

Student epistemological framing on paper-based assessments

Kelli Shar 

Department of Chemistry, Biochemistry, and Physics, University of Tampa, Tampa, Florida 33606, USA

Rosemary S. Russ

Department of Curriculum and Instruction, University of Wisconsin, Madison, Wisconsin 53706, USA

James T. Lavery *

Department of Physics, Kansas State University, Manhattan, Kansas 66506, USA



(Received 25 September 2019; accepted 19 June 2020; published 7 July 2020)

Assessments are usually thought of as ways for instructors to get information from students. In this work, we flip this perspective and explore how assessments communicate information to students. Specifically, we consider how assessments may provide information about what faculty and/or researchers think it means to know and do physics, i.e., their epistemologies. Using data from students completing assessment questions during one-on-one think aloud interviews, we explore how assessment features did (or did not) impact student engagement with the assessment problems. We analyze video recordings and transcripts to infer the epistemological framings and resources students use while completing introductory-level physics problems. Students' framings tended to be fairly stable, but when shifts occurred, they were triggered by a shift in epistemological resource, which can be activated by assessment feature. This work extends existing work on epistemological framing into the realm of assessment and allows us to consider the effects of assessments on our students' understanding of physics teaching and learning.

DOI: [10.1103/PhysRevPhysEducRes.16.020101](https://doi.org/10.1103/PhysRevPhysEducRes.16.020101)

I. INTRODUCTION

As physics educators at undergraduate institutions, we are all aware of the importance of assessment in our classes. Individual assessments help us and our students understand whether students learned the content and skills we painstakingly taught them. Additionally, we assume that how a student performs on assessments throughout the semester helps us evaluate and track student progress. Students in introductory physics frequently take multiple, high-stakes assessments or exams each semester along with weekly homework assignments and lab reports. Assessment occurs often and in a variety of ways.

Consider a common assessment item on introductory physics exams—a numerical problem that requires calculation to solve. Imagine a question such as

Diego is standing on a scale in an elevator and the elevator starts to accelerate upwards at 3 m/s^2 . If Diego

weighs 71 kg on his scale at home, how much will the scale read while the elevator is accelerating?

We suspect this type of question is familiar to many of our readers. We suspect the following solution is also familiar.

$$\begin{aligned} F &= ma \\ F_{\text{gravity}} &= ma_{\text{elevator}} \\ mg &= ma_{\text{elevator}} \\ (71)(9.8) &= m(3) \\ 232 \text{ kg} & \\ \text{Units match!} & \end{aligned}$$

As instructors, we might find ourselves horrified by this student's solution. It is definitely wrong. They may have matched units but they entirely forgot the normal force. Many of us can relate to seeing such solutions on our exams and being confused that the students did not learn what we taught them.

In these moments, our default is often to become frustrated—either with students for not learning enough content to solve the problem or with ourselves for not teaching it well enough. That is, we attribute their lack of performance either to them or to our teaching. In this paper, we suggest that there may be a mechanism other than poor content knowledge that accounts for what looks like “poor” performance.

*Corresponding author.
laverty@ksu.edu

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.

Specifically, if we look closer, we can understand what the student is doing here as a type of pattern matching, or numerical plug and chug using formulas she is familiar with [1]. She is solving the problem using the symbols and mathematical formalisms we use in class and teach students to use. She has found the force of gravity and appropriately substituted in the acceleration of the elevator. In that way, she is solving the problem in a way that looks like other problems she has done.

Viewing student performance through this lens assumes that students are working to make sense of the assessment *in the terms they think we want*. The logic is something like: “I have seen my teacher use this formula in problems like this, so I will do the same thing here.” This perspective on student engagement in assessment removes the assumption that students are being foolish when they do things that appear outlandish to experts. Instead, it assumes they are trying to do what we have asked them—even taught them—to do.

In this paper, we explore this perspective to move away from our traditionally deficit-oriented model of assessment. Specifically, using the theoretical machinery of epistemological framing and resources, we seek to understand student performance as a reasonable—rather than misguided—interpretation of what assessment writers have asked them to do. We draw on data from think-aloud interviews with undergraduate students in introductory physics to (a) explore how students frame assessment, (b) unpack the knowledge resources that underlie those frames, and (c) document stabilities and dynamics in resources and frames within the context of assessment items. Specifically, we describe how assessments may send students messages about what types of knowledge and knowledge production activities are appropriate to use in the assessment context. We discuss implications for these findings for educational research and assessment practice.

II. LITERATURE REVIEW

A. Epistemology

Within psychology and education, researchers have long been interested in how people understand the nature of knowledge and learning [2]. These understandings are referred to as a person’s epistemology. A person’s epistemology involves their knowledge and beliefs about the nature of knowledge itself and how knowledge is built and evaluated. Since the 1950s, scholars have proposed a variety of dimensions of epistemology. For example, epistemological knowledge includes knowledge about the goal or aims of knowledge construction [3], how knowledge is structured [4], how knowledge is justified [5], and the appropriate activities to use in constructing knowledge [6]. Scholars have explored personal epistemologies, which describe how people view their own knowledge and learning, and scientific epistemologies

which describe how professional scientists construct knowledge and learning [7].

Much of the work understanding epistemology has taken place in educational settings [8]. The focus on educational settings arises because of the particular importance of epistemology for learning. Specifically, there is the assumption that “epistemological premises are a part of and an influence on the cognitive processes of thinking and learning [8]. In his cognitive model of learning physics specifically, Redish calls epistemology a control structure in that it “interact[s] strongly with (and often controls) the [knowledge] resources students have for creating knowledge” [9] (p. 30). That is, a person’s epistemology can impact the way they engage in learning by dictating the kinds of knowledge they use and the ways they use it.

Of particular importance for us in this work is the finding that what has often been labeled learning “difficulties” [10–12] may instead “stem in part from [...] epistemology” [13]. For example, Lising and Elby present the case of a student Jan, who, despite possessing all the knowledge and skills needed to make sense of a physics tutorial, does not do so because her epistemology, or her understanding of what knowledge and knowledge building she is supposed to do in physics class, “gets in the way” [13] (p. 381). Epistemology mediates content knowledge in introductory physics.

Despite substantial interest in epistemology in physics learning since Hammer’s initial introduction of it [14], we do not yet know of any who have directly examined epistemology in the context of assessment in undergraduate physics. The work has been constrained mostly to classroom or classroomlike contexts. In our work, we explore the hypothesis that the same mechanism at work in the case of Jan—the same “epistemological effect” [13]—might be at play in assessments. Specifically, we explore what epistemologies are active when students use their content knowledge to complete assessments.

B. Assessment design

What types of assessments do our undergraduate students typically engage in? Assessment is a central pillar of our current education system and is often divided into two types: formative and summative. Here, we will use the definition of formative assessments as “intended to provide feedback to the system to inform next steps for learning” and summative assessments as measures “of individual achievement” [15]. Both of these definitions focus on assessments as ways to get information about what students know. In this paper, we focus on summative assessments. Beyond these two definitions, current theories of summative assessment design focus explicitly on how we get information about students’ current understanding.

The “assessment triangle,” described in *Knowing What Students Know*, has been central to assessment development for nearly twenty years. It focuses on three

interconnected models for developing assessments: cognition, interpretation, and observation [15]. Taken together, these three models (aka, the assessment triangle) are designed to treat summative assessment as an evidentiary argument, focused on designing tasks that allow the instructor or researcher to gather evidence to make claims about student knowledge.

More recent examples of assessment design theories build on the assessment triangle model. Approaches such as evidence-centered design [16–20] and the BEAR assessment system [21] also focus on obtaining evidence to support claims of student knowledge. Both of these approaches have been highlighted as promising ways to assess the Next Generation Science Standards [22,23].

The *Standards for Educational and Psychological Testing* states, “Test development is the process of producing a measure of some aspect of an individual’s knowledge...” [24]. In physics in particular, Adams and Wieman have argued that the development of concept inventories in PER typically follows the steps outlined by this document [25].

In each of these approaches to designing assessments, the focus is on obtaining information about student knowledge and little to no attention is paid to the messages that these assessments send to students. Assessments developed in these ways are assumed to be measurement instruments to get data about the knowledge of students in the same way a thermometer is a measurement instrument to get data about the temperature of a water bath; that is, that the measurement does not affect the system (or affects it minimally). Given this state of assessment design theories, we suggest that instructors and other individuals designing tests for physics courses probably do not think about the messages being sent to students either.

C. Assessments in PER

The history of PER includes the development of many standardized assessments. As of this writing, the website Physport currently lists 93 research-based assessments, divided into 6 categories: content knowledge (63), problem solving (2), scientific reasoning (2), lab skills (6), beliefs and attitudes (14), interactive teaching (6) [26]. These assessments provide straightforward, off-the-shelf ways to evaluate student learning [27]. Because of this, they have been used to evaluate different learning environments, instructional strategies, and curricula (among other things) [28–35].

Within the assessment culture of undergraduate physics, most of the assessments are focused exclusively on evaluating student content knowledge. Further, many do so in a way that tacitly neglects students’ understanding of either physics generally or the assessment in particular. There are a few assessments that attempt to directly measure student epistemologies. For example, Redish, Saul, and Steinberg declare, “[W]e describe the

Maryland Physics Expectations survey; a 34-item Likert-scale agree/disagree survey that probes student attitudes, beliefs, and assumptions about physics” [36] and Adams *et al.* start their abstract, “The Colorado Learning Attitudes about Science Survey (CLASS) is a new instrument designed to measure student beliefs about physics and about learning physics” [37]. However, we are not aware of any assessments that are designed to study or explicitly understand the connection between student epistemology and assessment of content knowledge.

III. THEORETICAL FRAMEWORK

As described above, researchers interested in epistemology have not yet examined the ways in which student understandings of knowledge and learning are present in their engagement in assessments. Similarly, researchers interested in assessment overlooked the ways in which assessments are understood from the perspective of knowledge and learning. Here, we turn our analytic attention to bridging the gap between those two literatures. In what follows, we describe our theoretical stance on epistemology and its implications for our research on assessments.

A. Epistemological framing

Understanding the ways in which epistemology interacts with student engagement in assessment requires a precise conceptualization of epistemology. When it was first conceptualized, epistemology was understood as a set of categories that students adopt and then progress through over the course of their lifetime. These categories applied to all aspects of their learning at any given age [2].

However, research shows significant “flexibility and variability in student reasoning” in K–16 science and physics classrooms [38]. As such, rather than a stable model of epistemology, we adopt a model of epistemology rooted in the sociolinguistic and anthropological construct of framing [39,40]. Framing is a person’s sense of “What is it that is going on here?” In recent work in physics education, scholars describe epistemological framing [41] as students’ answer to “How should I approach knowledge?” [38]. Epistemological framing, then, is the tacit stance students take toward learning-based activities [42].

A key feature of framing in general, and epistemological framing in particular, is that it is contextual and dynamic rather than stable across time and place [43]. For example, when a student enters a science class learning about electrostatics, she likely thinks very differently about knowledge than when she is in a discussion with her friends about what pizza to order. Even more, she is likely to think differently about learning during a portion of science class that is a lecture versus a small group discussion [44]. Even within small group discussion, students’ sense of what knowledge and knowledge building activities should be used can shift dramatically [41]. This

contextuality means that framing is highly dynamic. Because people shift their understandings of knowledge in different contexts, framing must also change over the timescale of contextual change (hours, minutes, and seconds) rather than over the scale of a lifetime.

Existing research on framing in undergraduate physics education suggests that students adopt a variety of frames when engaged in learning physics. For example, several scholars have explored the sensemaking frame in which students reason about physical phenomena by constructing an explanation and then filling in a gap in that explanation [38,42,45].

Bing and Redish [46] identify four common epistemological framings students adopt during physics problem solving: Calculation, physical mapping, invoking authority, and math consistency. In their work exploring framing in quantum mechanics, Modir, Thompson, and Sayre [47] describe a set of frames that differ along two dimensions—whether students draw on mathematics or physics, and whether they are engaged in algorithmic or conceptual thinking. Each of these framings involves different—though not necessarily better or worse—understandings of knowledge and knowledge construction in physics class.

In addition to identifying framings that are prevalent in physics, scholars have also focused on whether and how students move between framings. Hammer and colleagues describe transitions between mathematical manipulation and intuitive sensemaking that are both short lived and lasting [38]. Bing and Redish identified their four frames by explicitly looking for and unpacking shifts [46] and have suggested that the frequency and fluency at moving between framings is part of becoming and expert [48]. This work highlights the dynamic nature of epistemological framing.

The theory of epistemological framing suggests the need to refine our question of interest even further. Specifically, this framework suggests that students may not adopt a single epistemology during assessment. Instead, they may transition between multiple framings. As such, we now ask: How do student epistemological framings influence their engagement in assessment tasks?

B. Epistemological resources

As we have described, existing scholarship within physics education highlights the dynamics of framing. To examine and document those dynamics, much of the literature focused on identifying observable shifts in behaviors. Scherr and Hammer [41] pioneered focus on behaviors with their careful analysis of students' verbal, nonverbal, and para-verbal behaviors in small group interactions. They describe how “different behavioral clusters are evidence of—and in dynamic interaction with—student epistemologies” (p. 148). This finding led to an explosion of work in PER that identified behavioral clusters and their associated framings.

For example, consider more closely the work of Modir, Thompson, and Sayre who identified four epistemological framings in upper level student problem solving [47]. To identify framings, they describe how they “reflected on [...] episodes [of student activity], seeking to answer ‘what’s going on?’ for each of them [...] we sought to capture changes in students discussion or behavior that might indicate a shift in the students problem solving processes” (p. 020108). Here, they focus on the behaviors associated with their framing.

This focus on behavior makes sense and has given the field substantial traction in identifying moments when epistemological framing shapes student engagement and learning. However, a feature of the theory of epistemological framing that is commonly left out of the literature in physics education is that framings are local collections of epistemological resources [38,49]. Specifically, our model of epistemology—grounded in the work of Hammer and Elby [6]—assumes that rather than being unitary entities that stably exist as a coherent unit, framings are more like networks of many smaller elements that are all activated together in context [38].

The small elements that make up epistemological frames (or what Rosenberg, Hammer, and Phelan call “multiple local coherences”) are epistemological resources [49]. These resources, which are finer-grained elements of knowledge, are assembled anew in each context to create a person’s in-the-moment epistemology. This finer-grained structure affords contextuality in a person’s epistemology.

A number of potential epistemological resources have been proposed that govern student behavior in knowledge-production contexts. Some deal with the nature of knowledge itself: the form of the knowledge product [50] or the source of the knowledge [6]. Others deal with how knowledge is produced or constructed: knowledge production activities [44] or the goals of those activities [3].

Our model, as drawn from Hammer, Elby, Redish, and Scherr, for the relationship between epistemological resources and frames is that frames are local coherences of epistemological resources [38]. This means that, when we identify frames by behaviors, we can expect to find a fairly stable set of resources associated with that frame. Alternatively, if we see a group of resources reliably coming up together, we can expect there to be a set of associated behaviors, or frames, each time. Some resources are likely to be more central to the frame [49]. Others may be reliant on those more central resources. In our work, we focus specifically on the resources other empirical work has found relevant in learning science and physics [44,49].

In our work, we hypothesize that understanding the specific resources that make up the framings will give us more insight into how and why students engage in learning in particular ways than merely identifying their epistemological framings as a whole. This assumption follows the work of Rosenberg, Hammer, and Phelan [49] and Russ

and Luna [44] who each go an analytic level below framings to the level of resources. Specifically, we apply this assumption to the context of assessments and ask, How are the epistemological resources that make up student epistemological framings evident in their (potentially dynamic) engagement in assessment tasks?

IV. RESEARCH QUESTIONS

Our everyday experience as physics instructors leads us to question why students engage with assessments in the way they do. Further, our underlying asset-oriented perspective toward student learning leads us to reject the existing deficit-based explanations in the literature [51]. As such, we began this research with the question: Why might students reasonably engage with assessments in the way that they do?

As a result of our theoretical framework, we refine and extend our question:

1. What epistemological resources make up the framings students adopt during assessments?
2. How do these epistemological resources influence their dynamic engagement in assessment tasks?

Specifically, this work seeks to understand whether existing analyses and assumptions of how students dynamically bring different forms of knowledge into physics learning can be extended into the realm of assessment.

V. METHODS

A. Participants

In this article, we analyze data collected as part of a different study on how to assess scientific practices in introductory physics courses [52]. The participants in this study were predominantly engineering majors enrolled in the first or second semester of a calculus-based, introductory level physics course. The interviews were conducted near the end of the semester. Students volunteered to take part in the study and were compensated for their time (equivalent to \$20). No preference was given to their performance in the course. Pseudonyms are used throughout the paper. The study collected data from two groups of ten students, with each group of ten receiving a separate exam.

B. Data collection

The data were collected as individual think-aloud interviews designed to replicate the context of a summative assessment. To do so, interviewers printed each list of questions as a formatted test, had subjects write their names at the top, answered only clarifying questions, and told students to complete the questions “as a test” with no permission needed to move on to another question. Each interview took each student 45 to 90 min to complete. The students had unlimited time to complete the questions.

The questions on the tests were developed using the 3D-LAP to elicit evidence of students abilities to engage in the scientific practices [53]. One of the exams focused on the scientific practice of using mathematics and the other focused on developing and using models. Students solved physics problems on a variety of first-semester physics subjects. The exam included both selected and constructed response items. Only the constructed response items were used in this analysis.

C. Data selection

From the recorded interviews, there were two sets of 10 students. One student from each group of 10 did not have usable audio, resulting in 9 recorded interviews for both problem sets. The assessment given to the first group of 9 students featured 3 constructed response questions. The assessment given to the second group of 9 students featured 5 constructed response questions. This yields a total of 72 instances of students answering a constructed response question. Of these 72 responses, 5 were not used. Reasons for not using a response included the student not attempting that specific problem or stating that they have seen the problem before and know the answer (and, thus, do not engage in the problem solving process). Below is an excerpt from Mark.

Mark: Alright, so we had this exact problem on our test last year.

Here, Mark states he knew the answer and goes on to repeat the answer he remembered from his exam. Because we were specifically interested in how students engaged with the features of this specific assessment question, we decided not to use his response or similar responses where students do not engage in problem solving.

The examples we use throughout the paper come from specific assessment questions we will refer to as the car problem and the ferris wheel problem, shown in Figs. 1 and 2, respectively. We also transcribed the audio from what we will refer to as the Gravitron problem, which we chose as an example of students ardently staying in their problem solving strategy. The full Gravitron question is shown in Fig. 3. Although we largely reference those problems in the cases we present, all constructed response questions were analyzed via the process described in the data analysis section.

D. Data analysis

After reducing our dataset to only instances where students engaged in problem solving with usable audio, we were left with 67 attempts at solving a physics problem. Each instance corresponds to a specific student solving a specific problem. Of the remaining 67 instances, all were coded for epistemological frames. A sample of 20 instances were transcribed and coded for epistemological resources. After coding each full sentence expressed by the subject,

4-A. The Ford manufacturers are testing their new version of the Ford Focus before releasing it to the market. The car has a mechanism for calculating its velocity, acceleration and position and displaying those values on the dashboard. The manufacturers want to test whether this feature is functioning as intended. The case study is, the car will be travelling at a speed of 5.2 m/s and after 2 s, it starts accelerating at a constant rate of 2.2 m/s² until it reaches maximum speed of 16.6 m/s. If the dashboard readings for the velocities are 5.2, 5.2, 7.4, 9.6, 11.8 (in m/s) in each second (until 5 s), is the feature working correctly?

4-B. If the dashboard readings for the positions are 5.2, 10.4, 16.7, 25.2, 35.9 (in m) in each second (until 5 s), is the feature working correctly? Assume position is considered to be zero at time zero (0 s).

FIG. 1. The car problem.

6. Consider a Ferris wheel in an amusement park in California. A Ferris wheel is a large circular machine with seats attached to the rim of it. The seats can freely rotate so that when the Ferris wheel is spinning, the seats hang downwards at all times. Assume the wheel is rotating with angular velocity ω and the diameter of the wheel is D . At what point in the motion does a rider feel “heaviest” and “lightest”?

FIG. 2. The ferris wheel problem.

7. You are asked to design a Gravitron for the county fair, an amusement park ride where the rider enters a hollow cylinder, radius of 4.6 m, the rider leans against the wall and the room spins until it reaches angular velocity, at which point the floor lowers. The coefficient of static friction is 0.2. You need this ride to sustain mass between 25-160 kg to be able to ride safely and not slide off the wall. If the minimum ω is 3 rad/s will anyone slide down and off the wall at these masses? Explain your reasoning using diagrams, equations, and words.

FIG. 3. The Gravitron problem.

we then compared when frame shifts occurred to when shifts in individual resource shifts occurred. In this section, we describe how we coded the data and offer an example of coding one instance.

1. Coding epistemological frames

For this particular project, we opted to use epistemological frames that had already been identified in the literature. In particular, we wanted to use frames that could be identified by looking at behaviors, a process originally described by Scherr and Hammer [41]. We chose the frames described in Chari *et al.* [54]. We selected this set of frames and behaviors primarily because they were already identified and because they mapped well onto our data. Chari *et al.* used their frames to analyze group problem-solving interviews, but the language used to define the frames does not limit them to group or individual work. Below are the characteristics and behaviors associated with each frame as described in Chari *et al.* [54].

Conceptual physics: Students and instructors are in this frame when they discuss physics scenarios and phenomenon, about properties of physics quantities related to the task at hand. They may also exploit

the symmetry of a physical system by investigating related concepts.

Algorithmic physics: In the algorithmic physics frame, students recall physics equations or apply physics knowledge to rearrange known equations using math. Students may also derive expressions for specific cases from a general physics equation or validate an expression via dimensional analysis.

Algorithmic math: This frame refers to performing mathematical computation by following well-established protocols without questioning the validity of those protocols, e.g., solving an equation or computing an integral.

Conceptual math: Students are in this frame when they exploit properties of mathematical constructs to quickly obtain a result without diving into algorithmic manipulation; e.g.; noticing that all the odd terms in a sum are equal to zero.

Using these frames and definitions, we analyzed the recorded video interviews using a top-down coding scheme to identify frames and noted each time a student switched from one set of behaviors to another and identified this as a frame shift.

2. Coding epistemological resources

As we were specifically interested in what causes a frame shift, we transcribed 20 student solutions: the three examples we found of a clear frame shift, and all responses to two questions where almost all students solved the problem in the same frame. These were coded for resources to investigate what prompted students to stay in this frame.

Next, we coded all transcribed attempts for epistemological resources. Each full sentence of the transcribed interview was sectioned, and we inferred epistemic resources according to definitions given in Table I. We grouped by each full sentence because that was as large as we could go while still determining a resource shift and as small as

we could go while still having the context to determine a resource.

To code the resources, we used categories defined in Hammer and Elby's paper, *On the form of a personal epistemology* [6]. We specifically focused on nature of knowledge, source of knowledge, epistemic activity, and epistemic source because we found clear indicators of these in the words and actions of the students.

The individual resources we used also came from Hammer and Elby (top-down coding). In addition to repurposing codes from the literature on epistemology, we also generated codes from the data in an emergent fashion (bottom-up coding) [55]. These codes are equation

TABLE I. Definitions used for each epistemological resource.

Resource type	Resource definitions	Example from data
<i>Knowledge type</i> If you asked the student how they know this, they will say...	<i>Propagated</i> ... because someone told me.	Okay so centripetal force is going to be equal to $r w$ squared
	<i>Fabricated</i> ...because I figured it out using my prior knowledge.	Your potential energy is going to be the highest when you're at the top, lowest at the bottom, so I think those are going to be the points at which you're going to be feeling heaviest
	<i>Directly perceived</i> ...because it is apparent or obvious.	The radius is 4.6
	<i>Intuitive</i> ...because I feel like this is true.	Um, like any equation doesn't seem to give me any intuition
<i>Knowledge source</i> If you asked the student where they got this knowledge from, they would say...	<i>Equation sheet</i> ...from my equation sheet	Okay, so, we look at our rotational kinematics equations
	<i>Physical experience</i> ...from my memory of an experience I had or observed	So that's based off my life experiences, not any equations
	<i>Authority</i> ...from something my professor or textbook told me.	I'm trying to recall what the, uh, like a similar problem we did
	<i>Memory</i> ...from my memory of solving another problem.	So I know there will be a friction force which is going to be equal to $m g \mu$
<i>Epistemic activity</i> The student obtained this knowledge by...	<i>Accumulating</i> ...gathering information from external sources.	So this is just off to the side, I just kind of like to write down some of the stuff for later use
	<i>Forming</i> ...creating new knowledge about the specific problem from knowledge they already had.	I guess all that matters is the direction of the force
	<i>Checking</i> ...reconciling their thoughts with another piece of information.	If the person weighs more, it's going to be easier for them to drop straight down and harder for them to go... Is that right? (Checks equation sheet)
	<i>Causal storytelling</i> ...attributing a cause to each effect.	Alright, so this force has to keep them from falling down
<i>Epistemic form</i> The student believes the answer to their question will be indicated by...	<i>Number</i> ...a quantity.	Okay, we need to find alpha
	<i>Reasoning</i> ...an application of conceptual knowledge.	It doesn't tell you the mass. So, I guess because this is a conceptual question

sheet as a knowledge source; and number or reasoning as epistemic forms. The working definitions and an example from the data for each resource are listed in Table I. Only resources that appeared in the data are listed.

We did not limit epistemic form to the answer type requested by the assessment question. Instead, we decided there could be different epistemic forms possible for one question and for different portions of the problem-solving process. For example, suppose a multiple-choice question asks students to find an acceleration, given a mass and net force, and lists four different values for acceleration. To solve, students could plug values into Newton's second law and choose the acceleration closest to their answer. They could also reason through the options listed and pick the most probable value and look for the answer with the correct units. We believe these two students, although their ultimate answers were a letter, had a difference in epistemic form because they were working towards different ending conditions, a number or reasoning, that they later translated into a letter.

3. Interrater reliability

The first author coded all 67 transcripts for their frames and frame shifts. The second author again coded a subset (8) of the transcripts. The two coders agreed on all codes (100%), yielding perfect agreement.

Additionally, the first author coded each full sentence of all 20 transcripts for each of the four dimensions of epistemological resources. To check the reliability of this coding, the second author coded a 10% subset of the transcripts (2). We then calculated percent agreement across the doubly coded transcripts for each dimension. The percent agreement for each epistemological dimension (aka coding category) was greater than or equal to 80%. Specifically, agreement was 81%, 84%, 81%, and 98% for knowledge type, knowledge source, epistemic activity, and epistemic form, respectively, which is considered acceptable agreement for qualitative coding in research [56].

4. Example analysis

Here we include an example of how this analysis is done in practice. Figure 3, or "The Gravitron problem" features a problem that gives students the radius, coefficient of friction, and angular velocity of a Gravitron ride and asks them to determine if the ride is safe for riders within a range of masses. Spoiler alert: The safety of the rider is not dependent on their mass. Below is a transcript of Erica beginning to solve the problem:

Erica: Okay, so I'm going to get the rotational ones (picks up equation sheet, starts copying equations onto paper). Um, so you just want to find like your minimum 25 kilograms and your maximum 160 kilograms and then any answer you get between that is that going to, um, stay on the wall. (Looks at problem) Um, 3 radians

a second, so I'm just going to write that up here (writing on paper) radians per second (typing into calculator) Let me get that number 60 seconds for one minute and its 2 pi radians for one revolution (types into calculator) or something like that.

First, we used behaviors to determine the epistemic frame each student was working in. Erica starts the problem by rearranging known physics equations that she can use to input given values. She believes the output of her function will tell her whether some people will slide off the Gravitron. For these reasons, we determined that Erica began the problem in an algorithmic physics frame.

Second, we grouped coherent thoughts and coded each thought using the four categories of epistemic resources defined in Table I. Erica's first full sentence is "Okay, so I'm going to get the rotational ones." We coded her type of knowledge as propagated because the knowledge of the physics equations is communicated directly to Erica through the equation sheet. We coded her source of knowledge as her equation sheet. We coded her epistemic activity as accumulating because Erica is gathering knowledge from an external source. Erica's epistemic source is ambiguous here because she could be intending to use the equation to consider how the different quantities given affect each other and reason her way to an answer. We used additional context from later in the problem to determine that Erica intended to use the equation to input her given values and output a number that would indicate which riders would be safe, so we coded her epistemic form as number.

We repeated this process across all 20 solution attempts that we transcribed.

VI. RESULTS

A. Different frames are made up of different sets of resources

Existing literature examining student epistemologies describes how frames are made up of resources [38]. Specifically, frames are "local coherences" of resources [49]; that is to say groups of several different resources tend to co-occur and change together. For example, Russ and Luna [44] identified distinct sets of resources one teacher drew on when she engaged in behaviors from different framings. In one frame, she engaged in particular epistemological activities with particular epistemological goals and in another frame she used different activities and goals.

Our independent coding of frames by behaviors and resources by utterance allows us to look for local coherences of resources within and across frames identified by other researchers [54]. That is, we can engage in an analysis similar to that done by Russ and Luna [44] and Rosenberg, Hammer, and Phelan [49]. Although this result is not theoretically "new" (by definition frames are made up of resources), we begin with it here for two reasons.

First, other researchers in physics education research have yet to unpack frames in terms of their underlying resources. This result demonstrates the feasibility of that work. Second, knowing the resources underlying the frames for assessment is essential for arguments we will make later in the results section.

We begin by looking at the framings themselves. Of the 67 responses used, three students' behaviors indicated they started in one frame and transitioned to another. These students are featured as case studies in Secs. C and D. Of these three students, two began a problem in an algorithmic physics frame and transitioned to a conceptual physics frame, and one student started a problem in a conceptual physics frame and transitioned to an algorithmic physics frame. This yields a total of 70 frame instances observed.

Table II shows the resources we identified in each of the four frames. Within each frame, we identified a subset of resources often used when students were working in each frame. Across frames, different combinations of resources are used. Particular frames are associated with particular groups of resources different from other frames. This provides empirical evidence for our claim as predicted from our theoretical perspective.

B. Students seldom shift frame

The introductory students in this study almost always began problems in an algorithmic physics frame. Looking across the 67 total initial frame instances, 54 students started in the algorithmic physics frame, 11 students started in the conceptual physics frame, one student started in an algorithmic math frame, and one student started in a conceptual math frame.

For some questions, beginning and staying in the algorithmic frame is an effective way to solve the problem. An example of this would be the car problem that asked students to verify odometer readings at several times for a car. Students were given the car's acceleration and initial velocity. See Fig. 1 for the full problem. Below is a quote from Lisa as she solves the problem:

Lisa: So we can use this other equation if we want to confirm positions. So our V naught is going to be 5 point 2, 10 point 4, 16 point 7, 25 point 2, 35 point 9. So, using the equation to double check for position... (On her paper, Lisa isolates d , then types numbers into calculator) 16.7. Should I do it again? It might be fun. (Types into calculator) The next one is 25.2, which is as far as I want to check.

Lisa finishes by indicating on her paper that the feature is working correctly. It was determined that Lisa was working in the algorithmic physics frame because Lisa first accumulates information given in the question and equation sheet, then manipulates a physics equation to isolate the desired variable, then inputs her given values into her new equation and obtains a number that she interprets as an answer.

In this case, the algorithmic physics frame serves Lisa well, and she has no need to transition. Other questions are more easily solved in a conceptual frame. Ideally, students could switch fluidly when they get stuck on a problem in their current frame, as experts do [48]. However, our data show a fairly stable initial frame, meaning that students seldom switch out of their initial frame after beginning the problem. We found the algorithmic physics frame implements a specific set of resources, as shown in Table II.

The question that showcases this best from the problem set is the Gravitron problem (Fig. 3). Students were specifically asked to show their reasoning using diagrams, equations, and words. The problem may have been solved easily in a conceptual physics frame by noticing that the force of friction and the force of gravity should be equal so as not to let the rider slide down, which results in mass canceling out. The resulting equation shows that gravity will overcome the force of friction and all riders will fall down. Here is a quote from Amanda as she works to complete the Gravitron problem:

Amanda: So, right now I'm thinking about what to do for the Torque, to try to find or use some of the kinematic equations. I'm trying to remember how to do that.

TABLE II. Total instances of each frame observed and the resources associated with them.

Epistemological frames $N = 70$	Epistemological resources			
	Knowledge source	Knowledge nature	Epistemic activity	Epistemic form
Conceptual physics $N = 13$	Self Physics concepts	Fabricated	Forming Comparing Storytelling	Reasoning
Algorithmic physics $N = 55$	Equation sheet	Propagated	Accumulating	Number
Algorithmic math $N = 1$	Math protocol	Propagated	Accumulating Computing	Number
Conceptual math $N = 1$	Math concepts	Fabricated	Comparing Forming	Reasoning

Find the alpha, but... So I know there will be a friction force (Draws or writes on paper) which is just going to be $mg \mu$, and then... and that will keep you put at the wall minimum. So then, (Writes on paper) you can find torques for the minimum and the maximum mass. (Types into calculator, writes on paper) Okay, so now we have two torques. Hmm. I'm not sure where to go from there.

Amanda begins the problem by attempting to rearrange given rotational kinematics equations such that she could plug her given values into an equation and get an answer. For those reasons, we determined she was in the algorithmic physics frame.

Of the nine students who attempted the Gravitron problem, we observed all students beginning the problem in the algorithmic physics frame as indicated by their listed behaviors. Overall, students found the problem very challenging in this frame. Most did not solve the problem, but no students shifted their initial frame.

Across problems, our data show that students often started in and did not shift from the algorithmic physics frame. Analysis of students' statements in the Gravitron problem shows that students did not attempt another frame even when making little or no progress in their current frame. We provide possible reasons for this in Sec. VII.

C. Kicking a resource can shift a frame

While students seldom shifted frames, we observed three instances of a frame shift. All were caused by a change in a single epistemological resource. This was the result of two main influences: interviewer intervention and assessment feature.

Our first example of this comes from a student named Luke working on the car problem (Fig. 1). Below is the transcript of Luke beginning the problem:

Luke: Okay, so the positions are... these make sense. So, as its moving, it's accelerating at 2.2 meters per second every second. It's accelerating. So... 5.2... Since it's not accelerating during this time. Well, I guess I proved here that the difference in between is one second, so I guess that is the correct position. I feel like there is a simpler way of doing this that I'm overlooking. Um... (re-reads the problem silently)

This segment spans 2 min. Luke begins the problem by thinking conceptually about acceleration and the quantities he was given. Notice how he does not start any mathematical calculation or manipulation, but rather expects to obtain an answer by reasoning through the problem. The interviewer, noticing Luke was on the wrong path, intervenes:

Interviewer: When you find the positions, you can compare. The question is giving position. It is asking for validation.

Luke: Oh, I think I assumed the wrong thing when I did it this way. Because I assumed that... Well, I guess... So I assumed that... yeah. So, assuming that how fast it's going, starting at that velocity, I found that it would take one second to get from this point to this point, which verifies that this is correct. So then going to this point, I guess I would have to do this same thing again. That's just a lot of math.

This segment spanned 1 min. After the intervention by the interviewer, Luke realizes that he can solve the problem by substituting his givens into a known physics equation, which is quite the relief. Initially, Luke attempts to solve the problem by contemplating the properties of the quantity acceleration he was given in the question. After the intervention, Luke finishes the question by substituting the given quantities into a physics equation on his equation sheet. We interpret this, first, to be a shift from a conceptual physics frame to an algorithmic physics frame after the intervention by the interviewer.

A closer look at the resources Luke uses reveals a shift in epistemic form at that moment. Before the intervention, Luke believes he can reason his way to an answer to the yes or no question of whether the odometer is functioning properly on the car. We interpret his epistemic form at that moment to be reasoning. During the intervention, Luke realizes that he can easily calculate a number that will reveal the answer to the yes or no question. We argue that the intervention at the level of epistemic form is what caused Luke to shift from a conceptual physics to an algorithmic physics frame, and thus a new set of resources.

D. Assessment features can shift a resource, which shifts the frame

Outside of intervention, instructors may be able to influence their students' resource use, and thus their framing, through features of an assessment question.

To show this, we have two examples of a student (Lisa and Jack) completing a problem where they were asked to determine where a rider feels heaviest and lightest on a ferris wheel. See Fig. 2 for full problem. Students could solve the question in a conceptual physics frame by thinking about which direction the normal force points at different locations on the ferris wheel. Notice that the question does not give the students any quantities, and the variables given in terms of letters are constant. As you will read, this prevents students from successfully solving the problem in the algorithmic physics frame.

A total of nine students attempted this problem. Seven attempted the problem in what we determined to be a conceptual physics frame. Although we cannot conclude what made the students treat the ferris wheel problem differently, some admitted to having seen the problem before, so it is possible they already knew it could be solved conceptually. The remaining two students, Jack and Lisa,

treated the problem very differently. First, let us look at how Lisa solves the ferris wheel question. Below is the transcript of Lisa starting the problem:

Lisa: Um... (picks up equation sheet) Where's centripetal force? (Writes down centripetal force equation) Hmm... (looks at equation sheet) not given any numbers. It's just weird to me because it seems like nothing is changing. I mean, this is just going to be D over 2. The mass of the person isn't changing. Angular velocity is not changing.

Lisa begins the problem by attempting to manipulate equations from her equation sheet and plug in known quantities that she can translate into a position on the ferris wheel as her final answer. For these reasons, we determined that Lisa starts in the algorithmic physics frame. She becomes frustrated when she realizes that this method will not yield an answer for her. After pausing, she changes her approach:

Lisa: All I'm thinking about is that your potential energy is going to be the highest when you're at the top, lowest at the bottom, so I think those are going to be the points at which you're going to be feeling heaviest, and those are the points... I don't know why I'm thinking this but I just keep thinking of a clock and a pendulum swinging or anyone being on a swing and your... the points where you feel like you're accelerating the fastest are the ones where you're crossing this vertical axis. But I couldn't tell you why.

After realizing that she cannot answer the problem by manipulating given physics equations, Lisa changes her approach by contemplating her physical experience of being on a swing and applying her knowledge of potential energy. For these reasons, we determined that Lisa switches to the conceptual physics frame.

Initially, Lisa believes she can solve mathematically for an expression which she can translate into a position on a Ferris Wheel as her final answer. We interpret Lisa's epistemological form here as Number. Not providing numbers or varying quantities is a feature of the assessment that interrupts Lisa's initial epistemological form, thus forcing her to try something new to finish the question.

Next, Lisa believes she can determine the answer by reasoning through it, so we interpret her epistemic form to be Reasoning. This is a clear case of an assessment feature shifting an epistemic resource, thus shifting Lisa's epistemic frame.

We saw a very similar pattern in how Jack completed the problem as well. Below is the transcript of Jack starting the problem:

Jack: So this... for you to have this (points upward) that means you have to have a greater force acting on you

because, like, which is like, you have to have a greater force acting on you because like... you'll have a greater force acting on you, and here you'll have a smaller force acting on you. (Picks up equation sheet) We'll see, the angular velocity equation. I should look for the angular velocity equation. I should, yeah. (Flips through equation sheet) It's this one. Moment of inertia. Um, it doesn't tell you the mass.

Jack begins the problem by thinking about which forces are acting on the rider at different points on the ferris wheel. He concludes that he needs the angular velocity equation to solve and finds it on his equation sheet. Because Jack attempted to solve the problem by plugging values into known physics equations, we conclude that he was initially in an algorithmic physics frame. Because Jack thought plugging values into the angular velocity equation should give him an answer he could translate into a position on a Ferris Wheel, we interpret his initial epistemic form to be a Number. Jack hits a roadblock when he realizes he cannot solve the problem this way, and decides to change his strategy:

Jack: So, I guess because this is a conceptual equation... I am going to guess, for that, um, okay, so the greatest force you would be feeling (points upward) since you're going up, that means you would feel the lightest at the top, and the heaviest at the bottom. Because, like, you're going, accelerating faster at the bottom, and your mass is the same, so like the force would be greater at the bottom than it would be at the top. Because you're accelerating faster at the bottom than at the top. So, yeah, that's my answer.

Jack finishes the problem by applying what he knows about the physics phenomenon of force. We determined Jack finished the problem in a conceptual physics frame. Jack also switched from trying to obtain a mathematical expression as his answer to trying to reason his way to an answer, so we determined his final epistemic form to be reasoning. As Jack expresses, the assessment question not including numbers causes him to believe the question is conceptual in nature, which causes him to change what he is doing by switching from plugging numbers into an equation to thinking about the direction in which several forces are pointing. We interpret this as the assessment feature (no numbers) shifting his epistemic form, which shifts his epistemic frame.

VII. DISCUSSION

Our first claim is that epistemological frames (as identified by behaviors) can be defined by the epistemological resources used. Our analysis is consistent with much of the other work on problem solving. This suggests we can make sense of students' problem-solving work

during exams in much the same way we can make sense of their reasoning in other situations.

When students engage in problem solving during an assessment, they draw on a range of epistemological resources to do so. In Rosenberg *et al.*'s paper [49], researchers noticed groups of epistemological resources that consistently appeared together when students engaged in specific problem solving strategies that they called "epistemological coherences." They described that each group of resources was stable and reinforced the others. In our interpretation of these data, we understand their epistemological coherences to be epistemological frames as defined by student behaviors.

Additionally, defining epistemological frames by epistemological resources makes some intuitive sense. For example, "causal storytelling" as an activity would not make sense in the algorithmic math frame where students are computing and executing well established protocol. Alternatively, one could imagine that students displaying the behavior associated with the algorithmic math frame must engage in accumulating as they gather given information for their computation. Another example is that "number" as an epistemological form would not make sense in a conceptual physics frame where students are thinking about physics phenomenon and reasoning their way to an answer, whereas "reasoning" certainly would make sense.

Defining frames in terms of resources also changes the way we think about their relationship to one another. This could provide a useful tool to researchers. For any given coherent thought given by a student, it may prove easier at times to determine three to four epistemological resources that can be matched with a particular frame than attempt to interpret sometimes ambiguous student behavior. Defining frames in terms of both resources and behaviors gives researchers on small teams an efficient way to check their frame coding for accuracy and consistency.

Our second claim is that novice physics students seldom shift frames. Our results agree with previous findings. As discussed in Sec. II, Bing, and Redish found that one marker of expertise in physics is the ability to switch between frames as necessary [48]. Ideally, students could shift frames when one fails to help them complete the task at hand. Students did not shift frames when their current frame proved unhelpful, such as in the specific example of the Gravitron. As discussed in Sec. VI B, the Gravitron problem could have been solved very quickly in a conceptual physics frame, yet most students stayed in an algorithmic physics frame and struggled to arrive at an answer.

Rather than believing students perform poorly at switching frames, we believe students may be consciously choosing to begin and stay in the algorithmic physics frame. This could be due to some physics textbooks which specifically call for students to solve physics problems algorithmically [57–59]. This could also be due to the course instructors who sometimes teach students an

algorithm for solving physics problems [60,61]. It is possible that some part of the assessment itself cues students to begin a problem in a particular frame. Students could also begin problems in an algorithmic physics frame because that is what they generally do in physics assessment. Our data do not reveal why students start in the algorithmic physics frame. This idea is discussed further below.

Our third and fourth claims are that epistemological frame shifts happen on the level of an individual epistemological resource and that assessment features can prompt that resource shift. Our data imply that researchers could focus on how to shift resources as the means to the desired ends of frame shifting. We did not observe simply telling students to solve the problem conceptually or algorithmically to be effective in shifting frames. Notice that in the Gravitron problem (Fig. 3) students are explicitly asked to show their reasoning through words, but this failed to shift students towards reasoning in a conceptual physics frame. The frame shifts we did observe happened due to a shift at the level of a single resource. Therefore, we argue that trying to shift an individual epistemological resource may be an effective way to shift a student's epistemological frame.

Frame shifting as the result of a shift in a single epistemological resource is consistent with the findings of Rosenberg *et al.* In the Rosenberg *et al.* paper, students began completing their worksheet about the rock cycle by trying to accumulate as much information as possible from the external sources of information presented to them. After hearing them struggle, their teacher intervened by telling the students to "start with what you know." Their teacher suggested the students shift their source of knowledge, and the way the students engaged with their assignment changed. Rosenberg *et al.* noted the underlying meaning of her statement which was that students have many ways of engaging with assignments. More importantly, Phelen, the teacher, implied that the epistemological coherence students work in can be prompted to change by a single phrase when she asked students to "start with what [they] know."

A. Future research

Future work should explore the ability to do the reverse and make assessment questions that prompt resource shifts. It may be possible to create a list of recommendations for prompting resource and frame shifts in assessment writing.

Additionally, future work should explore why students often begin and stay in the algorithmic physics frame. It is possible that students begin in the algorithmic physics frame due to expectations of what "doing physics" means from before they entered the course. For example, physicists on television often discuss numbers and equations, so students may believe that they should solve problems by plugging numbers into equations. Students may begin most problems in an algorithmic physics frame due to their

professor or textbook explicitly telling them this or from success with previous problems when using this frame. To determine which, if any, of these ideas is the case, future research should explore students' perceptions of what doing physics is throughout an introductory physics course. A more extensive dataset across courses, professors, and student populations could also help us identify additional coherences of resources and observe different sets of epistemic frames.

VIII. CONCLUSION

In this work, we explored the relationship between epistemological frames and epistemological resources during student engagement in assessment.

Students in this study typically began problems in the algorithmic physics frame by writing down their givens, determining what they are solving for, and searching through their equation sheet to see what they can manipulate to return a value. To those of us who have taught in introductory physics courses, this finding is not surprising. In fact, it implies the algorithmic frame is well-rehearsed in the students' physics courses, either in lecture, homework, assessments, or some blend of the three.

While students had a fairly stable initial frame, we found influences like intervention from authority (i.e., the interviewer) and assessment features are effective at shifting an individual epistemic resource, which shifted the students' epistemic frame. This finding suggests that although the algorithmic physics frame is persistent, student epistemological framing during assessment can be fluid in much the same way as it is during other classroom activities [62].

These findings suggest a need to attend even more closely to assessment than we already do. We often think of assessment as occurring after instruction as something "added on" at the end of instruction for us as instructors to get information about our students. This work highlights that students are also getting information from assessments. This information might impact their understandings of how they should engage in physics learning. If that is the case, then we as instructors and researchers must pay more

attention to the kinds of messages we might be sending with our assessments to be sure they align with our goals for the course. Specifically, we must pay attention that we are cuing specific epistemological resources that will be productive for students. Additionally, we must be intentional about how we do that cuing so that it leads to substantive reframings.

Further, in this work, we used the perspectives of epistemological framing and resources to make sense of student engagement in assessment. The fact that this analysis "worked" at all is in and of itself an important result. It was possible that the cognitive mechanisms that have been used so productively to understand learning just would not have been useful in this different context. However, we found that framing analysis is consistent with student behavior during assessments. Additionally, the fact that our findings align with other work on problem solving outside of the assessment context is also noteworthy. Again, it was possible that what students did during assessment would be entirely discontinuous with their engagement in learning. Thus, our findings that students do rely on problem-solving-like framings and their associated resources during assessment are noteworthy. Specifically, it encourages us to reexamine our own assumptions about whether and how assessment is different from other classroom activities—both for ourselves and for our students.

ACKNOWLEDGMENTS

We would like to thank Katherine Ventura and Amali Jambuge for creating and administering the tests and recording the interviews analyzed in this paper. We would like to thank the Physics Department at Kansas State University for their support and the Research Experiences for Undergraduates (REU) program at KSU. This program is funded by the National Science Foundation (NSF) Grant No. 1757778. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

-
- [1] J. Tuminaro and E. Redish, Elements of a cognitive model of physics problem solving: Epistemic games, *Phys. Rev. ST Phys. Educ. Res.* **3**, 020101 (2007).
- [2] P. M. King and K. S. Kitchener, The reflective judgment model: Twenty years of research on epistemic cognitions, in *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (Lawrence Erlbaum Associates Publishers, Hillsdale, NJ, 2002).
- [3] L. Berland and K. Cruet, Epistemological trade-offs: Accounting for context when evaluating epistemological sophistication of student engagement in scientific practices, *Sci. Educ.* **100**, 5 (2016).
- [4] C. A. Chinn, L. A. Buckland, and A. Samarapungavan, Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology, *Educ. Psychol.* **46**, 141 (2011).

- [5] W. A. Sandoval and A. Çam, Elementary children's judgments of the epistemic status of sources of justification, *Sci. Educ.* **95**, 383 (2011).
- [6] D. Hammer and A. Elby, On the form of a personal epistemology, in *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, edited by B. K. Hofer and P. R. Pintrich (Lawrence Erlbaum, Hillsdale, NJ, 2002).
- [7] R. S. Russ, Epistemology of science vs epistemology for science, *Sci. Educ.* **98**, 388 (2014).
- [8] B. K. Hofer and P. R. Pintrich, The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning, *Rev. Educ. Res.* **67**, 88 (1997).
- [9] E. Redish, A theoretical framework for physics education research: Modeling student thinking, in *Proceedings of the International School of Physics*, edited by E. F. Redish and M. Vicentini (IOS Press, Varenna, Italy, 2003), Vol. 156 of *Research on Physics Education*, pp. 1–64.
- [10] L. Bollen, P. van Kampen, C. Baily, M. Kelly, and M. De Cock, Student difficulties regarding symbolic and graphical representations of vector fields, *Phys. Rev. Phys. Educ. Res.* **13**, 020109 (2017).
- [11] T. I. Smith, W. M. Christensen, D. B. Mountcastle, and J. R. Thompson, Identifying student difficulties with entropy, heat engines, and the Carnot cycle, *Phys. Rev. ST Phys. Educ. Res.* **11**, 020116 (2015).
- [12] B. R. Wilcox and S. J. Pollock, Validation and analysis of the coupled multiple response Colorado upper-division electrostatics diagnostic, *Phys. Rev. ST Phys. Educ. Res.* **11**, 020130 (2015).
- [13] L. Lising and A. Elby, The impact of epistemology on learning: A case study from introductory physics, *Am. J. Phys.* **73**, 380 (2004).
- [14] D. Hammer, Epistemological beliefs in introductory physics, *Cognit. Instr.* **12**, 151 (1994).
- [15] N. R. Council, *Knowing What Students Know: The Science and Design of Educational Assessment* (National Academies Press, Washington, DC, 2001).
- [16] R. J. Mislevy, Validity by design, *Educ. Res.* **36**, 463 (2007).
- [17] R. J. Mislevy and H. Geneva, Implications of evidence-centered design for educational testing, *Educ. Measurement: Issues Practice* **25**, 6 (2006).
- [18] R. J. Mislevy, L. S. Steinberg, and R. Almond, Design and analysis in task-based language assessment, *Language Testing* **19**, 477 (2002).
- [19] R. Almond, L. Steinberg, and R. Mislevy, Enhancing the design and delivery of assessment systems: A four-process architecture, *J. Technol. Learn. Assess.* **1** (2002), <https://ejournals.bc.edu/index.php/jtla/article/view/1671>.
- [20] L. S. Steinberg, R. J. Mislevy, R. G. Almond, L. V. Dibello, H. Chernick, and A. C. H. Kindel, Introduction to the Biomass project: An illustration of evidence-centered assessment design and delivery capability, CRESST Technical Report 2009.
- [21] M. Wilson, Constructing measures: An item response modeling approach, <https://www.amazon.com/Constructing-Measures-Response-Modeling-Approach/dp/0805847855> (2005).
- [22] NGSS Lead States, *Next Generation Science Standards: For States, By States* (National Academies Press, Washington, DC, 2013).
- [23] National Resources Council, *Developing Assessments for the Next Generation Science Standards* (National Academies Press, Washington, DC, 2013).
- [24] AERA, *Standards for Educational and Psychological Testing*, 2014 ed. (AERA, Washington, DC, 2014).
- [25] W. K. Adams and C. E. Wieman, Development and validation of instruments to measure learning of expert-like thinking, *Int. J. Sci. Educ.* **33**, 1289 (2011).
- [26] Physport: Supporting Physics Teaching with Research-Based Resources, URL: <https://www.physport.org/>.
- [27] A. Madsen, S. B. McKagan, M. S. Martinuk, A. Bell, and E. C. Sayre, Research-based assessment affordances and constraints: Perceptions of physics faculty, *Phys. Rev. Phys. Educ. Res.* **12**, 010115 (2016).
- [28] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *Am. J. Phys.* **66**, 64 (1998).
- [29] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 84108415 (2014).
- [30] J. Von Korff, B. Archibeque, K. A. Gomez, T. Heckendorf, S. B. McKagan, E. C. Sayre, E. W. Schenk, C. Shepherd, and L. Sorell, Secondary analysis of teaching methods in introductory physics: A 50 k-student study, *Am. J. Phys.* **84**, 969 (2016).
- [31] E. Brewster, V. Sawtelle, L. H. Kramer, G. E. O'Brien, I. Rodriguez, and P. Pamel, Toward equity through participation in modeling instruction in introductory university physics, *Phys. Rev. ST Phys. Educ. Res.* **6**, 010106 (2010).
- [32] M. Kohlmyer, M. Caballero, R. Catrambone, R. Chabay, L. Ding, M. Haugan, M. Marr, B. Sherwood, and M. Schatz, Tale of two curricula: The performance of 2000 students in introductory electromagnetism, *Phys. Rev. ST Phys. Educ. Res.* **5**, 020105 (2009).
- [33] M. D. Caballero, E. F. Greco, E. R. Murray, K. R. Bujak, M. Jackson Marr, R. Catrambone, M. A. Kohlmyer, and M. F. Schatz, Comparing large lecture mechanics curricula using the Force Concept Inventory: A five thousand student study, *Am. J. Phys.* **80**, 638 (2012).
- [34] A. Madsen, S. B. McKagan, and E. C. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020121 (2013).
- [35] M. Lorenzo, C. H. Crouch, and E. Mazur, Reducing the gender gap in the physics classroom, *Am. J. Phys.* **74**, 118 (2006).
- [36] E. F. Redish, J. M. Saul, and R. N. Steinberg, Student expectations in introductory physics, *Am. J. Phys.* **66**, 212 (1998).
- [37] W. Adams, K. Perkins, N. Podolefsky, M. Dubson, N. Finkelstein, and C. Weiman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010101 (2006).

- [38] 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. Mestre (2005), pp. 89–120, <https://www.amazon.com/Transfer-Learning-Modern-Multidisciplinary-Perspective-ebook/dp/B01FNA3756>.
- [39] G. L. MacLachlan and I. Reid, *Framing and Interpretation* (Melbourne University Press, Carlton, Victoria, 1994).
- [40] D. Tannen, *Framing in Discourse* (Oxford University Press, New York, 1993).
- [41] R.E. Scherr and D. Hammer, Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics, *Cognit. Instr.* **27**, 147 (2009).
- [42] T.O.B. Odden and R.S. Russ, Defining sensemaking: Bringing clarity to a fragmented theoretical construct, *Sci. Educ.* **103**, 187 (2019).
- [43] D. Tannen and C. Wallat, Interactive frames and knowledge schemas in interaction: Examples from a medical examination/interview, *Social Psychol. Quarterly* **50**, 205 (1987).
- [44] R.S. Russ and M.J. Luna, Inferring teacher epistemological framing from local patterns in teacher noticing, *J. Res. Sci. Teach.* **50**, 284 (2013).
- [45] P. Hutchison and D. Hammer, Attending to student epistemological framing in a science classroom, *Sci. Educ.* **94**, 506 (2010).
- [46] T. Bing and E. Redish, Analyzing problem solving using math in physics: Epistemological framing via warrants, *Phys. Rev. ST Phys. Educ. Res.* **5**, 020108 (2009).
- [47] B. Modir, J.D. Thompson, and E.C. Sayre, Students' epistemological framing in quantum mechanics problem solving, *Phys. Rev. Phys. Educ. Res.* **13**, 020108 (2017).
- [48] T.J. Bing and E.F. Redish, Epistemic complexity and the journeyman-expert transition, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010105 (2012).
- [49] S. Rosenberg, D. Hammer, and J. Phelan, Multiple epistemological coherences in an eighth-grade discussion of the rock cycle, *J. Learn. Sci.* **15**, 261 (2006).
- [50] A. Collins and W. Ferguson, Epistemic forms and epistemic games: Structures and strategies to guide inquiry, *Educ. Psychol.* **28**, 25 (1993).
- [51] R.E. Scherr, Modeling student thinking: An example from special relativity, *Am. J. Phys.* **75**, 272 (2007).
- [52] National Research Council, *A Framework for K-12 Science Education* (National Academies Press, Washington, DC, 2012).
- [53] J.T. Laverty, S.M. Underwood, R.L. Matz, L.A. Posey, J.H. Carmel, M.D. Caballero, C.L. Fata-Hartley, D. Ebert-May, S.E. Jardeleza, and M.M. Cooper, Characterizing college science assessments: The three-dimensional learning assessment protocol, *PLoS One* **11**, e0162333 (2016).
- [54] D.N. Chari, H.D. Nguyen, D.A. Zollman, and E.C. Sayre, *Am. J. Phys.* **87**, 875 (2019).
- [55] A. Coffey, H. Beverley, and A. Paul, Qualitative data analysis: Technologies and representations, *Sociological Res. Online* **1**, 8091 (1996).
- [56] M.L. McHugh, Interrater reliability: the kappa statistic, *Biochem. Med. (Zagreb)* **22**, 276 (2012).
- [57] R.D. Knight, *Physics for Scientists and Engineers: A Strategic Approach with Modern Physics and Mastering Physics*, 2nd ed. (Benjamin Cummings, San Francisco, CA, 2007).
- [58] W. Bauer and G. Westfall, *University Physics with Modern Physics*, 2nd ed. (McGraw-Hill Education, New York, 2013).
- [59] E. Mazur, *Principles & Practice of Physics*, 1st ed. (Pearson, Boston, 2014).
- [60] D.S. Wright and C.D. Williams, A wise strategy for introductory physics, *Phys. Teach.* **24**, 211 (1986).
- [61] P. Heller and M. Hollabaugh, Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, *Am. J. Phys.* **60**, 637 (1992).
- [62] M. Lau, Understanding the dynamics of teacher attention, <https://drum.lib.umd.edu/handle/1903/10897>.