *Tools of Chemistry Education Research*, American Chemical Society, pp. 219–239.
- Cooper M. M., Caballero M. D., Ebert-May D., Fata-Hartley C. L., Jardeleza S. E., Krajcik J. S., Lavery J. T., Matz R. L., Posey L. A. and Underwood S. M., (2015a), Challenge faculty to transform STEM learning. *Science*, **350**(6258), 281–282.
- Cooper M. M., Williams L. C. and Underwood S. M., (2015b), Student understanding of intermolecular forces: A multi-modal study. *J. Chem. Educ.*, **92**, 1288–1298.
- Cooper M. M., Kouyoumdjian H. and Underwood S. M., (2016), Investigating Students' Reasoning about Acid–Base Reactions. *J. Chem. Educ.*, **93**(10), 1703–1712.
- Cooper M. M., Posey L. A. and Underwood S. M., (2017a), Core ideas and topics: Building up or drilling down? *J. Chem. Educ.*, **94**(5), 541–548.
- Cooper M. M., Stieff M. and DeSutter D., (2017b), Sketching the Invisible to Predict the Visible: From Drawing to Modeling in Chemistry. *Top. Cognit. Sci.*, **9**(4), 902–920.
- Corcoran T., Mosher F. A. and Rogat A., (2009), *Learning progressions in science: An evidence based approach to reform (RR-63)*, Consortium for Policy Research in Education.
- Crandell O. M., Kouyoumdjian H., Underwood S. M. and Cooper M. M., (2019), Reasoning about Reactions in Organic Chemistry: Starting It in General Chemistry. *J. Chem. Educ.*, **96**(2), 213–226.
- Crooks T. J., (1988), The impact of classroom evaluation practices on students. *Rev. Educ. Res.*, **58**(4), 438–481.
- DeFever R. S., Bruce H. and Bhattacharyya G., (2015), Mental rolodexing: Senior chemistry majors' understanding of chemical and physical properties. *J. Chem. Educ.*, **92**(3), 415–426.
- Driver R., Newton P. and Osborne J., (2000), Establishing the norms of scientific argumentation in classrooms. *Sci. Educ.*, **84**, 287–312.
- Duschl R., Maeng S. and Sezen A., (2011), Learning progressions and teaching sequences: A review and analysis. *Stud. Sci. Educ.*, **47**(2), 123–182.
- Harris C. J., Krajcik J. S., Pellegrino J. W. and DeBarger A. H., (2019), Designing knowledge-in-use assessments to promote deeper learning. *Educ. Meas.: Issues Pract.*, **38**(2), 53–67.

- Henderleiter J., Smart R., Anderson J. and Elian O., (2001), How do organic chemistry students understand and apply hydrogen bonding? *J. Chem. Educ.*, **78**(8), 1126–1130.
- Hora M. T. and Oleson A. K., (2017), Examining study habits in undergraduate STEM courses from a situative perspective. *Int. J. STEM Educ.*, **4**(1), 1.
- Kararo A. T., Colvin R. A., Cooper M. M. and Underwood S. M., (2019), Making Predictions and Constructing Explanations: An investigation into introductory chemistry students' understanding of structure-property relationships. *Chem. Educ. Res. Pract.*, **20**(1), 316–328.
- Kohn K. P., Underwood S. M. and Cooper M. M., (2018), Connecting Structure–Property and Structure–Function Relationships across the Disciplines of Chemistry and Biology: Exploring Student Perceptions. *CBE—Life Sci. Educ.*, **17**(2), ar33.
- Kuhn D., (1991), *The skills of argument*, Cambridge, UK: Cambridge University Press.
- Laverty J. T., Underwood S. M., Matz R. L., Posey L. A., Carmel J. H., Caballero M. D., Fata-Hartley C. L., Ebert-May D., Jardeleza S. E. and Cooper M. M., (2016), Characterizing college science assessments: The Three-Dimensional Learning Assessment Protocol. *PLoS One*, **11**, e0162333.
- Maeyer J. and Talanquer V., (2010), The role of intuitive heuristics in students' thinking: Ranking chemical substances. *Sci. Educ.*, **94**, 963–984.
- Matz R. L., Fata-Hartley C. L., Posey L. A., Laverty J. T., Underwood S. M., Carmel J. H., Herrington D. G., Stowe R. L., Caballero M. D., Ebert-May D. and Cooper M. M., (2018), Evaluating the extent of a large-scale transformation in gateway science courses. *Sci. Adv.*, **4**(10), eaau0554.
- McNeill K. L. & Krajcik J., (2008), Chapter 11: Inquiry and Scientific Explanations: Helping Students Use Evidence and Reasoning, *Science as Inquiry in the Secondary Setting*, pp. 121–134.
- Momsen J., Offerdahl E., Kryjevskaja M., Montplaisir L., Anderson E. and Grosz N., (2013), Using Assessments to Investigate and Compare the Nature of Learning in Undergraduate Science Courses. *CBE-Life Sci. Educ.*, **12**(2), 239–249.
- Murphy K., Holme T. A., Zenisky A., Caruthers H. and Knaus K., (2012), Building the ACS Exams Anchoring Concept Content Map for Undergraduate Chemistry. *J. Chem. Educ.*, **89**(6), 715–720.
- National Research Council, (1999), *How people learn: Brain, mind, experience, and school*, National Academies Press.
- National Research Council, (2012a), *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*, National Academies Press.
- National Research Council, (2012b), *Education for life and work: Developing transferable knowledge and skills in the 21st century*, National Academies Press.
- NGSS Lead States, (2013), *Next Generation Science Standards: For States, By States*, The National Academies Press, <http://www.nextgenscience.org/ngss-high-school-evidence-statements>.
- Noyes K. and Cooper M. M., (2019), Investigating Student Understanding of London Dispersion Forces: A Longitudinal Study. *J. Chem. Educ.*, **96**(9), 1821–1832.
- Osborne R. J. and Cosgrove M. M., (1983), Children's conceptions of the changes of state of water. *J. Res. Sci. Teach.*, **20**(9), 825–838.
- Peterson R. F. and Treagust D. F., (1989), Grade-12 students' misconceptions of covalent bonding and structure. *J. Chem. Educ.*, **66**(6), 459.
- Rivet A. E., Weiser G., Lyu X., Li Y. and Rojas-Perilla D., (2016), *What Are Crosscutting Concepts in Science? Four Metaphorical Perspectives*, <https://repository.isls.org/handle/1/356>.
- Sadler T. D., (2004), Informal reasoning regarding socioscientific issues: A critical review of research. *J. Res. Sci. Teach.*, **41**, 513–536.
- Sandoval W. A. and Millwood K. A., (2005), The quality of students' use of evidence in written scientific explanations. *Cognit. Instr.*, **23**, 23–55.
- Schmidt H.-J., Kaufmann B. and Treagust D. F., (2009), Students' understanding of boiling points and intermolecular forces. *Chem. Educ. Res. Pract.*, **10**, 265–272.
- Scouller K., (1998), The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Educ.*, **35**(4), 453–472.
- Scouller K. and Prosser M., (1994), Students' experiences in studying for multiple choice question examinations. *Stud. Higher Educ.*, **19**(3), 267–279.
- Shane J. W. and Bodner G. M., (2006), General chemistry students' understanding of structure-function relationships. *Chem. Educ.*, **11**, 130–137.
- Sienko M. J. and Plane R. A., (1966), *Chemistry: Principles and Properties*, McGraw-Hill.
- Snyder B., (1973), *The Hidden Curriculum*, The MIT Press.
- Stowe R. L., Herrington D. G., McKay R. L. and Cooper M. M., (2019), The Impact of Core-Idea Centered Instruction on High School Students' Understanding of Structure–Property Relationships. *J. Chem. Educ.*, **96**(7), 1327–1340.
- Struyven K., Dochy F. and Janssens S., (2005), Students' perceptions about evaluation and assessment in higher education: a review. *Assess. Eval. Higher Educ.*, **30**(4), 325–341.
- Talanquer V., (2018), Progressions in reasoning about structure–property relationships. *Chem. Educ. Res. Pract.*, **19**(4), 998–1009.
- Tansey J. T., Baird T., Cox M. M., Fox K. M., Knight J., Sears D. and Bell E., (2013), Foundational concepts and underlying theories for majors in “biochemistry and molecular biology.” *Biochem. Mol. Biol. Educ.*, **41**(5), 289–296.
- The College Board, (2009), *College Board Standards for College Success: Science*, The College Board, <https://apcentral.collegeboard.org/pdf/cbscs-science-standards-2009.pdf?course=ap-physics-c-electricity-and-magnetism>.
- Toulmin, S. E., (1958), *The uses of argument*, Cambridge, UK: Cambridge University Press.
- Underwood S. M., Reyes-Gastelum D. and Cooper M. M., (2015), Answering the questions of whether and when student learning occurs: Using discrete-time survival analysis to

- investigate how college chemistry students' understanding of structure-property relationships evolves. *Sci. Educ.*, **99**(6), 1055–1072.
- Underwood S. M., Reyes-Gastelum D. and Cooper M. M., (2016), When do students recognize relationships between molecular structure and properties? A longitudinal comparison of the impact of traditional and transformed curricula. *Chem. Educ. Res. Pract.*, **17**(2), 365–380.
- Underwood S. M., Posey L. A., Herrington D. G., Carmel J. H. and Cooper M. M., (2018), Adapting Assessment Tasks To Support Three-Dimensional Learning. *J. Chem. Educ.*, **95**(2), 207–217.
- Villafañe S. M., Bailey C. P., Loertscher J., Minderhout V. and Lewis J. E., (2011), Development and analysis of an instrument to assess student understanding of foundational concepts before biochemistry coursework. *Biochem. Mol. Biol. Educ.*, **39**, 102–109.
- Williams L. C., (2015), *Students' Understanding of Structure-Property Relationships and the Role of Intermolecular Forces*, PhD Dissertation, Michigan State University.
- Williams L. C., Underwood S. M., Klymkowsky M. W. and Cooper M. M., (2015), Are Noncovalent Interactions an Achilles Heel in Chemistry Education? A Comparison of Instructional Approaches. *J. Chem. Educ.*, **92**(12), 1979–1987.