

Variability in student reasoning about radioactive decay as a stochastic processMichael M. Hull¹, Alexandra Jansky², and Martin Hopf³¹*University of Alaska Fairbanks, Department of Physics, Fairbanks, Alaska, USA*²*University of Zurich, Educational Development, Zurich, Switzerland*³*University of Vienna, Physics Education Research, Vienna, Austria*

(Received 24 January 2025; accepted 3 December 2025; published 8 January 2026)

For several years, we have been investigating student conceptual understanding of radioactivity, particularly regarding half-life. Although it is considered to be random when an individual nucleus undergoes the sudden process of radioactive decay, the duration of time within which half of a radioactive sample decays (i.e., the half-life) is a predictable quantity. Many students generally struggle with the idea that random behavior of individual agents can give rise to predictable patterns in the collective; specific to radioactivity, many students have said both in interviews and surveys that, if you are looking at an individual nucleus, half of the nucleus will have decayed after one half-life. Our findings have indicated, however, that this idea (of individual nuclei decaying in a predictably continuous manner) is often not a robust and intact mental structure; rather, in other contexts, the same students correctly discuss radioactive decay as being instantaneous and unpredictable. We developed a survey to measure this context dependency and administered it both to high school students in Austria and to college students at the University of Alaska Fairbanks. Using both cluster analysis and alluvial diagrams to organize our data, we find a great variety in answering patterns across the survey prompts, indicating that student ideas have not yet crystallized. Our paper thus provides experimental evidence in support of a “knowledge in pieces” model of student cognition about radioactivity as opposed to a “misconceptions” model.

DOI: [10.1103/PhysRevPhysEducRes.22.010102](https://doi.org/10.1103/PhysRevPhysEducRes.22.010102)**I. INTRODUCTION**

There are several reasons why high school children learn about radiation and radioactivity. These topics pervade our everyday life in the medical, industrial, energy, and entertainment sectors. As we all know, the Simpsons live in a city that lies very close to a nuclear power plant. During the episodes of this TV show, this fact sometimes leads to interesting plots of the stories. Due to the popularity of this show, we have good reason to assume that students of all age groups have been—at least in a comical form—introduced to radiation and radioactivity. In the real world, medical procedures that utilize ionizing radiation, such as gamma scintigraphy, x-rays, and CT scans, are becoming more and more prevalent. The topic of half-life—the characteristic time in which half of a radioactive sample undergoes radioactive decay—allows researchers to estimate the age of objects used by ancient civilizations and even the age of the Earth itself. An understanding

of half-life is also crucial for making informed decisions about radioactive waste.

There is scientific consensus that the temperature of the Earth’s surface is increasing and that this is caused by human activity, primarily CO₂ emissions. In their 2030 Agenda for Sustainable Development, the United Nations described climate change as “one of the greatest challenges of our time” [1]. They listed 17 “sustainable development goals,” with goal number 7 being to “ensure ... sustainable and modern energy for all,” particularly for least developed countries, which are expected to require significant increases in energy per capita as they develop. Energy sources that are “sustainable,” in the sense that they do not exacerbate climate change, are those that emit minimal greenhouse gases. The International Atomic Energy Association, a subgroup of the United Nations, describes nuclear fission power as being a candidate for such an energy source, and their 2023 projection “now sees a quarter more nuclear energy capacity installed by 2050 than it did as recently as 2020, underscoring how a growing number of countries are looking to this clean and reliable energy source to address the challenges of energy security, climate change and economic development” [2]. Over the source life cycle, nuclear power plants produce fewer greenhouse gases per kilowatt-hour than photovoltaic cells, and air pollution is similarly marginal (IAEA 2017) [3].

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

Many argue that green energy alone will be insufficient to handle the additional energy needs of developing China, India, and African countries, even if it could replace fossil fuels to meet today's needs. As a result, nuclear energy must play a role in the energy mix if we prioritize alleviating the effects of climate change [4].

As with any energy source, nuclear energy has disadvantages. Fission reactors rely upon rare-earth elements, so named because they are relatively scarce. Although the World Nuclear Association estimates a sufficient uranium supply for over 90 years, that is assuming that growth follows the restricted trajectory described above [5]. The upfront capital investment in building a nuclear power plant is prohibitively high. Even more prohibitive is the struggle to find waste storage sites that can withstand the extremely long time required for the waste's radioactivity to reach acceptable levels. The primary ingredient in spent fuel pins, as in the mined uranium ore itself, is U-238, which has a half-life of 4.5 billion years. As such, these repositories must outlive not only political changes and shifting territory boundary lines, but also dramatic geological changes (such as the next ice age). There are currently no operating permanent repositories for high-level nuclear waste (although Finland may change that in the near future [6]). While part of the difficulty in finding appropriate waste storage locations is technological, much of the challenge comes from a "not in my backyard!" mindset from the public. Generally, the public has a very negative image of nuclear energy, which includes the above reasons as well as a host of other ideas that are either exaggerations or falsehoods (for example, that a nuclear power plant can explode with the devastating power of a nuclear bomb [7]). When it comes to making decisions about the future of energy, the public should be educated and informed about nuclear energy, particularly regarding the issue of where to store nuclear waste. Central to understanding this issue is an understanding of the concept of half-life.

Prior work by the authors [8–12] has investigated student understanding of half-life by interviewing and administering surveys to Austrian high school students who had (ostensibly) already learned the topic in school. This paper expands upon our prior work by comparing the former data with survey responses from Alaskan college students enrolled in introductory physics. By comparing the diverse populations via cluster analysis and alluvial diagrams, we address two research questions: (RQ1) To what extent do diverse populations of students learning introductory physics compare in their understanding of half-life? and (RQ2) How can we characterize student understanding of half-life?

II. THEORETICAL BACKGROUND

A. Student reasoning about radioactivity

To date, researchers have documented a number of student ideas regarding radioactivity, including several

ideas regarding the time it takes for radioactive nuclei to decay and what the implications are for that decay [13–16]. There were many media reports from the Chernobyl and Fukushima nuclear disasters, which described half-life as "the period in which a radioactive isotope loses its radioactivity" [14] or which used statements like "Iodine-131 has a short life" [15]. Furthermore, a Delphi study revealed that some radiation experts have encountered the lay idea that "after the half-life there is no danger left" [7]. It has been found that many students think that a radioactive sample loses half of its mass after one half-life (in fact, the daughter nuclei, which generally remain in the sample, are essentially the same mass) [16]. Henriksen [17] argued that the issue stems from the term "half-life" itself, as the term "invites the invention of explanations: whereas it is impossible to guess what a becquerel might be for one who has not been given the definition, the very word 'half-life' suggests that something is halved. For someone lacking a formal level of knowledge in the area of radiation, it might seem reasonable to suppose that it is the mass that is halved—mass being a very tangible concept with which most people are familiar." It has also been argued that part of the underlying difficulty could be a failure to understand the stochastic nature of radioactivity [13,18]. Indeed, studies examining student reasoning about the decay of individual nuclei have found that many students assume that if half of a radioactive substance has decayed after one half-life, then it follows that half of each individual nucleus making up the sample must have decayed [8,19,20].¹

We have argued that, particularly regarding ideas about the timing of the decay of an individual nucleus, the difficulty could arise in part because of a failure to understand radioactivity as an emergent process [8]. Wilensky *et al.* [21] discussed how students demonstrate a "level confusion" when they assume that the agent level (in this case, individual nuclei) and the system level (the overall radioactive sample) share the same property (being half-transformed after one half-life). In our prior work with Austrian high school students, we showed that this "level confusion" is not necessarily a stable and rigid cognitive structure; students often seem to have myriad ideas available to them simultaneously, including ideas about the random nature of individual nuclei, and different ideas come to the fore depending upon the specific context that students are considering [8,9,22].

B. Methods to assess the structure of student ideas

Much of physics education research concerns the investigation of student ideas and how they interact with what students are meant to learn in the classroom. Various theoretical models have been proposed to describe these

¹In fact, for a single nucleus, there is a 50% chance for decay to occur in a time period of one half-life, provided, of course, that the nucleus has not yet decayed at the start of that time period.

student ideas, and it remains a question of debate which model is most effective in which situation. One point in which these models differ is the underlying assumption about the robustness and rigidity of the student ideas themselves. For example, whereas the Theory Theory model of Vosniadou and colleagues [23] attributes relative stability and difficulty to change to student ideas, the Knowledge in Pieces model of diSessa and others views student ideas as context dependent and fluid, shifting moment by moment [24,25]. Scherr [26] identified eight properties of student ideas in setting up a dichotomy between the “misconceptions” and “pieces” models of student thinking, including that a misconception is context independent, stable, and difficult to change. This distinction is not merely interesting from a theoretical perspective; rather, it has direct implications for effective instruction. If student ideas are stable, context independent, and resistant to change (that is, a “misconception” or a “theory” [26,27]), then a cognitive conflict strategy is recommended [28,29]. If, on the other hand, student ideas are better described as being composed of smaller pieces of knowledge, then instruction may be more effective if instructors have students capitalize on the knowledge pieces embedded within their ideas that are shared with the expert view of the material [24,26,27]. diSessa maintains that it is important to build upon prior knowledge: “Students have a richness of conceptual resources to draw on. Attend to their ideas and help them build on the best of them” (p. 41 of Ref. [25]). In this paper, we use the term “misconception” to indicate a stable cognitive structure that is activated in a wide variety of contexts and that is difficult to change. We use the phrase “student idea” as a general term that includes not only misconceptions but also ideas that are transient in nature, composed of knowledge pieces that are only loosely connected and thus context sensitive.

A significant amount of research has documented fluidity in student reasoning through think-aloud interviews and in the classroom as students respond to various contextual cues [24,30–36]. Recently, education researchers have turned their attention to developing surveys that can gain insight into the structure of student ideas [37–44]. Free-response survey prompts allow students to describe a wide range of ideas that they have pertaining to a given situation, rather than selecting the one that feels most salient to them, as they generally must do with multiple-choice surveys. On the other hand, by constraining the range of responses a student can give, a multiple-choice format can make it easier for students to understand what the prompt is asking. In general, multiple-choice prompts require less time for students to complete the survey and less time for education researchers or teachers to assess the responses.

When administering an assessment, the shorter the test time, the better. Many instructors feel pressure to cover a demanding expanse of content by the end of the semester, and they are likely to be reluctant to carve

time out of their crowded schedules to administer the assessment [45]. The fact that many assessments are to be administered twice (pre- and post-test) exacerbates the situation. To avoid using class time for assessments, some instructors have students complete surveys online [45,46]. Although this preserves in-class time, it does not change the fact that it takes time away from students that could otherwise be spent preparing for the course. Efforts to reduce testing time with minimal loss of accuracy and precision are an active area of research in physics education. For example, Han *et al.* [47,48] divided the Force Concept Inventory (FCI, a 30-item survey that consumes most of a class period) into two half-length tests containing different subsets of the original assessment. Recently, Yasuda *et al.* [49] demonstrated the potential use of computerized adaptive testing (CAT) to reduce test time with the FCI.

It is possible to “split the difference” by having part of the survey be multiple choice and part free response. Weinlader *et al.* [38] had students choose one option on a multiple-choice item that they thought was the correct answer and also provide alternative explanations for the other possible responses on the item. Weinlader *et al.* concluded that “By having students give their alternative explanations to a question, we can examine how ‘close to the surface’ those resources are, and future research can investigate whether the availability of these resources (even if they are not believed) predicts the success of resource-focused instruction” [38].

Our work so far [9] has found the use of isomorphic problems [50] to be promising. Isomorphic problems are problems that require the same conceptual understanding to answer but have different surface features that may result in a given student answering correctly on only some of the problems. Singh argued that the reason for this fluidity in reasoning is that “problem context with distracting features can trigger the activation of knowledge that a student thinks is relevant but which is not actually applicable in that context” [50]. Similarly, physics education researchers operating within frameworks of dual-process theories discuss how a “salient distracting feature” (SDF) in a prompt can cue students into responding with a gut reaction that they would have been less likely to activate in a version of the prompt without the SDF [51–54]. In such a case where knowledge is triggered in some problem contexts but not others, it would be inappropriate to think of students as having a stable and resistant “misconception.” We see, therefore, that isomorphic multiple-choice questions are a relatively time-saving means to collect information about the context sensitivity of student reasoning. A survey that is purely multiple choice has the greatest potential to reduce time demands on both respondents and those who assess those responses. Surveys that utilize isomorphic prompts are able to probe for fluidity while maintaining this efficient multiple-choice format.

The concept of isomorphic problems can be traced back to Simon and Hayes [55–57], who, in their investigation of problem solving, defined two problems as isomorphic if they had problem spaces with the same structure. As such, the similarity of isomorphic problems can be as superficial as being the same problem but with different numbers used for relevant parameters; similarly, the same problem posed twice, once asking for a numerical answer and once asking for a qualitative relationship, constitutes an isomorphic pair of problems. Singh [58] wrote that two prompts are isomorphic if they “require the same physics principle to solve them.” In the most extreme case of isomorphic prompts, one can make “the surface features of the problems very different as in the problem pair chosen by Simon and Hayes [the ‘Tower of Hanoi’ and ‘the cannibal and the missionary problem’] or by introducing distracting features into one of the problems.”

Singh [58] investigated student responses to two isomorphic prompts: (i) a spinning wheel that is dropped on a surface with friction and (ii) a pool ball that is struck with a force passing through the center of mass. In both scenarios, students were asked whether or not the final speed of the object would depend on the coefficient of friction and, if so, how. The two prompts require the same physics principles to solve, and they happen to have the same answer (that the final speed does not depend upon the coefficient of friction). However, the surface features of the two prompts are dramatically different. Whereas the linear speed of the pool ball decreases due to friction, the wheel’s linear speed is actually increased by the friction force. Singh found that, as a result of the very different surface features creating the contexts of the two prompts, students “sometimes used different knowledge resources.” With the problem of the pool ball, many students reasoned that a higher coefficient of friction would result in a lower linear speed. Only about half as many students reasoned this way with the spinning wheel problem (where the linear speed starts at zero). If one would give only the pool ball problem to one of these students, the researcher or teacher might see only evidence of the student thinking that friction always reduces the speed of an object and might label that a misconception. By using the second isomorphic prompt, however, the examiner can see that the student can, in certain problem contexts (an object beginning “at rest” with no other forces present that could plausibly increase the linear speed), reason about friction increasing the speed of an object.

Although the isomorphic problems discussed in Ref. [58] were free response, Singh discusses in Ref. [50] other pairs of problems that are multiple choice. For example, on one item of Singh’s survey (Q.21), which concerns two people pulling in opposite directions on a stationary box, many students chose the correct answer “(a) $F = f$ because the mass is not accelerating.” On an isomorphic prompt (Q.22), however, when asked how large

friction must be on a table such that the table does not move when you push it, many students reasoned in a different way, saying that the friction force is equal to the coefficient of static friction times the normal force (i.e., larger than the force of the push). Even in this multiple-choice format, we can see that Newton’s second law is not necessarily a rigidly held idea that students will consistently use. In the presence of an SDF (such as a coefficient of friction provided in the problem statement), students may resort to formulaically setting the friction force to equal the coefficient of static friction times the normal force. As we discuss in the next section, we utilize isomorphic prompts to investigate the extent to which student ideas about radioactivity are rigidly held and intact, and the extent to which they are fluid and context dependent.

III. METHODOLOGY

Our research discussed in this paper is quantitative in nature, and we address our research questions with descriptive statistics on data collected with our survey, the Stochastic World of Radioactive Decay Evaluation (SWORDE). SWORDE was motivated by a pilot study of seven interviews that utilized an approach similar to Brown and Clement’s bridging strategy [59] to explore the fluidity of student reasoning about half-life. We created SWORDE to see how widespread and to what degree the context-dependency we observed is. For details about these seven interviews and the development of SWORDE, see Refs. [8,9,22]. For this study, we examine two prompts from SWORDE: (i) a decision of which day would be best to watch an individual unstable nucleus trapped in a cage decay (“CAGE”) [20], and (ii) the “many vs one” (“MvO”) prompt, asking students to compare the amount of a radioactive substance that has not yet decayed after an integer number of half-lives, when starting with many atoms (“MvO:MANY”) and when starting with just one atom (“MvO:ONE”). Both of these prompts specify that the radioactive substance is I-131, and an explanation of half-life is provided that includes the information that, after eight days, half of the I-131 sample will “transform into a different atom” if one begins with a large number of atoms. Figures 1 and 2 show the final English version of the MvO and CAGE prompts, respectively, used at the University of Alaska Fairbanks (UAF).

We conceptualize both prompts as isomorphic in that both can be answered correctly if one understands that it is random when an individual nucleus undergoes radioactive decay. In both prompts, the information about the half-life serves as an SDF: students are more likely to answer MvO: ONE and CAGE correctly if they ignore the information about half-life. In addition to asking students to select an answer, the questions also ask students to select a justification in the second tier of the question.

Iodine-131 is an example of a radioactive atom. It has a half-life of 8 days, so if one begins with a large number of these atoms, then half of the atoms will have transformed into a different atom after 8 days. The following questions relate to various situations with Iodine-131. Imagine that there are 100 million Iodine-131 atoms in the beginning. How much Iodine-131 will have not yet transformed ... i. ... after 8 days? A) 100 million atoms B) 50 million atoms C) 0 atoms D) 100 million OR 0 atoms E) 65.5 atoms ii. ... after 16 days? A) 100 million atoms B) 25 million atoms C) 0 atoms D) 100 million OR 0 atoms E) 32.8 atoms Explain briefly how you came to your answers. Select all statements that led to your answer. ☐ After the half-life, half of the atoms have transformed. ☐ It is random. ☐ One cannot have half an atom. ☐ Atoms do not disappear. ☐ After two half-lives, all of the atoms have transformed. ☐ In the end, all atoms have transformed	<i>[MvO:MANY appeared later in the survey, separated from MvO:ONE. The same explanation about half-life was given at the start of MvO:ONE]</i> Imagine that there is just a single Iodine-131 atom in the beginning. How much Iodine-131 will have not yet transformed ... i. ... after 8 days? A) One atom B) One half atom C) No atoms D) One atom OR no atoms E) 65.5 atoms ii. ... after 16 days? A) One atom B) One quarter atom C) No atoms D) One atom OR no atoms E) 32.8 atoms Explain briefly how you came to your answers. Select all statements that led to your answer. ☐ After the half-life, half of the atom will have transformed. ☐ It is random. ☐ One cannot have half an atom. ☐ Atoms do not disappear. ☐ After two half-lives, the whole atom will have transformed. ☐ In the end, the atom will have transformed
---	--

FIG. 1. The MvO prompt administered to college students at UAF. MvO:MANY is on the left and MvO:ONE is on the right. The Austrian high school students had a German version of the prompts.

A. Survey administration and coding

Between November 9, 2020, and March 24, 2021, SWORDE collected data from 527 respondents in Austria via Survey Monkey. Of these, 80 respondents were instructors who were considering administration of the survey to their students. Since our research question involves looking for consistency of student responses across two isomorphic prompts, we removed all survey respondents (157) who did not answer both prompts. The survey begins by asking students if they have learned about half-life before. The survey was administered only to students in 11th and 12th grades (17- to 18-year-olds). Although the topic of half-life is part of the national mathematics curriculum in tenth grade, 24 of the remaining students answered that they had not learned about half-life previously. Their responses were removed. Finally, the first

part of MvO:MANY (asking for how much of the radioactive substance will remain after one half-life) served as a screening question to remove an additional 32 respondents. Specifically, despite the explanation about half-life just prior to the prompt, a number of students nevertheless selected either “100 million atoms” (the starting amount), “0 atoms,” or “100 million or 0 atoms” for the amount that would not have yet decayed after 8 days. Since we had encountered no difficulty in understanding this prompt in any of the nine survey validation interviews that were involved with the survey creation, we assumed that these responses were due to not taking the survey seriously, and the respondents were removed. After these measures were taken, a total of $N = 234$ responses remained, and these responses were used for all analyses discussed in this article. We justified grouping these data into one dataset,

Iodine-131 is an example of a radioactive atom. It has a half-life of 8 days, so if one begins with a large number of these atoms, then half of the atoms will have transformed into a different atom after 8 days. The following questions relate to various situations with Iodine-131.

Imagine that your friend has created a single Iodine-131 atom and is keeping it locked in a cage. You would like to see the transformation of the atom. Your parents, however, will allow you to miss school for only one day. Would you go on the day that the atom was made? Or would you go on a later day? On which day?

- A) On the day the atom was made
- B) On the next (second) day
- C) After the second day, but before the 7th day
- D) On the 7th day
- E) On the 8th day
- F) On the 9th day
- G) Later than the 9th day
- H) It does not matter when I go. I can choose the timing completely arbitrarily.

Explain briefly how you came to your answer. Select all statements that led to your answer.

- ☐ It is unpredictable when the atom transforms.
- ☐ The atom transforms continuously.
- ☐ The transformation goes very quickly after the atom is made.
- ☐ After the half-life, half of the atom will have transformed.
- ☐ If I wait too long, the atom will have already transformed.
- ☐ The atom transforms on the day of the half-life.

FIG. 2. The CAGE prompt administered to college students at UAF. The Austrian high school students had a German version of the prompts.

despite being collected from multiple classes, based on results from a cluster analysis [60]. Specifically, although three clusters were identified, these clusters did not coincide with classes of students.²

In Fairbanks, Alaska, the first author created an online asynchronous lesson on radioactivity [61,62] that is modeled after Johnson's Inquiry into Radioactivity curriculum [63]. His students in online algebra-based Physics 2 (primarily biology and wildlife majors) in 2023, 2024, and 2025, and his students in face-to-face calculus-based Physics 2 (primarily engineering majors) in 2024 and 2025, were invited to complete the survey for extra credit before and after the lesson. After data cleaning, a total of $N = 63$ from AlgPhys and $N = 111$ from CalcPhys remained.

²Since SWORDE was administered online, it was possible for students to take the survey on their own outside of regularly scheduled class time. To give maximum confidentiality to student responses, no demographic information, including teacher name or class, was recorded. However, by looking at the starting time of the survey responses, we were able to find students who may have been in the same class as each other. Specifically, before carrying out the data cleaning process described above, we arranged all respondents by starting time of the survey. Respondents who started within 5 min of each other during typical school hours were assumed to be in the same class.

In this paper, only pretest data from both student populations are presented, as this is most comparable to the Austrian data (post-high school instruction, precollege instruction). In particular, Austrian high school students “gain insights into changes in the atomic nucleus as the cause of radioactivity (properties of alpha, beta, and gamma rays); Recognize radioactive decay as a constantly occurring process; and understand the basic processes of energy conversion in the sun, stars, and nuclear reactions (nuclear fusion, nuclear fission)” [64]. These topics are generally found in high school physics textbooks in the United States as well [65]. For information about student learning with the online asynchronous lesson, readers are referred to previous publications [61,62].

1. Coding of responses

We assigned a label of “LC” to responses on the isomorphic prompts that indicated a level confusion and a label of “RAND” to responses that indicated awareness of the randomness of radioactive decay.

On CAGE, we focused exclusively on the reasoning tier for this prompt (see the Appendix of Ref. [9] for justification). We assigned a RAND code if the respondent selected “it is unpredictable when the atom transforms.”

TABLE I. Number of codes from the two isomorphic prompts from the Austria data ($N = 234$) and the UAF data ($N = 63$ from AlgPhys and $N = 111$ from CalcPhys). We assigned LC codes to responses that indicated a level confusion and RAND codes to responses that indicated awareness of the randomness of radioactive decay. Note that respondents could receive both a RAND code and an LC code on either prompt. For example, 38% of Austrian high school students received a RAND code for their response to MvO and 57% received an LC code to the prompt.

Prompt	Code	Count		
		Austria HS	UAF AlgPhys	UAF CalcPhys
CAGE	RAND	59 (25%)	26 (41%)	48 (43%)
	LC	154 (66%)	37 (59%)	43 (39%)
MvO	RAND	90 (38%)	37 (59%)	71 (64%)
	LC	133 (57%)	30 (48%)	29 (26%)

We assigned an LC code to responses of “the atom transforms continuously,” “after the half-life, half of the atom has transformed,” and “the atom transforms on the day of the half-life.”

On MvO, evidence for understanding of randomness can be found in MvO:ONE. We assigned a RAND code to the response of “1 atom OR no atoms” on each part of the answer tier. We also assigned a RAND code to the response of “It is random” on the reason tier. We assigned an LC code to the response of “half an atom” on the answer tier for 8 days. We also assigned an LC code for the response of “after the half-life, half of an atom will have transformed” on the reason tier. Finally, with a justification readers can find in the appendix of Ref. [9], we assigned an LC code if a respondent correctly chose “50 million atoms, 25 million atoms, 12.5 million atoms” as the answers to MvO:MANY but then chose either “1 atom, 0 atoms, 0 atoms” or “0 atoms, 0 atoms, 0 atoms” as the answers to MvO:ONE with the reason of “One cannot have half an atom.” Virtually all respondents who answered “50 million atoms, 25 million atoms” on the first two parts of MvO:MANY selected “12.5 million atoms” as their answer to the third part (likewise for MvO:ONE). As such, the prompt was improved when used at UAF by removing the third part of each question. The frequency of codes is shown in Table I. It is noteworthy and yet unsurprising that the UAF students performed better (were more likely to receive a RAND code and less likely to receive an LC code) than the Austrian high school students. Despite having learned about half-life in high school physics, the high school students represent a less biased population: some will go on to pursue a science, technology, engineering, and mathematics (STEM)-related career, and others will not. In what follows, we investigate this finding more carefully by separating the UAF students into the algebra-based physics learners (mostly biology and wildlife majors) and the calculus-based physics learners (mostly engineering

majors). It is also noteworthy that MvO is easier than CAGE for all students (more RAND codes and fewer LC codes for each student group). To gain insights beyond these, we invoke two tools, cluster analysis and alluvial diagrams, to see what patterns in respondent answers we can detect. For example, are the students receiving RAND codes different students than those receiving LC codes?

B. Cluster analysis

Based upon labels applied to responses, each respondent received a total of 4 binary codes (0 if “no” and 1 if “yes”):

1. Did the respondent receive at least one LC code in their responses to CAGE?
2. Did the respondent receive at least one RAND code in their responses to CAGE?
3. Did the respondent receive at least one LC code in their responses to MvO?
4. Did the respondent receive at least one RAND code in their responses to MvO?

The data were normalized using the mean and standard deviation for each of these four labels to account for the fact that it might be easier, say, for a respondent to receive a RAND code on MvO than on CAGE. To determine the number of clusters, the “elbow method” [66] was used by creating a scree plot in R Studio of within-group sum of squares (WSS) plotted vs number of clusters. WSS was seen to drop significantly in going from one cluster to two clusters, and again (less dramatically) from two to three clusters and from three to four clusters, but increase going from four to five clusters (see Fig. 3 below). It thus seems that the data are clustering into three or four clusters. As will be shown below, one of these clusters (cluster 1) corresponds to students who received codes of RAND on both MvO and CAGE but did not receive an LC code on either prompt. Cluster 2 is the opposite: students who were coded for LC answers on both survey prompts and who received no RAND codes. The remaining two clusters could be categorized as “other” or “mixed.” From our previous work [9], we already know that many students are in this “other” category, and our goal now is to examine more carefully what that looks like. Keeping two “other” categories (as well as utilizing the alluvial diagrams that we will discuss below) allows us to gain this insight.

We used kmeans in R-Studio to determine clustering of respondents. kmeans begins with randomly assigned cluster centers and runs an iterative process until the within-cluster sum of squares is minimized for that set of random initial values. Since results vary depending on the initially randomly chosen cluster centers, we ran kmeans 10 times and selected the run with the smallest total WSS (474). This corresponded to a ratio of between-cluster sum of squares to total sum of squares of 70.9%. The results are shown in Table II.

From these results, we see that high school learners in Austria are mostly found in cluster 2. UAF students in

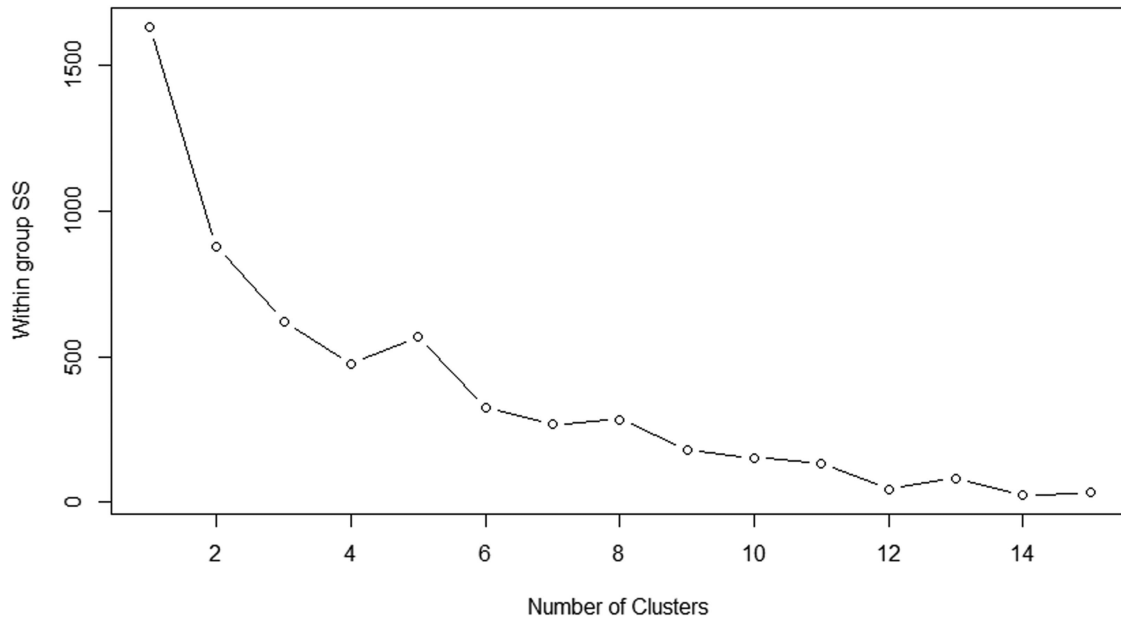


FIG. 3. Scree plot showing within-group sum of squares (WSS) vs number of clusters. WSS exhibits an increase going from four to five clusters, and so we maintain four clusters in our analysis.

face-to-face calculus-based introductory physics (CalcPhys) were predominantly grouped into cluster 1. UAF students in asynchronous algebra-based introductory physics (AlgPhys) were between these two extremes, populating both clusters 1 and 2. Although the sample size is small, it is apparent that CalcPhys students were not only more likely to be in cluster 1 than AlgPhys students; they were also more likely to be in cluster 4. Table III shows that cluster 4 is characterized by receiving neither a RAND label nor an LC label on CAGE.

C. Alluvial diagrams

Alluvial diagrams simplify and highlight structural changes between clusters so that transition rates can readily be seen. Compared with tables of transition rates, alluvial diagrams make the analysis more efficient by improving the readability and preventing oversight. Alluvial diagrams are a

type of Sankey diagram, which are used in a variety of research fields, including physics education research [67–69].

As with cluster analysis, we utilize alluvial diagrams to visually inspect patterns in student responses to the survey. We will compare findings with the alluvial diagrams to check for consistency with the cluster analysis and to see what new insights can be gained.

III. RESULTS

A. Cluster analysis

In the following figures, the top-right corner corresponds to the ideal combination of responses, and the bottom-left corner corresponds to the most novice responses. For example, in the top figure, the ideal response to CAGE receives a label of RAND and does not receive a label of LC.

TABLE II. Percentage of learners grouped into each cluster. Whereas the most prevalent cluster for Austrian high school students was cluster 2 (47% of the HS students), the most prevalent cluster for students in calculus-based physics was cluster 1. Algebra-based physics students were in-between these two extremes, with clusters 1 and 2 being roughly equally populated.

Cluster	Austria HS	UAF AlgPhys	UAF CalcPhys	Number
1	23%	40%	42%	125
2	47%	37%	14%	150
3	17%	16%	20%	72
4	13%	8%	23%	61
Number	234	63	111	408

TABLE III. Final cluster centers for the four clusters identified in Table II. Although many students were coded as both “RAND” and “LC” on a given prompt, the four dominant clusters were not characterized by these students. As seen in Table II, the most dominant cluster was cluster 2, which we see here as representing an LC code on both CAGE and MvO (and no RAND codes).

Cluster means				
Cluster	CAGE.RAND	CAGE.LC	MvO.RAND	MvO.LC
1	1.44	−0.96	0.66	−0.67
2	−0.60	0.86	−0.69	1.06
3	−0.66	0.86	0.22	−0.94
4	−0.69	−1.16	0.08	−0.12

TABLE IV. Summary of dominating alluvia in each population. *Italic (bold)* indicate (un)desired responses. The number in parentheses is the number of respondents in that alluvium.

Alluvium	Dominant in group	CAGE LC	CAGE RAND	MvO LC	MvO RAND
P	Austrian HS (93, 40%) UAF AlgPhys (16, 25%) UAF CalcPhys (13, 12%)	Yes	No	Yes	No
N	Austrian HS (19, 8%)	Yes	No	<i>No</i>	No
O	UAF AlgPhys (6, 10%)	Yes	No	Yes	<i>Yes</i>
M	Austrian HS (21, 9%) UAF AlgPhys (7, 11%) UAF CalcPhys (15, 13%)	Yes	No	<i>No</i>	<i>Yes</i>
E	UAF CalcPhys (9, 8%)	<i>No</i>	No	<i>No</i>	<i>Yes</i>
A	Austrian HS (37, 16%) UAF AlgPhys (15, 24%) UAF CalcPhys (34, 31%)	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>

At first glance, we notice that the red and green dots dominate in the bottom-left and top-right corners, respectively. Indeed, there are more respondents in the top-right and bottom-left corners than in the other two corners.

Recall that high school learners in Austria are mostly found in cluster 2 (red). UAF students in asynchronous algebra-based introductory physics are represented in both clusters 2 (red) and 1 (green). UAF students in face-to-face calculus-based introductory physics (CalcPhys) are primarily in cluster 1 (green). We will interpret these findings in the Discussion section of this paper.

B. Alluvial diagrams

In the diagrams that follow, the data are arranged such that optimal responses (students in alluvia A) are at the top. The first two columns are for CAGE, and the next two columns are for MvO. In each pair, the left column is for LC codes. Recall that there are a total of 408 respondents in the data pool. Notice a little more than half of the students indicate on at least one of the two prompts awareness of the random nature of radioactive decay. On the other hand, respondents in alluvia G (13 respondents, 3%), H (16 respondents, 4%), N (28 respondents, 7%), and P (122 respondents, 30%) consistently did not receive RAND codes.

Alluvia A contains students who received a RAND code on both CAGE and MvO and did not receive LC codes. Alluvia P contains students who consistently received an LC code on both CAGE and MvO and did not receive RAND codes. As expected and consistent with the cluster analysis above, we see a growth in the A alluvia going from high school to CalcPhys and a decrease in the P alluvia. To better describe the students who would normally be grouped together into a generic “other” category, we remove these two “consistent student” alluvia, as well as alluvia of three or fewer respondents, to produce Figs. 6–9

below (left-hand side). Alluvial diagrams including all alluvia, are in the Supplementary Material [70]. Other than alluvia A and P, comparing population percentages in alluvia E, M, N, and O provides insights. The frequencies and percentages of each population in these six alluvia are in Table IV, and a diagram of alluvia E, M, N, and O for each population appears in Figs. 6–9 below (right-hand side).

As shown in Fig. 6, notable alluvia for the entire sample of responses are M (43, 11%) and N (28, 7%).

As shown in Fig. 7, notable alluvia for the Austrian high school students are M and N with 21 and 19 respondents (9% and 8%), respectively.

As shown in Fig. 8, notable alluvia for the asynchronous algebra-based physics students are M(7) and O(6).

As seen in Fig. 9, notable alluvia for the calculus-based physics students were M (15) and E (9).

IV. DISCUSSION

In this paper, we have presented findings from two prompts from the Stochastic World of Radioactive Decay Evaluation (SWORDE). These two prompts are isomorphic, in that they can be answered appropriately with the same underlying physics principle that radioactive decay is considered to take place at random. Both prompts also allow respondents to answer in a way indicating the student idea that what is true for the radioactive sample is also true for the individual nucleus (that is, a level confusion).

A. Findings consistent with prior studies

The results we have discussed in this paper are consistent with what we and other researchers have documented in prior publications. In particular, many students think that after one half-life, since half of the

radioactive sample has decayed, it follows that half of each atom in that sample has decayed [8,16,19]. A total of $N = 408$ responses survived our data cleaning process (students who said they had not yet learned about half-life, for example, were removed). On one of the two isomorphic prompts (MvO:ONE), 161 (39%) said that, after a half-life had passed, half an atom would not have yet transformed. Across both isomorphic prompts, 275 (67%) showed evidence of this level confusion on at least one prompt. Clearly, the idea of “what is true for the radioactive sample is true for an individual unstable nucleus” is salient for many students. However, we found that less than half of these respondents (151 respondents, 37%) exhibited a level confusion consistently across both prompts. In terms of properties associated with misconceptions [26,27], context-dependency across the isomorphic prompts indicates that naïve ideas of learners have not yet crystallized into stable structures.

From both the cluster analysis and alluvial diagram construction, we find a result consistent with our prior findings: although many learners are consistent in their answer patterns, suggesting either a misconception (37% in cluster 2/30% in alluvium P) or a normative conception (31% in cluster 1/21% in alluvium A) of radioactive decay, the remaining students (49% in the other alluvia) are *not*, indicating reasoning about radioactivity that is fluid and context-dependent.

B. New findings obtained with cluster analysis and alluvial diagrams

We find additional commonalities between the results of the cluster analysis and alluvial diagram construction. For each of the three populations, we find that CAGE is a more difficult prompt than MvO. Many students answered MvO without a level confusion and with a demonstration of understanding randomness, but did not do so for CAGE (see, for example, the right side of Fig. 6, which is dominated by alluvium M). Looking across the three populations, we first note that high school learners were the most likely to answer both CAGE and MvO with a level confusion and not with a response indicating awareness of the stochastic nature of radioactive decay (47% were in cluster 2, and their alluvium P was the widest of the three populations). The calculus-based physics university students, on the other hand, were just the opposite (42% were in cluster 1, and their alluvium A was the widest of the three populations).

The use of both cluster analysis and alluvial diagrams allows not only to check for consistency between analysis methods but also to gain insights from one that are not apparent in the other. From the cluster analysis, we see again that Figs. 4 and 5 are dominated by two large dots, a red one in the lower-left corner (cluster 2, showing unfavorable responses) and a green one in the top-right corner (cluster 1, showing favorable responses). Let us

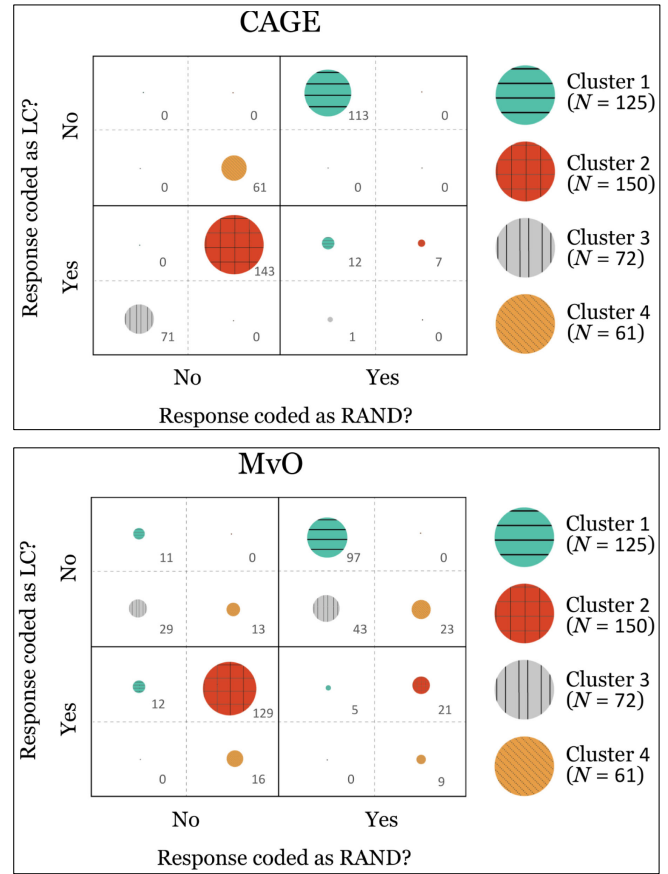


FIG. 4. Labels given to respondents in each cluster on CAGE (top) and MvO (bottom).

focus, however, on the other two clusters. Approximately 20% of each population was in cluster 3, and (as seen in Table III) this is associated with recognizing the stochastic nature of radioactive decay on MvO but not CAGE and, conversely, exhibiting a level confusion on CAGE but not on MvO (in other words, CAGE is harder for all three populations). Cluster 4 is particularly interesting because this cluster is associated with letting go of the level confusion on CAGE while not yet accepting the stochastic nature of radioactive decay (Fig. 4, top). Looking at the RAND and LC figures (Fig. 5), we see a consistent picture: the yellow dots are in quadrants corresponding to not recognizing radioactive decay as a stochastic process on CAGE, but also not having a level confusion. We note that a larger portion of CalcPhys learners (23%) are in this cluster compared to AlgPhys learners (8%) or high school students (13%).

The alluvial diagrams allow us to more precisely describe this observation. Combining all students in the alluvial diagrams (Fig. 6), we see that alluvium M is significant, consistent with CAGE generally being more difficult than MvO. This is a dominant alluvium in all three populations (Figs. 7–9). In addition, in the high school data, we find alluvium N to be dominant, and for AlgPhys,

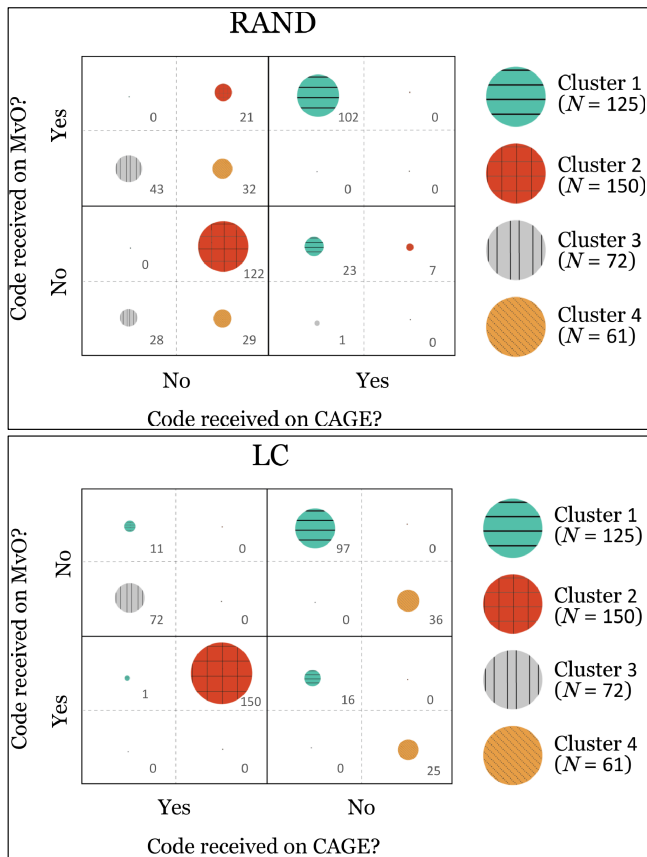


FIG. 5. Labels given to respondents in each cluster on CAGE (top) and MvO (bottom).

we find alluvium O to be dominant. These alluvia correspond to continued struggle with CAGE, but progress is being made with MvO. Finally, looking at the CalcPhys students, we see alluvium E being significant (Fig. 9), where these students have already mastered MvO and have started making progress in understanding CAGE, rejecting the level confusion-based reasoning, but still not yet accepting the stochastic nature of the situation. These results are summarized in Table IV.

It is interesting that, whereas the high school and AlgPhys students made progress with MvO both by accepting the idea of randomness (alluvium O) and by resisting the level confusion (alluvium N), it was largely one-sided for the CalcPhys students in progressing with CAGE. Not many students struggled with a level confusion on CAGE while accepting the idea of randomness. The opposite, however, as seen in alluvium E, was true. It may be that these predominantly engineering majors, who comparatively value and enjoy physics, are more put off by the idea of randomness in science, insisting that “God doesn’t play dice.” Future work involving student interviews of CalcPhys students, as well as the collection of additional survey data, could help determine if there is indeed a difference and, if so, the cause of this difference.

Our findings allow us to address our research questions: RQ1: To what extent do diverse populations of students learning introductory physics compare in their understanding of half-life? Broadly speaking, we find that many students, even those in calculus-based physics, are

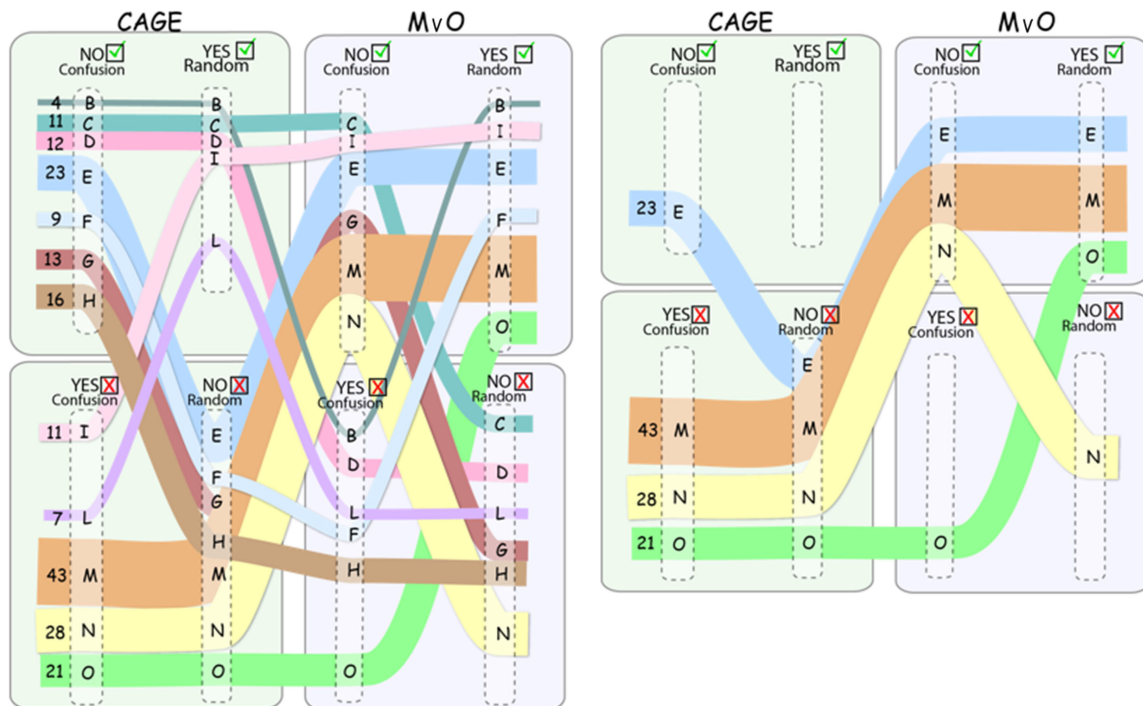


FIG. 6. Alluvial diagram consisting of all alluvia with four or more respondents (except alluvia A and P) from the full set of 408 respondents (left) and of alluvia E, M, N, and O (right).

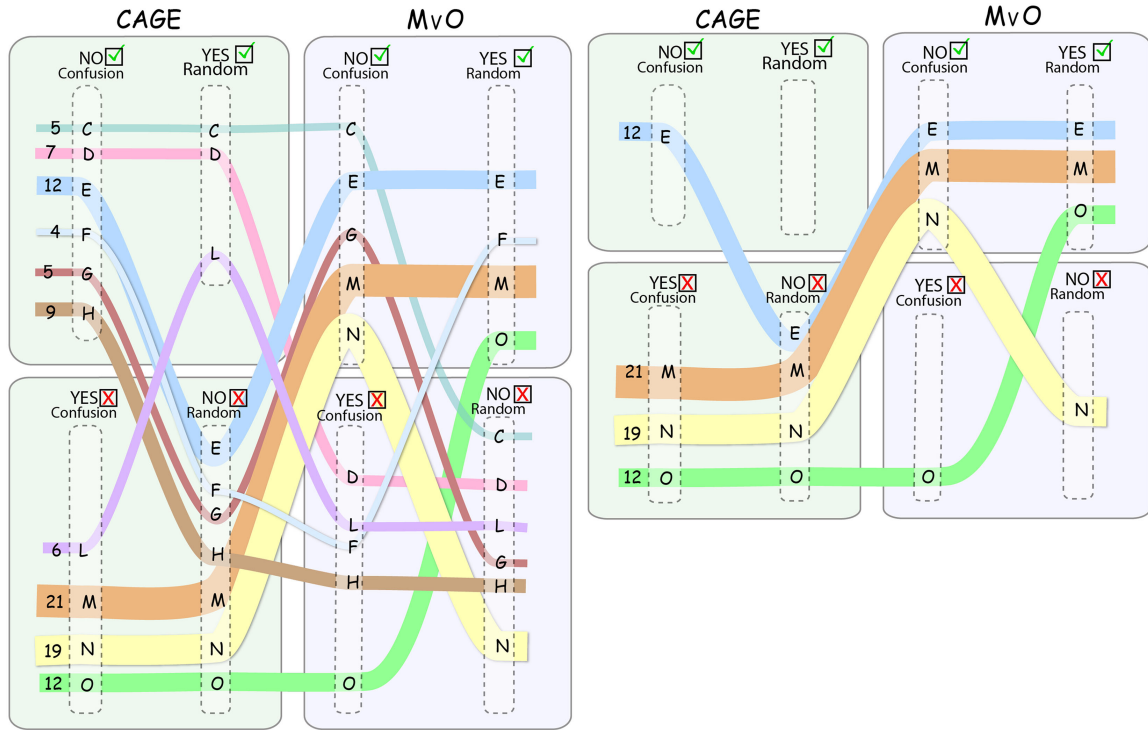


FIG. 7. Alluvial diagram consisting of all alluvia with four or more respondents (except alluvia A and P) from the Austrian high school students (left) and of alluvia E, M, N, and O (right).

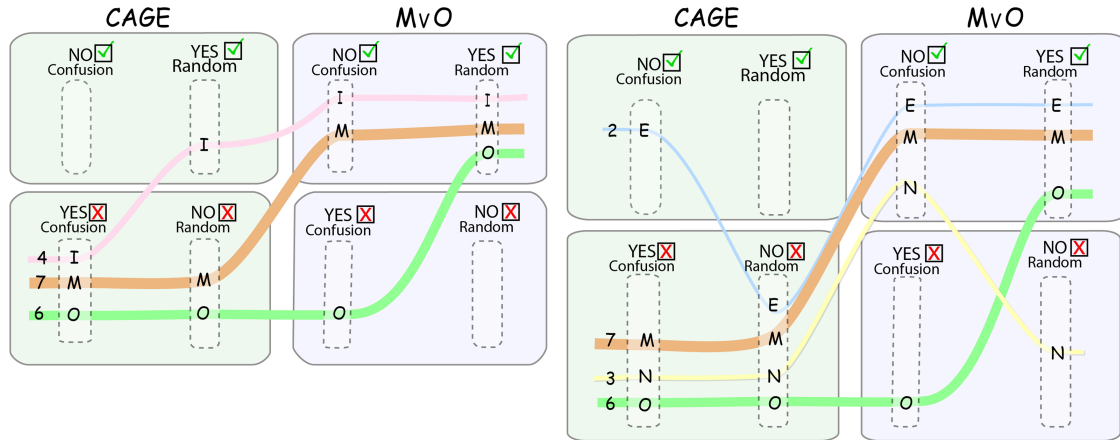


FIG. 8. Alluvial diagram consisting of all alluvia with four or more respondents (except alluvia A and P) from the algebra-based asynchronous physics students at UAF (left) and of alluvia E, M, N, and O (right).

inconsistent in their reasoning about radioactivity. Furthermore, it is perhaps the case that students find it easier to think in terms of levels (in this case, that the properties of an individual atom may be different from those of a sample of radioactive material) than to accept the stochastic nature of radioactive decay, and that this is particularly true for calculus-based physics students. RQ2: How can we characterize student understanding of half-life? In our prior work [9], we focused on the high school population and documented considerable context dependency in student answers. Here we have extended our study to two additional samples of students: algebra-based physics

students in an online asynchronous course and calculus-based physics students who are predominantly engineering majors. We found, again, that, for many students (200 out of 408, or 49%), ideas about half-life are not robust and intact cognitive structures; rather, they are better described as being temporary constellations of smaller knowledge pieces that fluidly shift in response to context.

IV. LIMITATIONS AND FUTURE WORK

There are several limitations that should be kept in mind when interpreting the findings presented in this paper.

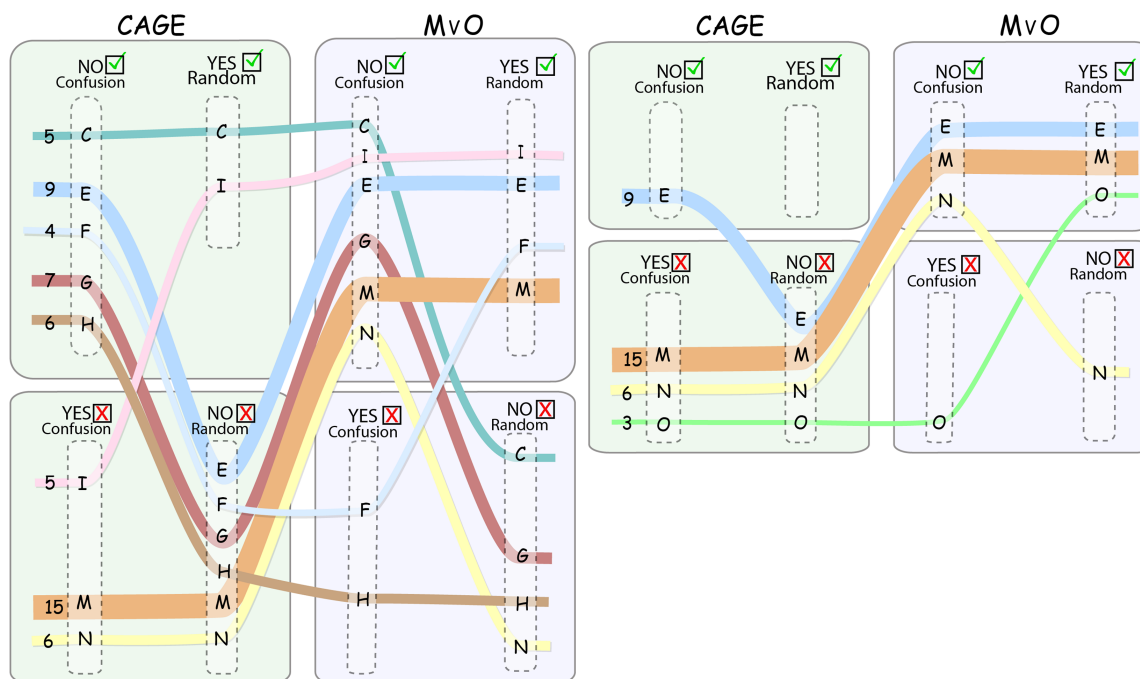


FIG. 9. Alluvial diagram consisting of all alluvia with four or more respondents (except alluvia A and P) from the calculus-based physics students at UAF (left) and of alluvia E, M, N, and O (right).

First, it should be noted that the data collected from Austrian high school students were collected in 2020/2021. Although this is only a few years prior to the data collected at UAF, with a topic like radioactivity, that could make a difference. Societal attitudes toward radioactivity may have shifted during those years. For instance, during the COVID-19 pandemic, widespread discussions about randomness and probability, particularly in the context of vaccination strategies, may have influenced students' perceptions of stochastic processes. Additionally, the ongoing global discourse surrounding alternative energy sources and the resurgence of interest in nuclear energy, as major companies advocate for reinstating power plants, may also have shaped attitudes toward radioactivity. As such, if Austrian high school students were surveyed again today, we may see different patterns in responses. Furthermore, the UAF dataset is relatively small, and it constitutes several years of data; as such, it is possible that there may have been differences in student cohorts from year to year that have been averaged over in our methodology. This limits the generalizability of the findings and suggests the need for larger sample sizes in future studies. Although accumulation of additional (and, in the case of the Austrian students, more recent) data may affect the specific answer to our first research question (that it is easier to think in terms of levels than to grasp stochastic processes), we expect that it would not change our answer to the second research question (that even diverse student populations—different educational systems, cultural attitudes toward science, and prior exposure to concepts such

as randomness and radioactivity—share the feature of having a context-dependent understanding of radioactivity). In the early stages of this manuscript, we analyzed our data for the first time in 2024 (prior to the most recent year of data collected) and found the same answers to both research questions.

A. Future work

In this paper, we have expanded upon our prior findings that many learners do not have misconceptions about radioactivity; rather, their student ideas change fluidly in response to the context of the question being asked. A longitudinal study as learners progress through their study of radioactivity, for example, using the microgenetic approach [71,72], would allow us to see the stability (or lack thereof) of student ideas over shorter time intervals in which instruction does and does not take place. During the validation process of SWORDE, for example, the eighth survey validation interviewee completed the survey in class and then again, ten days later, for the survey validation interview. The respondent's answers differed between the two administrations, especially on MvO:ONE. Specifically, in class, the respondent correctly answered MvO:ONE that it is random when the individual atom would decay, and so, after any number of half-lives, there would be "one OR zero" atoms remaining. During the interview, however, the student selected the option that there would be no atoms remaining after one half-life, with the justification that "one cannot have half an atom." Since the interviewee expressed

a lack of confidence in this answer, the interviewer asked if there were any other options the interviewee could think of, and the participant said that the competing idea is that there would be “half an atom” remaining after one half-life. When the interviewer shared the student’s former response to the prompt, the interviewee reacted with surprise. When asked if he could remember why he chose that option originally, the interviewee responded “Yes, that is a good question. Um, to be frank, I might have had a different way of thinking about it... if I think about it again now though, I would actually not choose that... I’m a little surprised myself, frankly speaking, that I chose that” and the interviewee maintained that, after one half-life, “everything has transformed, so there are zero that have not yet transformed” (translated from German by the authors).

In our earlier work, we showed that more confident students were not more likely to be consistent in their answering patterns [9]. It would be interesting to investigate other characteristics, such as scientific beliefs, that may correlate with the stability of student reasoning. For example, students who voice acceptance in general about the idea of randomness in natural processes may be more likely to receive a RAND code on both of the two survey prompts. Similarly, it would be interesting to compare responses from university students who are not STEM majors with those who are biology and wildlife majors (in AlgPhys) with those who are engineering majors (in CalcPhys).

An additional interesting question to pursue in future research is whether students are aware when they are being self-contradictory. Students who answer differently on the isomorphic prompts can be asked how the prompts are similar or different from each other, and whether the student thinks that his or her responses are self-consistent. Such an approach was used by Singh [58], whose “students were asked an open-ended question about whether the two problems are similar or different and they had to explain their reasoning ... 28% explicitly said that the problems are different.”

B. Developing a learning progression

Although it is interesting that the reasoning patterns seen in the three populations differed, it is not entirely surprising. Although all students have received comparable prior education on radioactivity, learners in CalcPhys are pursuing careers most closely interwoven with physics. Although AlgPhys learners are not physics and engineering majors, they are nevertheless pursuing STEM-related careers, primarily in biology and wildlife. The high school students can be conceptualized as a mix of these students, as well as many students who do not pursue STEM-related careers. In consideration that the CalcPhys students exhibited somewhat greater pretest FCI scores (46% compared to 28% in AlgPhys), it is plausible that we are seeing indicators of a

learning progression [73,74] in our data from heterogeneous high school students to AlgPhys students (who perhaps did not learn physics as thoroughly in high school as the CalcPhys students) to the CalcPhys students. From Table II, we would visualize such a progression as going from cluster 2 to cluster 1 (high school students were more likely to be in cluster 2, CalcPhys students were more likely to be in cluster 1, and AlgPhys students were equally likely to be in both clusters).

In examining Fig. 4 (top) for CAGE, the more difficult of the two prompts, we see a path from the big red dot (cluster 2, undesirable responses) to the big green dot (cluster 1, desirable responses) through the smaller yellow dot (cluster 4). This suggests that a learning progression in mastering the content is to first be rid of the level confusion (moving from the bottom-left quadrant to the top-left quadrant) and then to understand the stochastic nature of radioactive decay (moving to the top-right quadrant). Recalling also that a relatively large percentage (23%) of CalcPhys students were in cluster 4 (Table II), in comparison to the 13% of high school learners, we find additional support that the learning progression from cluster 2 to cluster 1 is through cluster 4.

We can also look to the alluvial diagrams for indicators of a learning progression. Combining all students in the alluvial diagrams (Fig. 6), we see that alluvia M is significant, consistent with CAGE generally being more difficult than MvO. This is a dominant alluvium in all three populations (Figs. 7–9). In addition, for the high school learners, we find alluvium N to be dominant, and for AlgPhys, we find alluvium O to be dominant. These alluvia correspond to continued struggle with CAGE, but progress is being made with MvO. It would make sense for students to transition from alluvia N and O to alluvium M. Finally, looking at the CalcPhys students, we see alluvium E being significant (Fig. 9), where these students have already mastered MvO and have started making progress in understanding CAGE, rejecting the level confusion-based reasoning, but still not yet accepting the stochastic nature of the situation. It is plausible that students progress from alluvium M to alluvium E before arriving at alluvium A. In summary, this data, though tentative, is consistent with a learning progression in which students start in alluvium P (consistently thinking with a level confusion and not with stochastic reasoning), progress to alluvia N and O (where students can begin to make sense of situations like MvO), further progress to alluvium M (where they have mastered MvO but continue to struggle with CAGE), further progress to alluvium E (where they make progress on CAGE by rejecting level confusion-based reasoning), and finally reach alluvium A (where they adopt a stochastic reasoning process on CAGE).

Both the cluster analysis and the alluvial diagrams suggest that teaching about emergence and levels, for example, through ontology training [75,76], may be a

fruitful place to start before teaching about stochastic processes. The potential longitudinal study described above would help here as well, to strengthen or refute this tentative claim.

V. CONCLUSIONS

In this paper, we have compared responses of three student populations to two survey prompts (MvO and CAGE) asking about their understanding of half-life. Consistent with our prior findings, we see that, although many students indicate at least some acceptance of the idea that it is random when an individual nucleus decays, relatively few of these students respond this way on both prompts. Similarly, although most students indicate a level confusion (that if the radioactive sample decays in a predictable, continuous manner, then the individual nuclei comprising that sample must do likewise) on one of the two prompts, relatively few do so consistently across both prompts. This suggests that students do not generally have robust and intact misconceptions about half-life, but rather pieces of ideas that ebb and flow depending upon the context and that can be drawn upon for productive sensemaking.

We have built upon our previous work, which looked only at the Austrian high school students by (i) adding data from two student populations at the University of Alaska Fairbanks and (ii) using two additional tools to look for similarities in answering patterns between the three populations. Although Table I is sufficient to conclude that UAF CalcPhys students were the most likely to recognize the stochastic nature of radioactive decay on either prompt and were the least likely to have a level confusion, it does not tell us, for example, if students in CalcPhys who did not receive a RAND code on CAGE instead received an LC code on the prompt. From the cluster analysis, on the other hand, we see that many students in CalcPhys did not receive the LC code and also did not receive the RAND code on the prompt. Specifically, Table II tells us that

CalcPhys students fell primarily into cluster 1 (42%), but that a sizable portion also fell into cluster 4 (23%). Table III tells us that cluster 4 is associated with not receiving an LC code on CAGE but also not receiving a RAND code. We see directly in the alluvial diagrams that alluvium E was a noteworthy alluvium for CalcPhys, associated with mastery on MvO, but not accepting the stochastic nature of radioactive decay on CAGE (Fig. 9). The use of tools such as these provides fine-grained details that allow for nuanced comparisons in answering patterns across populations and throughout a learning progression. Although our sample size is small, such analyses of the preliminary data we have presented in this paper, for example, suggest that students might generally find it easier to accept the idea of the agent level having different properties than the collective level (that is, let go of their level confusion) than to accept the idea of randomness in science.

In terms of instructional implications, our findings combined suggest that something similar to ontology training [75,76] to help students learn about emergence may be important as an early step for their developing an understanding of half-life, provided that instruction guides students to “build on the best of [their ideas]” (p. 41 of Ref. [25]) as opposed to taking a cognitive conflict strategy [28,29].

ACKNOWLEDGMENTS

The authors wish to thank Shizuka Nakayama for the illustrations that accompany our papers and presentations.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available because of legal restrictions preventing unrestricted public distribution. The data are available from the authors upon reasonable request.

-
- [1] United Nations, Resolution adopted by the General Assembly on 25 September 2015 70/1. Transforming our world: the 2030 agenda for sustainable development (2015), https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf.
 - [2] IAEA, IAEA annual projections rise again as countries turn to nuclear for energy security, and climate action, <https://www.iaea.org/newscenter/pressreleases/iaea-annual-projections-rise-again-as-countries-turn-to-nuclear-for-energy-security-and-climate-action> (2023).
 - [3] International Atomic Energy Agency (2017). How Nuclear Power Helps Meet Global Energy Demand The Role of the IAEA, <https://www.iaea.org/sites/default/files/17/02/how-nuclear-power-helps-meet-global-energy-demand-the-role-of-the-iaea.pdf>.
 - [4] R. Rhodes, A sensible climate change solution, borrowed from Sweden, NY Times (2019), <https://www.nytimes.com/2019/02/05/books/review/bright-future-joshua-s-goldstein-staffan-a-qvist.html>.
 - [5] WNA, Supply of uranium, <https://world-nuclear.org/information-library/nuclear-fuel-cycle/uranium-resources/supply-of-uranium#:~:text=the%20world's%20present%20measured%20resources,is%20normal%20for%20most%20minerals> (2024).
 - [6] S. El-Showk, Final resting place, *Science* **375**, 806 (2022).

- [7] H. M. C. Eijkelhof, C. W. J. M. Klaassen, P. L. Lijnse, and R. L. J. Scholte, Perceived incidence and importance of lay-ideas on ionizing radiation: Results of a Delphi-study among radiation-experts, *Sci. Educ.* **74**, 183 (1990).
- [8] M. M. Hull and M. Hopf, Student understanding of emergent aspects of radioactivity, *Int. J. Phys. Chem. Educ.* **12**, 19 (2020), <https://ijpce.org/index.php/IJPCE/article/view/125>.
- [9] M. M. Hull, A. Jansky, and M. Hopf, Does confidence in a wrong answer imply a misconception?, *Phys. Rev. Phys. Educ. Res.* **18**, 020108 (2022).
- [10] M. M. Hull, A. Jansky, and M. Hopf, Radioactivity as “quintessentially eternal”: Two survey prompts, in *Proceedings of the GDCP (Gesellschaft für Didaktik der Chemie und Physik) 2021 Conference 184* (2022), https://www.gdcp-ev.de/wp-content/tb2022/TB2022_184_Hull.pdf.
- [11] M. M. Hull, A. Jansky, and M. Hopf, Two approaches to analyzing fluidity of student reasoning on a multiple choice survey about half-life, in *2021 ESERA Proceedings* (2021), p. 62, https://www.esera.org/wp-content/uploads/2023/02/CNF21-Proceedings_Strand-1_p.4-89.pdf.
- [12] M. M. Hull, A. Jansky, and M. Hopf, Reasoning fluidly about half-life on a two-tier multiple-choice survey, in *GDCP Conference Proceedings (Online)* (2020), https://gdcp-ev.de/wp-content/uploads/2021/07/GDCP_Band41_050721.pdf.
- [13] H. M. C. Eijkelhof, *Radiation and Risk in Physics Education* (CDß Press, Utrecht, 1990).
- [14] P. L. Lijnse, H. M. C. Eijkelhof, C. W. J. M. Klaassen, and R. L. J. Scholte, Pupils’ and mass-media ideas about radioactivity, *Int. J. Sci. Educ.* **12**, 67 (1990).
- [15] H. M. C. Eijkelhof and R. Millar, Reading about Chernobyl: The public understanding of radiation and radioactivity, *Sch. Sci. Rev.* **70**, 35 (1988).
- [16] E. E. Prather, Students’ beliefs about the role of atoms in radioactive decay and half-life, *J. Geosci. Educ.* **53**, 345 (2005).
- [17] E. K. Henriksen, Laypeople’s understanding of radioactivity and radiation, *Radiat. Protect. Dosim.* **68**, 191 (1996).
- [18] M. M. Hull, A. Jansky, and M. Hopf, Probability-related naïve ideas across physics topics, *Stud. Sci. Educ.* **57**, 45 (2021).
- [19] C. W. J. M. Klaassen, H. M. C. Eijkelhof, and P. L. Lijnse, Considering an alternative approach to teaching radioactivity, in *Relating Macroscopic Phenomena to Microscopic Particles: A Central Problem in Secondary Science Education* (1990), pp. 304–316, https://www.academia.edu/download/50665841/Considering_an_alternative_approach_to_t20161201-19533-jvt2py.pdf.
- [20] A. Jansky, *Die Rolle von Schülervorstellungen Zu Wahrscheinlichkeit Und Zufall Im Naturwissenschaftlichen Kontext* (University of Vienna, 2019), https://www.researchgate.net/publication/340686195_Die_Rolle_von_Schulervorstellungen_zu_Wahrscheinlichkeit_und_Zufall_im_naturwissenschaftlichen_Kontext.
- [21] U. Wilensky and M. Resnick, Thinking in levels: A dynamic systems approach to making sense of the world, *J. Sci. Educ. Technol.* **8**, 3 (1999).
- [22] M. M. Hull, Emergent aspects of radioactivity: Creation of a survey on half-life, in *GDCP Conference Proceedings* (2019), pp. 590–593, https://www.gdcp-ev.de/wp-content/tb2020/TB2020_590_Hull.pdf.
- [23] S. Vosniadou and I. Skopeliti, Conceptual change from the framework theory side of the fence, *Sci. Educ.* **23**, 1427 (2014).
- [24] D. Hammer, Misconceptions or p-prims: How may alternative perspectives of cognitive structure influence instructional perceptions and intentions, *J. Learn. Sci.* **5**, 97 (1996).
- [25] A. A. diSessa, A bird’s-eye view of the “pieces” vs. “coherence” controversy (from the “pieces” side of the fence), in *International Handbook of Research on Conceptual Change*, edited by S. Vosniadou (Routledge, 2009), p. 35, 10.4324/9780203154472.
- [26] R. E. Scherr, Modeling student thinking: An example from special relativity, *Am. J. Phys.* **75**, 272 (2007).
- [27] E. F. Redish, *Teaching Physics with the Physics Suite* (New York, NY, 2003), <https://www.per-central.org/items/detail.cfm?ID=3707>.
- [28] G. J. Posner, K. A. Strike, P. W. Hewson, and W. A. Gertzog, Accommodation of a scientific conception: Toward a theory of conceptual change, *Sci. Educ.* **66**, 211 (1982).
- [29] L. C. McDermott, Millikan lecture 1990: What we teach and what is learned—Closing the gap, *Am. J. Phys.* **59**, 301 (1991).
- [30] A. A. diSessa, Toward an epistemology of physics, *Cognit. Instr.* **10**, 105 (1993).
- [31] D. Hammer, More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research, *Am. J. Phys.* **64**, 1316 (1996).
- [32] D. Hammer, Student resources for learning introductory physics, *Am. J. Phys.* **68**, S52 (2000).
- [33] D. Hammer, A. Elby, R. E. Scherr, and E. F. Redish, Resources, framing, and transfer, in *Transfer of Learning from a Modern Multidisciplinary Perspective*, edited by J. P. Mestre (IAP, 2006), pp. 89–121, https://dhammer.phy.tufts.edu/home/publications_files/rft.pdf.
- [34] S. Kapon and A. A. diSessa, Instructional explanations as an interface—The role of explanatory primitives, *AIP Conf. Proc.* **1289**, 189 (2010).
- [35] J. Minstrell, Facets of students’ knowledge and relevant instruction, in *Research in Physics Learning: Theoretical Issues and Empirical Studies, Proceedings of an International Workshop* (1992), pp. 110–128, https://www.researchgate.net/profile/Hans-Niedderer/publication/330993365_Research_in_Physics_Learning-Theoretical_Issues_and_Empirical_Studies/links/5c5f8abea6fdccb608b40ca2/Research-in-Physics-Learning-Theoretical-Issues-and-Empirical-Studies.pdf#page=110.
- [36] J. P. I. Smith, A. A. diSessa, and J. Roschelle, Misconceptions reconceived: A constructivist analysis of knowledge in transition, *J. Learn. Sci.* **3**, 115 (1994).
- [37] M. Lemmer, Nature, cause and effect of students’ intuitive conceptions regarding changes in velocity, *Int. J. Sci. Educ.* **35**, 239 (2013).
- [38] N. K. Weinlader, E. Kuo, B. M. Rottman, and T. J. Nokes-Malach, A new approach for understanding student

- resources with multiple-choice questions, *presented at PER Conf. 2019, Provo, UT*, [10.1119/perc.2019.pr.Weinlader](#).
- [39] H. Eshach, T. C. Lin, and C. C. Tsai, Taiwanese middle school students' materialistic concepts of sound, *Phys. Rev. Phys. Educ. Res.* **12**, 010119 (2016).
- [40] S. Hasan, D. Bagayoko, and E. L. Kelley, Misconceptions and the certainty of response index (CRI), *Phys. Educ.* **34**, 294 (1999).
- [41] P. Klein, S. Küchemann, S. Brückner, O. Zlatkin-Troitschanskaia, and J. Kuhn, Student understanding of graph slope and area under a curve: A replication study comparing first-year physics and economics students, *Phys. Rev. Phys. Educ. Res.* **15**, 020116 (2019).
- [42] M. Planinic, W. J. Boone, R. Krsnik, and M. L. Beilfuss, Exploring alternative conceptions from Newtonian dynamics and simple DC circuits: Links between item difficulty and item confidence, *J. Res. Sci. Teach.* **43**, 150 (2006).
- [43] M. Potgieter, E. Malatje, E. Gaigher, and E. Venter, Confidence versus performance as an indicator of the presence of alternative conceptions and inadequate problem-solving skills in mechanics, *Int. J. Sci. Educ.* **32**, 1407 (2010).
- [44] I. Testa, A. Colantonio, S. Galano, I. Marzoli, F. Trani, and U. Scotti Di Uccio, Effects of instruction on students' overconfidence in introductory quantum mechanics, *Phys. Rev. Phys. Educ. Res.* **16**, 010143 (2020).
- [45] B. R. Wilcox and S. J. Pollock, Investigating students' behavior and performance in online conceptual assessment, *Phys. Rev. Phys. Educ. Res.* **15**, 020145 (2019).
- [46] D. MacIsaac, R. P. Cole, L. McCullough, and J. Maxka, Standardized testing in physics via the World Wide Web, *Electron. J. Sci. Educ.* **6** (2002), <https://ejrsmc.icrsme.com/article/download/7681/5448>.
- [47] J. Han, L. Bao, L. Chen, T. Cai, Y. Pi, S. Zhou, Y. Tu, and K. Koenig, Dividing the Force Concept Inventory into two equivalent half-length tests, *Phys. Rev. ST Phys. Educ. Res.* **11**, 010112 (2015).
- [48] J. Han, K. Koenig, L. Cui, J. Fritchman, D. Li, W. Sun, Z. Fu, and L. Bao, Experimental validation of the half-length Force Concept Inventory, *Phys. Rev. Phys. Educ. Res.* **12**, 020122 (2016).
- [49] J.-i. Yasuda, N. Mae, M. M. Hull, and M.-A. Taniguchi, Optimizing the length of computerized adaptive testing for the Force Concept Inventory, *Phys. Rev. Phys. Educ. Res.* **17**, 010115 (2021).
- [50] C. Singh, Assessing student expertise in introductory physics with isomorphic problems. II. Effect of some potential factors on problem solving and transfer, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010105 (2008).
- [51] J. C. Speirs, M. R. Stetzer, B. A. Lindsey, and M. Kryjevskaja, Exploring and supporting student reasoning in physics by leveraging dual-process theories of reasoning and decision making, *Phys. Rev. Phys. Educ. Res.* **17**, 020137 (2021).
- [52] J. C. Speirs, R. Leuteritz, T. K. Lê, R. Deng, and S. W. Ell, Investigating the efficacy of attending to reflexive cognitive processes in the context of Newton's second law, *Phys. Rev. Phys. Educ. Res.* **19**, 010108 (2023).
- [53] T. Fittswood, D. J. Rosen, and M. R. Stetzer, Insights from an intervention designed to support consistent reasoning, *presented at PER Conf. 2022, Grand Rapids, MI*, [10.1119/perc.2022.pr.Fittswood](#).
- [54] M. Kryjevskaja, M. R. Stetzer, B. A. Lindsey, A. McInerney, P. R. Heron, and A. Boudreaux, Designing research-based instructional materials that leverage dual-process theories of reasoning: Insights from testing one specific, theory-driven intervention, *Phys. Rev. Phys. Educ. Res.* **16**, 020140 (2020).
- [55] H. A. Simon and J. R. Hayes, The understanding process: Problem isomorphs, *Cognit. Psychol.* **8**, 165 (1976).
- [56] J. R. Hayes and H. A. Simon, Psychological differences among problem isomorphs, in *Cognitive Theory*, edited by N. J. Castellan, D. B. Pisoni, and G. R. Potts (Lawrence Erlbaum, Hillsdale, NJ, 1977).
- [57] K. Kotovsky, J. R. Hayes, and H. A. Simon, Why are some problems hard? Evidence from the Tower of Hanoi, *Cognit. Psychol.* **17**, 248 (1985).
- [58] C. Singh, Assessing student expertise in introductory physics with isomorphic problems. I. Performance on nonintuitive problem pair from introductory physics, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010104 (2008).
- [59] D. E. Brown and J. Clement, Overcoming misconceptions via analogical reasoning: Abstract transfer versus explanatory model construction, *Instr. Sci.* **18**, 237 (1989).
- [60] R. L. Thorndike, Who belongs in the family?, *Psychometrika* **18**, 267 (1953).
- [61] M. M. Hull and K. R. Pokharel, Learning about radioactivity via synthetic groupwork in introductory physics, *presented at PER Conf. 2024, Boston, MA*, [10.1119/perc.2024.pr.Hull](#).
- [62] M. M. Hull, M. Jeidler, and E. S. Holzinger, Simulated groupwork in an asynchronous course learning about radioactivity, *presented at PER Conf. 2023, Sacramento, CA*, [10.1119/perc.2023.pr.Hull](#).
- [63] A. Johnson, IBI series winner, radiation and atomic literacy for nonscientists, *Science* **342**, 436 (2013).
- [64] M. Jeidler, M. Wintersteller, and M. M. Hull, The influence of "the radiation around us" on student conceptual understanding and interest, in *PhyDid B—Didaktik der Physik—Beiträge zur DPG-Frühjahrstagung* (2022), <https://ojs.dpg-physik.de/index.php/phydid-b/article/view/1225>.
- [65] P. G. Hewitt, *Conceptual Physics*, 13th ed. (Pearson Higher Ed, 2022), <https://www.pearson.com/en-us/subject-catalog/p/conceptual-physics/P200000006941/97801373-94975?srsltid=AfmBOor9z3wo-2jRP2fLgWzJb5-e3dh-0Ks5pmBykFITPDNjMZVq91Rs>.
- [66] FutureLearn, How to determine the optimal number of clusters: scree plots, and the elbow method, <https://medium.com/@tarangds/determining-the-optimal-number-of-clusters-using-the-elbow-method-how-when-where-what-and-why-cd1e94b82c1e> (2024).
- [67] A. L. Traxler, T. Suda, E. Brewre, and K. Commeford, Network positions in active learning environments in physics, *Phys. Rev. Phys. Educ. Res.* **16**, 020129 (2020).
- [68] E. Sowles, D. J. Rosen, and M. R. Stetzer, Using meta-cognitive prompts to explore student reasoning trajectories, *presented at PER Conf. 2022, Grand Rapids, MI*, [10.1119/perc.2022.pr.Sowles](#).

- [69] A. Werth, J. R. Hoehn, K. Oliver, M. F. J. Fox, and H. J. Lewandowski, Instructor perspectives on the emergency transition to remote instruction of physics labs, *Phys. Rev. Phys. Educ. Res.* **18**, 020129 (2022).
- [70] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/5yyz-hpx8> for alluvial diagrams that contain all alluvia.
- [71] K. S. Brock and R. Taber, Making claims about learning: A microgenetic multiple case study of temporal patterns of conceptual change in learners' activation of force conceptions, *Int. J. Sci. Educ.* **42**, 1388 (2020).
- [72] R. Brock and K. S. Taber, The application of the microgenetic method to studies of learning in science education: Characteristics of published studies, methodological issues and recommendations for future research, *Stud. Sci. Educ.* **53**, 45 (2017).
- [73] I. Neumann, G. W. Fulmer, and L. L. Liang, Analyzing the FCI based on a force and motion learning progression, *Sci. Educ. Rev. Lett.* **8** (2013), <https://repository.nie.edu.sg/bitstream/10497/15718/1/SERL-2013-8.pdf>.
- [74] A. C. Alonzo and J. T. Steedle, Developing and assessing a force and motion learning progression, *Sci. Educ.* **93**, 389 (2009).
- [75] M. T. H. Chi and J. D. Slotta, The ontological coherence of intuitive physics, *Cognit. Instr.* **10**, 249 (1993).
- [76] J. D. Slotta and M. T. H. Chi, Helping students understand challenging topics in science through ontology training, *Cognit. Instr.* **24**, 261 (2006).