Identifying student conceptual resources for understanding physics: A practical guide for researchers

Amy D. Robertson,[1,*] Lisa M. Goodhew,[1,†] Lauren C. Bauman,[2,‡] Brynna Hansen,[1,∥] and Anne T. Alesandrini[2,§]

[1]*Department of Physics, Seattle Pacific University, Seattle, Washington 98119, USA*
[2]*Department of Physics, University of Washington, Seattle, Washington 98125, USA*

(Received 30 January 2023; accepted 7 July 2023; published 4 October 2023)

[This paper is part of the Focused Collection on Qualitative Methods in PER: A Critical Examination.] Identifying student ideas about particular physics topics is one of the earliest and longest-standing foci of physics education research. This paper presents a method for identifying common conceptual resources for understanding physics, using large numbers of written student responses to conceptual questions. We walk researchers step by step through how we have done this ourselves, from collecting data, to identifying candidate resources, to turning our lists of candidate resources into a coding scheme that can then be applied to datasets of >200 responses. The outcome of these methods are lists of common conceptual resources for understanding particular topics in physics, such as forces, energy, circuits, and heat and temperature.

DOI: 10.1103/PhysRevPhysEducRes.19.020138

I. INTRODUCTION

One of the earliest and most enduring strands of physics education research (PER) has been identifying student ideas about particular physics topics [1–3], from forces to energy to circuits to waves (see, e.g., Refs. [4–10]). For many years, the vast majority of this work focused on student difficulties, alternative conceptions, and/or misconceptions—what it is about physics that is especially difficult for students to learn, and/or what are ways in which students routinely misunderstand or misconceive formal physics [3,11,12]. Docktor and Mestre [2], in their "comprehensive synthesis of physics education research at the undergraduate level," published in 2014, list "identifying common misconceptions" and "developing and evaluating instructional strategies to address students' misconceptions" as two of the three primary research questions pursued by those studying conceptual understanding in physics. This work has had a variety of instructional impacts, including the development of research-based instructional materials for physics (see, e.g., Refs. [13–15]) and conceptual inventories that diagnose and assess student understanding (see, e.g., Refs. [5,6,16,17]). Methodologically, this work has typically focused on *recurrence* and *commonality*, looking for themes in student responses to conceptual questions, foregrounding a model of generalizability that seeks reproducibility across multiple sources of heterogeneity [18].

Much of this research has been practical, focusing specifically on (1) patterns in student responses that point to ways in which students have difficulty with or misconceive formal physics and then on (2) developing instructional strategies and materials to elicit and address these difficulties or misconceptions. Early on, some researchers (see, e.g., Ref. [19]) worked to develop *explanatory models* that would help make sense of why students were answering questions incorrectly, even after instruction; in many cases these researchers modeled misconceptions as stable and student thinking as theorylike, requiring direct and targeted intervention to change [4,19,20]. Other researchers took a less overtly theoretical stance, instead focusing on patterns in empirical data, emphasizing that in many cases findings suggested that students "lacked a consistent conceptual system" [21].

Alongside and in response to misconceptions- and difficulties-oriented research was the emergence of resources theory, which modeled student thinking as the context-dependent activation of pieces of knowledge [22–24]. Resources theory frames student thinking as generative, having been derived from a person's experience and then used to (helpfully) make sense of the material world [23–29]. In resources theory, students' existing ideas are framed as continuous with (and as the basis for) more sophisticated or canonical understandings.

*robertsona2@spu.edu, she/her/hers
†goodhewl@spu.edu, she/her/hers
‡lcbauman@uw.edu, she/her/hers
∥hansenb5@spu.edu, she/her/hers
§aadrini@uw.edu, they/them/theirs

Early scholarship on resources largely focused on developing theory and then on creating curriculum and instructional strategies that are consistent with a view of students' thinking as context dependent and potentially productive (see, e.g., Refs. [22,24]). Few researchers, if any, used the tools that were developed to identify common misunderstandings to go on to identify *common* conceptual resources, even as there were overt calls to do so. For example, Hammer [25] wrote:

> …whereas the physics education research community has devoted substantial attention to studying the nature of student difficulties, it has paid little attention to documenting and systematizing extant ideas about student resources. Without that attention, this knowledge remains mostly tacit and unexamined. I am arguing that it should become a primary agenda of the physics education research community to develop explicit accounts of student resources, to allow their exchange, review, and refinement.

Though it took almost 15 years from the time of this call, identifying common conceptual resources for understanding specific physics topics is a growing focus in physics education research (see, e.g., Refs. [30–34]). Our own position is that this kind of research has significant potential to transform introductory physics instruction at the college and university levels, particularly if it is accompanied by tools and support for instructors, such as instructional materials, illustrative examples of students using particular conceptual resources, and guidance on how to implement resources-oriented instruction.

Our research team is among those who have focused on identifying common conceptual resources in students' written responses to conceptual questions, in order to complement existing, topic-specific misconceptions- and difficulties-oriented research. The aim of this paper is to share our methods for doing so: we walk readers step by step through our process, developed and iterated multiple times over the past eight years. This is meant to be a pragmatic paper, with the aim of supporting researchers who want to do something similar. We try to be as transparent as possible about what shapes our method and how our values and preferences are informing our decisions.

Our methods are best suited to address questions like, "What are some of the common conceptual resources that students use when reasoning about *topic x*?," where *topic x* could be forces, momentum, energy, current, mechanical waves, and so on. Though some have posed connections between framing students' ideas as resources and equity-oriented pedagogy (see, e.g., Ref. [35]), we do not think our methods are well suited to answer questions about identity, affect, equity, or even models of cognition.

The remainder of the paper is organized to provide interested researchers with enough information about our methods—and their affordances and limitations—to implement and adapt them for their own work and/or to replicate existing studies (like ours) that seek to identify common conceptual resources. Toward this end, Sec. II defines what we mean by "resource," and Sec. III lays out some of the values that shape the specific form our methods take. Section IV walks readers through our method, step by step. Section V poses some final reflections, and Supplemental Material offers a workbook for interested readers to try out the method on some fresh data [36].

II. WHAT DO WE MEAN BY "RESOURCE"?

Throughout this paper, we will use the term "resource" to mean an idea that is expressed by a student, in this case in writing, that can be framed as continuous with formal physics. (Though not the focus of this paper, parts of our methods are also useful for identifying resources in video or interview data.) As an example, the airplane question from the Force Concept Inventory [6] tells students that an airplane is moving to the right and drops a bowling ball out of its cargo bay, and then asks them which of the shown trajectories represents the motion of the ball after it is dropped (see Fig. 1). A student in one of our previous studies chose trajectory A and wrote, "The bowling ball

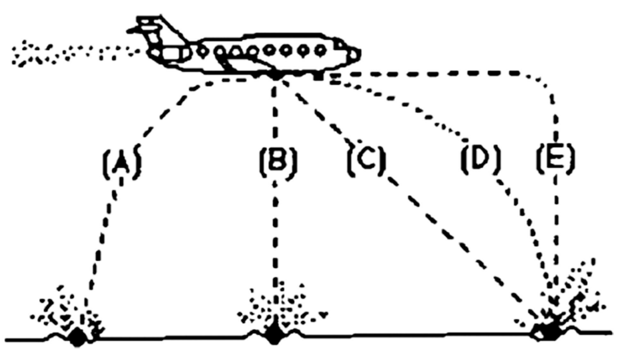


FIG. 1. Airplane question from the Force Concept Inventory. Reprinted from Hestenes *et al.*, Force Concept Inventory, Phys. Teach. **30**, 141 (1992), with the permission of AIP Publishing. Image description: Sketch of an airplane moving to the right, with five different dotted lines emerging from the bottom center of the plane. The lines represent trajectories of a bowling ball. The leftmost trajectory is a curved line that shows the bowling ball moving backward and landing on the ground behind the plane. The next trajectory is a straight line to the ground, showing the bowling ball landing directly under the plane. The next trajectory is a straight diagonal line showing the bowling ball landing in front of the plane. The next trajectory is a curved line also showing the bowling ball landing in front of the plane. The rightmost, final trajectory looks like a curved letter L that has been rotated 180 deg, such that the ball moves horizontally forward at first and then abruptly falls to the ground. The trajectories are labeled from left to right: A, B, C, D, and E.

would most likely follow the path A, because the ball will experience air resistance as it is dropped. The air resistance will pull the ball back making it follow this trajectory." Here, the student attributes the backward motion of the ball to a "pull" from "air resistance"; they seem to be thinking that the ball will move backward because it is pushed by the air as it falls. One way to highlight the connection between Newton's laws and the student's response is to say that the response uses the resource that "forces influence the motion of objects" [32]. This statement of the resource is neither *exactly* what the student said nor *exactly* Newton's laws; it is meant to capture the relationship between the two while preserving the essence of what students are saying.

Generativity.—Our use of "resource" to mean a student idea that is continuous with formal physics intentionally ascribes generativity, from a physics perspective, to student ideas. In other words, we want our work to highlight that student ideas—even canonically incorrect ones—can be generative for student learning in introductory physics. This choice is consistent with existing resources frameworks, which emphasize that conceptual resources are derived from a person's experience, including prior learning, and are helpful to them in making sense of the material world [23–29]. For example, diSessa [23] says that phenomenological primitives ("p-prims," which we consider to be a kind of resource) such as "closer means stronger" are best understood as "serv[ing] individuals well in dealing effectively with the physical world," e.g., in making sense of it, interacting with it, etc. Smith *et al.* [24] define resources as "any feature of the learner's present cognitive state that can serve as significant input to the process of conceptual growth," emphasizing the continuity between students' intuitive ideas and learning objectives.

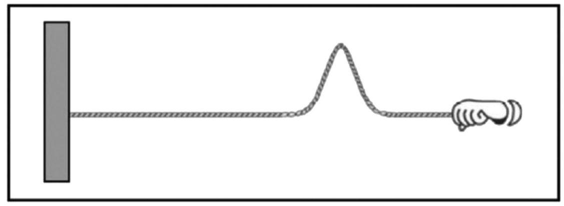
Type and grain size.—Our methods focus on *conceptual* resources—over and above, for example, procedural or epistemological resources [37–39]—because one of our primary goals is to support the development of instructional materials where the objective is conceptual understanding. Our methods also focus on resources at the grain size of a typical learning goal in introductory physics—rather than, e.g., small cognitive elements that together form ideas in context—because we want instructors to as immediately as possible see the connection between the resources we name and conceptual learning objectives in introductory physics (often stated in "idea" terms such as "objects in motion stay in motion").

Context sensitivity.—In resources theory, a *resource* is a piece of knowledge that gets activated in real time, in context-sensitive ways [22–29,40]. Consistent with this theory, we expect the ideas we report to be context dependent [22,23], and are likely composed of smaller "pieces" of varying grain size and levels of organization that may activate differently in a different context [22,23,28]. In other words, while our model of generalizability [18] suggests that resources that we identify as recurring across contexts and samples may be likely to come up in new contexts, we do not expect stability or coherence for a particular student or sample.

Relationship to other resourcelike constructs.—Our use of the word "resource" is expansive enough to include constructs like p-prims [23] and FACETs [41], but is not a one-to-one mapping of either. P-prims, as we understand them, are elements of knowledge that are "minimal abstractions of common events," such as "bouncing" or "closer means stronger" [23]. In naming and defining p-prims, diSessa's goal was "to develop a framework for understanding the origins and development of commonsense knowledge about the physical world, particularly as it influences the learning of school physics," and in particular to "chart the structure" and deployment of commonsense knowledge, asking questions like how knowledge elements are organized and how to intervene in this organization. Because he was looking to identify knowledge *elements*, the set of criteria that diSessa used to identify p-prims included a focus on principles like "ready availability," "impenetrability," and "unproblematic genesis." diSessa's aim is different than ours, which has been to identify common conceptual resources for understanding introductory physics topics. It is possible (though we think unlikely) that the resources we identify represent cognitive elements, and it is also possible that some of the p-prims diSessa identified are common and could be framed in ways that highlight the continuity between the idea and formal physics. But this has not been either of our emphases—we have not used methods that would establish the resources we identify as p-prims, and diSessa did not, to our knowledge, use methods that would establish the commonality of p-prims.

Our goals overlap more with those of Minstrell [41] in identifying FACETs, which he defines as "the pieces of knowledge or reasoning that students seem to be applying in problem situations." Minstrell says that in naming FACETs, he and his colleagues "use the language of the students" and seek to "capture the intention of each expressed idea." He elaborates that they chose the name FACET "to avoid the assumptions associated with terms like misconceptions, alternative conceptions, naïve theories or beliefs, etc.," because "much of the students' knowledge is valuable," and sometimes this knowledge "can be used as anchors of understanding." Like our work, Minstrell's work aims to provide instructors with knowledge of their students' specific physics ideas. Unlike our work, Minstrell's FACETs do not, in our view, intentionally seek to highlight continuities between formal physics and student thinking; many of the FACETs are stated in ways that instructors would immediately recognize as incorrect and act on as such (e.g., "current is proportional to potential difference, regardless of resistance" [41]).

Consider the following two scenarios:
In scenario 1, your Teaching Assistant (TA) creates a pulse by flicking the end of a spring, as in the figure at right. In scenario 2, your TA pulls the spring so that it is more taut (*i.e.*, increases the tension in the spring) and then creates a pulse by flicking the end of the spring in the same way.



The pulse in scenario 2 travels down the spring faster (*i.e.*, has a larger speed) than the pulse in scenario 1.

Why would it make sense for a pulse to move faster on a higher-tension spring? (We're trying to understand your intuition, not whether or not you can remember particular equations. In other words, we want to know *how you make sense of* this phenomenon.)

FIG. 2. Tension pulse-flick question, originally published in Goodhew *et al.*, Student resources for understanding mechanical wave propagation, Phys. Rev. Phys. Educ. Res. **15**, 020127 (2019). Image description: Text of the tension pulse-flick question and a figure illustrating a string attached to a wall, held by a hand to the right of the wall. The hand is horizontal, positioned halfway between the top and bottom of the wall, and a pulse moves to the left along the top of the string. The question text reads, "Consider the following two scenarios: In scenario 1, your Teaching Assistant (TA) creates a pulse by flicking the end of a spring, as in the figure at right. In scenario 2, your TA pulls the spring so that it is more taut (i.e., increases the tension in the spring) and then creates a pulse by flicking the end of the spring in the same way. The pulse in scenario 2 travels down the spring faster (i.e., has a larger speed) than the pulse in scenario 1. Why would it make sense for a pulse to move faster on a higher-tension spring? (We're trying to understand your intuition, not whether or not you can remember particular equations. In other words, we want to know *how you make sense of* this phenomenon)".

III. VALUES GUIDING OUR MEANINGS AND METHODS: INSTRUCTIONALLY SIGNIFICANT, THEORETICALLY CONSISTENT, ACCOUNTABLE TO THE DISCIPLINE

In this section, we articulate three values that guide our characterization of resources (as outlined in Sec. II) and the specific methods we developed to identify, name, and share them (described later in Sec. IV). In particular, we want our methods to produce *instructionally significant* insights, to be *consistent with resources theory*, and to be *accountable to the discipline of physics education research*.

Instructionally significant.—The methods described in this paper are meant to produce a list of topic-specific, common resources that are specific, conceptual, and continuous with learning goals in introductory physics. We want these lists to be "instructionally significant," by which we mean that we hope to produce work that at least some instructors interpret as relevant to their classroom. In particular, we hope our work supports instructors in noticing and building on student thinking in physics. Given that classrooms are often a "blooming, buzzing confusion of sensory data" ([42], quoting James [43]), we expect that lists of resources like ours will serve as *one input* to a complex process of both emergent and planned instructional decision making.

Instructional significance shapes our efforts to design questions that we think instructors could use as formative assessment tools [44,45]—i.e., that will help instructors understand their students' thinking in relation to topics commonly covered in introductory physics courses. We speak at greater length about how we design our questions in a forthcoming paper [46], but we provide an example question in Fig. 2. This question, the *tension pulse-flick* question, was designed to elicit thinking related to the introductory physics learning objective that students will understand that the propagation speed of a pulse depends on the properties of the medium (in this case, the tension and the linear mass density). The key design feature of this question is that it tells students what happens—when the spring is under more tension, the pulse travels faster—and asks for students to explain how they make sense of the observation.[1] In doing so, this question targets students' ideas about why or how the tension of the medium affects pulse speed [33].

Our goal of producing instructionally significant research also shapes what we notice in written responses. In particular, we look for ideas that sound like "seeds" or "beginnings" of formal physics understandings [48], that we expect could be generative in a learning environment focused on particular conceptual targets. In Sec. II, we shared a student response to the "airplane question" from the Force Concept Inventory [6] (Fig. 1): "The bowling ball would most likely follow the path A, because the ball will

[1]The style of this question shares features with questions in Ref. [47].

experience air resistance as it is dropped. The air resistance will pull the ball back making it follow this trajectory." Our use of a resources analytic lens means that we are attuned to the part of this response that is continuous with Newton's laws, namely, that a "pull" force will change the bowling ball's motion. It is not as significant, in this resources analysis, to focus on the student's mislabeling of air resistance as a pull rather than a push, or to focus on their overestimation of the effect of air resistance on the ball's motion. In many cases, our focus on conceptual resources means we ignore procedural, epistemological, analogical, etc., resources [37–39] that we could report; again, this choice is shaped by the particular form of instructional significance we are aiming for.

Finally, our goal of producing instructionally significant research shapes how we name the resources we identify. The names we assign to resources seek to clarify the connections we perceive between the ideas and conceptual targets in introductory physics. This usually means that the wording of the resource accounts for both (i) the language students use in their writing and (ii) language that resembles conceptual targets in physics. For example, "forces influence the motion of objects" is neither exactly Newton's laws nor exactly what students say, but it is close enough to both that (we think) it can support instructors in identifying it in student discourse and in perceiving the connection between this discourse and formal physics.

Theoretically consistent.—In addition to answering to our value of instructional significance, our methods are consistent with resources theory, and in particular the instructional orientation that the theory is meant to encourage—that of attending to and building from students' generative ideas [49,50]. Though an orientation toward student ideas as generative is our primary emphasis, we have also sought consistency with the theory more broadly, which includes a focus on context dependence and a particular definition of learning as changing the structure or activation of resources, by reorganizing, refining, properly activating, increasing the degree of formality of, or changing the role of resources [22–25,28,29]. The dynamics of learning are not a primary focus of the methods in this paper (which identify resources from written responses), but are a central part of how we think about developing curriculum that builds from the resources we identify.

The context dependence of resource activation is a guiding assumption that shapes how we report our findings. Not only do we tend to report frequencies of resource use, showing that different samples of students draw on particular resources at different frequencies, but also we frame our findings as useful in pointing instructors to the kinds of ideas students *may* use, over and above being helpful in *predicting* the frequency of a particular idea in any given context. We see our work as most useful for illustrating *that* students have generative ideas for reasoning about physics, and we seek to provide as much detail as possible about how we perceive that to be happening in our data to ease the process of identifying similar ideas in other contexts.

Accountable to the discipline of PER.—PER has a long history of research that identifies common, topic-specific difficulties, misconceptions, and misunderstandings, with attendant methods that have been refined over time and have some degree of consensus among researchers. These methods typically include collecting large (>200) numbers of written student responses to conceptual questions, and then seeking themes that point to misunderstandings, difficulties, or misconceptions.

Our work answers to calls (like the one in the Introduction) for resources-oriented research that draws on time-tested tools and methods from misconceptions- and difficulties-oriented research. We use similar methods, in that we collect large numbers of student responses to conceptual questions and look for themes; however, our focus is on what is continuous with formal physics in students' responses rather than what about those responses indicates that students do not understand. One point of divergence from some misconceptions- and difficulties-oriented research is that we have often used two coders instead of one, so that we can establish interrater agreement.

These methods choices are consistent with more general qualitative thematic coding methods [51,52], but they draw on disciplinary knowledge and a resources orientation to specifically identify conceptual resources for understanding introductory physics. Our goal in using established, patterns-oriented methods of this type is to locate our work in the same methods landscape as existing misconceptions- and difficulties-oriented research on student ideas.

IV. METHODS: HOW WE IDENTIFY COMMON CONCEPTUAL RESOURCES FROM STUDENTS' WRITTEN RESPONSES TO PHYSICS QUESTIONS

In this section, we walk readers through our methods for identifying common conceptual resources using students' written responses to physics questions. We name seven steps, starting with data collection and ending with calculating interrater agreement on coded responses. In addition to this step-by-step method with examples, we offer Supplemental Material where interested readers can "try out" our method on a small dataset [36]. Our experience is that the overarching practice we are honing as we employ these steps is assuming that students have good reason for answering the way they do, and then articulating a relationship between that good reason and introductory physics learning objectives.

Step 1: Collect students' written responses to open-ended conceptual questions.—We tend to ask questions that rely on models, mechanisms, and concepts, and rarely on calculations, because we are most interested in these aspects of conceptual understanding [53]. (Another researcher who is more interested in mathematical or procedural reasoning—including the conceptual reasoning therein—may elect to use very different questions.) Figure 2 shows the tension pulse-flick question, one of the questions we used in our study of conceptual resources for understanding mechanical wave propagation [33]. Notably, this question takes the form of telling students what happens in a particular scenario and asking them to sense-make about why it happens that way. We call questions like this "explain" questions (in contrast to "predict" questions) [53]. Most of our questions take this form, but any question that asks students to explain their reasoning can work for a resources analysis.

We are seeking to make claims about *common* conceptual resources, which for us means that the resources are common across samples of students and common across questions (within a particular topic, like mechanical wave propagation or electric current). That is, we are seeking to make the claim that the resources we identify come up for many students in multiple instructional contexts, in response to multiple questions (and thus span a conceptual space). For this reason, when gathering students' written responses, we try to (a) gather responses to a single question from multiple colleges and universities and (b) gather responses to at least a few different questions for each topic. In many cases we gather much more data than we analyze because some questions turn out not to elicit particularly rich student responses. We rarely perform a detailed analysis for questions that elicit short responses or little variety among responses. For a recent study on student conceptual resources for understanding forces [32], we collected data on 17 questions, from 13 colleges and universities. We ultimately analyzed five of the questions, with responses from nine universities.

Related to seeking commonality by surveying students from multiple universities, Kanim and Cid [54] have problematized the homogeneity of student populations in PER studies, showing that most research in PER has been done on white, wealthy, high-mathematics-SAT-scoring populations of students. This homogeneity limits the generalizability of claims in PER; claims made about "common" conceptual resources where the sample is exclusively or primarily from, e.g., predominantly white universities, both hyperuniversalizes and sets up a norm against which all other students are compared. We are responding to the critique of Kanim and Cid in our own work by making strides to identify what constitutes a representative sample of introductory physics students [55] and in part by purposefully sampling from two-year colleges and minority-serving institutions, whose students are underrepresented in existing PER studies. Kanim and Cid encourage researchers to report the demographics of their samples, so that we as a community can work to diversify the student populations represented in PER.

Step 2: Identify candidate student conceptual resources specific to individual questions.—In this step, multiple (usually 2–3) researchers read a subset of student responses to an individual question (such as the tension pulse-flick question) and identify conceptual resources specific to that question. Typically, we either (i) look at 10% of the sample (often 15–30 responses, given the sizes of our datasets) or (ii) stop looking when we have reached saturation, in this case meaning that we are no longer seeing resources that we have not already noted.

Researchers read each response to first try to understand what students mean by what they are saying and then identify or imagine connections between what we think students mean and formal physics concepts as we understand them. Returning again to the example from the airplane question above (Fig. 1), the student chose the trajectory where the ball moves backward from the plane and wrote, "The bowling ball would most likely follow the path A, because the ball will experience air resistance as it is dropped. The air resistance will pull the ball back making it follow this trajectory." Here, we interpret the student to be imagining that as the ball is released from the plane, it encounters air resistance (we imagine them thinking of something like a rush of wind), and that air resistance pushes the ball backward. It is not clear whether the student thinks that air resistance "overcomes" the forward motion of the ball, or whether the student is not considering the forward motion of the ball (instead imagining the ball in the reference frame of the airplane), but in either case, the student response implies a connection between a force exerted on an object and a change in its motion, making this response continuous with Newton's laws.

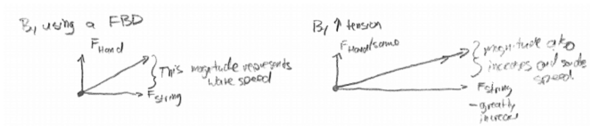
As we researchers move through step 2, we make preliminary (candidate) lists of resources we perceive in student responses. At this stage we pay little attention to whether we think we will find the resources in other questions or other responses, and we tend to stay close to the students' wording, e.g., "gravity makes things fall" or "friction slows things down" (rather than "forces affect objects' motion"). We illustrate each candidate resource with student responses.

At the end of step 2, each researcher has a bullet-pointed list of fine-grained, candidate resources, with a series of examples of each one. For example, Fig. 3 is a sample (partial) bulleted list of resources that one researcher identified in a small subset of responses to the tension pulse-flick question (Fig. 2). Researchers may compare or combine lists and refine individual candidate resources to come closer to consensus, either at this point or later in the process.

Importantly, a single response can contain multiple resources, particularly at this stage. For example, consider the response in Fig. 3 that reads,

> From just thinking about the scenario, I think it makes sense for a pulse to move faster on a higher tension spring, because as the pulse travels through, each part of the higher tension spring

- **Lower mass density means it takes less energy for the pulse to travel.**
  Pulling a spring (creating tension) decreases the number of coils in a given distance. This in turn lowers the mass density of the same length of spring. With a lower density and the same energy a wave uses less energy to travel through the medium and more energy to "create" its speed.
- **Increasing the tension of the spring increases the potential energy (which can then be transformed into kinetic energy, and the pulse can move faster).**
  If the tension of the string is greater, the system would have more potential energy, then once the system is given a pulse, that potential energy is converted to kinetic energy which is a function of velocity. Therefore it would make sense for a pulse to move faster on a higher tension spring by the law of conservation of energy.
- **The restoring force is bigger in a higher tension spring.**
  From just thinking about the scenario, I think it makes sense for a pulse to move faster on a higher tension spring, because as the pulse travels through, each part of the higher tension spring will [be] displaced less than the normal spring would and would return to equilibrium faster. The pulse would not create such a large amplitude, which would slow it down, which is what happens when you decrease the tension.
- **The wave speed is determined by a force that is the sum of the tension and the hand force.**



- **Increasing tension means energy is transferred more quickly.**
  Higher string tension = quicker transfer of energy. I think a pulse would move faster with higher tension. The string would move less, therefore leading to less displacement, overall leading to the ability for faster pulses.
- **The speed of the wave is related to how fast the particles of the spring return to equilibrium after being disturbed.**
  A wave is basically a disturbance in a medium, in this case, the string. The speed of the wave moving through the medium is directly related to how fast the particles of the string can return to equilibrium position after being disturbed. The more tense the string, the faster the ability of the particles to return to equilibrium position. This is why wave speed increases with increase in tension. Greater tension = greater restoring force = greater wave speed.
- **Greater tension means a greater force between the particles of the spring.**
  The wave pulse speed would increase when the tension of the spring increases because higher tension means there's a higher force between the particles in the string. When the wave is displaced, there is a greater restoring force to bring the particles back to equilibrium. Therefore, there's a greater tendency to pass the displacement along, resulting in a higher wave speed.

FIG. 3. Sample (partial) list of question-specific candidate resources for the tension pulse-flick question (Fig. 2). Bolded text is the list of fine-grained, candidate resources, and subsequent plain text shows examples from student responses. Image description: List of bolded, bullet-pointed sentences, representing candidate resources, and italicized student responses underneath each one. The first bullet reads, "Lower mass density means it takes less energy for the pulse to travel," and the italicized example underneath it reads, "Pulling a spring (creating tension) decreases the number of coils in a given distance. This in turn lowers the mass density of the same length of spring. With a lower density and the same energy a wave uses less energy to travel through the medium and more energy to "create" its speed." The second bullet point reads, "Increasing the tension of the spring increases the potential energy (which can then be transformed into kinetic energy, and the pulse can move faster)." The italicized example below this bullet point reads, "If the tension of the string is greater, the system would have more potential energy, then once the system is given a pulse, that potential energy is converted to kinetic energy which is a function of velocity. Therefore it would make sense for a pulse to move faster on a higher tension spring by the law of conservation of energy." The third bullet point reads, "The restoring force is bigger in a higher tension spring." The example student response underneath this bullet point reads, "From just thinking about the scenario, I think it makes sense for a pulse to move faster on a higher tension spring, because as the pulse travels through, each part of the higher tension spring will [be] displaced less than the normal spring would and would return to equilibrium faster. The pulse would not create such a large amplitude, which would slow it down, which is what happens when you decrease the tension." The fourth bullet point reads, "The wave speed is determined by a force that is the sum of the tension and the hand force." Underneath this bullet is a drawing with two sets of vector diagrams. The vector diagram on the left is labeled, "By using a FBD," and there are three vectors, one pointing straight up, one pointing horizontally to the right, and one that is the sum of the two. The vector pointing straight up is labeled, "F_{Hand}," the horizontal vector is labeled, "F_{string}," and there is a bracket between the horizontal vector and the vector sum labeled, "This magnitude represents wave speed." The vector diagram on the right is labeled, "By [up arrow] tension," and has similar vectors to the diagram on the left, except the horizontal (F_{string}) vector is longer, and thus so is the vector representing the sum of the vertical and horizontal vectors. On this diagram, the student has indicated that the vertical vector is the same, the horizontal vector greatly increases, and so the magnitude of the resultant vector also increases. The fifth bullet point reads, "Increasing tension means energy is transferred more quickly." The example response below it reads, "Higher string tension = quicker transfer of energy. I think a pulse would move faster with higher tension. The string would move less, therefore leading to less displacement, overall leading to the ability for faster pulses." The sixth bullet point reads, "The speed of the wave is related to how fast the particles of the spring return to equilibrium after being disturbed." The example response below it reads, "A wave is basically a disturbance in a medium, in this case, the string. The speed of the wave moving through the medium is directly related to how fast the particles of the string can return to equilibrium position after being disturbed. The more tense the string, the faster the ability of the particles to return to equilibrium position. This is why wave speed increases with increase in tension. Greater tension = greater restoring force = greater wave speed." The seventh (final) bullet point reads, "Greater tension means a greater force between the particles of the spring." The example response below it reads, "The wave pulse speed would increase when the tension of the spring increases because higher tension means there's a higher force between the particles in the string. When the wave is displaced, there is a greater restoring force to bring the particles back to equilibrium. Therefore, there's a greater tendency to pass the displacement along, resulting in a higher wave speed".

> will [be] displaced less than the normal spring would and would return to equilibrium faster. The pulse would not create such a large amplitude, which would slow it down, which is what happens when you decrease the tension.

This response could illustrate both “the restoring force is bigger in a higher tension spring” and “the speed of the wave is related to how fast the particles return to equilibrium after being disturbed,” or a single resource that includes these ideas together. In this case, we kept the candidate resources separate (and at a finer grain size) at this stage, because ideas about restoring force were often expressed more macroscopically and ideas about particle displacement more microscopically; some students who referenced a restoring force did not refer to the transverse motion of *parts* of the spring.

Almost all of the candidate resources in Fig. 3 are phrased as answers to the tension pulse-flick question. That is, each candidate resource is a reason that a pulse moves faster in a higher-tension spring: the mass density is less and so it takes less energy for the pulse to move through, the restoring force is bigger (and so the spring returns to equilibrium faster), and greater tension means a greater force between the particles. Though not always quite this specific, it is typically the case that our candidate resources sound like at least part of a reason that someone might give for why something happens the way it does. “Air resistance affects the motion of the bowling ball,” “friction slows things down,” and “gravity makes things fall” are all similar to these wave examples in this way.

As we said earlier, our aim in naming candidate resources is to try to capture both what students mean by what they are saying and how what they are saying is continuous with formal physics. This is a subjective and creative process. For example, the first student quote in Fig. 3 reads,

> Pulling a spring (creating tension) decreases the number of coils in a given distance. This in turn lowers the mass density of the same length of spring. With a lower density and the same energy a wave uses less energy to travel through the medium and more energy to ‘create’ its speed.

This response includes many ideas for researchers to consider. For example, we might focus on the relationship this student is identifying between lengthening a spring and changing its mass density, suggesting that they seem to understand (in at least a rudimentary way) that μ is m/L. In our choosing to use this response to articulate the resource, “lower mass density means it takes less energy for the pulse to travel,” we are honing in on why the student thinks higher tension means a faster pulse, which is that doing so makes it “easier” for the pulse to move through the spring (because its mass density is lower). We interpret them to be saying that using less energy to move through the medium means more energy can be “spent” on speed (to “create” the wave’s speed). This ties the candidate resource we wrote to energy conservation, even if misapplied.

Step 3: Reduce the preliminary set of (specific, often fine-grained) candidate resources into a smaller set of (often coarser-grained) common conceptual resources for understanding topic X.—This step might be characterized as theme seeking, with the fine-grained candidate resources serving as cases of the coarser-grained resources. For example, “gravity makes things fall” and “friction slows things down” might be cases of (and thus combined into) the resource “forces affect the motion of objects.” During this step, candidate fine-grained resources are often left out if they are either uncommon or specific to a single question. There are a number of considerations that matter for how we engage in this process, which often operate simultaneously, subconsciously, and in a dynamically evolving way. In particular, we are

- *Engaging our own conceptual understandings of the physics topic*. We usually begin step 3 with long lists of candidate resources that highlight continuities between (i) formal physics and (ii) students’ answers to *particular* questions. In performing step 3, we are engaging in the physics practice of asking ourselves, “What are the *fundamental* physics ideas that these candidate resources represent?” This practice relies on our being able to identify, for example, gravity, air resistance, and friction as cases of forces, and speeding up, slowing down, and turning as cases of changes in motion, so that we can say that “gravity speeding things up” and “friction slowing things down” are both instances of a force changing an object’s motion. The fine-grained candidate resources we include in our sense making about the coarser-grained resources need not be canonically correct—e.g., we may include the idea that “the centrifugal force changes the direction the car moves” in our thematic grouping toward “forces affect the motion of objects.” Notably it is in step 3 that we discard candidate resources that we deem less relevant to the specific topic we want to investigate—e.g., cutting resources that are fundamentally about speed, if the study is about forces.
- *Staying as close to students’ meaning as possible*. As we engage our own physics understandings, we are committed to neither overgeneralizing nor overinterpreting students’ responses, and our aim is to stay as close to students’ meaning as possible in the names we assign to the coarser-grained resources. A principle we often use is to extrapolate no more than one step away from the candidate (finer-grain-sized) resources (e.g., from “gravity” to “forces” and from “makes

things fall" to "affects the motion of objects"). We select names for resources that we think would feel familiar and affirming to students if they saw them.

- *Being responsive to introductory physics learning goals*. The way that we ultimately choose to name our resources, at a coarse-grain size, in step 3, is shaped by our sense of introductory physics learning goals. That is, we try to name resources in ways that highlight not only continuities between students' ideas and formal physics, but also (and specifically) the ways in which students' ideas are continuous with the physics we want them to learn. In other words, we select names for resources that we think will feel familiar and affirming to instructors of introductory physics.
- *Being practical about coding*. In practice, we have found that with datasets as large as those we tend to analyze (usually exceeding 1000 written responses), a coding scheme that has more than five resource codes is difficult to use consistently. As we collapse our set, we aim for a final coding scheme that includes as many of our common fine-grained candidate codes as possible, which often means that the more candidate codes we have, the more abstracted the final codes are from our original set. Alternatively, we may decide at this stage to narrow the scope of the kind of resource we report—e.g., to a smaller conceptual domain—in order to avoid further abstracting our final codes.
- *Working to maintain fidelity with the original data*. When we have a preliminary list of five (or so) coarse-grained resources, we return to the data (usually the example student responses we used to generate candidate resources in step 2). For each candidate resource that we collapsed into the coarser-grained code, we ask if, in encountering the example student response, we would code that response as the coarser-grained resource. If it is repeatedly the case that our answer is no, we work to refine our coarse-grained resource codes.

Figure 4 shows two lists. The first (orange box on the left) is a subset of the fine-grained resources that we identified from student responses to a series of conceptual questions about mechanical wave propagation, including (but not limited to) the tension pulse-flick question. The second (blue box on the right) shows how we grouped these ideas to articulate coarser-grained resources that became codes in our first-pass coding scheme. The candidate resources (c) "The restoring force is bigger in a higher tension spring," (f) "Increasing tension means amplitude is less and energy is transferred more quickly," (g) "The speed of the wave is related to how fast the particles of the spring return to equilibrium after being disturbed," and (j) "Greater tension means a greater force between the particles of the spring" were collapsed into the coarser-grained resource "The speed or duration of transverse motion affects pulse speed." All of these question-specific, finer-grained ideas seemed to reflect the idea that a pulse travels faster on a higher-tension or lower-mass-density spring because the spring (or pieces of it) moves up and down faster (or the time it takes is less). Students offered a variety of reasons: that the restoring force is higher, the amplitude is less, and the particles pull on one another more. In another example, the candidate resources (b) "A pulse has to move further in a slack string, so increasing tension will allow it to move faster," (d) "The speed of the pulse depends on the medium through which it travels/properties of the spring," and (h) "A lighter string has less inertia or resistance to the motion of the pulse, so a pulse will move faster" were collapsed into the coarser-grained resource, "The properties of the medium impede or facilitate pulse movement." Here, as above, students offer a variety of properties that make it easier or harder for the pulse to move: the length or slackness of the spring and the "lightness" or weight of the spring. Finally, the candidate resources (a) "Increasing the tension of the spring increases the potential energy (which can then be transformed into kinetic energy), and the pulse can move faster," (i) "It takes less energy to move a pulse through a lighter spring," and (e) "Lower mass density means it takes less energy for the pulse to travel" were collapsed (with others) into the coarse-grained resource, "Pulse speed is affected by the kinetic energy of the pulse" (the wording of this resource was refined in step 4). In all of these finer-grained cases, students seemed to be treating the pulse as an object whose speed is related to its kinetic energy, where the more energy that was left over (from potential energy or from doing "work" on the spring), the more that could be "used" as kinetic energy.

As we constructed these coarser-grained codes, we could have grouped ideas differently. For example, the ideas that we grouped as "the speed or duration of transverse motion affects pulse speed" also share a focus on motion of parts of the medium and/or interactions between parts of the medium. We might have chosen two codes of slightly larger grain size to capture this (e.g., "propagation involves transverse motion" and "propagation involves interactions between neighboring parts of the medium"). We landed on the codes in Fig. 4 because, in our judgement, they more accurately captured the meaning of the student responses elicited by the wave propagation questions we used for this analysis. The alternative codes we suggest here would capture ideas we see students expressing in their responses, but are more abstracted from the responses themselves. These alternative codes may be useful for different instructional or research goals (e.g., these may be more common, may be applicable to a broader range of questions, and/or more available to students at an earlier stage in physics learning).

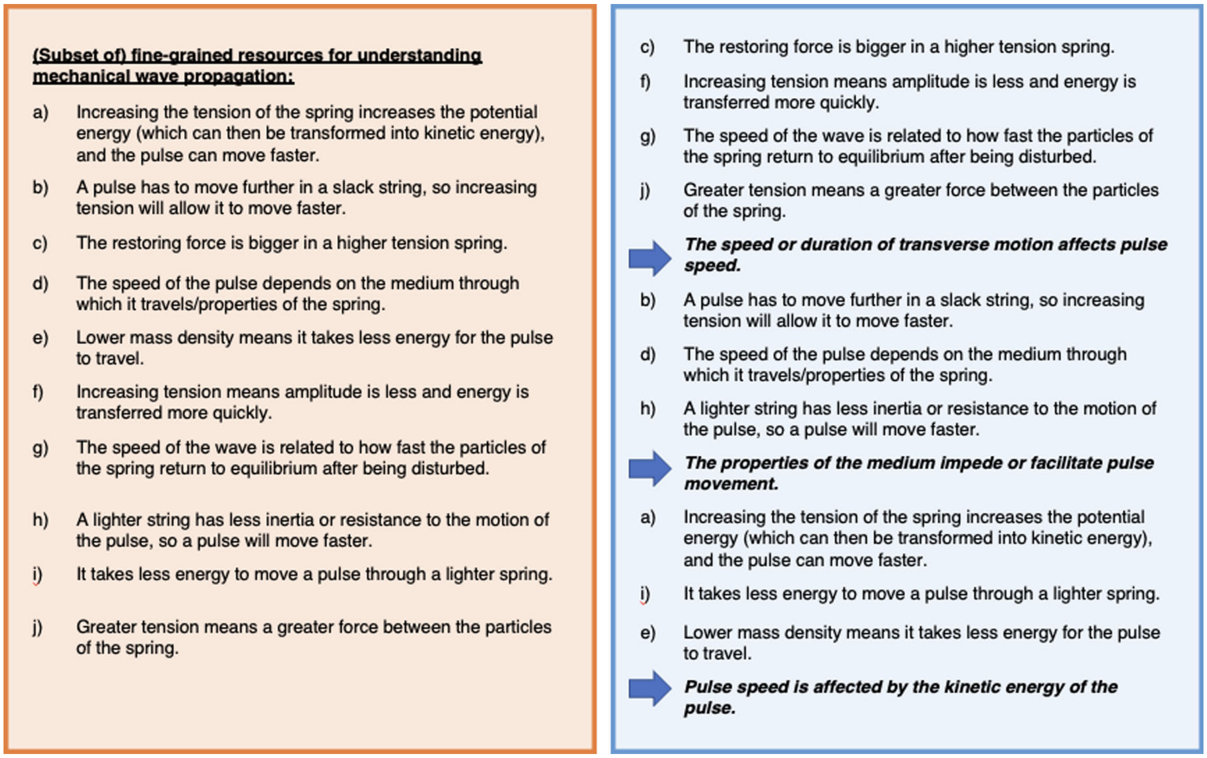


FIG. 4. Sample reduction of preliminary list of candidate resources (left, orange box) into a smaller set of common conceptual resources for understanding mechanical wave propagation (bolded sentences in blue box on right). Image description: On the left is an orange box with a bold, underlined title at the top that reads, “(Subset of) fine-grained resources for understanding mechanical wave propagation:.” Under this title is a list of resources including, “a) Increasing the tension of the spring increases the potential energy (which can then be transformed into kinetic energy) and the pulse can move faster.”; “b) A pulse has to move further in a slack string, so increasing tension will allow it to move faster.”; “c) The restoring force is bigger in a higher tension spring.”; “d) The speed of the pulse depends on the medium through which it travels/properties of the spring.”; “e) Lower mass density means it takes less energy for the pulse to travel.”; “f) Increasing tension means amplitude is less and energy is transferred more quickly.”; “g) The speed of the wave is related to how fast the particles of the spring return to equilibrium after being disturbed.”; “h) A lighter string has less inertia or resistance to the motion of the pulse, so a pulse will move faster.”; “i) It takes less energy to move a pulse through a lighter spring.”; and “j) Greater tension means a greater force between the particles of the spring.” On the right is a blue box that includes the resources from the orange box [a) through f)], grouped into clusters that represent coarser-grained resources, with each coarse-grained resource listed under each cluster. At the top, resources “c) The restoring force is bigger in a higher tension spring”; “f) Increasing tension means amplitude is less and energy is transferred more quickly”; “g) The speed of the wave is related to how fast the particles of the spring return to equilibrium after being disturbed”; and “j) Greater tension means a greater force between the particles of the spring” are listed one after another. Underneath this cluster of fine-grained resources is a horizontal blue arrow pointing to the right, with text beside it that reads, “The speed or duration of transverse motion affects pulse speed.” Under this is another list of fine-grained resources: “b) A pulse has to move further in a slack string, so increasing tension will allow it to move faster”; “d) The speed of the pulse depends on the medium through which it travels/properties of the spring”; and “h) A lighter string has less inertia or resistance to the motion of the pulse, so a pulse will move faster.” This cluster of resources is also followed by a blue arrow, next to which is the text, “The properties of the medium impede or facilitate pulse movement.” Following this is a final list of fine-grained resources: “a) Increasing the tension of the spring increases the potential energy (which can then be transformed into kinetic energy) and the pulse can move faster”; “i) It takes less energy to move a pulse through a lighter spring”; and “e) Lower mass density means it takes less energy for the pulse to travel.” Below this cluster is a final blue arrow, next to which is text that reads, “Pulse speed is affected by the kinetic energy of the pulse”.

Step 4: Test and refine the preliminary coding scheme.— In step 4, we select a subset of responses from the full dataset, and two coders practice using the set of 4–5 coarse-grain-sized resources that were named in step 3 to code the subset. Typically, this subset is 30 responses from the set of questions that were used in step 2; we choose responses from multiple samples, and we usually pick the first several responses from each dataset.

In steps 4 and 5, we use a coding spreadsheet that includes links to the written data (leftmost column), the response being coded (second column), a series of columns where coders can enter their codes (a response will receive a “1” in a code column if we think the response merits that code), and columns that indicate whether two coders agree. The layout of this spreadsheet illustrates that a single response can receive no code, one code, or multiple codes, since the categories are not independent or mutually exclusive. We offer this as an example tool that makes it easy for us to calculate percentages of responses that evidence each resource and to calculate interrater agreement, but others need not use the same tool. We have investigated but not yet used qualitative data analysis software such as MaxQDA or NVivo.

Coders use these spreadsheets independently and only combine after finishing their pass through the data. After both coders have coded all 30 (or more) responses, the two coders discuss their results. In some cases, coders use the spreadsheet to identify where they disagree and choose only to discuss those responses. In other cases, coders go through the responses one by one and discuss how they made the coding assignments they did.

A central practice within step 4 is finding edge cases—places of disagreement that are difficult to resolve, or responses that are hard to agree on even with discussion—and quintessential cases—responses that both coders agree confidently are good examples of the code. These cases tend to help us clarify and refine the coding scheme and come to some agreement about the breadth of responses that we would “count” as a particular resource. During this process, we often refine the wording of codes or resources, and in some cases we collapse two codes into a single one. We also try to come to consensus about how we will code edge cases that we know are common.

Our choices within step 4—having two independent coders, using a subset of the data to clarify and refine the coding scheme, etc.—are largely shaped by our “accountable to the discipline” value. These choices reflect recommendations for qualitative research in education [51,56] and practices of concept-focused, difficulties- and misconceptions-oriented research in PER [2,3].

Step 5: Reach sufficient percentage agreement between coders.—Standard statistical measures of interrater agreement, such as Cohen’s κ, require that codes are independent or mutually exclusive [57]. The coding schemes produced by our methods, where it is possible for a single response to be assigned multiple (or no) resource codes, reflect the theoretical stance that multiple resources can be activated as students reason about a physical phenomenon. Thus, codes are not independent and a single response can be assigned multiple resource codes. For this reason, standard statistical methods are rarely appropriate for this coding method. As a measure of percentage agreement, we take the normalized difference between the total number of possible codes and the total number of disagreements between the two coders in step 4:

$$\%\ \text{agreement} = \frac{(n_{\text{possible codes}})(n_{\text{coded responses}}) - (n_{\text{total disagreements}})}{(n_{\text{possible codes}})(n_{\text{coded responses}})}.$$

If the percentage agreement is less than 85%, the two coders recode the 30 responses, plus 10 additional responses, chosen in the same way as the original 30 (across questions and samples). We again discuss, as in step 4, and repeat a maximum of 2 times to try to reach 85% agreement. When 85% agreement has been reached, we move to step 6. (If 85% agreement cannot be reached after two additional tries, it likely means that we need to go back to earlier steps.) We chose 85% agreement as our threshold because it is typical for about 15% of student responses to be idiosyncratic and not well matched to our coding schemes.

Step 6: Independently code the full dataset for the study. This dataset usually includes three questions about a particular topic, for 2–3 college or university samples each. By “independently code” we mean that coders should not share coding workbooks or discuss how they are coding individual responses and should mostly not discuss as they code. We allow conversation if a coder identifies a recurring theme that is not captured by previous discussions (e.g., they are noticing a new resource that they would like to see included, or a recurring idea that lives at the boundary of two codes). When coders do engage in conversation, that conversation stays general, and coders do not pull up and discuss specific responses. Having applied these methods multiple times now, we are beginning to allow a single coder to code the dataset after 85% agreement has been reached in step 5.

Some of the practices that we have found useful for coding large datasets are coding one question at a time; highlighting responses that are especially good (quintessential) examples of a resource or especially confusing or edge examples as we code, in the code book; and keeping our code books sorted in a way that makes it possible to observe differences in frequency that depend on context (e.g., by keeping institutional datasets and questions separate from one another—different tabs, bolded horizontal lines, etc.).

Step 7: Calculate percentage agreement for the fully coded dataset.—We report this number in the methods sections of our publications. When we have two coders, we also choose to report a conservative estimate of the frequency of responses that we consider to be evidence of a particular resource. In particular, for a given resource, we report the percentage of responses for which there is full agreement—i.e., if we say that 75% of responses use resource A, we mean that both coder 1 and coder 2 identified those 75% of responses as resource A. (Coder 1 may have identified an additional 4% of responses as resource A, but these were not also coded as such by coder 2 and so are left out of our final percentage.)

Resources identified in previous studies by our team

Energy (Sabo, Goodhew, & Robertson, 2016)
- Students account for energy transfers and transformations in a scenario.
- Students associate (i) forms of energy with indicators and (ii) changes in energy with indicators of change.
- Students relate energy to forces and/or work.
- Students implicitly use the second law of thermodynamics.
- Students quantitatively represent energy scenarios.

Mechanical wave propagation (Goodhew et al. 2019)
- Properties of the medium either impede or facilitate the motion of the pulse.
- The speed or duration of transverse motion affects pulse speed.
- The speed of the pulse is affected by its energy.

Superposition (Bauman, Goodhew, & Robertson, 2019)
- Localization: Superposed effects must be co-located.
- Independence: Effects must be independent of each other.
- Quantifiability: Superposed effects can be described quantitatively.
- Additiveness: Effects add together.

Kinematics (Broadfoot, et al., 2020)
- Acceleration is present when there is a change in the velocity of an object.
- The magnitude of acceleration is related to the magnitude of the change in velocity.
- The directional relationship between an object's velocity and acceleration determines what is observed about its speed.
- Gravity changes the speed of objects.

Forces (Robertson et al. 2021)
- Moving objects keep moving.
- Forces influence the motion of objects.
- Imbalanced forces change the motion of objects (and balanced forces do not).
- It takes more effort to overcome a given force than to match it.
- It takes more effort to overcome a bigger (net) force (and less effort to overcome a smaller one).
- It takes more effort to change the motion of an object than to sustain it.

Linear momentum (Hansen et al. 2021)
- Conservation: Momentum is conserved.
- Direction: Momentum has a direction.
- Collisions: The type of collision an object undergoes matters.
- Properties: The properties of an object matter.

Heat and temperature: macroscopic (Abraham et al. 2021)
- Heat transfer is directional
- An object's physical properties matter in thermal processes.
- Hotter objects have more energy.

Heat and temperature: microscopic (Alesandrini et al. 2022)
- Differences will eventually even out.
- Macroscopic changes connect to microscopic collisions.
- When something is hotter (colder), its molecules are moving faster (slower).

Circuits (Bauman, et al., under review)
- Current is responsive.
- Voltage drives current.
- Resistance opposes current.
- The way elements are connected within the circuit matters.

FIG. 5. Resources identified in previous work from our research team, by topic. Image description: A series of nine underlined phrases (all topics in physics, followed by journal citations) and bullet-pointed lists to accompany each one. The title of the box is "Resources identified in previous studies by our team." The first underlined topic is energy, with an accompanying citation of Sabo, Goodhew, & Robertson, 2016. Under this topic are five bullet-pointed resources: "Students account for energy transfers and transformations in a scenario"; "Students associate (i) forms of energy with indicators and (ii) changes in energy with indicators of change"; "Students relate energy to forces and/or work"; "Students implicitly use the second law of thermodynamics"; and "Students quantitatively represent energy scenarios." The second underlined topic is mechanical wave propagation, with an accompanying citation of Goodhew *et al.*, 2019. Under this topic are three bullet-pointed resources: "Properties of the medium either impede or facilitate the motion of the pulse"; "The speed or duration of transverse motion affects pulse speed"; and "The speed of the pulse is affected by its energy." The third underlined topic is superposition, accompanied by a citation to Bauman, Goodhew, & Robertson, 2019. Under this topic are four bullet-pointed resources: "Localization: Superposed effects must be co-located"; "Independence: Effects must be independent of each other"; "Quantifiability: Superposed effects can be described quantitatively"; and "Additiveness: Effects add together." The fourth underlined topic is kinematics, with an associated citation of Broadfoot *et al.*, 2020. Under this topic are four bullet-pointed resources: "Acceleration is present when there is a change in the velocity of an object"; "The magnitude of acceleration is related to the magnitude of the change in velocity"; "The directional relationship between an object's velocity and acceleration determines what is observed about its speed"; and "Gravity changes the speed of objects." The fifth underlined topic is forces, accompanied by the citation Robertson *et al.*, 2021. Under this topic are six bullet-pointed resources: "Moving objects keep moving"; "Forces influence the motion of objects"; "Imbalanced forces change the motion of objects (and balanced forces do not)"; "It takes more effort to overcome a given force than to match it"; "It takes more effort to overcome a bigger (net) force (and less effort to overcome a smaller one)"; and "It takes more effort to change the motion of an object than to sustain it." The sixth underlined topic is linear momentum, with the associated citation Hansen *et al.*, 2021. Under this topic are four bullet-pointed resources: "Conservation: Momentum is conserved"; "Direction: Momentum has a direction"; "Collisions: The type of collision an object undergoes matters"; and "Properties: The properties of an object matter." The seventh underlined topic is heat and temperature (macroscopic), with the associated citation Abraham *et al.*, 2021. Under this topic are three bullet-pointed resources: "Heat transfer is directional"; "An object's physical properties matter in thermal processes"; and "Hotter objects have more energy." The eighth underlined topic is heat and temperature (microscopic), with the associated citation Alesandrini *et al.*, 2022. Under this topic are three bullet-pointed resources: "Differences will eventually even out"; "Macroscopic changes connect to microscopic collisions"; and "When something is hotter (colder), its molecules are moving faster (slower)." The ninth and final underlined topic is circuits, with the citation Bauman *et al.*, under review. Under this topic are four bullet-pointed resources: "Current is responsive"; "Voltage drives current"; "Resistance opposes current"; and "The way elements are connected within the circuit matters".

This practice means that the percentage agreement we report in the methods section of our publications is about the reliable use of the coding scheme (where by "reliable" we mean "there is consistent agreement between coders in its application"), whereas the frequencies of particular resources we report represent 100% agreement between coders (increasing the "trustworthiness" of our reported frequencies in a model of generalizability that emphasizes recurrence).

If we find that percentage agreement in step 7 is at least 85%, we consider the analysis complete. If percentage agreement is less than 85%, we will consider other methods of reaching agreement, such as resolving disagreements through discussion.

Product of steps 1–7.—At the end of this series of steps, researchers have a set of common (coarse-grained) conceptual resources for understanding topic X, a series of finer-grained resources that highlight some of the variability in how these resources show up in different contexts and for different students, a coding workbook that makes it possible to identify example student responses, frequencies with which students use these resources in their responses to a set of questions, and a measure of coder agreement. In our case, the resources are shaped to be instructionally significant, and the process is meant to produce results that are trustworthy, theoretically and disciplinarily.

V. DISCUSSION

In this paper, we have described our method for identifying common conceptual resources for understanding physics, offering a step-by-step guide for researchers who may want to do something similar. Our methods are designed to produce lists of common conceptual resources that are instructionally significant, using techniques that are theoretically grounded and accountable to existing methods for identifying common student ideas in physics. We have used these methods to identify common conceptual resources for understanding energy, mechanical wave propagation, superposition, kinematics, forces, linear momentum, heat and temperature, and circuits [32–34,58–63]. Notably, the resources we identified in these studies (Fig. 5) are framed differently; this is in part an artifact of an evolving method (the studies span 2016 through forthcoming work), and in part a reflection of the diverse team that has worked on the project.

Our aim in sharing this method is to invite other researchers to participate in identifying common conceptual resources for understanding introductory physics topics—to build an *enterprise* that might collectively respond to the call cited in the Introduction: "to develop explicit accounts of student resources, to allow their exchange, review, and refinement" [27]. If we did this collectively, we might be able to systematically answer questions such as the following.

- Does a method of identifying common conceptual resources produce different insights than those documented in already-existing misconceptions- and difficulties-oriented studies? Does it produce insights that are different than what we would get by reinterpreting existing studies through a resources lens? (We think the answer to both questions is yes, as we have said elsewhere [32], but more systematic studies would lend additional credence to this position.)
- In what ways do insights from this kind of research shape instructional practice? Is this instructional practice effective according to traditional measures? Do *students* experience instruction differently when it is informed by research on common resources than when it is informed by research on common difficulties or misconceptions?
- How common are the ideas we report? How context dependent and in what ways?

In our view, one of the primary affordances of this method is its appropriateness for undergraduate researchers and/or researchers who are new to our team. The steps are well defined, and the primary practice is sense making about others' ideas and connecting those ideas to introductory physics learning objectives. We have found that the most difficult part of this work for most new team members is to choose to engage with the parts of students' responses that are generative, even when there are parts of those responses that are canonically incorrect, and we have also found that every undergraduate researcher who has decided to participate in our work has excelled. This is in part due to the collaborative structure of our team, where experienced (sometimes undergraduate) researchers partner with new members in every step of the method.

Finally, we want to call attention to the unsettledness and dynamicity of our methods and the language we are using to describe them. Though our methods have stabilized through repeated use, they continue to evolve as we learn from and with new team members and the broader PER community. Significantly, the language we use to describe our methods and to define what we mean by "resource" continues to shift—an artifact both of our deepening understandings of the literature and theory and of the ongoing negotiations of meaning in our field. We (the authors) do not yet have (confident) answers to questions like, "How do the resources we are identifying relate to p-prims? To conceptual difficulties?" or "What makes a question 'good' for eliciting conceptual resources? Epistemological resources? Mathematical or procedural resources? Are these questions the same or different, and how so?" We hope this paper can contribute to the articulation of questions like these, and to add to community conversation about eliciting and understanding student ideas in introductory physics.

ACKNOWLEDGMENTS

This work was supported in part by NSF Grants No. 1608510, No. 1608221, No. 1914603, and No. 1914572. All opinions are those of the authors and not the granting agency. We are grateful for the collaboration and contributions from current and past Conceptual Resources team members, including Yohannes M, Abraham, Cheyenne Broadfoot, Jonathan Corcoran, Paula R. L. Heron, Tra Huynh, Katie Marvin, Clausell Mathis, Rachel E. Scherr, Olin Sorby, Marcella Su, and Mikayla Valentin.

[1] R. J. Beichner, *An Introduction to Physics Education Research* (American Association of Physics, 2009), https://www.per-central.org/items/detail.cfm?ID=8806.

[2] J. L. Docktor and J. P. Mestre, Synthesis of discipline-based education research in physics, Phys. Rev. ST Phys. Educ. Res. **10**, 0202119 (2014).

[3] L. C. McDermott and E. F. Redish, Resource letter: PER-1: Physics education research, Am. J. Phys. **67**, 755 (1999).

[4] J. Clement, Students' preconceptions in introductory mechanics, Am. J. Phys. **50**, 66 (1982).

[5] R. Thornton and D. Sokoloff, Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula, Am. J. Phys. **66**, 338 (1998).

[6] D. Hestenes, M. Wells, and G. Swackhamer, Force Concept Inventory, Phys. Teach. **30**, 141 (1992).

[7] D. M. Watts, Some alternative views of energy, Phys. Educ. **18**, 213 (1983).

[8] J. Solomon, Messy, contradictory, and obstinately persistent: A study of children's out-of-school ideas about energy, Sch. Sci. Rev. **65**, 225 (1982).

[9] M. Kryjevskaia, M. R. Stetzer, and P. R. L. Heron, Student understanding of wave behavior at a boundary: The relationships among wavelength, propagation speed, and frequency, Am. J. Phys. **80**, 339 (2012).

[10] 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).

[11] L. McDermott, A view from physics, in *Toward a Scientific Practice of Science Education*, edited by M. Gardner, J. G. Greeno, F. Reif, A. H. Schoenfeld, A. diSessa, and E. Stage (Lawrence Erlbaum Associates, Hillsdale, NJ, 1990), pp. 3–30.

[12] R. Duit, Bibliography—Students' and teachers' conceptions and science education, http://www.ipn.uni-kiel.de/aktuell/stcse/, accessed 01/30/23.

[13] L. C. McDermott, P. S. Shaffer, and Physics Education Group at the University of Washington, *Tutorials in Introductory Physics*, Preliminary edition (Prentice-Hall College Division, 2011).

[14] E. Mazur, *Peer Instruction* (Pearson, 2013).

[15] D. R. Sokoloff and R. K. Thornton, *Interactive Lecture Demonstrations: Active Learning in Introductory Physics* (Wiley, New York, 2006).

[16] P. Engelhardt and R. Beichner, Students' understanding of direct current resistive electric circuits, Am. J. Phys. **72**, 98 (2004).

[17] L. Ding, R. Chabay, B. Sherwood, and R. Beichner, Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, Phys. Rev. ST Phys. Educ. Res. **2**, 010105 (2006).

[18] T. D. Cook, Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them, Educ. Eval. Policy Anal. **24**, 175 (2002).

[19] 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).

[20] M. McCloskey, Intuitive physics, Sci. Am. **248**, No. 4, 124 (1983).

[21] L. C. McDermott, Research on conceptual understanding in mechanics, Phys. Today **37**, No. 7, 24 (1984).

[22] 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 (Information Age Publishing, Inc., Greenwich, CT, 2005), pp. 89–119.

[23] A. A. diSessa, Toward an epistemology of physics, Cognit. Instr. **10**, 105 (1993).

[24] J. P. Smith III, A. A. diSessa, and J. Roschelle, Misconceptions reconceived: A constructivist analysis of knowledge in transition, J. Learn. Sci. **3**, 115 (1993).

[25] D. Hammer, Student resources for learning introductory physics, Am. J. Phys. **68**, S52 (2000).

[26] B. W. Harrer, Identifying productive resources in secondary school students' discourse about energy, Ph.D. thesis, University of Maine, 2013.

[27] D. E. Brown and D. Hammer, Conceptual change in physics, in *International Handbook of Research on Conceptual Change*, edited by S. Vosniadou (Taylor and Francis, Inc., New York and London, 2008), pp. 127–154.

[28] E. F. Redish, A theoretical framework for physics education research: Modeling student thinking, in *Proceedings of the Enrico Fermi Summer School Course CLVI*, edited by E. Redish and M. Vicentini (Italian Physical Society, 2004), https://en.sif.it/books/series/proceedings_fermi.

[29] D. Hammer, Misconceptions or p-prims: How may alternative perspectives of cognitive structure influence instructional perceptions and intentions?, J. Learn. Sci. **5**, 97 (1996).

[30] D. E. Young and D. C. Meredith, Using the resources framework to design, assess, and refine interventions on pressure in fluids, Phys. Rev. Phys. Educ. Res. **13**, 010125 2017).

[31] Y. Cao and B. Brizuela, High school students' representations and understandings of electric fields, Phys. Rev. Phys. Educ. Res. **12**, 020102 2016).
[32] A. D. Robertson, L. M. Goodhew, R. E. Scherr, and P. R. L. Heron, University student conceptual resources for understanding forces, Phys. Rev. Phys. Educ. Res. **17**, 010121 (2021).
[33] L. M. Goodhew, A. D. Robertson, P. R. L. Heron, and R. E. Scherr, Student resources for understanding mechanical wave propagation, Phys. Rev. Phys. Educ. Res. **15**, 020127 (2019).
[34] 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).
[35] B. Warren, C. Ballenger, M. Ogonowski, A. S. Rosebery, and J. Hudicourt-Barnes, Rethinking diversity in learning science: The logic of everyday sense-making, J. Res. Sci. Teach. **38**, 529 (2001).
[36] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.19.020138 for series of activities where researchers can try out our methods on a new set of data.
[37] J. Tuminaro and E. F. Redish, Elements of a cognitive model of physics problem solving: Epistemic games, Phys. Rev. ST Phys. Educ. Res. **3**, 020101 2007).
[38] D. Hammer and A. Elby, Tapping epistemological resources for learning physics, J. Learn. Sci. **12**, 53 (2003).
[39] M. C. Wittmann and K. E. Black, Mathematical actions as procedural resources: An example from the separation of variables, Phys. Rev. ST Phys. Educ. Res. **11**, 020114 2015).
[40] R. E. Scherr, Modeling student thinking: An example from special relativity, Am. J. Phys. **75**, 272 (2007).
[41] J. Minstrell, Facets of students' knowledge, and relevant instruction, in *Research in Physics Learning: Theoretical Issues, and Empirical Studies*, edited by R. Duit, F. Goldberg, and H. Niedderer (Institute for Science Education, Kiel, Germany, 1992), pp. 110–128.
[42] B. Sherin and J. R. Star, Reflections on the study of teacher noticing, in *Mathematics Teacher Noticing: Seeing through Teachers' Eyes*, edited by M. G. Sherin, V. R. Jacobs, and R. A. Philipp (Routledge, New York, 2011), pp. 66–78.
[43] W. James, *The Principles of Psychology* (Henry Holt and Company, New York, 1890).
[44] F. Erickson, Some thoughts on 'proximal' formative assessment of student learning, Yearhb. Natl. Soc. Study Educ. **106**, 186 (2007).
[45] P. Black and D. Wiliam, Assessment and classroom learning, Assess. Educ. **5**, 7 (1998).
[46] L. M. Goodhew, A. D. Robertson, and L. C. Bauman, Designing questions to probe students' scientific explanations, Phys. Teach. (to be published).
[47] E. Etkina, D. T. Brookes, and G. Planinsic, *Investigative Science Learning Environment: When Learning Physics Mirrors Doing Physics* (Morgan & Claypool Publishers, San Rafael, CA, 2019).
[48] D. Hammer and E. van Zee, *Seeing the Science in Children's Thinking: Case Studies of Student Inquiry in Physical Science* (Heinemann, Portsmouth, NH, 2006).
[49] D. Hammer, F. Goldberg, and S. Fargason, Responsive teaching and the beginnings of energy in a third grade classroom, Rev. Sci. Math. ICT Educ. **6**, 51 (2012).
[50] D. Hammer, Discovery learning and discovery teaching, Cognit. Instr. **15**, 485 (1997).
[51] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed. (SAGE Publications, Inc., Thousand Oaks, CA, 1994).
[52] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology* (Sage, Thousand Oaks, CA, 2013), 3rd ed.
[53] L. M. Goodhew, A. D. Robertson, P. R. L. Heron, and R. E. Scherr, Students' context-sensitive use of conceptual resources: A pattern across different styles of question about mechanical waves, Phys. Rev. Phys. Educ. Res. **17**, 010137 (2021).
[54] S. Kanim and X. C. Cid, Demographics of physics education research, Phys. Rev. Phys. Educ. Res. **16**, 020106 (2020).
[55] R. Mondesir and A. D. Robertson, Towards characterizing the demographics of introductory physics courses, in *Proceedings of the 2020 Physics Education Research Conference*, edited by S. Wolf, E. Bennett, and B. Frank (2020).
[56] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology* (SAGE Publications, Inc., Los Angeles, 2019).
[57] M. Banerjee, L. Capozzoli, and D. Sinha, Beyond kappa: A review of interrater agreement measures, Can. J. Stat. **27**, 3 (1999).
[58] L. C. Bauman, L. M. Goodhew, and A. D. Robertson, Students' use of conceptual resources for understanding superposition, in *Proceedings of the 2019 Physics Education Research Conference*, edited by Y. Cao, S. F. Wolf, and M. B. Bennett (2020), pp. 38–43.
[59] C. Broadfoot, B. Hansen, and A. D. Robertson, Identifying student resources for understanding kinematics, in *Proceedings of the 2020 Physics Education Research Conference* (Ref. [53]), pp. 57–62.
[60] B. Hansen, L. C. Bauman, Y. A. Abraham, M. Valentin, and A. D. Robertson, Identifying student resources for understanding linear momentum, in *Proceedings of the 2021 Physics Education Research Conference*, edited by M. B. Bennett, B. W. Frank, and R. Vieyra (2021).
[61] Y. A. Abraham, M. Valentin, B. Hansen, L. C. Bauman, and A. D. Robertson, Exploring student conceptual resources for understanding heat and temperature, in *Proceedings of the 2021 Physics Education Research Conference*, edited by S. Wolf, E. Bennett, and B. Frank (2021).
[62] L. C. Bauman, B. Hansen, L. M. Goodhew, and A. D. Robertson, Student conceptual resources for understanding circuits, Phys. Rev. Phys. Educ. Res. (to be published).
[63] A. T. Alesandrini, T. Huynh, L. C. Bauman, and A. D. Robertson, Identifying student resources for reasoning microscopically about heat and temperature, in *Proceedings of the 2022 Physics Education Research Conference* (2022), pp. 37–43.