Leveraging understanding of energy from physics to overcome unproductive intuitions in chemistry

Beth A. Lindsey,[1,*] Megan L. Nagel,[2] and Brandi N. Savani[2]

[1]*Physics Department, Penn State Greater Allegheny, McKeesport, Pennsylvania 15132, USA*
[2]*Chemistry Department, Penn State Greater Allegheny, McKeesport, Pennsylvania 15132, USA*

(Received 4 October 2018; published 8 April 2019)

Electrostatic potential energy is a topic of great difficulty for many students. In this paper, we empirically test the utility of two approaches for laying a foundation for developing an understanding of energy in an electrostatic context, with interdisciplinary relevance. We examine student responses to a question about how the potential energy of a system of two attracting ions varies with distance (the "ions" task), and investigate how these responses change after students have been exposed to either a question designed to help them think about gravitational potential energy or the potential energy of a system of two attracting magnets. We found that performance on the ions task improved for those students who were prompted to think about the gravitational context, while it did not change for those who considered the magnets. The results are interpreted using dual-process theories of reasoning and decision making, and implications for instruction are discussed.

DOI: 10.1103/PhysRevPhysEducRes.15.010120

I. INTRODUCTION

Many students who take physics courses do so in preparation for other fields of study, such as chemistry, biology, or engineering. Thus, topics that are common across these disciplines are of particular importance within the physics curriculum, as they provide much-needed opportunities for students to apply physics concepts in other disciplines [1]. Energy is an example of a topic of fundamental importance to many scientific disciplines. In particular, the energy changes associated with electrostatic interactions provide a foundation for several topics throughout introductory curricula in many sciences. For instance, numerous important topics throughout the chemistry curriculum involve the interaction of charged or partially charged species. An understanding of the origin of the energy changes associated with these chemical interactions provides the foundation to explain observations during processes such as phase changes and chemical bonding [2]. Potential energy is of particular importance in interdisciplinary contexts. As Cooper and Klymkowsky assert, "an accurate working understanding of potential energy is a prerequisite for understanding chemical energy or indeed any energy changes associated with bonding or intermolecular forces." [3]. A functional understanding of these energy changes can also be extended to more complex processes including those in biological systems [3]. It is well known that supporting students' understanding of electrostatics and particularly energy in electrostatics can lead to improvements in performance in physics and chemistry courses [4,5]. However, research in physics education as well as in chemistry education reveals that many students do not begin with a solid understanding of energy in general and electrostatic potential energy in particular [2,6–9].

The work described in this paper draws upon other physics concepts to help students answer a question about electrostatic potential energy in the context of a college chemistry course. The importance of interdisciplinary connection making is highlighted by the Next Generation Science Standards (NGSS), which treats energy as a both a core concept and one that cuts across the STEM disciplines [10]. The NGSS emphasize the development of a coherent understanding of energy changes across different scales and disciplinary contexts, treating all energy as either energy of motion (kinetic energy) or energy stored in fields (potential energy). Researchers in the fields of physics and chemistry education research have similarly expressed a need for interdisciplinary collaboration in the investigation of student understanding of energy as well as in the development of curricula to support student learning [11–13].

In previous work, we had demonstrated some success at improving student performance on questions about the potential energy of a system of two interacting charged particles by drawing student attention to a situation

*bal23@psu.edu

involving potential energy in the context of gravity [7]. The attempt to develop an understanding of new content by making connections to existing knowledge is a technique well established in the educational research literature [14,15]. A gravitational approach, such as the one used in our previous work, is commonly employed as an introduction to electrostatic potential energy by both introductory physics and general chemistry textbooks [16–19]. However, the translation between the near-Earth macroscopic scale at which gravitational potential energy is typically treated and the microscopic scales of interacting charged particles is not necessarily obvious for many students [3]. An alternate choice to connect to many students' lived experience would be the potential energy associated with a system of two magnets. Although "magnetic potential energy" is not commonly taught explicitly in introductory physics courses, it is used to introduce ideas related to potential energy in some curricula for teachers [20]. Moreover, many reports from instructors and students, as well as our own informal observations, note that magnets are commonly used as an analogy for electrostatic or chemical interactions [5,21,22].

Different representations and analogies used to teach can lead to measurable differences in student responses and reasoning [23]. While both gravity and magnets are commonly used to introduce students to the energy ideas necessary for building an appropriate understanding of the energy of electrostatic interactions, we are not aware of any study in either the physics or chemistry education research literature that formally analyzes the utility of either the gravitation or the magnets analogy for understanding the potential energy of an electrostatic interaction, nor are we aware of any studies that attempt to empirically compare the effectiveness of the gravitational and magnetic analogies. The focus of this paper is to do just that. In this work, we attempt the following:

(1) to empirically determine whether near-Earth gravitational examples can serve as an appropriate foundation for the learning of potential energy in electrostatic contexts;
(2) to compare the effectiveness of examples involving magnetism to those involving near-Earth gravity; and
(3) to begin to examine why specific examples might be helpful using the theoretical framework of Dual Process Theories of Reasoning.

II. BACKGROUND

A. Research on energy instruction

Potential energy is a complicated concept. Previous work [6,7] has shown that a question as simple as relating changes in potential energy to the distance between two oppositely charged particles can be very difficult for students. Many students conflate the potential energy of the system with the force that the ions exert on one another, or apply an equation ($U = k_e q_1 q_2 / d$) to conclude that increasing the distance between two objects must decrease the potential energy, without recognizing that for attracting ions this potential energy has a negative value. Formally, the potential energy of the system becomes *less negative* as the separation distance increases. Research has already shown that many students reject the possibility of negative potential energy [24], which is difficult to reconcile with a substance metaphor for energy [25].

One way students can be supported in learning these often new energy concepts is by building on pieces that students already understand [7,14,26,27]. In many cases, this may be done by analogical mapping of the unfamiliar domain onto a familiar domain [28]. Indeed, the analogies that students use have been shown to relate to the reasoning and inferences that they generate in physics [29–31]. However, when analogies are used unsystematically, they can drive confusion and misconceptions [31,32]. Evidence suggests that it is primarily the relationships between objects, rather than attributes of the objects themselves, that are mapped between domains [33]. Thus in the teaching of the atom, a solar system analogy is frequently used, with a main focus being on the attraction between the sun and the planet and its similarity to the attraction between a nucleus and an electron [29,31,33].

With respect to introducing the concept of electrostatic potential energy in both physics and chemistry, many textbooks draw an analogy between electric potential energy and gravitational potential energy [16–19,31]. For example, a common macroscopic scenario, like a ball on a hill, that illustrates near-Earth gravitational interactions is presented. Not only is this type of scenario familiar to students, but students who have taken a basic mechanics course would have received formal instruction on the kinetic and gravitational potential energy of such scenarios [34]. It should be noted that the mathematical treatment of potential energy for universal gravitation ($U = -Gm_1 m_2 / d$) and electrostatic attraction ($U = k_e q_1 q_2 / d$) is nearly identical. However, for near-Earth scenarios, students are often drawn to the equation $U = mgh$ when considering potential energy thus obfuscating the similarity in mathematical treatment between these two scenarios [7]. Thus, it is unclear what benefit, if any, students derive in terms of their understanding of microscopic, atomic-level electrostatic interactions from the analogy to near-Earth gravitation [3].

Research in chemistry education has demonstrated improvements in understanding and the application of energy topics when students are engaged in a revised curriculum built upon scaffolded progressions and emphasizing energy as a core course topic, reflective of the NGSS recommendations [5,35,36]. These reformed courses, in part, take advantage of the "bottom-up framework" by Nahum [37]. This framework introduces a coherent view of energy from the atomic to the macroscale and emphasizes the Coulombic forces between individual atoms prior to introducing bonding. By emphasizing electrostatic forces as the basis of chemical interactions, students are provided a simple framework upon which further learning can be built [15,38]. In this bottom-up approach, bond types are introduced as a continuation of this single concept that can then be used as a foundation for explaining molecular structure and chemical properties. However, even at its most fundamental level, this approach requires some understanding of electrostatic interactions and their relationship to force and potential energy [38]. Glossing over these details in instruction can lead students to adopt alternative explanatory frameworks that are inconsistent with much of the content they are expected to learn later [39]. Thus, by seeking to help students improve their understanding of electrostatic potential energy, we aim to lay a solid foundation on which a deeper understanding of a great many phenomena can be built.

B. Theoretical framework: Dual-process theories of reasoning

In this study, we draw upon dual-process theories of reasoning [40–42] to understand patterns of student responses and guide approaches to instruction. These theories are receiving increased interest as a way of making sense of inconsistent student response patterns observed to physics questions involving many topics, including capacitance [42,43], static friction [44], and buoyancy [45]. According to dual-process theories, people engage in two distinct cognitive processes when reasoning or making decisions. The first process, known as “type 1” or the “heuristic process,” is fast, automatic, and intuitive. The other process, known as “type 2” or the “analytic process” is slow, deliberate, and reflective. Type 1 reasoning does not require controlled attention, while type 2 reasoning makes heavy demands on the working memory [46]. Research has shown that in many cases, students will give a quick, intuitive response even in situations in which they have previously demonstrated mastery of the necessary content [42–45]. In many cases, the response generated by type 1 reasoning will be strongly cued by specifics of the task, including how the question is posed and how students are framing the task at hand. In this sense, the response generated by the intuitive process may be consistent with the “knowledge in pieces” framework by diSessa [47]. Dual-process theories account for this behavior by suggesting that when attempting to answer a question, the heuristic process will generate a response, which the analytic process evaluates. However, the analytic process tends to accept the response generated by the heuristic process unless it finds a compelling reason to reject it (the “satisficing principle”) [40]. Thus, responses generated by the heuristic process—which may involve reasoning shortcuts and other reasoning biases—tend to be the default responses, particularly in cases of limited time or limited knowledge [41]. In defining a set of ten heuristics that frequently arise when students reason in chemistry, Talanquer writes: “we must find ways to help students build a more robust and coherent knowledge structure on which they can rely when making decisions.” [41].

It should be emphasized that only if the analytic system is placed on alert, and recognizes a compelling reason to reject the first available response, will the first available response be rejected. Evans explains, “In studies of concept learning and inductive reasoning, it is clear that people formulate one hypothesis at a time and test it positively until or unless evidence is encountered that forces them to give it up.” [40]. Thus, the process of helping students to build more appropriate knowledge structures, as Talanquer advocates, can be seen as the process of identifying effective approaches for helping students reject incorrect first-available responses. In this work, we use student responses to a set of online questions to identify an intuitively appealing (but incorrect) idea about potential energy that is common among a large proportion of introductory students. We then empirically test two different interventions to see if either is more successful at causing students to revise their original response. Thus, this work provides empirical insights into what content can serve as the best foundation for the knowledge structures about energy across the science disciplines.

III. METHODOLOGY

A. Research setting and student population

Data were collected from students enrolled in Problem Solving in Chemistry, an optional one credit support course taken concurrently with the first semester of General Chemistry at a large public university in central Pennsylvania. While this course is intended for students at risk for failing General Chemistry, those students who complete the support course have a similar average and distribution on the cumulative final exam to the full General Chemistry population [48], suggesting that at least by the end of the semester (when our study was conducted) the population of students in Problem Solving in Chemistry is comparable to of the General Chemistry population. Approximately 60% of students had completed a high school physics course. Only 20% had completed a college-level physics course (including 15% who had taken AP or IB physics in high school). Approximately 20% of the sample had never taken any prior physics course. This population thus can provide insights into the learning of students who are entering their first college physics course.

B. Research instruments

This study uses a pre- and posttest design with a target question (referred to as the “ions” question) asked twice, with an extremely minor educational intervention presented in between. The entire package (pretest, intervention, and post-test) was administered to students in the form of an online “survey” during the thirteenth week of a fifteen-week semester. All course topics with the exception of calorimetry and an introduction to equilibrium had been completed and tested. Administration of the surveys in an online format allowed us to sample a larger student population than had been available in our prior study. Furthermore, since each question (pretest, intervention, and post-test) was posed on a single page of the survey, students could be required to provide a response before continuing to the next question. Once any of these questions had been answered, students were not able to return to or edit their previous responses. From the fall of 2014 to the spring of 2017, several versions of this online survey were developed to address different research questions. In each of these semesters, students were randomly assigned to a single version of the survey through a link in the course management system. The surveys were made available for one week, and a small amount of homework credit was awarded for survey completion. Providing a small amount of credit has been shown to help ensure that the performance of the subset of students completing an online assessment is similar to what would be obtained if the assessment were administered on paper in class [49]. All research activities for this study were deemed "exempt" by the Penn State Institutional Review Board, and those students who did not consent to have their data included in the study have been excluded from the analysis. The surveys were typically completed by about 75% of students enrolled in the course; of those, typically between 5% and 15% of students completing the survey did not give consent to be included in the data sample. Thus approximately 65% of students enrolled in Problem Solving in Chemistry during the time in which the study took place completed one of these surveys.

Altogether, seven distinct surveys were developed with the primary goal of investigating whether student responses to a question about potential energy in an electrostatics context could be swayed by exposure to questions about potential energy in other contexts. All of the questions on these surveys had been used in interviews, most of which are reported in detail in a prior study [7]. On the basis of these interviews, we have evidence that questions were being interpreted as intended.

All seven survey versions had a similar form including an identical first question, the ions question, that is used to establish a baseline for student performance (Fig. 1) [7]. For this question, students were provided multiple-choice options about which ion configuration has a greater potential energy, and were also required to explain the reasoning for their answer in a text box. The correct response is that the ion configuration has a greater potential energy in configuration (b). One way to arrive at this response is to recognize that if the ions were released from rest in each position and allowed to accelerate toward one another, they could gain more kinetic energy if released from the position in (b) than they could from the position in (a). This first instance of the ions question will be referred to as the "initial ions" question.

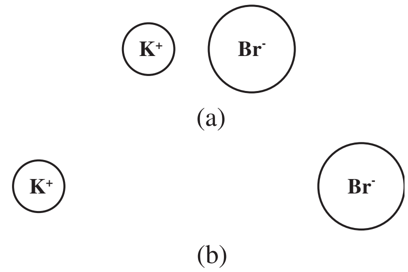

FIG. 1. The ions question. Students were asked, "Which figure has a greater potential energy?" with answer choices "Figure a has a greater potential energy"; "Figure b has a greater potential energy"; "Both figures have the same potential energy"; and "Impossible to determine."

Students then answered intervening questions (representing an educational intervention), which differed depending on the survey version they were assigned. Indeed, the only difference between different survey versions lies in these intervening questions. Each of the intervening questions on every survey version was designed to allow students to reflect on the idea that the potential energy for a system of two attracting objects is greater when the objects are at a larger separation.

This paper will focus on two versions of the survey: the "gravitational" version and the "magnetic" version, which address our primary research question. Other versions of the survey will be discussed in a future manuscript. In the gravitational survey, the initial ions question was followed by the "books" question, shown in Fig. 2. In this question, two identical books are on shelves at different heights above the Earth. Students were again asked a multiple-choice question as to which configuration has a greater potential energy and also to explain the reasoning for their choice in a text box. The correct response is that the book on a higher shelf, book C, represents the configuration with a greater potential energy. In the magnetic version, rather

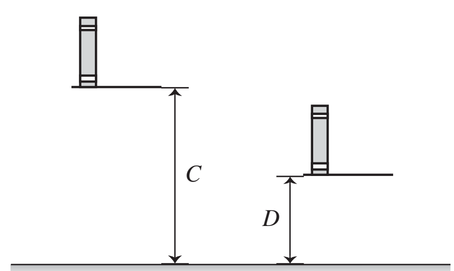

FIG. 2. The books question. Students were asked, "Which configuration, if either, has a greater potential energy?" with answer choices "The configuration with the book on the shelf at height C has greater potential energy"; "The configuration with the book on the shelf at height D has greater potential energy"; "Both configurations have the same potential energy"; and "Impossible to determine."

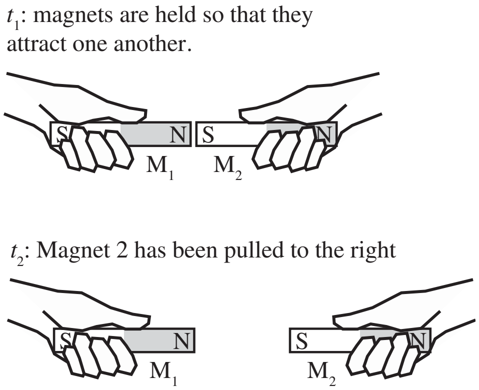

FIG. 3. The magnets question. Students were asked, "Between t_1 and t_2, does the potential energy of the two-magnets system *increase*, *decrease*, or *remain the same?*" with answer choices "Between t_1 and t_2, the potential energy increases"; "Between t_1 and t_2, the potential energy decreases"; "Between t_1 and t_2, the potential energy does not change"; and "Impossible to determine."

than seeing the books question students were presented with the "magnets" question, shown in Fig. 3. In this case, students were asked whether the potential energy of the two-magnet system increases, decreases, or remains the same between t_1 and t_2. (They were also given the option of indicating that it was not possible to determine an answer.) In this case, the correct response was that the potential energy increases as the magnets are moved farther apart. Students were again asked to explain the reason for their response to the previous question. We hoped that students would recognize that each of the questions (ions, books, and magnets) involved potential energy in a situation where attractive forces are acting.

As a post-test, students then had an opportunity to revisit the ions context and again decide which configuration had a greater potential energy. (This will be referred to as the final ions question). In the very first semester that the survey was used, the survey ended at this point. In all subsequent semesters, after completing the final ions question, students were asked to indicate explicitly whether their answer *or reasoning* to the question had changed. If they indicated that their answer or reasoning had changed, survey logic directed them to a multiple-choice (yes or no) question as to whether the intervening question (books or magnets) had been the cause of that change. If, on the other hand, they indicated that their answer had not changed, they were directed to a multiple-choice (yes or no) question about whether the ions question bore any similarities to the intervening questions. In all cases, they were asked to explain their answer in a text box. The overall survey flow for survey versions included in this manuscript is represented in Fig. 4.

The books and magnets questions are not isomorphic. Indeed, there are several differences between them that we expect would be meaningful to students. For one, the magnets question involves one system at two different instants, whereas the books question involves two separate systems. In the magnets question, there is an active agent inputting energy into the system, whereas the mechanism by which the books come to be resting on their shelves is obscured. Finally, in the magnets question, the two interacting objects that make up the system for which potential energy is being analyzed are explicit. In the books question, the potential energy correctly is associated with a system consisting of each book and the Earth [50]. The role of the Earth in this system, however, is unlikely to be obvious to students [8]. Examining the wording of the questions, however, it is worth noting that the language used in the books question was very similar to that used in the ions question, whereas the language of the magnets question differed substantially.

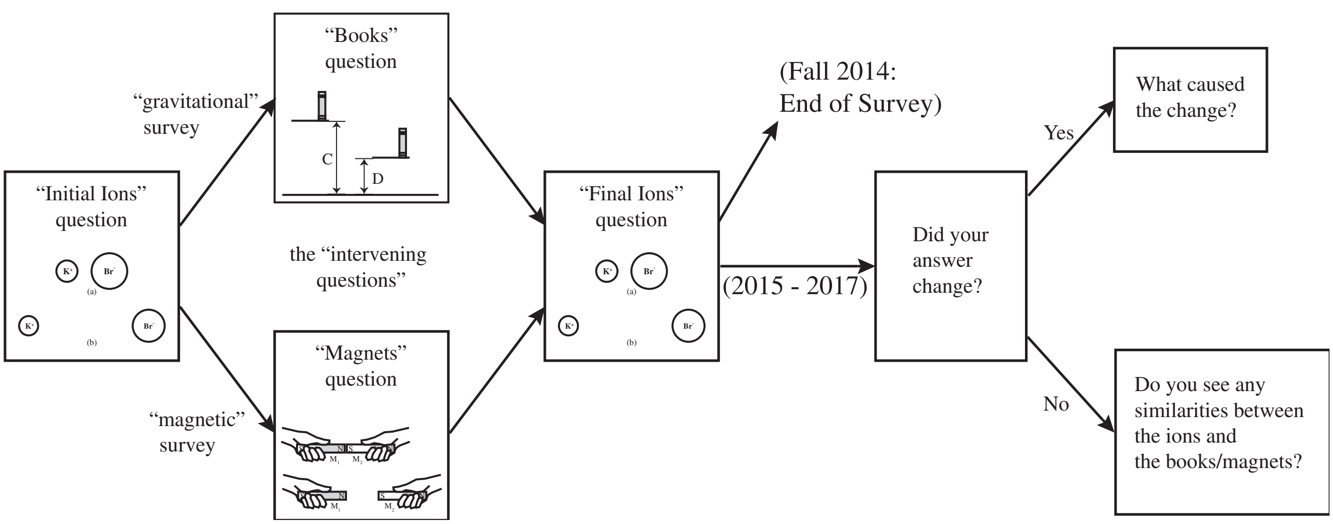

FIG. 4. Overview of the survey flow for the two versions addressed in this paper.

TABLE I. Results from the initial and final ions questions for both the gravitational and the magnetic survey versions.

		% of students responding correctly (95% Wald C.I.)		
Survey version	N	(a) Initial ions question	(b) Final ions question	(c) Shift from initial to final (95% Wald C.I.)
Gravitational	157	29% (22%–36%)	45% (37%–53%)	+16% (+9%– + 23%)
Magnetic	105	36% (27%–45%)	33% (24%–42%)	−3% (−11%– + 5%)

IV. RESULTS

A. The "initial ions" question

Student performance on the initial ions question was typically poor, indicating that some intervention on this topic is warranted. This is consistent with results previously reported for a similar question [7] and provides the primary motivation for this study. Over six semesters and all survey versions ($\underline{N} = 850$ students), on average only 33% of students respond correctly that the potential energy of the system is greater in configuration (b) (95% C.I. based on the variances of the six samples: 30%–36%). This performance is somewhat better than has previously been observed in a General Chemistry course at a small regional campus of a large state university [7], but on par with results reported in an Introductory Physics course [6]. Students responding incorrectly generally selected configuration (a) as having the highest potential energy. This response was typically given by 64% of respondents (95% C.I.: 61%–67%). When asked to explain their reasoning, 50% of the students cited only the distance between the ions with no further explanation. However, those students who did say more followed previously reported trends [6,7], with both the force between the ions and the Coulomb's law equation being commonly referenced. Students also had the option of responding that the potential energy was the same for both configurations, or that it was not possible to determine which configuration had a greater potential energy, but the percentage of students selecting either of these options was typically small ("both": mean 3%, 95% C.I. 2%–5%; "impossible to determine": mean 1%, 95% C.I. 0%–1%).

Performance on the initial ions question, which was identical in every version of the survey, was similar over the course of six semesters. To verify this, a chi-square test of independence was conducted [51]. The chi-square test examines whether there is an association between some measure (in this case, correctness on the initial ions question) and some grouping of the subjects. In this case, students were grouped by semester and version of the survey they had completed (e.g., one such group would be "Fall 2015, gravitational"). The chi-square test assumes that groups are independent, that the measure being tested has at least two independent and exhaustive values (in this case, correct or incorrect on the initial ions question are considered independent because no student could be both correct and incorrect in their response, and they are exhaustive because no student could be classified as something other than "correct" or "incorrect"). Finally, the expected population in any given category (such as, incorrect students on the Fall 2015 gravitational survey) cannot be smaller than 5 students. All of these assumptions were met, and the test reveals no significant difference between the groups in performance on the initial ions questions. Thus, the student populations were deemed identical and students who had responded to the same version of the survey in different semesters were grouped together where applicable.

To compare only students completing the gravitational and magnetic versions of the survey, we treated each instance of the ions question as a binary-response variable (coded as "1" if a student responded correctly and "0" if a student responded incorrectly to the question), and calculated the 95% Wald C.I. [51] (based on the normal approximation to the binomial distribution) for the performance on the initial ions question for each survey version. Results are shown in column (a) of Table I.

B. The final ions question

Student performance on the final ions question differs by survey version. Performance on the final ions question on the gravitational survey represents an improvement over their performance on the initial ions question [see column (b) of Table I]. An exact McNemar test (a nonparametric test which tests for changes in proportion of paired dichotomous nominal data) [52] was used to assess the significance of this difference. The McNemar test indicates that the difference is significant, $p < 0.001$, with an odds ratio of 6 (i.e., students were 6 times more likely to shift from an incorrect answer on the initial ions to a correct answer on final ions than vice versa). On the magnetic survey, performance on the final ions question appears slightly worse than on the initial ions question [see column (b) of Table I]. An exact McNemar test suggests, however, that this difference is not statistically significant ($p = 0.629$), with an odds ratio of 0.7.

C. Comparing changes from the initial to the final ions question

The overall performance of students on both the initial and final ions questions is shown in Fig. 5. Although it is clear that a large majority of students respond identically to both the initial and final ions (responding either incorrectly

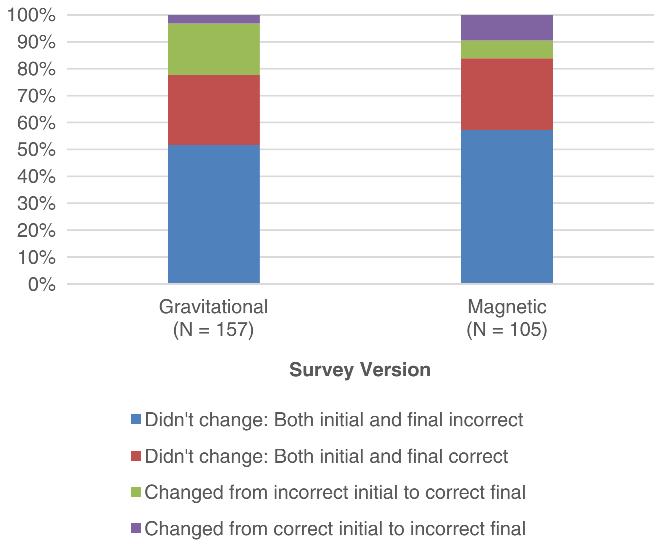

FIG. 5. Overall performance of students on the combination of the initial and final ions questions, indicating how individual students changed their answers (or failed to do so) on each version of the survey.

on both or correctly on both), there is evidence from the figure and from Table I for a difference in the final outcomes based on survey version. To verify this difference, we computed a "change in answer" score for each student. Each student was scored as "0" if they answered either correctly on both initial and final ions questions, or incorrectly on both questions. A student was scored as "1" if they answered incorrectly on the initial ions question, but gave a correct response to the final ions question. Finally, a student was scored as "−1" if they initially responded correctly but changed their response to an incorrect answer on the final ions question. On average, the change for the gravitational survey was 0.16 and for the magnetic survey was −0.03.

By applying the Mann-Whitney U test (a nonparametric test that compares the rank order of scores between two independent samples), [53] we are able to determine if the distribution of the "change in answer" scores differs based on survey version. Results of the test suggest that the survey versions do indeed differ (Mann-Whitney $U = 6776$; $p = 0.001$, $r = 0.2$—a small to medium effect).

To estimate the size of the effect, we calculated separate confidence intervals for the shift in performance on the ions question for the gravitational and magnetic survey versions, treating the results for the initial and final ions questions as dependent samples. These shifts are shown in column (c) of Table I. The difference in shifts between the two survey versions, treating the groups of students completing each survey version as an independent sample, is 18% (95% C.I.: 8%–29%). This suggests that students completing the gravitational version of the survey are more likely to shift from an incorrect answer to a correct answer than students completing the magnetic version of the survey.

TABLE II. Performance on the books and magnets questions. Values in parentheses indicate the associated number of students.

	% of students responding correctly (N)		
	Overall	Among those who responded correctly on initial ions	Among those who responded incorrectly on initial ions
Books question ($N = 157$)	78% (122)	85% (39)	75% (83)
Magnets question ($N = 105$)	36% (38)	74% (28)	15% (10)

D. The intervening questions: books and magnets

Performance on the intermediate questions may provide insights into the performance differences on the final ions questions. On the gravitational survey, 78% of students responded correctly that the system has a larger potential energy when the book is on a higher shelf (see Fig. 3). To test whether the books task is independent from the initial ions task, we use Fisher's exact test. This test is an exact (i.e., not based on approximating a distribution) test of the independence of two different classifications for data (i.e., data that can be represented in a 2×2 contingency table such as the two halves of Table II) [51]. Students were equally likely to give the correct response for the books question regardless of whether or not they had responded correctly on the initial ions question (Fisher's exact test, $p = 0.21$). These results are represented in Table II.

In contrast with the books version, on the magnetic survey only 36% of students recognize that the potential energy of the system increases as the magnets are moved farther apart (see Fig. 4). As is evident in Table II, the students who responded correctly on the magnets question were significantly more likely to have responded correctly on the initial ions question (Fisher's exact test, $p < 0.001$), with an odds ratio of 16.0 (95% C.I.: 6.0–42.8). In other words, performance on the magnets question tended to track with performance on the initial ions question—students who responded correctly on the initial ions question were about 16 times more likely to respond correctly on the magnets question than those who had not. This is somewhat surprising given differences between the ways in which the ions and magnets questions were posed. In particular, the interacting system of objects was made obvious in the magnets question, as was the presence of an active agent inputting energy into the system. These features might be expected to translate to better performance on the magnets question, however, no such effect was observed.

E. Questions regarding causes and similarities

As noted above, students who indicated that their answers or reasoning to the ions question had changed from initial to final were asked whether the change had been caused by the intervening questions (books or magnets). On both surveys, of students who indicated that their answer or reasoning had changed ($N = 27$ for gravitational, $N = 26$ for magnetic) the majority (78% for gravitational, 85% for magnetic) selected the option that this change was due to the intervening questions. However, this change affected students differently depending on which survey they completed. Of the students who answered "yes" when asked if their reasoning had been influenced by the books question, 76% gave a correct answer to the final ions question. On the other survey, however, students who indicated that their answer or reasoning had been influenced by the magnets question were much less likely to arrive at a correct answer to the final ions question. Only 41% of the students who indicated that the magnets had influenced their thinking went on to answer the final ions question correctly.

Even among students who did not change their answer on the ions question, the two survey versions reveal differences. As shown in Table III, overall, students were significantly more likely to choose "yes" when asked if the ions question and the magnets question are similar than they were to choose yes when asked about the ions question and books question. 92% of students indicated that there was a similarity between magnets and ions, whereas only 64% of students recognized any similarity between the books and ions questions. This represents a significant difference (Fisher's exact test, $p < 0.001$), with an odds ratio of 6.8 (95% C.I.: 2.5–18.3). This suggests that the difference in wording between the books and the magnets, with the wording of the books question more nearly mirroring the wording of the ions question, did not prevent students from recognizing the similarities between ions and magnets, nor did the similarity in wording between the books and ions questions provide any particular advantage.

TABLE III. Percent of students indicating that the ions question and the books or magnets questions are similar in some way. This question was posed only to students who did not change their answer from initial ions to final ions, and was not used in the earliest version of the survey (Fall 2014), which is why total N values differ from those elsewhere in the paper. Values in parentheses indicate the associated number of students.

	Percent of students who recognized similarity to ions question.		
Survey version	Total (N)	Among students responding correctly to ions (N)	Among students responding incorrectly to ions (N)
Books ($N = 56$)[a]	64% (36)	84% (21)	48% (15)
Magnets ($N = 79$)[b]	92% (73)	100% (25)	89% (48)

[a]Among these students, 25 (45%) responded correctly on both the initial and final ions question.
[b]Among these students, 25 (32%) responded correctly on both the initial and final ions question.

The differences become even more apparent when data are separated based on the students' responses to the final ions question. Students who responded to the magnets survey uniformly tended to recognize that there was a connection between the magnets question and the ions question, regardless of whether or not they responded correctly to the final ions question (Fisher's exact test reveals no difference between the groups who responded correctly or incorrectly to the final ions, $p = 0.17$). On the gravitational survey, on the other hand, students who responded correctly to the final ions question were likely to recognize that there was a connection between the books and the ions, whereas students who did not respond correctly to the final ions were significantly less likely to recognize this connection (Fisher's exact test, $p = 0.01$), with an odds ratio of 5.6 (95% C.I. 1.6–20.2).

V. DISCUSSION

In many respects, a question about gravitational potential energy in the context of two books on shelves appears more effective in improving student performance on a question about the potential energy of a system of two ions than does a question regarding potential energy in the context of two magnets. Students who had the opportunity to consider the books question were more likely to change from an incorrect answer on the initial ions question to the correct answer on the final ions question, whereas students who considered only the magnets question were unlikely to shift in their answer to the ions question.

Much of the shift in answers on the gravitational survey may be due to the fact that students have a better understanding, in general, of gravitational potential energy than they do of other forms of potential energy, as evidenced by our prior work [7]. This is further supported by the much higher correct response rates on the books question than on the magnets question (78% vs 36% correct). The majority of students (81%) gave consistent answers on the magnets and initial ions questions, suggesting that most students were already thinking similarly about the ions and the magnets. We suspect that due to this similarity, the magnets question did not prompt many students to reconsider their thinking about potential energy. The books question, however, may have reminded students that potential energy can increase with distance. Many students did not explicitly describe what allowed them to successfully transfer this idea to the ions scenario. For example, one stated simply "in figure b the ions are farther apart, there is more distance in between them. My answer has changed because I realized that the greater the distance in between the objects the greater the potential energy." Some students, however, articulated their thinking more clearly: "I believe the higher book will have more potential energy because it has a greater distance to fall, so my answer to the ion question changed based on the distance between." "The ones that are farther from the object they're attracted to (book and floor) (K and Br particle) have higher potential energy." These student responses suggest that students are successfully developing a more correct general rule or heuristic for questions about potential energy. This resulted in a larger fraction of correct responses on the final ions task.

In one respect, however, the magnets task holds an advantage over the books task—students are more likely to recognize that the questions about potential energy in the context of the magnets and the ions are similar than they are for the books and the ions. This occurs despite surface similarities in the wording of the books and the ions which are absent in the magnets task. If students do not see a similarity between the potential energy clearly, they will be unlikely to translate their correct understanding of gravitational potential energy to a correct understanding of electrostatic potential energy. This is consistent with research into student learning using analogies, which suggests that "the instructional value of an analogy decreases if it is difficult to identify and map the important features shared by the analog and the target." [31]. This suggests that to support student learning of electrostatic potential energy, which is so fundamental to their understanding in other STEM courses, [5] instructors should find ways to aid students in drawing connections between gravitational contexts and electrostatic contexts.

We do not claim that after completing the books question students have "learned" electrostatic potential energy. Viewed through the lens of dual-process theories of reasoning, [40,42] it seems likely that when initially confronting the ions question, a common initial response generated by the heuristic process is that potential energy is greater when the ions are closer together. This may be connected to a "closer means stronger" heuristic or *p*-prim [47] tied to the force in these situations, which is indeed greater when the ions are closer. In the absence of any threat to this intuitive response, the analytic process would tend to accept this response in accordance with the "satisficing principle." We speculate that consideration of the books question causes the analytic process to reengage and in some cases reject the initial intuitive response when students are once again exposed to the ions question. The magnets question, on the other hand, cues the same intuitive response as the ions, and thus even if the analytic system engages, it would have no reason to intervene and reject the initial intuitive response. Supporting students in developing an understanding of electrostatic potential energy can thus be viewed as a process of helping them to build new intuitions about potential energy. The books task may provide one helpful element in this process.

The instrument described in this study represents a "microintervention" that was administered near the end of the semester to a self-selected group of students who had enrolled in a general chemistry support course. These students had already been exposed to concepts relevant to the intervention at least once. As a result, we do not claim that the surveys administered in this study will have a lasting impact on student understanding of electrostatic potential energy. Indeed, research suggests that quick and isolated fixes are unlikely to leave long-term effects [41]. We recognize that, with at most 45% of students responding correctly to the final ions question, and with only a relatively small proportion of students changing their answer from the initial to the final ions question, there is still much room for improvement in student performance. Given the extremely minor intervention that our survey represents, however, it is noteworthy that any answers changed, much less that we observed distinct differences in the changes between the two survey versions.

This study is a small piece of a larger initiative that seeks to enable students to become more aware of their heuristic and analytic processes. We ultimately hope that this study can inform the design and implementation of research-based instructional revisions that could lead to lasting improvements in student reasoning. While a single "quick fix" may not have a lasting impact on student reasoning, a strategy involving repeated exposure to questions with demonstrated effectiveness for some students in the short term may help students to build productive and durable new intuitions about potential energy that will help them in the long term.

VI. CONCLUSIONS AND IMPLICATIONS

A solid foundational understanding of the energy associated with interactions between electrostatic charges is fundamental to understanding many phenomena across the STEM disciplines, particularly chemical phenomena. We have empirically demonstrated that student performance on a question about electrostatic potential energy improves when students consider a related question about gravitational potential energy, but does not improve when they instead consider a question about the potential energy of a system of two magnets. Many students, however, still respond incorrectly after either intervention. Results suggests that appealing to students' understanding of gravitational potential energy may be more effective for building an understanding of electrostatic potential energy than invoking the potential energy of systems of magnets, even though students are more likely to recognize a connection between magnets and the ions than between the books and the ions.

We interpret these results through the lens of dual-process theories of reasoning. Students may answer the ions question incorrectly at first due to an initial mental model involving the intuitively appealing idea that "closer means stronger." Reasoning about the gravitational potential energy associated with books on shelves may cause the analytic process to engage on the ions question in a way that reasoning about magnets does not. The mechanisms by which the books question lessens the intuitive appeal of the closer means stronger approach, or renders the correct response more appealing, are not yet well understood. In particular, we wish to investigate the connections that students perceive between the books and ions questions to identify which connections are associated with an effective intervention of the analytic process. While the current study does not examine our data through the lens of an analogical reasoning framework [33], we plan in the future to investigate whether such a framework might help us to better understand why the analytic system is more likely to engage productively when presented with the books scenario rather than with the magnets scenario. In particular, we expect that students who recognize that the gravitational potential energy in the books scenario is associated with the attractions within a system of two objects (the book and the Earth) might be more likely to apply their knowledge of gravitational potential energy to the ions scenario. The testing of this hypothesis will be the subject of future work.

A previous study [7] had shown that a longer instructional sequence involving both the books question and a related question about the potential energy of a system including the Earth and an orbiting space shuttle could help students to better understand electrostatic potential energy; we now find that even a more minimal intervention can have an effect. We have demonstrated that the effect of exposing students to a gravitational question is different than the effect of exposing them to a magnetic question. While questions about the longevity and aftereffects of this effect remain open, this finding can inform the development of instructional strategies that will benefit physics students from many STEM disciplines.

ACKNOWLEDGMENTS

The authors are extremely grateful to Mary Jo Bojan for allowing us access to her course for the duration of this study. Thanks are also due to Mila Kryjevskaia and Andrew Boudreaux for their helpful commentary on this article. This material is based upon work supported by the National Science Foundation under Grants No. DUE-1431857, No. DUE-1431541, No. DUE-1431940, No. DUE-1432765, and No. DUE-1432052.

[1] The Joint Task Force on Undergraduate Physics Programs, *Phys21: Preparing Physics Students for 21st-Century Careers* (American Physical Society, College Park, MD, 2016).

[2] N. M. Becker and M. M. Cooper, College chemistry students' understanding of potential energy in the context of atomic-molecular interactions, J. Res. Sci. Teach. **51**, 789 (2014).

[3] M. M. Cooper and M. W. Klymkowsky, The trouble with chemical energy: why understanding bond energies requires an interdisciplinary systems approach., CBE Life Sci. Educ. **12**, 306 (2013).

[4] B. W. Dreyfus, J. S. Gouvea, B. D. Geller, V. Sawtelle, C. Turpen, and E. F. Redish, Chemical energy in an introductory physics course for the life sciences, Am. J. Phys. **82**, 403 (2014).

[5] B. Venkataraman, Emphasizing the significance of electrostatic interactions in chemical bonding, J. Chem. Educ. **94**, 296 (2017).

[6] B. A. Lindsey, Student reasoning about electrostatic and gravitational potential energy: An exploratory study with interdisciplinary consequences, Phys. Rev. ST Phys. Educ. Res. **10**, 013101 (2014).

[7] M. L. Nagel and B. A. Lindsey, Student use of energy concepts from physics in chemistry courses, Chem. Educ. Res. Pract. **16**, 67 (2015).

[8] B. A. Lindsey, P. R. L. Heron, and P. S. Shaffer, Student understanding of energy: Difficulties related to systems, Am. J. Phys. **80**, 154 (2012).

[9] B. A. Lindsey, P. R. L. Heron, and P. S. Shaffer, Student ability to apply the concepts of work and energy to extended systems, Am. J. Phys. **77**, 999 (2009).

[10] NGSS Lead States, Next Generation Science Standards: For States, By States (The National Academies Press, Washington, DC, 2013).

[11] K. Bain, A. Moon, M. R. Mack, and M. H. Towns, A review of research on the teaching and learning of thermodynamics at the university level, Chem. Educ. Res. Pract. **15**, 320 (2014).

[12] B. W. Dreyfus, B. D. Geller, D. E. Meltzer, and V. Sawtelle, Resource letter TTSM-1: Teaching thermodynamics and statistical mechanics in introductory physics, chemistry, and biology, Am. J. Phys. **83**, 5 (2015).

[13] K. P. Kohn, S. M. Underwood, and M. M. Cooper, Energy connections and misconnections across chemistry and biology, CBE—Life Sci. Educ. **17**, ar3 (2018).

[14] D. P. Ausubel, *Educational Psychology: A Cognitive View* (Holt, Reinhart, and Winston, Inc., New York, 1968).

[15] K. S. Taber, Building the structural concepts of chemistry: Some considerations from educational research, Chem. Educ. Res. Pract. **2**, 123 (2001).

[16] S. S. Zumdahl and S. A. Zumdahl, *Chemistry*, 9th ed. (Cengage Learning, Belmont, CA, 2013).

[17] T. L. Brown, H. E. LeMay, B. E. Bursten, C. J. Murphy, and M. Woodward, Patrick, *Chemistry: The Central Science*, 12th ed. (Prentice-Hall, Boston, 2012).

[18] D. C. Giancoli, *Physics for Scientists and Engineers with Modern Physics*, Vol. II, 3rd ed. (Prentice-Hall, Upper Saddle River, NJ, 2000).
[19] J. D. Cutnell and K. W. Johnson, *Physics*, 7th ed. (John Wiley & Sons, New York, 2007).
[20] F. M. Goldberg, V. K. Otero, and S. Robinson, *Physics and Everyday Thinking*, 2nd ed. (It's About Time, Armonk, NY, 2007).
[21] R. A. Kjonaas, An analogy for solubility: Marbles and magnets, J. Chem. Educ. **61**, 765 (1984).
[22] B. W. Dreyfus, E. F. Redish, and J. Watkins, Student views of macroscopic and microscopic energy in physics and biology, 2011 Physics Education Research Conference, AIP Conf. Proc. **1413**, 179 (2012).
[23] N. S. Podolefsky and N. Finkelstein, Analogical scaffolding and the learning of abstract ideas in physics: An example from electromagnetic waves, Phys. Rev. ST Phys. Educ. Res. **3**, 010109 (2007).
[24] B. Stephanik and P. S. Shaffer, Examining student ability to interpret and use potential energy diagrams for classical systems, AIP Conf. Proc. **1413**, 367 (2012).
[25] B. W. Dreyfus, B. D. Geller, J. S. Gouvea, V. Sawtelle, C. Turpen, and E. F. Redish, Ontological metaphors for negative energy in an interdisciplinary context, Phys. Rev. ST Phys. Educ. Res. **10**, 020108 (2014).
[26] J. D. Novak, Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners, Sci. Educ. **86**, 548 (2002).
[27] H. C. Sabo, L. M. Goodhew, and A. D. Robertson, University student conceptual resources for understanding energy, Phys. Rev. Phys. Educ. Res. **12**, 010126 (2016).
[28] A. Collins and D. Gentner, How people construct mental models, in *Cultural Models in Language and Thought*, edited by D. Holland and N. Quinn (Cambridge University Press, New York, 1987), pp. 243–265.
[29] N. S. Podolefsky and N. Finkelstein, Use of analogy in learning physics: The role of representations, Phys. Rev. ST Phys. Educ. Res. **2**, 020101 (2006).
[30] N. S. Podolefsky and N. Finkelstein, Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies, Phys. Rev. ST Phys. Educ. Res. **3**, 020104 (2007).
[31] S. M. Glynn, *Teaching science with analogies: A strategy for teachers and textbook authors, Reading Research Rep. No. 15* (National Reading Research Center, Athens, GA, 1994).
[32] M. Orgill and G. Bodner, What research tells us about using analogies to teach chemistry, Chem. Educ. Res. Pract. **5**, 15 (2004).
[33] D. Gentner, Structure-mapping: A theoretical framework for analogy, Cogn. Sci. **7**, 155 (1983).
[34] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 10th ed. (Wiley, New York, 2013).
[35] S. M. Underwood, D. Reyes-Gastelum, and M. M. Cooper, 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**, 365 (2016).
[36] M. M. Cooper, M. W. Klymkowsky, and N. Becker, Energy in chemical systems: An integrated approach, in *Teaching and Learning of Energy in K–12 Education*, edited by R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, and A. Scheff (Springer International Publishing, Cham, 2014), pp. 301–316.
[37] T. L. Nahum, R. Mamlok-Naaman, A. Hofstein, and K. S. Taber, Teaching and learning the concept of chemical bonding, Stud. Sci. Educ. **46**, 179 (2010).
[38] J. Joki, J. Lavonen, and M. Aksela, Research and Practice Coulombic interaction in Finnish middle school chemistry: A systemic perspective on students' conceptual structure of chemical bonding, Chem. Educ. Res. Pract. **16**, 901 (2015).
[39] K. S. Taber, The Mismatch between Assumed Prior Knowledge and the Learner's Conceptions: A typology of learning impediments, Educ. Stud. Math. **27**, 159 (2001).
[40] J. St. B. T. Evans, The heuristic-analytic theory of reasoning: Extension and evaluation, Psychon. Bull. Rev. **13**, 378 (2006).
[41] V. Talanquer, Chemistry education: Ten heuristics to tame, J. Chem. Educ. **91**, 1091 (2014).
[42] M. Kryjevskaia, M. R. Stetzer, and N. Grosz, Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics, Phys. Rev. ST Phys. Educ. Res. **10**, 020109 (2014).
[43] M. Kryjevskaia and M. R. Stetzer, Examining inconsistencies in student reasoning approaches, AIP Conf. Proc. **1513**, 226 (2013).
[44] M. Kryjevskaia, M. R. Stetzer, and T. K. Le, Failure to engage: Examining the impact of metacognitive interventions on persistent intuitive reasoning approaches, *Proceedings of the 2014 Physics Education Research Conference 2014, Minneapolis, MN* (AIP, New York, 2015), p. 143.
[45] C. R. Gette, M. Kryjevskaia, M. R. Stetzer, and P. R. L. Heron, Probing student reasoning approaches through the lens of dual-process theories: A case study in buoyancy, Phys. Rev. Phys. Educ. Res. **14**, 010113 (2018).
[46] J. St. B. T. Evans and K. E. Stanovich, Dual-process theories of higher cognition, Perspect. Psychol. Sci. **8**, 223 (2013).
[47] A. A. DiSessa, Knowledge in pieces, in *Constructivism in the Computer Age*, edited by G. Forman and P. Pufall (Lawrence Erlbaum Associates, Hillsdale, NJ, 1988), pp. 49–70.
[48] M. J. Bojan (private communication).
[49] J. M. Nissen, M. Jariwala, E. W. Close, and B. Van Dusen, Participation and performance on paper- and computer-based low-stakes assessments, Int. J. STEM Educ. **5**, 21 (2018).
[50] J. W. Jewett, Energy and the confused student II: Systems, Phys. Teach. **46**, 81 (2008).
[51] A. Agresti, *Categorical Data Analysis*, 3rd ed. (John Wiley & Sons, Hoboken, NJ, 2013).
[52] M. A. Morrison, McNemar's test, Encycl. Res. Des. **780**, (2012).
[53] P. R. Hinton, Mann-Whitney U Test, Encycl. Res. Des. **748**, (2010).