JOURNAL OF CHEMICAL EDUCATION

Article

Cite This: J. Chem. Educ. 2019, 96, 1821−1832

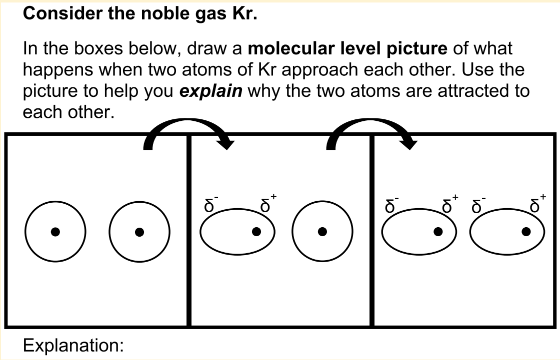
pubs.acs.org/jchemeduc

Investigating Student Understanding of London Dispersion Forces: A Longitudinal Study

Keenan Noyes and Melanie M. Cooper*

Department of Chemistry, Michigan State University, 578 South Shaw Lane, East Lansing, Michigan 48824, United States

Supporting Information


Consider the noble gas Kr.

In the boxes below, draw a **molecular level picture** of what happens when two atoms of Kr approach each other. Use the picture to help you ***explain*** why the two atoms are attracted to each other.

Explanation:

ABSTRACT: London dispersion forces (LDFs) are the most fundamental of the noncovalent interactions in that they are interactions that act between all molecular species and require only that students can use the "electron cloud" model of atomic structure. In this longitudinal study we investigate how students respond to prompts, designed to elicit causal mechanistic reasoning about the origins of LDFs, using a modified coding scheme from a previous study. We find that including explicit scaffolding improves students' representations of the mechanism compared to student responses to an unscaffolded prompt. When, on a later administration of a similar prompt, the scaffolding is removed, the sophistication of the students' representations decreases, but does not decline to the level of students who had never been exposed to a scaffolded prompt. We also explore the connection between student drawing and writing and find that there is a significant association between the reasoning level of the drawing and the written explanation. Finally, we show that there appears to be no significant difference between performance on ungraded (formative) and graded (summative) items that are administered in the second semester when LDFs are not explicitly covered. That is, students respond to ungraded homework activities similarly to examinations at time points where LDFs are not explicitly discussed in the course.

KEYWORDS: *First-Year Undergraduate/General, Chemical Education Research, Curriculum, Testing/Assessment, Noncovalent Interactions*

FEATURE: Chemical Education Research

■ INTRODUCTION

Understanding the causes and consequences of intermolecular forces (IMFs) (or more generally noncovalent interactions) is crucial to an understanding of the macroscopic behavior of molecular species. From phase changes to the behavior of complex biological systems, the interactions and concomitant energy changes that arise from transient or permanent electron density differences play a central role in predicting and explaining how molecular level systems behave. Unfortunately, there is ample evidence that the idea that molecules (or parts of molecules) can "stick together" without being permanently bonded appears to be difficult for students.[1,2] In prior work we found that students in traditional general chemistry courses rarely were able to draw and explain the types of IMFs typically taught in general chemistry in a coherent fashion.[3,4] While students could describe IMFs in textbook terms, they were unable to translate these ideas into a visual representation. That is, many students drew not only hydrogen bonding, but also dipole−dipole interactions and London dispersion forces (LDFs), as being located within a (small) molecule, rather than between molecules.[3] This may well be one source of the problematic ideas about how substances undergo phase changes. For example, if students are told that when water boils hydrogen bonds are broken, then we cannot be surprised when they interpret phase changes as breaking covalent bonds.[5−7]

While the mechanisms by which the electron density of participating molecules becomes unequally distributed are somewhat different for each of these IMFs, the ultimate consequence is the same. For example, if the temperature of the system is reduced, partial charges on different molecules (or different parts of the same molecule) will attract each other, and when the kinetic energy of the molecules is reduced sufficiently, this will eventually lead to a phase change (e.g., liquid to solid) during which IMFs are formed with the release of energy to the surroundings. Overcoming such IMF interactions requires an input of energy, resulting in a phase change (e.g., solid to liquid). Extending this reasoning to chemical reactions leads to an understanding of how the reacting molecules are attracted to initiate the bond breaking/bond formation steps of a chemical reaction. For this reason, we believe that helping students understand both the causes and consequences of IMFs is crucial for students to have a deep and useful knowledge of chemistry.

Received: May 14, 2019
Revised: July 17, 2019
Published: August 1, 2019

DOI: 10.1021/acs.jchemed.9b00455
J. Chem. Educ. 2019, 96, 1821−1832

There are two main mechanisms by which molecules acquire an unequal distribution of electrons, which then results in attractions between molecules. Understanding how molecules acquire a permanent dipole, resulting in hydrogen bonding and dipole−dipole interactions, requires a long chain of inference involving the relative effective nuclear charge of the participating atoms and the overall shape of the molecule. That is, to understand these IMFs students must have a working knowledge of bonding, VSEPR, molecular shape, and electronegativity.[8] Although students can (and do) take cognitive shortcuts such as identifying electronegative atoms to determine whether a molecule is polar, these shortcuts may lead to erroneous ideas such that carbon dioxide is polar.[9] The language used to refer to hydrogen bonding adds another problematic issue, since hydrogen "bonds" are not covalent bonds, which is also confusing to students.[1−3]

In contrast, understanding how LDFs arise requires less supporting knowledge of molecular structure, shape, and electronegativity. While students must be able to use an "electron cloud" atomic model, they need only understand that this electron density can fluctuate producing transient charges which may then induce fluctuations in nearby molecules, causing these molecules to interact. For this reason, LDFs are introduced early in the transformed general chemistry curriculum, "Chemistry, Life, the Universe & Everything" (CLUE).[10,11] LDFs are important noncovalent interactions; they operate between all molecular species, and while they are often discussed as being the "weakest" of all the IMFs, in fact they are dependent on molecular size and provide a significant contribution to the total interaction forces for large molecular species.[12]

Causal Mechanistic Reasoning and LDFs

Mechanistic reasoning "is a powerful thinking strategy that allows one to explain and make predictions about phenomena."[13] Russ and co-workers propose that "...mechanisms account for observations by showing that underlying objects cause local changes in the system by acting on one another".[14] Krist et al. in their proposed framework provide epistemic heuristics to support mechanistic reasoning.[13] They emphasize that an essential piece of a mechanistic explanation is the term "underlying objects"; that is, mechanistic reasoning involves thinking across scalar levels, and the "underlying objects" are at least one scalar level below the phenomenon to be explained. In chemistry, we are often discussing phenomena at the atomic−molecular level (e.g., why do molecules react?); in our own work then we have taken the underlying objects at a scalar level below molecules to be electrons. That is, a mechanistic explanation in chemistry should involve the behavior of electrons. While this may align with the organic chemistry idea of mechanisms, it involves much more than simply describing (or drawing) how electrons move during molecular level events but should also encompass the causal reasoning for this behavior.[14] Designing activities that require such mechanistic reasoning can help students connect the chain of inferences that are required to use molecular level structure to predict and explain observable macroscopic events.[8,15,16]

While some researchers use the term mechanistic thinking to encompass both the causes and the underlying mechanisms by which students can reason, in our work on chemical phenomena we have seen that students sometimes provide causes without mechanisms and mechanisms without causes.[15,16] For example, in our previous work on the development of student reasoning about LDFs,[15] we saw that some students explained the idea that nonpolar substances can "stick together" as the attraction between oppositely charged parts of the individual molecules or atoms. That is, the students are able to identify the cause of the attraction, but not the mechanism by which the charge separation in the nonpolar molecule arises. For example, "Atoms attract to each other [because] the [London] dispersion forces make the partially negative end of one atom attract to the partially positive end of another atom."[15] In contrast a causal mechanistic explanation involves the development of an instantaneous dipole as electron density rapidly fluctuates. The electrons are the "underlying entities" that produce the cause of the interaction. For example, "The atoms attract because of the London dispersion forces between the two atoms in which an atom with [instantaneous] dipole causing another atom to become an induced dipole. In this scenario, one atom has its electrons shift to certain region of the electron cloud, giving that region a negative charge. This negative region then induces the movement of electrons in another atom, being attracted to the new present positive region of the second atom."[15]

Alternatively, in our studies on student understanding of acid−base chemistry, some students are able to articulate a mechanism using electron movement (that is they use a Lewis acid−base model) but do not explain why the electrons move in that particular way as a result of the electrostatic interactions.[16] For example, "The O in the H_2O gives its electrons to the H in the HCl bond, and simultaneously the HCl bond breaks, placing those electrons onto the Cl. This reaction happens because it is more favorable."[16] In contrast, a causal mechanistic explanation explains why the electrons move in that way, "The lone pair on the water molecule attracts the Hydrogen from the HCl. The H−Cl bond is broken and forms a new bond with oxygen. The reaction occurs because the partial negative charge on the oxygen attracts the partial positive charge on the hydrogen."[16]

Because of these nuances in which causality and mechanism may appear separately in student explanations, we have chosen to separate these ideas in the coding of the explanations. So, while some researchers include causality when they discuss mechanistic explanations, in our work we refer to causal explanations, mechanistic explanations, and causal mechanistic explanations where appropriate, with an implicit assumption that causal mechanistic explanations are the most sophisticated.[15,16] In the case of student reasoning about LDFs, we see only causal and causal mechanistic reasoning, because in our earlier work[15] we did not see any students discussing the mechanism for the separation of charge without also discussing the consequences: That is, the resulting dipoles are attracted to each other. In contrast, in our characterization of acid−base reaction reasoning,[16] we did see students discussing a mechanism, including electron movement, without discussing the cause of that movement.

Designing the Task

As we develop approaches to helping students' reason about molecular level phenomena, we have developed tasks that prompt students to think about the causes for, and the mechanisms by which, these phenomena occur. There are several reasons why we have chosen to focus on tasks that involve causal mechanistic reasoning:

(1) Structuring the prompt so that students have to think both about how a phenomenon occurs and why that phenomenon occurs sends a signal that these are the thinking skills that are valuable and necessary to understand chemistry. If we consider causal mechanistic reasoning to be a valued form of scientific thinking, then it is important that we assess it, as that is what students will find important.[17]

(2) When used as an assessment (either formative or summative), these types of tasks can provide more persuasive evidence about what students know and can do. According to the principles of evidence centered design,[18,19] assessment can be considered as an evidentiary argument. That is, assessments should be designed to elicit evidence about what students know and can do with regard to the target of the assessment. Student responses to these tasks can be characterized with an appropriate coding scheme to identify how students reason about the phenomenon under consideration.

(3) These tasks provide polytomous student responses. That is, the reasoning students provide is not right or wrong, but can be scored or coded along a range of sophistication. By using such tasks at different time points in a students' education, we are able to detect how student reasoning changes over time.

However, designing an activity that elicits causal mechanistic reasoning is not a simple task. We know that if we ask students to simply predict or rank without providing an explanation, they default to system I type thinking.[20] That is, they use readily available heuristics rather than thinking about the underlying causes and mechanisms of the phenomena they are being asked about.[9] In order to help students reason causal mechanistically, we look to Hammer's resources framework.[21]

This framework posits that students do not necessarily have coherent conceptions (or misconceptions), but rather a wealth of resources that are activated in particular moments.[21] By focusing on what students know and can do (rather than what they do not know), we can identify the resources that students are using to construct models and explanations. Ideally this might allow us to help students produce more sophisticated and complete explanations if we can help students link their ideas in productive ways.

In this context, if we are asking students to explain LDFs causal mechanistically, we need to design the prompt to activate the appropriate conceptual resources (i.e., subatomic particles and electrostatic interactions). In chemistry, this may provide additional challenges. This framework was originally developed in the context of macroscopic physics ideas, and therefore, the resources required to provide a coherent explanation are ideas they may have experienced in their everyday life.[21] In chemistry, though, the phenomena are often taking place at the unseen atomic−molecular level governed by ideas of forces and energy that are notoriously difficult.[22−25] Therefore, we are asking students to leverage resources that are not only counterintuitive (forces and energy) but also must have been developed via some kind of instruction in chemistry rather than through everyday experiences (electrons, atoms molecules, electron distributions). That is, providing causal mechanistic reasoning in explanations is particularly difficult for chemical phenomena because students must use entities and ideas that they cannot experience at the macroscopic level and that require a good deal of prior instruction to engage with. To help students reason causal mechanistically about LDFs, we need to scaffold the prompts carefully to activate the appropriate resources.

Scaffolding is a process introduced initially by Wood et al.[26] as the process by which someone more knowledgeable is able to help someone less knowledgeable accomplish a task they would otherwise be unable to. While Wood et al.[26] were describing this process in the context of tutoring, since then many have applied the principles of scaffolding, which also aligns with Vygotsky's Zone of Proximal Development ideas,[27,28] in a variety of related fields.[29] Scaffolding an activity is usually supported by (1) identifying what the learner can do without assistance to determine what exactly the scaffolding should target and (2) eventually fading the scaffolding corresponding to an increase of cognitive responsibility on the student.[29,30]

While many use these principles to improve one on one or small group interactions,[26,29,31] others have used these principles to help design assessments/activities.[30,32,33] For example, Ge and Land[32] acknowledge that the scaffolding can be informed by research in the area to specifically target areas that have been found to be problematic. Reiser[33] suggests that since we cannot interact with students individually to provide the scaffolding, we can structure the question and take cue's from Wood et al.'s[26] work to reduce degrees of freedom, focusing students in on the areas they need to focus on. McNeil et al.[30] shows how fading can be applied in a similar situation as they reduced the presence of the scaffold over the course of the semester. However, there are few studies on the removal of scaffolding and its effects on the way students respond.

We have previously developed an assessment task and coding scheme that allows us to elicit evidence about how students are thinking about LDFs,[15] which involves both a written explanation and drawn model for how nonpolar species can "stick together", that is, how LDFs originate and operate in nonpolar species. In this paper we adapt and simplify the previous coding scheme so that we can investigate several research questions that focus on the evidence that we can elicit from such questions and the implications for learning over time. We also adapted the original assessment task to incorporate a scaffolded drawing prompt that is intended to support students as they think about the mechanism of LDF formation.

We used the revised coding scheme and assessment task to perform a longitudinal study (over the course of two semesters) to investigate how students' models and explanations change over time, what happens when scaffolding is withdrawn, and how students' responses change between formative and summative assessment tasks. We also performed a cross-sectional study comparing students from prior semesters who were not exposed to scaffolded tasks.

■ RESEARCH QUESTIONS

The study is guided by these research questions:

(1) How does scaffolding impact student responses to prompts about LDFs (study 1)?

(2) How do students' explanations and models of LDFs change over a two-semester sequence (study 2)?

(3) How do students' written explanations and drawn models correlate with each other (study 3)?

DOI: 10.1021/acs.jchemed.9b00455
J. Chem. Educ. 2019, 96, 1821−1832

■ METHODS

In this section we describe the methodology for all of the studies in this paper. All of the participants in these studies were undergraduate students at a large research-intensive Midwestern university. All participants were informed of their rights as human subjects and gave their permission to have their responses used for research purposes. All of the participants' responses were deidentified before analysis.

Additionally, following analysis we frequently use statistical tests to help understand the significance of patterns in the data. For all statistical tests we use a significance level of $p < 0.05$.

Studies 1−3 involve the analysis of a longitudinal study, following ways in which a cohort of general chemistry students draw and explain LDFs over the two-semester sequence. Specifically, these students were enrolled in Chemistry, Life, the Universe and Everything (CLUE), a transformed general chemistry curriculum that is organized around progressions of core ideas in the context of scientific practices.[10,11] In these studies, we examine the impact of the scaffolding of the activity (study 1), how their responses changed over time (study 2), and how their text and drawing responses were related to one another (study 3).

Participants

For the major longitudinal study reported here, four assessment items involving understanding of LDFs were administered to general chemistry students over the course of the academic year 2015−2016. Of the 290 students enrolled in the sections in which our activities were administered, 260 completed all four of those assessments and gave consent to have their work used for research purposes. Using a random number generator, we randomly selected 150 of those students to analyze their responses to the tasks (described below) across all four assessments. We selected only 150 of the 260 possible students for practical reasons as this would give us a group large enough to use the appropriate statistical tests while reducing the number of responses that would need to be coded (by almost 900 text and drawings). This group is referred to as the 15−16 cohort in this paper. This group is majority female (64%) and white (71%) and in their first year of college (84%). The mean ACT score for this group is 26.2, and the mean grade (on a 4-point scale) in the course is 3.43 and 3.22 for the first and second semester, respectively (see Supporting Information S1).

Additionally, we carried out a cross-sectional study in which we compared the responses of this cohort with those from 129 students from our previous study[15] (referred to as the 14−15 cohort in this study, corresponding to the 2014−2015 academic year in which they completed the general chemistry sequence). The 14−15 cohort is a subset of the 250 students whose data were initially analyzed in the previous study[15] who then went on to take General Chemistry 2 the following semester. This cohort of students is similar to the 15−16 cohort based on their grade level, sex, race/ethnicity, and ACT score (see Supporting Information S1). Therefore, we claim that the only meaningful difference between these two groups of students with regards to this study is the assessment items they received.

Data Collection

For the longitudinal study we administered four activities over the two-semester general chemistry (GC) sequence to the 15−16 cohort to explore how students explain and draw LDFs. In this paper we name the items 1−4 corresponding to the formative (item 1), summative (item 2), homework posttest (item 3), and exam posttest (item 4) assessments (Figure 1).

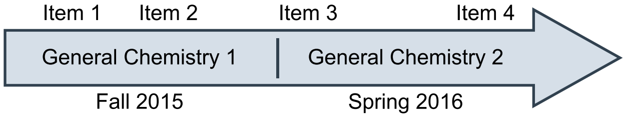


Figure 1. Relative timing of the four assessment items relative to the two semesters in the general chemistry sequence in the 2015−2016 academic year. The four assessments consist of a formative homework assessment (item 1), a summative exam assessment (item 2), a homework posttest assessment (item 3), and an exam posttest assessment (item 4).

These items vary in terms of the timing of their delivery, the stakes of the activity, the form of the assessment, and the structure of the prompt. We discuss the details of these activities in the sections that follow.

The items were administered as either a homework or on an examination. Homework activities in this course were administered via beSocratic, a web-based system that allows the collection of both written and drawn responses.[34] This homework is graded for completion rather than correctness to encourage students to explore their ideas rather than look up answers online. Students are not provided with copies of the correct answers, but the questions are discussed in class and often used to anchor the next instructional unit. In contrast, all of the exams in this course were given on paper, and credit was awarded to the students on the basis of their correctness. Students are provided with keys to the multiple choice items but not the open response items, but they are free to consult with graduate teaching assistants or undergraduate learning assistants. We consider the homework items "low stakes" assessments while the exam items are "high stakes" assessments.

The text and drawing prompts for all the LDF assessments administered in this study are outlined below (Table 1), and the full prompts are shown in Supporting Information S2. In the following section we further describe the context surrounding the individual items.

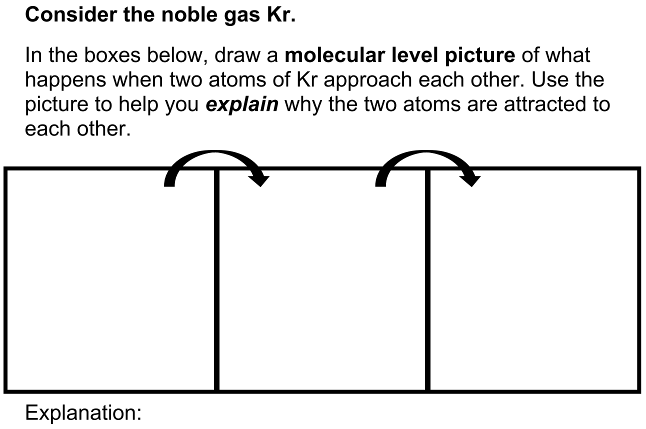
Item 1 (formative) was an instructional homework activity and formative assessment that was administered to students as they were learning about LDFs in class during the first 2 weeks of the first semester of general chemistry (GC1). This formative activity was developed iteratively over the course of several semesters using feedback from analysis of student responses including the work presented in our previous study.[15] In this administration (to the 15−16 cohort) the prompt is scaffolded by (1) asking students to think about the electrons to help them think about the scalar level needed to reason about the attraction, and (2) including three drawing boxes that are intended to prompt students to think about the sequence of events that lead to the attractive interaction (as shown in Figure 2).

In our previous study,[15] the students in the 14−15 cohort were provided with a prompt that was similar but less specific than the activity described here. In particular, the drawing prompt only had 1 box, rather than the 3 boxes in item 1. In this study, we want to provide as much scaffolding as possible to help students learn to construct mechanistic explanations, and it was clear from our previous study[15] that further scaffolding was needed for the drawing prompt.

Table 1. Information about the LDF Prompts Used throughout All Four Studies

Item		Prompt		Additional Administration Details			
Name	Type	Text	Drawing	Cohort	Type	Timing	Boxes for Drawing Prompt
Previous study[a]	Formative	What do you think causes the helium atoms to move toward each other? Hint: Think about what's happening with the protons and electrons.	Sketch a molecular level picture to show what you think is going on as the atoms move together. Use this space to describe in words what you are trying to represent in your drawing.	14–15	Online homework	Beginning of GC1	1
Item 1	Formative	As the [helium] atoms get closer they attract one another and the potential energy decreases as shown by the circled region. Please explain why the atoms attract and the process by which it occurs. Hint: Think about what's happening with the electrons.	Now, draw a series of molecular level pictures of what's happening within the atoms that causes the atoms to move together. Be sure to include the subatomic particles in your drawing. If there's anything you cannot explain in your drawing, explain it in the box below.	15–16	Online homework	Beginning of GC1	3
Item 2	Summative	Consider the noble gas Kr. In the boxes below, draw a molecular level picture of what happens when two atoms of Kr approach each other. Use the picture to help you explain why the two atoms are attracted to each other.		15–16	Paper exam	Beginning/middle of GC1	3
Item 3	Homework posttest	Is there an attraction between two neutral helium atoms? (Yes/No/I do not know) Below, explain in detail why or why not. Be sure to include any evidence you have for your claim and how this supports your reasoning.	Draw a molecular level picture of two helium atoms as they approach one another in the blue box, and explain what is happening in the black box.	15–16	Online homework	Beginning of GC2	1
Item 4	Exam posttest	Consider the noble gas Ar. Draw a molecular level picture of what happens when two atoms of Ar approach each other. Use the picture to help you explain why the two atoms are attracted.		15–16	Paper exam	End of GC2	1

[a]See ref 15.


Consider the noble gas Kr.

In the boxes below, draw a **molecular level picture** of what happens when two atoms of Kr approach each other. Use the picture to help you ***explain*** why the two atoms are attracted to each other.

Explanation:

Figure 2. Item 2 drawing prompt including the scaffolded 3 box drawing prompt to encourage students to draw the process by which the interaction formed.

Item 2 (summative) was included on the first midterm exam administered about 4 weeks into GC1. Mirroring the format of the formative assessment, three boxes were provided to encourage students to draw the mechanism of the attraction. However, the hint reminding the student to think about the electrons was removed.

Item 3 was administered on a homework activity given the first week of the second semester. The assessment was not scaffolded in the same way as the previous two so that we could see the effect of removing the scaffolding. While they were still asked to write and draw about what happens when two atoms of a noble gas approach, they were only provided one box in which to draw the atoms approaching.

Item 4 was administered on the last midterm exam in the second semester. Students were informed prior to the exam that there would be first-semester material on the exam, though they were not told specifically what that material would be. In this case, students were asked about how and why another neutral atom, argon, can attract each other, and again, the scaffolding was removed from the drawing portion of the activity.

Data Analysis

In our previous study,[15] we developed a holistic coding scheme to characterize how students write and draw about the mechanism by which LDFs form and act. This scheme consisted of six categories (0−6) that emerged from student interview data and analysis of a relatively small number of responses. However, as we continued coding larger numbers of student responses it became clear that the original scheme was too fine-grained to be broadly applicable. Therefore, we simplified our coding scheme by collapsing the codes into three broader categories based on the type of reasoning shown by students: nonelectrostatic, electrostatic causal, or causal mechanistic. Both authors worked to combine these codes through discussion and test coding student responses. Examples and definitions of the three categories of responses are provided in Tables 2 and 3.

Level 1 remained the same, capturing responses that do not provide any kind of electrostatic causal reasoning.[15] Typical responses include simply naming an interaction or describing that the atoms attract one another without providing a reason (essentially restating the question). The second category of responses, "electrostatic causal", combines levels 2 and 3 in

Table 2. Coding Scheme for Text LDF Responses

Type of Response	Previous Level[a]	Salient Text Features	Examples
Nonelectrostatic (NE) text response	0 and 1	The response does not include reasonable electrostatic evidence of the interaction. Instead, the response provides nonelectrostatic evidence or does not address the intermolecular interaction between molecules.	"The two atoms are attracted because of the electromagnetic forces that exist between them" "This happens because the two Ar atoms need 8 valence electrons so between the two of them they start pulling at the other to fill its shell. This causes the two to combine, making an Ar2 molecule"
Electrostatic causal (EC) text response	2 and 3	The response indicates that electrostatic charges cause the interaction. Examples of electrostatic causal evidence include subatomic particles, overall charge of the atom, partial charges, etc. These responses do not include a mechanism by which a separation of charge occurs.	"The atoms become attracted to [each other] because of the attractions between the protons & electrons. Protons have a (+) charge, while electrons have a (−) charge. Opposite charges attract, and the closer the atoms appear the stronger the attraction becomes" "Atoms attract to each other [because] the [London] dispersion forces make the partially negative end of one atom attract to the partially positive end of another atom."
Causal mechanistic (CM) text response	4 and 5	The response indicates that the interaction occurs due to electrostatic charges and includes the mechanism by which the instantaneous and/or the induced dipole forms.	"The atoms attract because of the London dispersion forces between the two atoms in which an atom with [instantaneous] dipole causing the another atom to become an induced dipole. In this scenario, one atom has its electrons shift to certain region of the electron cloud, giving that region a negative charge. This negative region then induces the movement of electrons in another atom, being attracted to the new present positive region of the second atom." "As the 2 atoms approach each other one of them becomes [instantaneously] dipole due to fluctuation in its electron cloud as most of the electrons are concentrated to one side of the atom making it have a partially positive and negative charges on opposite sides. This partially positive side of the attracts the [electrons] of the other atom making its electron cloud fluctuate and form an instantaneous dipole where both of them attract each other."

[a] See ref 15.

which students acknowledge that electrostatic charges are the cause of the interaction but do not discuss how these partial charges arise from neutral atoms.[15] The third category, "causal mechanistic", includes student responses that discuss how the unequal distribution of charge arises from electron movement to give temporary dipoles, which then cause the interaction corresponding with levels 4 and 5.[15] This causal mechanistic reasoning aligns with other researcher proposed definitions, in that the mechanism for a phenomenon lies at least one scalar level below the phenomenon being explained.[13,14] That is, mechanistic reasoning must involve a discussion of how the redistribution of electrons leads to the production of transient charges.

While collapsing the codes does reduce the selectivity of the coding scheme (for example, it removes the ability to distinguish between responses that provide the mechanism for forming one vs two dipoles), we believe that the three classifications are a more practical approach to coding large numbers of student responses while still capturing the presence of causal mechanisms. Another advantage of this new characterization scheme is that the classification previously applied to student drawings is now directly comparable to the written explanations. That is, drawings that do not show any kind of charge distribution are nonelectrostatic; those that show charge distribution are electrostatic causal, and those that show the sequence of events that produce the charge distribution are causal mechanistic.

Interrater reliability of the coding scheme was established between author K.N. and an undergraduate research assistant before beginning the analysis (see Supporting Information S1). Analysis of the responses was carried out by author K.N. using QSR's International Nvivo 10 Software[35] to assist in the coding process. Because of the nature of the prompt and the way data was collected, the authors were aware of the prompt from which the data came.

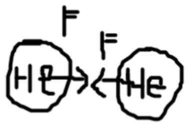
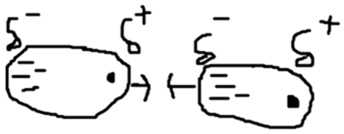
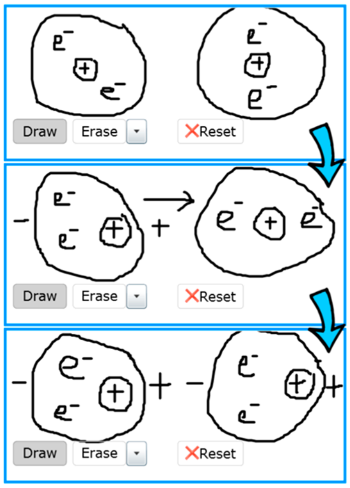
■ RESULTS AND DISCUSSION

Study 1: What Effect Does the Level of Scaffolding Have on Student Drawings and Explanations?

In this study we compare the student responses from the unscaffolded prompt administered in fall 2014 to those from the new scaffolded prompt, in which students were provided with three boxes to draw the mechanism of LDF formation, which was administered at the same time point in fall 2015 (item 1).

In the previously reported analysis of student responses administered in fall 2014, the drawing portion was not scaffolded; instead, students were provided with a blank drawing space.[15] Using the revised coding scheme on this previously collected data, we see (Figure 3) that while 87% of students from the 14−15 cohort provided an electrostatic causal drawing indicating the involvement of charges in attraction between the atoms, the number of students who drew a full causal mechanism was only about 2%.[15] In contrast, the formative prompt (item 1) from fall 2015 which included the scaffolded boxes for drawing the mechanism by which LDFs are generated produced an increase in the full causal mechanistic drawings: 55% of students now produce the full causal mechanism. That is, for these two equivalent cohorts of students who were completing a formative assessment, adding the scaffolded drawing prompt helped many more students to

Table 3. Coding Scheme for LDF Drawings

Type of Response	Previous Level[a]	Salient Drawing Features	Examples[b]
Non-electrostatic (NE) drawing	0 and 1	The drawing does not show any reasonable electrostatic interactions or interacting atoms.	
Electrostatic causal (EC) drawing	2 and 3	The drawing includes subatomic particles or charges, but does not include pictures showing the process by which partial charges develop.	
Causal mechanistic (CM) drawing	4	The drawing includes the subatomic particles or charges and includes several pictures showing the atoms transitioning into dipoles.	

[a] See ref 15. [b] Student drawings used with permission.

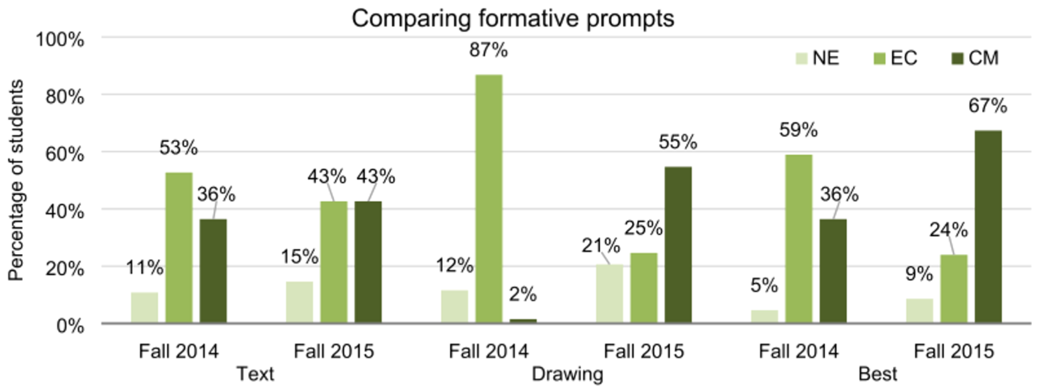


Figure 3. Distribution of text, drawing, and best (most sophisticated code for drawing or text) responses for the different formative assessments given during fall 2014 (N = 129) and fall 2015 (item 1, N = 150).

produce causal mechanistic drawings. However, the overall percent of students who provide answers including some electrostatic causal factor (that is, both EC and CM bins) is approximately the same as in fall 2014.[15]

We do not expect that scaffolding the drawing response should affect the written responses, because on the homework, the written prompt appears before the drawing prompt (item 1), and students cannot go back to prior questions. Indeed, there is no significant difference in the students' text responses from the different years (X^2 = 2.94, p = 0.230). In fall 2014, 36% of students provide a CM written response, while that number increases only slightly in fall 2015 (43%).[15]

When we compare their "best" responses (giving each individual student a score corresponding to their most sophisticated code for drawing or text response, CM > EC > NE) we see that the number of students producing a CM drawing or text response almost doubled from fall 2014 to fall 2015. While the text responses look the same for the two groups, by scaffolding the drawing responses we were able to elicit more CM students' responses, including many who did not produce a CM text response. That is, the presence of the scaffolded drawing prompt resulted in the activation of resources associated with the mechanism of the interaction. Since there was little difference in the instructional material that preceded the formative activity between the two years, it is probable that if the drawing prompt in fall 2014 had been similarly scaffolded we would have seen more causal mechanistic drawings than we saw (2%).

Study 2: How Do Students' Explanations and Models of LDFs Change over a Two-Semester Sequence?

Finding 1: A Large Percentage of Students Construct a Full CM Drawing on the First Summative Exam after Being Taught the Material. On the summative assessment (item 2), 86% of the students provide a drawing of a full causal mechanism. Including the EC responses, 95% showed separation of charge (Figure 4). There are several possible reasons for this improvement:

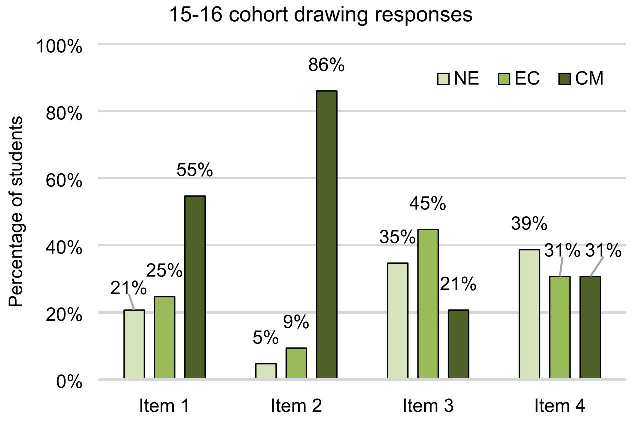


Figure 4. Percentage of 15−16 cohort drawing responses ($n = 150$) in each of the LDF categories for items 1−4 (see Table 1 for item descriptions).

(1) Students may be more likely to provide an effortful response on a summative item that is graded for correctness (item 2) than on the homework for which credit is given for completion (item 1).
(2) Item 1 was given when students were learning about LDFs; by the time of the exam, students would have more time to study and assimilate the ideas.
(3) Students may have simply memorized the three-pane drawing (although they had not received individual feedback on the homework, it had been reviewed in the next class) due to the close proximity of the assessment to the initial instruction.

Finding 2: Removal of the Scaffold Appears To Produce a Decline in Causal Mechanistic Drawings on Unstructured Posttest Assessments (Items 3 and 4). Items 3 and 4 remove the scaffolded 3 box drawing prompt, and there is a corresponding drop in the percentage of CM drawings. At the start of the second semester (item 3), the number of students who provided a drawn causal mechanism is down to 21%, and the percent who show any kind of electrostatic causal drawing is 66% as opposed to 95% on item 2. In addition to the removal of the scaffolding, there are other possible reasons for this decline including the following:

(1) These responses were collected several months after the explicit treatment of LDFs in the curriculum, and perhaps students had difficulty in recalling the details of the mechanism.
(2) Some students may have simply memorized the drawing for the exam (item 2) rather than deeply understanding the mechanism.
(3) The prompt was administered on the low stakes homework system.

Although we cannot be certain of the factors that contribute to the decline in CM drawings, on the final administration of the prompt (item 4) we do see an increase in CM from 21% on item 3 to 31%, though the total percentage of students providing any type of electrostatic causal drawing remains about the same (66% at item 4, 62% at item 3). The slight increase in CM responses on item 4 is probably because of the higher stakes nature of the exam. Despite this improvement, the percentage of CM drawings is still much lower than it was for the summative assessment from semester 1 (item 2). However, the relative stability of the percentage of CM drawings across the two posttests despite an additional 3 months passing may indicate that time from instruction is a less important factor than the lack of scaffolding.

Finding 3: The Initial Scaffolded Drawing Prompt Does Have an Effect on the Final Number of CM Drawings. Interestingly, the introduction of the scaffolded prompt in fall 2015 did appear to have a long-term impact on the students. Even though the decline in CM drawings after withdrawal of the scaffold is quite marked, in comparison to the responses from fall 2014, the percentage of students who do provide a full drawn causal mechanism on item 4 is significantly higher (31% to 2%) despite the fact that in fall 2014 the material was covered in the same semester that the assessment was given ($X^2 = 41.28$, $p < 0.001$, $\varphi = 0.38$).

Finding 4: The Text Responses Do Not Vary as Widely over the Four Time Points, but There Is a General Decrease in CM Written Explanations. For the first two time points, the percent of EC and CM explanations does not vary much (Figure 5). However, in the second semester (items 3 and 4), we see the same "drop off" in performance from the first semester as we did for the drawing

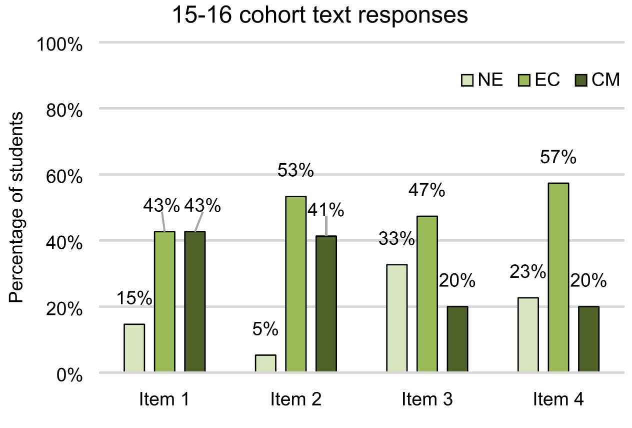


Figure 5. Percentage of 15−16 cohort text responses ($n = 150$) in each of the categories for items 1−4 (see Table 1 for item descriptions).

responses, perhaps because time passed and some students have forgotten: For item 4 only 20% of students provide causal mechanistic explanations, although in total 77% include an electrostatic causal factor.

Study 3: How Do Students' Written Explanations and Drawn Models Relate to One Another?

Finding 1: There Is Generally an Association between Text and Drawing Responses Except for EC Responses. While the text and drawing responses were coded separately, one might imagine that there should be some correspondence between the two types of responses. We used a series of Pearson's χ^2 tests[36] to analyze the relationship between students' text and drawing responses (Table 4). For results that showed a significant association, we determined the strength of the relationship using Cramer's V, a modified version of φ for contingency tables with more than 2 rows and columns,[36] using Cohen's suggestions for interpreting this value: small −0.1, medium −0.3, large −0.5.[37]

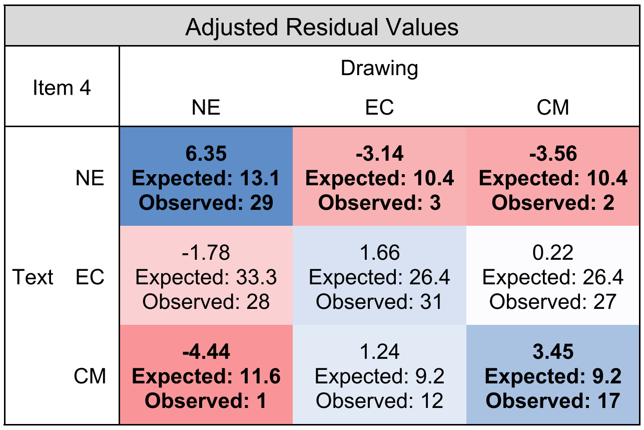
Table 4. Pearson's χ^2 Test Results Exploring the Relationship between the Text and Drawing LDF Responses

Activity	χ^2 Value	P Value	Cramer's V
Item 1	31.44	$P < 0.001$	0.32
Item 3	18.84	$P = 0.001$	0.25
Item 4	50.03	$P < 0.001$	0.41

For items 1, 3, and 4 there is a significant relationship between the text and drawing responses ($p \leq 0.001$) with Cramer's V values ranging from 0.25 to 0.41. For item 2, four of the nine cells (44%) had an expected count less than 5, which violates one of the assumptions of this test. This is because over 90% of the drawing responses were characterized as CM for item 2, and therefore, there were few responses in the remaining cells.

While the χ^2 test tells us if there is a relationship between the variables, post hoc analysis is necessary to determine what is driving that relationship.[38] To do this, we used the statistical software SPSS[39] to calculate the adjusted standardized residual (from here on called adjusted residual) for each of the cells in the contingency table. This value provides a standardized measure for how different the observed value is from the expected.[38] The sign of the adjusted residual indicates if the observed count is more (positive) or less (negative) than expected, while the magnitude of the value indicates how different the observed count is from the expected.[38] If the adjusted residual is large enough, then we reject the null hypothesis for that particular cell, indicating that particular relationship is driving the significant result of the initial χ^2 test.[38] To reduce the risk of type 1 error, we used the Bonferroni adjusted critical value (2.78 for a 9-cell table) as the threshold to determine if the cell is a primary driver of the relationship.[40]

For item 1, posthoc analysis revealed that positive associations between the CM text and drawings as well as the NE text and drawings are primary drivers of the significant result (see Supporting Information S3). There is also a negative association between NE text and CM drawing responses. A similar pattern is seen for item 3 with a positive association between NE text and drawings and a negative association between CM text and NE drawing responses driving the significant result (see Supporting Information S3). Item 4 not only had the strongest association from the initial χ^2 test but was also driven by associations in five of the nine cells: positive associations between CM text/drawings as well as NE text/drawings; negative associations between NE text and EC drawings, NE text, and CM drawings; and CM text and NE drawings (Figure 6).



Adjusted Residual Values

Item 4		Drawing		
		NE	EC	CM
Text	NE	**6.35** **Expected: 13.1** **Observed: 29**	**-3.14** **Expected: 10.4** **Observed: 3**	**-3.56** **Expected: 10.4** **Observed: 2**
	EC	-1.78 Expected: 33.3 Observed: 28	1.66 Expected: 26.4 Observed: 31	0.22 Expected: 26.4 Observed: 27
	CM	**-4.44** **Expected: 11.6** **Observed: 1**	1.24 Expected: 9.2 Observed: 12	**3.45** **Expected: 9.2** **Observed: 17**

Figure 6. Contingency table for the item 4 (exam posttest assessment) text and drawing LDF responses. In each cell the adjusted residual value is reported along with the observed and expected values. Adjusted residuals larger than the Bonferroni adjusted critical value (±2.78 for 9-cell table) are bolded. To visualize the sign and magnitude of the adjusted residuals, the cells are color coded from dark blue (most positive) to dark red (most negative).

From these posthoc analyses we see that the expected relationship between the students' text and drawing responses is present. There are positive associations between NE text/drawings and between CM text/drawings, and negative associations on mismatched bins, specifically between the NE and CM bins. We believe these associations between the text and drawing responses indicate that for items 1, 3, and 4 the students' drawings were not memorized without understanding, but instead reflective of their comprehension of the mechanism. We also see that the EC category represents the "messy middle": this may be a reflection of the wide range of responses captured in this category. For example, this category can represent students who have either drawn or written a CM response and do not see the need to provide a full response in their corresponding drawing or explanation. However, it can also capture students who are still trying to figure out the cause of a mechanism, so that a student who describes both atoms as having a single charge and a student who describes the separation of charge in each atom (but not the mechanism of dipole formation) would both be categorized as EC.

Finding 2: If We Use the "Best" Response, Drawing or Text, We See a Significant Increase in CM from Item 1 to 2, a Decrease from 2 to 3, and Then Little Change from 3 to 4. If we assume that some students will only provide the most detailed response in either the drawn or written response (but not both), we can investigate each student's "best" response over time. In this case we see an increase in CM from the first formative item (item 1) to the first summative item (item 2, Figure 7). As discussed previously this may be because of the increased understanding of students, or the increased effort that students might show on a graded summative exam. When the scaffolded prompt is

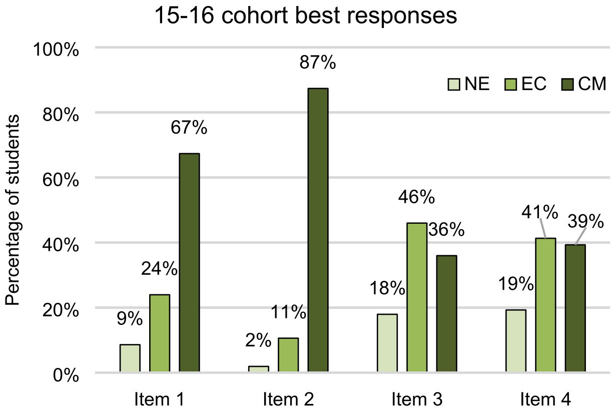


Figure 7. Percentage of 15−16 cohort's (*n* = 150) best responses (based on each students most highest level text or drawing response) in each of the categories for items 1−4 (see Table 1 for item descriptions).

removed in items 3 and 4, we now see that, while the combined percent of CM responses decreases from first to second semester, there is little difference between the homework or exam posttest response patterns (items 3 and 4). Therefore, we propose that this is evidence for the idea that student responses gathered from homework assessments are most likely representative of their best efforts at the time, despite the fact that they are not being graded for correctness.

■ SUMMARY

In these studies we have adapted and simplified a previously published coding scheme[15] that characterizes how students explain the origin of LDFs. This has allowed us to investigate the effects of adding and removing scaffolding designed to promote causal mechanistic thinking, the long-term impact of such scaffolding, and how and whether students respond differently to low and high stakes assessments.

We were able to perform these studies with large numbers of students because the instructors allowed us to design and administer four of assessment items, two as homework and two on semester exams. Because of the limitations of working with large numbers of students in real world settings, we were not able to design the studies as one might in a laboratory, particularly since half of the assessments were administered on an exam. This means that although we had a number of variables (scaffolding, time from instruction, and formative/summative administrations), we were not able to address each one separately. Nevertheless, we believe that there is sufficient evidence to tease out some more generalizable findings from these studies:

(1) Scaffolding of the drawing prompt has both an immediate effect on how students represent the origins of LDFs. Comparing responses on item 1 from fall 2014 and 2015, we see that adding the three box scaffold encouraged students to draw a causal mechanism. The percentage of CM drawings increased from 2% to 55%. Although it is probable that more than 2% of the fall 2014 students could also draw this interaction mechanistically, the appropriate resources were not activated as they answered the unstructured prompt.

(2) Scaffolding of the drawing prompt also appears to have a long-term effect on how students represent the origins of LDFs. While the percentage of CM drawings peaks dramatically on item 2, it also falls quite dramatically when the scaffolded prompt is removed. That being said, the percentage of CM drawings on both items 3 and 4 is still much higher than those in fall 2014. Although we do not have long-term data from fall 2014 to spring 2015, it is unlikely that the percent of causal mechanistic reasoning would have increased during this time.

(3) Student best responses to versions of the prompt are relatively unaffected by time, or by whether it is a high or low stakes assessment. While there is a large spike on CM drawing on item 2, we do not believe that this is an accurate representation of student understanding but instead is more likely to be a memorized set of representations. We believe that the comparable results from items 3 and 4 are a better representation of what students know about LDFs.

■ IMPLICATIONS FOR TEACHING

Use of Scaffolding in Formative Assessments

What seems clear from our findings is that scaffolding can greatly improve the immediate quality of student responses. When we compare the "best" response from each student from fall 2014 and fall 2015, at item 1 the percentage of CM responses nearly doubled. We also believe that scaffolding did have a long-term effect for some students. Certainly there was a marked drop in CM drawings between items 2 and 3, and it is not possible to determine from this study whether the drop is a function of the lack of scaffolding or the time from being taught. However, we believe that the relative stability seen in items 3 and 4 indicates that it is the loss of the scaffolding that is the main contributor to the drop in CM drawings. That being said, even by the end of the full year of general chemistry (items 3 and 4) there are many more students providing a CM response than students in the previous study (fall 2014) did just after learning about LDFs. While there is clearly more work to be done on the effects and removal of scaffolding, we do believe that adding scaffolding supports to formative assessments is productive and should be considered whenever possible

Formative and Summative Assessments

Interestingly, there is little difference between the levels of response for formative and summative assessments for items 3 and 4, despite the fact that students believed that item 4 would be graded and they knew item 3 would not. While some may be surprised that students appear to exert as much effort on the nongraded homework as on the exam, it has been our experience that most (but not all!) students do "try hard". This may be because the course expectations include the idea that the homework is for their benefit, to help them learn. The instructors attempt to create a culture that does not penalize mistakes on homework, but rather emphasizes that students use these activities to try out ideas. Most online systems are designed to provide immediate feedback, and because of this the nature of the tasks is mostly restricted to calculations, factual recall, or skills. Indeed, there is a danger that instruction and summative assessments will come to align with the restrictions placed upon homework systems rather than help students develop deep, connected, and useful knowledge. We believe that this study provides evidence that formative assessment systems such as beSocratic do not

DOI: 10.1021/acs.jchemed.9b00455
J. Chem. Educ. 2019, 96, 1821−1832

necessarily have to provide immediate feedback. This seems to be evidence for the idea that the very act of constructing the drawings (models) and written explanations actually helps students learn, since they have to reflect on what they know and make connections to answer the question.

We believe this finding provides a strong case for the use of more formative assessments, and for emphasizing their importance in student learning, both explicitly (by spending instructional time on them) and implicitly (by incorporating them into the grading scale for the course).

Importance of Long-Term Posttests: What Do Students Retain?

In many studies, the effect of interventions is often measured by pre/posttesting, where the posttest is often given immediately after the intervention. In this study our items are administered over a much longer time period. We we able to see whether responses to item 2, which occurred at the end of the instructional unit that covered LDFs (the first summative assessment), were truly indicative of the students having a deep understanding of these interactions. However, as we saw with items 3 and 4, many students did not maintain this level of performance, but these items are likely more indicative of student actual understanding. We believe that this kind of long-term testing is likely to provide more meaningful information about student learning than a posttest that immediately follows an intervention. After all, our goal is to support student learning, not to help them memorize material to pass tests and then forget it.

LIMITATIONS

As previously noted, this longitudinal study was conducted with students enrolled in a transformed curriculum that emphasizes scientific practices such as constructing models and explanations, in which students are routinely asked to provide mechanisms for a wide range of phenomena, and where students are not penalized for mistakes on homework. This approach is still quite rare, and therefore, the results here may not be transferrable to more traditional curricula. As we noted earlier, because this study was situated in the actual homework and course examinations, we were not able to vary each variable in a systematic way. Further studies that change, for example, the level of scaffolding in formative (but not summative) assessments may provide further insight into the role of scaffolding and how it affects long-term use of knowledge.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.9b00455.

> Additional information about the participant demographics of the 14−15 and 15−16 cohorts, establishing of interrater reliability, the full activity prompts, and additional contingency tables (PDF, DOCX)

AUTHOR INFORMATION

Corresponding Author

*E-mail: mmc@msu.edu.

ORCID

Keenan Noyes: 0000-0002-8587-1694

Melanie M. Cooper: 0000-0002-7050-8649

Notes

Any opinions, findings, conclusions, or recommendations expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Jennifer Cook for her help in the developing interrater reliability. This work is supported by the National Science Foundation under DUE 0816692 (1359818), DUE 1043707 (1420005), and DUE 1122472 (1341987).

REFERENCES

(1) Henderleiter, J.; Smart, R.; Anderson, J.; Elian, O. How Do Organic Chemistry Students Understand and Apply Hydrogen Bonding? *J. Chem. Educ.* **2001**, *78* (8), 1126.

(2) Peterson, R. F.; Treagust, D. F.; Garnett, P. Development and Application of a Diagnostic Instrument to Evaluate Grade-11 and −12 Students' Concepts of Covalent Bonding and Structure Following a Course of Instruction. *J. Res. Sci. Teach.* **1989**, *26* (4), 301−314.

(3) Cooper, M. M.; Williams, L. C.; Underwood, S. M. Student Understanding of Intermolecular Forces: A Multimodal Study. *J. Chem. Educ.* **2015**, *92* (8), 1288−1298.

(4) Williams, L. C.; Underwood, S. M.; Klymkowsky, M. W.; Cooper, M. M. Are Noncovalent Interactions an Achilles Heel in Chemistry Education? A Comparison of Instructional Approaches. *J. Chem. Educ.* **2015**, *92* (12), 1979−1987.

(5) Cooper, M. M.; Corley, L. M.; Underwood, S. M. An Investigation of College Chemistry Students' Understanding of Structure−Property Relationships. *J. Res. Sci. Teach.* **2013**, *50* (6), 699−721.

(6) Othman, J.; Treagust, D. F.; Chandrasegaran, A. L. An Investigation into the Relationship between Students' Conceptions of the Particulate Nature of Matter and Their Understanding of Chemical Bonding. *Int. J. Sci. Educ.* **2008**, *30* (11), 1531−1550.

(7) Pierri, E.; Karatrantou, A.; Panagiotakopoulos, C. Exploring the Phenomenon of "change of Phase" of Pure Substances Using the Microcomputer-Based-Laboratory (MBL) System. *Chem. Educ. Res. Pract.* **2008**, *9* (3), 234−239.

(8) Cooper, M. M.; Underwood, S. M.; Hilley, C. Z. Development and Validation of the Implicit Information from Lewis Structures Instrument (IILSI): Do Students Connect Structures with Properties? *Chem. Educ. Res. Pract.* **2012**, *13* (3), 195−200.

(9) Maeyer, J.; Talanquer, V. The Role of Intuitive Heuristics in Students' Thinking: Ranking Chemical Substances. *Sci. Educ.* **2010**, *94* (6), 963−984.

(10) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116−1122.

(11) Cooper, M. M.; Klymkowsky, M. W. CLUE: Chemistry, Life, the Universe and Everything. https://clue.chemistry.msu.edu/ (accessed Jul 17, 2019).

(12) Stone, A. J. *The Theory of Intermolecular Forces*; Clarendon Press: Oxford, 1996.

(13) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *J. Learn. Sci.* **2019**, *28* (2), 160−205.

(14) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Sci. Educ.* **2008**, *92* (3), 499−525.

(15) Becker, N.; Noyes, K.; Cooper, M. Characterizing Students' Mechanistic Reasoning about London Dispersion Forces. *J. Chem. Educ.* **2016**, *93* (10), 1713−1724.

(16) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid−Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703−1712.

(17) Shepard, L. A. The Role of Assessment in a Learning Culture. *Educ. Res.* **2000**, *29* (7), 4−14.

(18) Mislevy, R. J.; Almond, R. G.; Lukas, J. F. *A Brief Introduction to Evidence-Centered Design; Research Report RR-03-16*; Educatoinal Testing Service, 2003.

(19) Underwood, S. M.; Posey, L. A.; Herrington, D. G.; Carmel, J. H.; Cooper, M. M. Adapting Assessment Tasks To Support Three-Dimensional Learning. *J. Chem. Educ.* **2018**, *95* (2), 207−217.

(20) Kahneman, D. *Thinking, Fast and Slow*; Farrar, Straus and Giroux: New York, 2011.

(21) Hammer, D. Student Resources for Learning Introductory Physics. *Am. J. Phys.* **2000**, *68* (S1), S52−S59.

(22) Feynman, R. P.; Leighton, R. B.; Sands, M. *The Feynman Lectures on Physics*; Addison-Wesley: New York, 1964.

(23) Cooper, M. M.; Klymkowsky, M. W. The Trouble with Chemical Energy: Why Understanding Bond Energies Requires an Interdisciplinary Systems Approach. *Cell Biol. Educ.* **2013**, *12* (2), 306−312.

(24) Novick, S. No Energy Storage in Chemical Bonds. *J. Biol. Educ.* **1976**, *10* (3), 116−118.

(25) Boo, H. K. Students' Understandings of Chemical Bonds and the Energetics of Chemical Reactions. *J. Res. Sci. Teach.* **1998**, *35* (5), 569−581.

(26) Wood, D.; Bruner, J. S.; Ross, G. The Role of Tutoring in Problem Solving. *J. Child Psychol. Psychiatry* **1976**, *17* (2), 89−100.

(27) Vygotsky, L. S. *Mind in Society: The Development of Higher Psychological Processes*; Cole, M., John-Steiner, V., Scribner, S., Souberman, E., Eds.; Harvard University Press: Cambridge, MA, 1978.

(28) Stone, C. A. What Is Missing in the Metaphor of Scaffolding? In *Contexts for Learning: Sociocultural Dynamics in Children's Development*; Oxford University Press: New York, NY, 1993; p 15.

(29) van de Pol, J.; Volman, M.; Beishuizen, J. Scaffolding in Teacher−Student Interaction: A Decade of Research. *Educ. Psychol. Rev.* **2010**, *22* (3), 271−296.

(30) McNeill, K. L.; Lizotte, D. J.; Krajcik, J.; Marx, R. W. Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *J. Learn. Sci.* **2006**, *15* (2), 153−191.

(31) Hogan, K.; Pressley, M. Scaffolding Scientific Competencies within Classroom Communities of Inquiry. In *Scaffolding Student Learning: Instructional Approaches and Issues*; Brookline Books: Cambridge, MA, 1997; pp 74−107.

(32) Xun, G.; Land, S. M. A Conceptual Framework for Scaffolding Ill-Structured Problem-Solving Processes Using Question Prompts and Peer Interactions. *Educ. Technol. Res. Dev.* **2004**, *52* (2), 5−22.

(33) Reiser, B. J. Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work. *J. Learn. Sci.* **2004**, *13* (3), 273−304.

(34) Bryfczynski, S. Be Socratic: An Intelligent Tutoring System for the Recognition, Evaluation, and Analysis of Free-Form Student Input. Ph.D. Dissertation, Clemson University, Clemson, SC, 2012.

(35) *NVivo Qualitative Data Analysis Software*; QSR International Pty Ltd., 2012.

(36) Green, S. B.; Salkind, N. J. Two-Way Contingency Table Analysis Using Crosstabs. In *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data*; Pearson Education Inc.: Upper Saddle River, NJ, 2011; pp 366−376.

(37) Cohen, J. A Power Primer. *Psychol. Bull.* **1992**, *112* (1), 155−159.

(38) Agresti, A. Inference for Two-Way Contingency Tables. In *Categorical Data Analysis*; Wiley Series in Probability and Statistics; Wiley: Hoboken, NJ, 2013; pp 69−112.

(39) IBM Corp. *SPSS Statistics for Windows*; IBM Corp: Armonk, NY, 2017.

(40) MacDonald, P. L.; Gardner, R. C. Type I Error Rate Comparisons of Post Hoc Procedures for I j Chi-Square Tables. *Educ. Psychol. Meas.* **2000**, *60* (5), 735−754.