

Lessons from pendulums: A design comparison of three lab activities

Ian Descamps^{1,*} Roger Tobin¹ Paul Wagoner¹ David Hammer^{1,2}
 N. G. Holmes³ and Rachel E. Scherr⁴

¹*Department of Physics and Astronomy, Tufts University, 574 Boston Avenue, Suite 304, Medford, Massachusetts 02155, USA*

²*Department of Education, Tufts University, 12 Upper Campus Road, Medford, Massachusetts 02155, USA*

³*Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853, USA*

⁴*Physical Sciences Division, School of STEM, University of Washington Bothell, Bothell, Washington 98011, USA*



(Received 10 December 2025; accepted 20 March 2026; published 27 April 2026)

We present three versions of a pendulum lab activity to explore how theoretical commitments, motivations, and aspirations are reflected in curriculum design. In earlier work, Boudreaux and Elby [How curriculum developers' cognitive theories influence curriculum development, *Phys. Rev. Phys. Educ. Res.* **16**, 020144 (2020)] discussed how different theoretical perspectives led them to different designs of tutorials. Here, we discuss our finding that even starting from the same theoretical perspective and having similar goals for students can lead to differences in design. We give three interacting reasons why our labs diverge: We have different expectations of students at our respective institutions; we have different ancillary goals; and we draw different implications from our shared theoretical commitments. Our account demonstrates the complexity of the relationship between theory, goals, and curriculum design. In this way, it adds to prior arguments for the importance of designers' articulating the reasoning for their choices as well as for the possible value of instructors' responsive adaptations to curricula.

DOI: [10.1103/PhysRevPhysEducRes.22.010137](https://doi.org/10.1103/PhysRevPhysEducRes.22.010137)

I. INTRODUCTION

For longer than any of the authors of this paper have been learning physics, introductory labs have included a pendulum activity. Often, a goal of the activity is for students to demonstrate that the period of a pendulum has a small but measurable dependence on the amplitude of its swing, a result at odds with the simple harmonic oscillator model and Galileo's claim in *Two New Sciences* [1]. Such a lab provides a context for learning the practices and values of empirical science: to see the dependence, students need both to take care in their measurements and to be aware of the precision they have obtained.¹

Each of us has assigned pendulum labs at the start of our introductory courses, at our respective institutions, with the goals of introducing the practices and values of empirical science. We want the activities to promote students' attention to experimental error as well as reflection and

conversation about how science produces and assesses knowledge. These goals reflect our shared view of students: that they have mostly experienced step-by-step instructions for experiments that demonstrate or validate theoretical claims from lecture, and that they need support to engage in more deeply scientific practices.

The similarities in our interests, goals, and views motivated research on how students may shift their approach to labs, to be less about “doing school” and more about “doing physics” [2–4]. That is, we have asked: when and how do students come to frame introductory labs in ways that support progress toward the practices and values of empirical science? We were supported by the National Science Foundation for three years to collect and study data from our courses.² We paid particular attention to the pendulum labs, because they were the students' first engagement with activities different from their expectations. The project led to several publications analyzing student framings, including with respect to confirmation bias and epistemic agency [5–9].

Throughout, we authors were aware of differences in our lab designs. We present the task differently: as testing the simple harmonic oscillator model $T = 2\pi\sqrt{\frac{L}{g}}$ (Cornell), as testing Galileo's claims (Tufts), or simply as studying what

*Contact author: ian.descamps@tufts.com

¹They do not need sophisticated technology! Mechanical stopwatches work fine.

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.

²<https://www.physport.org/curricula/thinkingcritically/>.

affects the period (UWB). We guide it differently, providing instructions for students' experimentation and statistical analyses (UWB and Cornell) or leaving almost everything up to the students (Tufts). The project was funded to study students' thinking, but as it ended, we became interested in studying our own: How did we come to design our labs so differently, given the similarities in our perspectives?

A. Theoretical frameworks

We were inspired by Boudreaux and Elby [10], who described how their theoretical perspectives shaped their designs of tutorials on static friction. They demonstrated multiple ways in which their distinct theoretical orientations gave rise to differences in tutorials covering the same content. Boudreaux worked from a “difficulties framework,” which attributes “a degree of stability in student conceptions”: “The strategy of challenging and desettling non-normative ideas relies on those ideas maintaining coherence long enough to be held up, examined, and modified by the learner as they progress through the curriculum.” This perspective supported designing a tutorial to “elicit, confront, and resolve” incorrect conceptions.

Elby worked from a “resources framework,” according to which “students’ responses to physics questions do not always correspond to stable conceptions.” Rather, patterns in students’ thinking depend on the context, “with contextual cues tipping students” to different ways of thinking. This perspective supported designing a tutorial to manipulate contextual cues that could tip students into activating different patterns of resources and guide them to notice and work to reconcile the differences.

The authors were clear that explicit employment of theoretical frameworks did not fully determine their tutorial designs. They also drew on their intuitions and experiences, which they note are deeply entangled with their own theoretical dispositions: “Theoretical orientations and instructional intuitions are not disjoint streams of knowledge that independently inform curriculum writing, but rather that they interact as they develop over time.”

B. Our theoretical orientations

In contrast to Boudreaux and Elby, we see ourselves as starting from the same theoretical perspective. Like Elby, we all take a resource-based view of student knowledge, reasoning, and framing [11]. The theoretical idea of “resources” is of “pieces” [12] of knowledge that can be active or not, depending on the context. Some resources bear on physical mechanisms, such as a basic sense that a stronger cause will produce a stronger effect, or that motion “dies away” if there is nothing to sustain it [12]. Some bear on kinds of activities and situations, influencing which features people recognize and pay attention to, as well as their judgments about appropriate actions. These resources are often referred to as “frames” and their activation as “framing” [11,13,14].

A resources perspective does not presume stability. Resources may become active or not depending on the context. For example, in the context of cold weather, it is easy and intuitive to think “gloves keep hands warm,” while in the context of baking, “gloves keep hands cool.” Much of the work of physics is in connecting reasoning in contexts that start out as separate. Similarly, for example, someone framing what is taking place as following instructions may not notice or think to check whether the instructions make sense to them, as they would if they framed the moment instead as participating in a decision.

Taking this perspective, we expect the dynamics of student thinking to be significantly influenced by how tasks are presented [15–17], instructor or teaching assistant (TA) interactions [8,18,19], intragroup social dynamics [14,20–22], and practices established in the classroom community [23,24]. In our prior work, we have considered dynamics at multiple scales of time and system, from individual minds in specific moments, to small collaborative groups over semesters, to classroom and school communities [14,17,25–28].

Our work in instructional labs has documented how students can frame lab as a matter of “jumping through hoops” [6], and/or expect that success means confirming a known result [3,9]. These are laboratory versions of “doing school” [29]: following “cookbook” instructions to arrive at the intended results, the findings expert authorities have determined to be correct. Unfortunately, these expectations are often at odds with the practices and values of empirical science.

The frames highlighted above, and students’ framing in general, have profound consequences for the nature and quality of students’ participation in labs. That is to say, how students frame the lab shapes the ways and extent to which they take up *epistemic agency* [6,8,9,30]. Epistemic agency is the prerogative for individuals or groups to participate in “the whole range of components of knowledge building goals, strategies, resources, evaluation of results, and so on” [31] (p. 108). In our research and in our designs, epistemic agency is central to our understanding of what constitutes productive participation in scientific knowledge construction.

C. Similarities and differences across our labs

In designing our labs, we all anticipate that many students will arrive with expectations they have formed over years of experience in traditional labs, disposed to frame experimental work as a matter of following instructions to arrive at the correct (predetermined) conclusions [3]. Part of what we all hope to accomplish with the pendulum lab is to disrupt that framing and to promote students’ seeing lab as a place for genuine inquiry, such as to make sense of a surprising phenomenon [2]. We expect students have resources for the beginnings of empirical science. Given the chance,

students can and will be curious about the world, interested and able to try things to learn more about phenomena they notice.

The instructional materials for the Cornell curriculum can be found on PhysPort; instructional materials for the UWB curriculum can be found on QUBES CourseSource [32]; instructional materials for the Tufts curriculum are attached as Supplemental Material [33]. As would be expected given our similar views and goals, our labs are similar in several ways. We all see the main purpose of introductory labs as helping students develop abilities in empirical science rather than reinforcing concepts from lecture [34,35]. The instructions, assessment, and feedback we provide emphasize practices of empirical science and challenge expectations that success means arriving at a “correct” answer. They give students time and autonomy to design, test, and refine their experimental methods and analyses, and all offer them significant latitude to draw and defend their own conclusions. Finally, all of our designs depend to some extent on “responsive teaching” [36], that is the lab instructors’ close attention to and engagement with student thinking, in individual and small group interactions, as well as in classroom community-level idea sharing and discussions.

Still, as we noted above and will elaborate below, our labs differ in several significant ways. We decided to look into those differences, and in what follows, we report on what we have found: Variations in our designs arise from (i) differences in our expectations of students at our respective institutions, (ii) variations in our ancillary pedagogical goals, and (iii) nuances of our theoretical perspectives. We will try to make our reasoning transparent, to expose uncertainties and flexibilities in our designs, to argue for expectations of complexity, and to motivate views of instruction and instructional design as ongoing and dynamic.

We hope that our comparisons provide insight into and concrete examples of how learning theories guide curriculum development. Scholars have long argued, and empirical evidence indicates that it is important for teachers implementing curriculum to understand the purposes and aims of its design [37–43]. Our comparative analysis supports these arguments, exposing “hidden decisions” [41] and “pedagogical judgments” [40] that led us to different versions of the pendulum lab.

In the next section, we review the literature on the design of “nontraditional” instructional labs, with an emphasis on their theoretical bases, both implicit and explicit. Then we move on to our curricula: first, in Sec. III, we discuss each of the institutional and instructional contexts; in Sec. IV, we describe our respective approaches and their ongoing evolution in turn and in some detail; and in Sec. V, we compare these approaches. In Sec. VI, we step back to make sense and discuss the implications of our comparison before, finally, concluding in Sec. VII.

II. LITERATURE REVIEW: LAB CURRICULUM RESEARCH AND DEVELOPMENT

Physicists, educators, and education researchers have long expressed dissatisfaction with traditional laboratory instruction [35,44–47]. We understand “traditional” labs to be defined by (i) aims to reinforce content knowledge, (ii) activities that demonstrate or verify physical concepts, and (iii) prescriptive experimental protocols. Despite such labs’ goal of supporting conceptual learning, evidence indicates that they are generally unsuccessful at it [27,46,48,49]. Research also indicates that traditional labs negatively impact the development of expertlike epistemological attitudes and beliefs [50–53].

Reshaping laboratory instruction with an eye toward authentic disciplinary and experimental practices has notable historical precedent [44,47]. A number of research projects have developed and documented alternative curricula, often inspired by educational theory:

- RealTime Physics (RTP; [46,54,55]).
- Investigative Science Learning Environment (ISLE; [48,56]).
- Thinking Critically in Physics labs, first at the University of British Columbia [57–59] and continuing at Cornell [6,8,30,60].
- The instructional labs at CU Boulder [61,62].
- The Design, Analysis, Tools, and Apprenticeship lab (DATA lab; [63]).

Some of these learning environments have documented evidence that they support productive experimental activity in both quantity and quality [48,49,64] and improve students’ epistemological attitudes and beliefs [49,51,53,65]. At the same time, generalizing the motivations of these curricula as engagement in genuine experimental inquiry or skills-focused designs loses the specificity necessary for researchers to rigorously evaluate and compare them and, crucially, for educators to effectively adapt and implement them. These lab curricula vary in their specific goals, theoretical influences, and pedagogical features and structures. In the next section, we discuss these curricula in more detail by focusing on the differing roles of theory in their development and documentation.

A. Theoretical influences on lab curricula

1. *Semiempirical*

The DATA and CU Boulder lab curricula are presented as the conclusion of a largely empirical transformation process. In each case, the lab developers interviewed interested faculty about their goals for laboratory education, synthesized the themes of those interviews into consensus goals, and then iteratively refined the goals along with faculty members. These consensus goals provided the basis for the design of the curriculum. We refer to this process as semiempirical because, while the overall course

transformations incorporated empirical research, the design decisions were not explicitly grounded in empirical (or theoretical) explanations.

For the DATA lab, they settled on four learning goals: experimental process, data analysis, collaboration, and communication. Notably, the “design was divorced from the specific physics content.” This feature was explicitly motivated by their empirical findings rather than by theoretical considerations: “the learning goals developed from a faculty consensus design did not include specific content” [63]. To select and design instructional materials and course structures, Funkhouser *et al.* [63] applied constructive alignment [66]: the alignment of an instructional system based on constructivist principles. In this scheme, course objectives should articulate how students can demonstrate understanding through actions; students should be put in situations that can elicit such “performances of understanding,” and students should be assessed based on their performance.

The DATA lab is a stand-alone course and a two-semester sequence. The typical lab activity takes place over two weeks and students generally follow the same two-week trajectory: prelab homework informs the qualitative exploration and prediction phase, and the first week culminates in students’ designing their experimental procedure; in the second week, students conduct their investigation and iteratively improve their procedure as they analyze their data and generate interpretable results.

Details of the initial CU Boulder lab curriculum, also a stand-alone lab course, can be found in a series of PERC papers [61,67–69]. The design of this curriculum was motivated by five main learning goals:

- students’ epistemological attitudes should align with experts,
- students should have positive attitudes about the course,
- students should have positive attitudes about experimental science,
- students should be able to generate a graph showing model and data, and
- students should demonstrate “set-like” reasoning about data.

Additionally, in other work [62], the CU Boulder team refers to their initial lab as seeking to develop skills. The research team also notes that, due to various constraints, the course retained many features of a traditional lab [70].

Nevertheless, Pollard *et al.* [70] highlight several transformed features they expect to be salient for student learning: for a prelab assignment, students watch short, interactive videos; the experimental activities require students to measure a previously unknown quantity or outcome, as opposed to verifying values learned in lecture or listed in a textbook; students use these measurements to make a prediction related to their experiment and then, once they measure that prediction, compare results with other

groups. That comparison is expected to support the intentional use of uncertainty values. Additionally, most of the lecture sessions (which are infrequent) are dedicated to measurement uncertainty concepts.

Although presented as a mostly empirical process, we contend that there is nontrivial theoretical work involved in the development and design of these curricula. For example, a feature of the DATA lab is that, for a given activity, each lab group conducts a different experiment, with no set end point. Funkhouser *et al.* mention that this feature “has added benefits for [each] student’s ownership and agency of the work.” It could be that ownership and agency were understood by the designers to support productive engagement and thus were part of the decision-making process. Another possibility is that increased ownership and agency were serendipitous benefits afforded by this choice, which was influenced by other considerations.

This approach obscures the justification of design decisions. Curricular features are instead presented as naturally emerging from learning goals. As a result, the documentation of curriculum transformations does not provide actionable insight for educators enacting the curriculum or selecting among alternatives as they adapt design features. The work of specifying how to accomplish learning goals can be productively understood as a theory-laden process of making hypotheses or conjectures about how to best support learning [71,72]. Accounts of the theory-laden reasoning that motivates design decisions in a curriculum are especially important given our understanding of the variation in how educators take up research-based instructional methods [73–75], the consequential role of the broader (activity) system within which specific design features are implemented [76,77], and the ways in which an awareness of this reasoning can support responsive instructional moves [43]

2. Theoretically informed

The curricula described in this section all explicitly invoke learning theories (and related empirical research) as they describe their designs, albeit in different ways: the design principles of RTP were inspired by educational research; the researchers at CU Boulder outline expectations about how authenticity supports learning and list various features of their CURE expected to promote it; the designers of ISLE list their overarching instructional values and intentionalities that help them select theoretical frameworks to do design work.

As one of the first evidence-based or research-based reformations of laboratory instruction, the RTP lab curriculum is a significant scholarly and historical milestone for PER as a field [44,46,47]. In addition to learning goals, Sokoloff *et al.* [54] delineate design principles “based on education research,” that undergird the RTP labs. The lab activities

- are sequenced to provide students with a coherent observational basis for understanding a single topic area in one semester or quarter of laboratory sessions;
- provide activities that invite students to construct physical models based on observations and experiments;
- help students modify their common conceptions about physical phenomena that make it difficult for them to understand powerful general principles of physics;
- work well when performed in collaborative groups of two to four students;
- incorporate microcomputer-based laboratory (MBL) tools so that students can test predictions by collecting and graphing data in real time;
- incorporate a learning cycle consisting of prediction, observation, comparison, analysis, and quantitative experimentation;
- provide opportunities for class discussion of student ideas and findings; and
- integrate homework assignments designed to reinforce critical concepts and skills.

The central innovation of the RTP labs is the MBL tools: sensors and interfaces that enable the real-time collection, display, and analysis of various physical measurements. These MBL tools were designed “with the student learner in mind” [46]. They enable students to focus on the physical world, giving them the ability to explore and learn from the world; the immediate feedback of the devices extends the range of potential investigations, and the graphical display options can provide students a start in developing representational competence [55]. The design of the MBL tools, along with the RTP lab curriculum, builds on findings from cognitive science and education research at the time [78–80] as well as the 1998 AAPT laboratory recommendations [81].

ISLE similarly aims for students both to learn concepts and to engage in empirical inquiry, but it differs from RTP in how it implements experimentation. ISLE aims more explicitly at supporting students’ development in the practices of empirical science. As part of that, it shifts responsibility for experimental design onto students, trading a richer experience for the efficiency and reliability of successful, canonical findings. Like RTP, ISLE involves “cycles” of learning and experimentation, but the form differs. RTP emphasizes student predictions and hypotheses, in part to elicit and respond to student misconceptions. ISLE, in contrast, begins with observational experiments and “something that needs explaining” [56], in part because of the designers’ views of student epistemologies as essential features of their engagement and targets for instruction.

To describe the theoretical foundations of their curriculum, Brookes *et al.* [56] overview their “core intentionalities” [82]—their beliefs and values about what their teaching is trying to accomplish—and use them as criteria for evaluating potential theoretical frameworks [72]. With

these criteria in mind, they discuss the “bricolage” of theoretical perspectives that actualize their beliefs about how learning happens and how to support it. This discussion expands on their 2010 paper [48], which highlighted *interpretive knowing* as a desired form of cognition and activity that guided their design. Such theoretical descriptions provide not only specificity around important design decisions but also context for how and to what end the theories were implemented.

In response to the need to teach remotely during the beginning of the COVID-19 pandemic, CU Boulder redesigned their lab course to be a Course-based Undergraduate Research Experience (CURE) [62]. The authentic discovery that defines CUREs made this approach an attractive option for their redesign, as it supports engagement in key scientific practices, including collaboration and iteration; furthermore, the emphasis on authenticity (with respect to professional physics research) is expected to support motivation and persistence. In other words, *authenticity* as a theoretical construct shaped many of the design decisions in the construction of this curriculum.

The designers expected the authenticity of this research to support learning in two main ways: first, authentic research requires scientific practices such as reading published literature, developing a research plan, data analysis, collaboration, and iteration; second, the authenticity of this research goes *beyond* “offer[ing] students glimpses of what it means to do experimental physics” [35], which is expected to positively impact student motivation and affect [83]. Oliver *et al.* [83] identify six different aspects of the course designed to engage students in such authentic research: a scaffolded literature review, meetings with the PI of the project (twice per semester for each group), metacognitive reflections, teamwork, peer review, a final reflection structured as a “memo to future researchers,” and the publication of their research.

The education research from the first few semesters of this course highlights that engagement in scientific practices did occur [83]. Students felt they participated in authentic research [62,83], that they had achieved their teamwork goals [84], and that teamwork contributed to their success in the course [84].

As with learning goals, theory does not uniquely and completely determine design [10,42,72]. Still, theoretically grounded descriptions provide insight into the design process and the reasoning that links learning goals and pedagogical structures. Empirical study is necessary to understand what works when, for whom, and how, but the application of those findings to new situations and curricula necessarily involves theorizing. Explicating the reasoning for design decisions, especially theoretically grounded reasoning, can generate instructionally actionable insight [42,43] and empirically testable conjectures [71,72,85].

III. INSTRUCTIONAL CONTEXTS

Although throughout this paper, we refer to the authors' theoretical commitments, we should be clear that these labs were designed and are implemented collaboratively. At Cornell, Emily Smith, Phil Krasicky, Mark Lory-Moran, Michael Niemack, Jared Maxson, Cristina Schlesier, and Rebeckah Fussell all contributed to the curriculum design, in addition to coauthor Holmes. Holmes and her colleagues were supported by an internal grant to do the research and development for *Thinking Critically in Physics* [86]. At Tufts, Timothy Atherton and Hugh Gallagher contributed to the curriculum design, in addition to co-authors Hammer, Tobin, and Wagoner.

Cornell is a large, private, selective research university; Tufts is a small, private, selective research university; and UWB is one of two small campuses of a large, public research university. UWB runs on the quarter system, and the lab course is part of a three-term sequence; both Cornell and Tufts are on the semester system.

The introductory lab at Cornell is a stand-alone course, and the lab curriculum at UWB is an adaptation of the Cornell curriculum. Each includes a lecture, where students reflect on their labs and work through activities to develop statistical skills. At Tufts and UWB, introductory labs are sections connected to the main lecture courses. Tufts has no lab-specific lecture and lab instructions change in small ways from one semester to the next. Here we focus on the spring instantiation of the Tufts labs, taught by Hammer, specifically the Spring 2024. In fall semesters, the labs are shared across algebra-based and calculus-based courses, and the overall course management proceeds differently.

The labs at Cornell are instructed by graduate TAs, they last two hours each, and there are a maximum of 20 students per section. At UWB, professors instruct the labs, which last two hours and have a maximum of 24 students per section. The Tufts labs are instructed by graduate and undergraduate TAs, last two and a half hours, and there are a maximum of 14 students per section.

At both UWB and Tufts, the pendulum lab activity occurs over three lab sessions; at Cornell, the activity takes two lab sessions. The lab courses at Cornell and UWB assign homework; Tufts labs do not (students complete all notes and reports within the 2.5-hour sessions).

IV. THREE VERSIONS OF THE PENDULUM LABS

Our designs share foundations, orientations, and aspirations. All take as central the desire to engage students in doing empirical science, focusing on “what it means to do experimental physics: the approach, techniques, skills and ways of thinking when conducting authentic physics experiments” [35] (p. 1). Crucially, this doing must be active and intentional—it must be epistemically agentic. We design our labs in such a way as to encourage and support students in taking up roles as knowledge agents,

who have the ability to creatively and intentionally pursue goals relevant to knowledge production.

In the subsections that follow, we provide a more in-depth description of the pedagogical logic and an outline of the design itself, along with an explanation of the *reasoning* for the various design decisions. In these descriptions, we seek to make clear how differences arise in our designs from different prioritizations and emphases of that overarching goal as well as from different “secondary” considerations.

A. Cornell

At Cornell, the design of the pendulum activity creates a context for experimental practices to have purpose and relevance by setting students up to encounter an empirical-theoretical discrepancy. The instructional materials for the Cornell pendulum lab can be accessed through PhysPort (footnote 2). The lab instructions discuss the pendulum period equation derived from a model (with assumptions listed), and the guidance of the in-lab instructions sets students on a path to produce data that indicate a deviation from this model. Students are assessed on their process and the quality of their work, rather than on whether they arrive at the correct conclusion.

At the heart of this lab activity is the epistemological richness of evaluating a model in conversation with experimental work; the (potential) conflict creates a need not only for calculating and interpreting uncertainty, but also for experimental design and data visualization skills that work together to respond to the theoretical-experimental discrepancy. This rich context aligns with the overarching course goal of students developing a more sophisticated understanding of what it means to know things and to construct knowledge in physics.

As the first lab activity, the structure of this lab aims to *cultivate* student agency. The guidance of the in-lab instructions constrains the field of inquiry, but in ways that provide students with a sense of direction for their work so as to avoid the frustration of not knowing what to do. The instructions are open ended to create decision-making opportunities. For example, in the first lab, the experimentation begins with a direction to set up a pendulum and record period measurements at 10° and 20° but students are not told how to carry out this experimentation. In this way, the in-lab instructions begin to shift responsibility for experimentation and knowledge production onto the students.

Before the experimentation begins in the first lab, the first several questions of the in-lab instructions prompt groups to discuss and reflect on how they are going to approach collaboration. Students are asked to discuss their purpose and goals in this course and how they will support each other's goals; they are asked to synthesize a common or collective goal for their group and to generate a plan of

some sort for ensuring fairness and equity in the division of labor. Their recorded responses to these questions make up the groups' "partner agreement." In addition to the ways this discussion gets groups to proactively think about how they are going to collaborate, the written record can serve as a useful artifact for navigating tension and disputes between group members.

After this discussion, there is one more thing students must do before they begin their experimentation: properly connect their CoLab Or Jupyter Notebook to their utilities folder. The instructional team at Cornell recently updated their in-lab instructions to be written in Jupyter Notebook format, combining the in-lab instructions with students' lab notebook. Students can use text cells to write out responses to questions or record other notes; the document also includes some base code that students can use to input and process their data. While coding and computational skills are not explicit foci of this course, including Jupyter Notebook and CoLab as tools in this course allows students to become familiar with them and see them as useful.

Once the code is properly set up, students begin their experimentation. Students start by constructing a pendulum and recording period measurements at 10° and 20° . They then perform some calculations (computing the mean, standard deviation, and standard uncertainty) and answer some questions connecting those statistical measures to the physical causes of variability and uncertainty in their data. This work launches them into the main focus of the first session: reducing uncertainty. Groups can choose between two options for their main experimental work of the session: exploring how increasing swings per trial reduces uncertainty or comparing the approaches of using repeated single swing trials to a multiswing trial.

In between lab sessions, during the lecture portion of the course, students participate in a contrasting cases invention activity on quantifying distinguishability. The activity introduces students to the "t-prime" (t') statistic [87], similar to Student's t test, with the idea that the contrasting cases invention structure supports student's conceptual understanding of this statistical tool. In the following lab activity, students are prompted to use and interpret t' values for the data they produced in the first session. The goal of the contrasting cases invention activity is to provide students with a tool they understand (as opposed to a black box) that they can use to make meaning from their data.

At the beginning of the second lab session, students compute t' values for their data and are then asked about "reasonable next steps or new questions based on your set of t' values?" The main focus of the second lab session is pushing models to their limits, and they are expected to use t' to help them in that work. As in the first session, groups can choose between two options; the two options this week are "reducing uncertainty" or "pushing on assumptions." In both cases, students must explain how their task helps test the limits of the model and better understand the physical

phenomenon. All groups generate ideas for possible experiments and then must "systematically test multiple measurement methods" for their chosen task.

The instructions also prompt groups to consult and collaborate with other groups. First, when groups are generating ideas for possible experiments, the instructions suggest that students consult other groups. Then, once groups have conducted their systematic tests, the instructions direct students to find another group who chose the other task option and to exchange results with them. This exchange and collaborative discussion forms the basis for yet another round of iterating and improving their experimental design. A desired outcome of these activities, and of the first lab session as well, is that students use their ideas, data, and tools to make decisions about their experiment and to construct knowledge about the physical phenomenon.

The directions and prompts to reduce uncertainty and test the limits of the model increase the likelihood that students encounter an empirical discrepancy. This discrepancy establishes a purpose for their experimental practices; in this context, ideally, students perceive a genuine need for careful consideration of uncertainty, statistical analysis, data visualization, and experimental design. At the same time, recognizing and responding to this discrepancy is not a requirement of the lab, and the assessment practices reflect that. The assessment emphasis is not on the content of their conclusions (whether they produce "the correct answer") but rather on their process and explanation.

Instructors expect that this activity will surface and cue students' expectations to verify authoritative knowledge and follow prescribed instructions. This approach aims to elicit and then respond to, if not confront, misaligned or improper expectations. At the end of the activity, through instructor feedback on their lab report and discussion in the associated lecture, students can reflect on their process and approach. Altogether, the lab activity aims to gradually transform students' expectations about what instructional labs are about and cultivate their agency. During this lab unit and throughout the course, the specificity of the instructions is scaled back so that students take on more responsibility for figuring out what to do.

B. UWB

The lab curriculum at UWB is an adaptation of the *Thinking Critically in Physics* curriculum described above. The instructional materials for the UWB pendulum lab can be accessed through QUBES CourseSource [32]. The design of this pendulum lab emphasizes meaningful, intentional use of experimental practices and the importance of collaboration and community feedback. The overarching goal of this activity, and of the lab course in general, can be summarized as supporting students to engage in intentional, purposeful activity and perceive their work as consequential for producing knowledge about physical

phenomena. Moreover, science is done with people, in community, with established (but transformable) social and epistemic norms. Collaboration, within groups and among the whole class, is a central feature of the curriculum.

Lab groups are tasked with generating a “team agreement” in the first week, and revisiting and updating it in the second week. The presence and clarity of a teamwork agreement in their notes is an explicit dimension of the grading rubric for the first two weeks. In the first lab session, lab instructors are encouraged to highlight a handout that discusses various potential roles one *could* take on during group work (e.g., skeptic, theorist, and principal investigator). Students also use a collaborative lab notebook in which students color-code their contributions to the document as they progress through the in-lab instructions.

The experimentation begins with a single pendulum for the whole class, whose period students measure after it is released from 10° and 20° . Importantly, students are given no instructions on how to do so. They are also asked to make a prediction about how the periods at 10° and 20° will compare, “based on the physics that you know.” Each individual measurement is recorded in a shared data table, and then the students (in their lab groups) are tasked with creating a histogram of these measurements to visualize the variability. This task also involves determining the precision of their measurement and calculating the mean and statistical uncertainty (standard deviation) of the data. This specific activity seeks both to direct student attention toward particular concepts (the role of data visualization and the physical meaning of uncertainty) and to foster student agency.

There is a trade-off pertaining to students’ agency here: It is not the students’ choice to construct a histogram. At the same time, constructing a histogram involves choice and deliberation. Students have to figure out what an appropriate bin size is and what “appropriate” means in this context. They are responsible for making the measurements in the first place, and once they construct a histogram, they are responsible for interpreting it. There is no predetermined correct answer their histogram is supposed to reveal; the purpose of the task is to do data interpretation and make sense of this phenomenon. Additionally, by directing students to create a histogram, they have a starting point or foothold in beginning to think about data visualization and the relationship between variation and uncertainty. The agentive trade-off aims to forestall the inhibitive effects of discomfort or frustration with not knowing what to do and to encourage choice and agency within the loose constraints of the activity.

The second week begins with a short activity on comparing datasets quantitatively, which first asks lab groups to invent a way to distinguish example measurements and then summarizes some key points using real-life examples connected to the interests of the instructor (e.g., when Scherr teaches lab, she discusses datasets around

classroom achievement in different instructional conditions). The invention activity primes students to notice and attend to particular quantities when interpreting and comparing data. Furthermore, both as an in-class, small group discussion and as part of their homework, students are asked, “What are two datasets for which knowing whether they are the same or not would make a difference to your life or your community?” This reflection question aims to highlight the utility and meaningfulness of the data interpretation practices they will be engaging in during the second (and third) lab sessions.

Following the minilecture and invention activity, the bulk of the experimental activity for the second lab revolves around reducing uncertainty. The students are tasked with using their data and experience to iteratively refine their experimental approach. The in-lab instructions direct students’ attention and activity toward certain concepts, such as measuring consecutive periods and collaborating with other groups, while also requiring explanations and interpretations of this work.

Additionally, the goal of reducing uncertainty provides a way to show students that *their activity* is consequential and important in this learning environment: at the end of the first week, students are asked to brainstorm ideas about how to reduce experimental uncertainty in measuring pendulum periods, and then in the second week they are told to pick one of those methods to implement. Many students do not seem to expect that their answers could inform their activity in the lab: they often list unhelpful options such as “eliminate human error” or “reduce air resistance.” The follow-up in the second week provides another opportunity for their coming to see lab as a place where their ideas are consequential.

The third week begins with an introduction to the t' statistic [87], which the students apply to the datasets they produced in the previous week. Instructions prompt students to interpret their calculated t' values and make decisions based on those interpretations. This again constrains their activity, but they remain responsible for generating interpretations, decisions, and conclusions. By introducing the t' statistic in this way and scaffolding its use, the lab activities provide students with multiple opportunities to use the t' statistic to derive meaning from and interpret their data.

Throughout, the lab instructions provide students with tools (e.g., a histogram template, a statistics vocabulary sheet, and the t' statistic) and then ask them to make meaning through their use. Even if the students do not initially recognize these tools or their use of them as meaningful, the repeated written prompts directing them to improve their experimental work serve to shift them in that direction. During lab meetings, the ways in which instructors engage, question, and respond to students similarly work to reinforce this framing, that the students’ use of these tools should have purpose and meaning.

After calculating t' and identifying ways to improve their experiment, students return to the overarching experimental question:

At this point, you are done with structured questions and ready to take the reins of your own investigation. Pick one of your proposed experiments and work with your group to design and carry out your own high-precision experiment. In your design, identify the possible outcomes and generate multiple possible explanations for each outcome. The goal is to answer the question: Does the period of your pendulum depend on the amplitude of the swing (for 10° vs 20°)?

The experimental prompt is neutral with respect to outcome. It provides no theoretical context about simple harmonic motion or Galileo's claims that favor specific predictions. Partway through the third lab session, the instructor sets up a shared data table for groups to fill in with their period measurements, uncertainty values, t' values, and conclusions. This provides a communal artifact that sparks intergroup discussion about results and conclusions.

It is not until after they have turned in their lab notes for grading that students are introduced to the simple harmonic model and associated period equation. Homework questions ask them, first, to restate their conclusion about the distinguishability of periods at 10° and 20° , and second, whether their conclusion aligns with the model. They are then asked to reflect on how knowing the "correct" result of an experiment can create bias, with a specific prompt to generate strategies to mitigate it.

C. Tufts

At Tufts, the written instructions change in small ways from semester to semester; examples are attached as Supplemental Material [33]. They are designed to create an environment in which students recognize the need for experimental practices, in particular, attention to precision and questioning authority. This fits an overarching goal for the course, including the "content" focused lecture portion, that students learn how to learn, including to assess the quality of their knowledge, to make decisions, and to take actions they see as reasonable. This perspective views students, especially at the start of the semester, as poised to frame lab as following prescribed steps toward predetermined correct results, "doing school" [29] in ways most have experienced. The labs try to promote "doing science," to shift responsibility onto students for figuring out how to approach, organize, and carry out their experimentation and for generating knowledge claims from that experimentation. To that end, the written materials are minimal, just half a page long for each of the first two sessions. They present the empirical challenge but provide no instructions on how to accomplish it.

The hope is that students feel empowered by the agency and responsibility afforded to them. Part of that, instructors anticipate, is students' learning to manage feelings of uncertainty over what to do, without the clarity of steps to follow. Ideally, they come to enjoy the autonomy, including the challenges of not knowing. That is, the labs are designed and taught in ways that may promote "meta-affective learning" [7,25]. In labs as in the courses as a whole, instructors and teaching assistants try to recognize confusion as productive, responding to students favorably (e.g., "great!") when they express feeling unsure and prompting them to elaborate (e.g., "talk about what has you confused").

The first activity asks students: "*How precisely can you measure the period of a pendulum?*" The instructions clarify the answer should be quantitative, "e.g., to within 0.2 s, or 0.1, 0.0000001, whatever you can manage." The TA shows students the materials available; there is a rich assortment. From there, they are responsible for building their own pendulums and deciding both how to be precise and how to quantify that precision. The session includes time for students to share their techniques and results.

Throughout the building of the pendulum and data collection, the TAs ideally focus on listening and supporting the substance of student ideas. Often students turn to standard deviation to quantify uncertainty, based on their experience in labs, and, in these cases, TAs respond along the lines of "What does standard deviation mean to you?" or "What does it tell us about data?" Most students are unclear about what it means; they see it as a black box algorithm they are "supposed to" use, a part of framing lab as "doing school." The TA should discourage using algorithms in this way; students should approach labs in ways that make sense to them.

With limited time, and to avoid prompting students' expectations of following prescribed steps, there is no instruction or guidance toward established statistical measures of variation or representational forms. A tacit element of the reasoning for this choice, which this comparative analysis brought to light, is a belief that students have sufficient capabilities with statistics and data representation to generate sensible approaches. If there were more time available, it would be possible to follow students' inventions with explicit instruction, as in *Thinking Critically in Physics*.

The instructions for the second week present the task as follows:

[During the previous lab session], we heard conjectures that have history to them: There were some groups saying the angle (or the "amplitude") of the pendulum doesn't affect the period; others said it does. Galileo Galilei (1564–1642) studied this and claimed the period of a pendulum does not depend on amplitude, or on the mass of the weight (called the 'bob'). The challenge today

is to study these questions: To the precision you are able to measure, does the amplitude (angle) affect the period? Does the mass of the bob?

This is the Spring 2024 version of the instruction sheet. Every year, the lab teaching team reads and considers possible revisions. Indeed, the first sentence above, about student conjectures, was written in response to discussions TAs reported having in their sections. This allows the instructions to be responsive to student ideas and builds up the agency of the TAs themselves [77]. A year-over-year consistent feature of this experimental prompt is the mention of Galileo’s claims.

Semantically, the question is not confirmatory, and the following paragraph begins, “Please take care to be honest about your data and your findings!” (That sentence was added based on observations in previous semesters of students setting their own evidence aside.) Still, the inclusion of Galileo’s name in the instructional prompt typically leads students to expect to confirm his claims (the famous scientist must be correct!). That expectation sets up an empirical discrepancy. In most sections, some lab groups will produce data that indicate (to them and/or to a knowing instructor) that the period does change with amplitude. This potential discrepancy generates a need for experimental practices: data visualization, uncertainty estimation, and an understanding of the *production* of data have functional relevance to the genuine need to make sense of this discrepancy.³

In the third session, lab groups present and discuss their conclusions.⁴ This discussion is termed a “meaning-making session,” where the whole class attempts to come to a consensus regarding the experimental question. The instruction sheet for this session includes a translated quote from the 11th-century Arab scientist and mathematician al-Hasan ibn al-Haytham, describing how seekers of truth should only submit to argument and demonstration and “make himself an enemy of all that he reads” [88]. The instruction sheet explicitly says that the values of reasoning and empirical evidence, as opposed to deference to authority, are core to the course. In addition to presenting their work to their peers, students discuss and reflect on this quote in light of their work in the first two weeks.

Throughout the three-week activity, the instructional team emphasizes that they do not grade for the correctness of students’ results but for the quality of their engagement in conducting their investigation. When the labs are complete, with comments and scores ready for students, the instructor brings up the topic in the main course lecture, first to affirm the independence with respect to mass (and to

discuss how that coheres with Newtonian theory), and second to consider the variation with amplitude. Almost always, students’ data provide the basis for claiming there is a dependence, which sparks a discussion about the role of authority in science.

V. COMPARISONS

We all begin the semester with a pendulum lab we design (and redesign) based on a resources-based view of student knowledge, reasoning, and framing, with shared objectives regarding epistemological framing. Labs are where students can and should engage in constructing knowledge through empirical study of physical phenomena. We all emphasize to students that the goal is not “getting the right answer,” as they might expect, but rather learning to engage in high-quality, intellectually rigorous experimentation. However, we design these labs in different ways. Table I lists the differences across our curricula that we identified in this work.

In this section, we compare our approaches with respect to presentation, written guidance, and responsive teaching. We selected these aspects of the curricula to compare because they connect tightly to our goals of student agency and shared theoretical commitments regarding framing and agency. To make sense of the differences in our curricula, we focus on the *design reasoning* involved in our decisions.

A. Presentation of the task

At Cornell, the instructions present an equation for the period, described as part of the simple harmonic model, along with some explanation of what a model is. At Tufts, the instructions present Galileo’s claims that the period does not depend on amplitude or mass. Additionally, at Cornell, the lab unit as a whole is introduced as a model-testing unit from the outset, while at Tufts Galileo’s claims are introduced only in the second of three sessions, with reference to student ideas or findings from the first session. At UWB, the prompt is neutral, simply asking students to investigate whether the period of a pendulum depends on the amplitude.

Part of the reasoning for providing a model that students are meant to test, as at Cornell and Tufts, is to cue up expectations of confirmation. For Hammer, the mention of Galileo sets up an eventual reflective discussion about the role of authority in science and the importance of evidence and argumentation. For Holmes, the similar approach with a mathematical model sets up an epistemological version of “productive failure” [89]. The expected discrepancy between the provided model and students’ data generates an epistemologically rich context for underscoring the importance of empirical model-testing. One reason for the more granular in-lab instructions at Cornell is to guide students to find the discrepancy.

³The expectation to confirm Galileo may also reflect or cue students’ framing lab as doing school, but see Ref. [9] for evidence it can support productive experimental practices.

⁴In 2026, as we prepare this manuscript, the lab team has decided to keep the pendulum lab to two sessions.

TABLE I. Overview of differences among the three lab activities.

Curriculum feature	Cornell	UWB	Tufts
Experimental question	Model testing, asked at beginning of unit	Neutral phrasing, asked at beginning of unit	Galileo testing, asked in second session
Statistics and representation guidance	Instructions prompt students to make and use histograms and t'	Instructions prompt students to make and use histograms and t'	No explicit written guidance
Medium of materials	Instructions and lab notes in CoLab or Jupyter Notebook	Instructions in view-only documents, lab notes in jointly edited documents	Instructions in PDFs, lab notes as shareable documents
Lab homework	Tutorials on statistics and uncertainty	Reflection questions	None
Group discussions	Dedicated lecture time and at the discretion of lab instructor	Minilectures at the beginning of lab sessions and planned check-ins throughout	Third lab session is dedicated to a whole group discussion
Collaboration support	Partner agreement and roles handout	Partner agreement and roles handout	None
Instructors	Graduate TAs	Professors	Graduate and undergraduate TAs
Course logistics	20 students per section, sessions last two hours	24 students per section, sessions last two hours	14 students per section, sessions last two and a half hours

An additional aspect of the reasoning for Hammer and Holmes is an expectation, borne out by their years of teaching these labs, that many students will have some awareness of the pendulum period equation and/or Galileo. In Scherr's experience, that expectation does not apply for students at UWB. She chooses a neutral phrasing so that the curriculum emphasizes empirical science as constructing facts in a community with shared standards. When disagreements about conclusions arise—as is a common occurrence in all our labs—the UWB curriculum still supports students in discussing notions of trust, accountability, and authority. Absent a reference to external authority, that discussion remains rooted in the students' own data and experiences.

We all agree on the important epistemological expectations we aim to communicate to students in the first lab. We all present the task, with an elicit-confront-resolve approach or neutral phrasing, as a way to support students' framing the lab as centered on their reasoning. The differences reflect our perceptions of the student populations in our courses. They also reflect differences in personal preferences: model-testing is a central and specific idea for the designers at Cornell, while at Tufts and UWB the focus is on constructing scientific knowledge and “science as a refinement of everyday thinking.”

B. Written guidance

One of the main differences among our curricula—indeed the initial inspiration for this paper—is the

written guidance we provide (or do not provide) to students. The lab instructions at Cornell and UWB specify particular tasks and activities; the instructions at Tufts do not.

For Holmes and Scherr, their more detailed instructions address a concern that students may feel lost or overwhelmed to a degree that would inhibit their learning. The concern for Hammer, on the other hand, is that the length and specifications of instructions can cue students' “doing school” framing. If there are no instructions to parse, the possibility of the instructions cuing or stabilizing that framing is not possible.

Holmes and Scherr take the perspective that this common framing can be a potentially productive resource: use student expectations that doing lab means following instructions, and design those instructions carefully to direct them toward productive activity, such as to make a histogram. In subsequent sessions, instructions become sparser, giving students more responsibility for their decisions.

Having the students use t' to quantify distinguishability plays out differently at UWB and Cornell. At UWB, the instructions guide students to consider the possibility of a subtle dependence. At Cornell, finding that dependence conflicts with the model they likely expected to confirm. At both institutions, it is up to the students to decide what to conclude. In this way, both curricula use students' willingness to “do school” as a way to scaffold productive activities, reducing the guidance over time (within the first lab activity and across the term as a whole) to cultivate agency.

Here, we see a difference in our theoretical stances toward student behavior and learning, within the same general framework. We differ in our expectations about the dynamics of students' framing, in particular over when and how they may come to frame lab as "doing science" for themselves. Can students' willingness to adhere to instructions—as "jumping through hoops" [6] or "doing school" [29]—support their progress toward productive scientific work, or is it something to disrupt from the start? We also have different perspectives on the risks of students' feeling lost or frustrated: Scherr and Holmes view these emotions as a barrier to engagement and participation, whereas Hammer hopes that students grow from wrestling with such feelings, that the lab is an opportunity for meta-affective learning [25]. (We all, meanwhile, agree it is important that students feel safe to try things and explore, including to trust that their grades are not based on arriving at prescribed outcomes.)

Recent work in these labs has highlighted the complexity of the epistemic and affective dynamics of student engagement. Studying labs at Cornell, Sundstrom *et al.* [8] observed that successful instructor interventions to shift students away from confirmation framing were not always associated with an increase in measures of student agency. And, studying labs at Tufts, Descamps *et al.* [9] observed students' confirmation framing was not always associated with a lack of agency; in fact, it seemed to drive productive practices of investigation.

Additionally, Jeon *et al.* [90] highlight a student's complex affective experiences of confusion in the Tufts labs: she insisted that she "hates physics," during side conversation with her lab mates, because their confusing experimental results defied her expectations that physics should align with common sense. At the same time, this student's feelings of confusion motivated her to take up more epistemic agency during the lab. Thinking in terms of a trade-off between giving students autonomy and managing their feelings of uncertainty might be a useful heuristic for design decisions, at least within nontraditional labs. Data from these studies of student engagement, however, suggest that the effects of different choices are not so clear-cut as "trade-off" might imply.

As theoretical constructs, framing and agency provide insight into various kinds of design decisions, but they can be taken up in ways that lead to different outcomes. One might focus on decision-making agency [64] or on framing as a phenomenon of an individual's cognition, but these are far from the only ways to understand and apply framing and agency to pedagogy. Additionally, we cannot ignore the influence of secondary and contextual factors in the divergence of our in-lab guidance structures. At Cornell and UWB, the instructions provide guidance that allows students to do "miniactivities" aimed at building up collaboration skills. The minimal lab instructions at Tufts have the effect of shifting various aspects of classroom management onto the TAs, which is aligned with

Hammer's active interest in TA education, which we take up in the next section.

C. Responsive teaching

We all encourage responsiveness, but, with minimal written guidance, the Tufts labs depend more critically on the instructors' practices of attending, interpreting, and responding to their students' thinking. In part, the differences in our expectations of instructors reflect differences in expectations for the materials. *Thinking Critically in Physics Labs*, the core of the materials used at Cornell and UWB, is designed for wide dissemination and use. It is difficult in general to rely on instructor responsiveness, and so these materials provide more detailed in-lab instructions to guide students toward the intended, rich arenas of experimental inquiry: making useful plots, quantifying uncertainty, refining methods, generating conclusions, and sifting through empirical conflicts.

In contrast, the Tufts labs are only for use at Tufts. That has the advantages and disadvantages of the instructors' freedom and responsibility to use their judgment as they work with students. We see this as a second "agentic trade-off" that comes with the more detailed in-lab instructions, analogous to the trade-off for students: The instructors, too, are guided and constrained with respect to what they may notice and how they may respond.

At UWB, Scherr expects the instructors, who are all faculty, to act with autonomy and agency, including noticing, assessing, and responding to student work during implementation. Scherr's own modifications to the *Thinking Critically in Physics Labs* curriculum reflect this aspect of her context. For example, when Scherr introduces the activities around quantitatively comparing datasets, she uses examples from her personal and research interests, while other instructors use different examples. Still, the detailed in-lab instructions function to provide a clear representation of pedagogical intentions, while at the same time taking a familiar form, and in these ways may facilitate instructors' adoption.

At Cornell, the instructors are graduate TAs, who come in with a wide range of experience and interest in teaching, familiarity with research-based pedagogical approaches, beliefs about how learning does and should happen, and confidence in the classroom. Moreover, as with the faculty at UWB, there is limited room for their preparation or supervision, although student TAs are not free to modify the materials. The detailed in-lab instructions function to minimize variation across the TAs and to make it easier for them to implement the curriculum as intended.

At Tufts, in contrast, part of the objective for Hammer is TA education and preparation. The TAs are a mix of graduate students and undergraduates, the latter often coming to the position after successful experiences as learning assistants [91] in the highly interactive lectures. The minimal written instructions have the effect of shifting responsibility to the TAs for noticing, interpreting, and

responding to students. Often this involves helping the students to shift their framing of lab, such as when students look to the TA for prescriptive guidance on how to proceed.

Hammer's practice has been to meet with the team of lab TAs for several hours each Friday. The TAs recount snippets of what they noticed in their students' participation over the week, and they choose examples of students' written work. They compare interpretations, following practices of educator support for responsive teaching [36], and discuss ideas for how they might respond. They also discuss plans for the coming week, including suggesting refinements to the coming assignment's wording and design. Experience at Tufts makes it clear that these meetings play an important role in ensuring that TAs implement the curriculum as intended and develop skills in responsive teaching.

VI. REFLECTIONS AND DISCUSSION

As we discussed in the Introduction, the purpose of our work here is to wade into the gap between our theoretical perspectives and our curriculum designs. Like Boudreaux and Elby [10], the inspiration for our comparison, we compare our designs with the hope of elucidating how our theoretical orientations shape the curricula. While they sought to connect the differences in their theoretical frameworks to the differences in their designs, we have sought to make sense of the differences in our design despite the alignment of our theoretical perspectives.

First, we review the similarities, many of which reflect our theoretical alignment. We all aspire to promote student epistemic agency within the process of experimentation and to center the substance of student ideas. We all work to make sure the experiments in our labs are not exercises in confirming or verifying concepts from lecture; we try to shift students away from problematic framing; we expect instructors to be responsive to the substance of student ideas; we build into our labs the generation and use of data representations; we highlight the need for some way to quantify uncertainty, and, in general, we aim to communicate to students that their ideas and their data are of central importance.

These similarities reflect a shared understanding of how framing can influence the agentive element of student behavior, of how, when given the chance, students can be capable of productive scientific work, and of how a wide variety of contextual cues can tip students into different framings. These similarities and the underlying theoretical connective tissue between them punctuate the motivation to consider how and why we arrive at designs with such substantive differences.

A. Presentation of activity

Above, we highlight how each of our labs presents the activity to students: At UWB, the experimental question is

neutral; at Cornell, the experiment is presented as a model-testing activity; at Tufts, students are asked to investigate the claims of a renowned authority. In the Sec. VA, we connected these differences to variations in the expectations we seek to cultivate in students.

The neutral phrasing of the UWB instructions clearly focuses the experimental work on the students. We expect the references to the simple harmonic model (Cornell) and Galileo (Tufts) to elicit potentially problematic expectations about verifying already-known, authoritative theoretical claims, which allows us to then confront and disrupt these expectations. Moreover, the reference to these models sets students up to experience a disagreement between theory and experiment, which in turn generates a genuine (for the students) need to engage in experimental practices such as working to reduce uncertainties, use and interpret statistical tools to distinguish datasets, and extend experiments to test emergent questions.

B. Structure of guidance

The instructions at Cornell also have the effect of scaffolding students' collecting data that will indicate a theoretical-empirical discrepancy. Tufts takes a different, minimalist approach to written instructions. In Sec. VB, we highlight how our shared understandings of framing, agency, and affect inform our differing decisions about instructions. To an extent, the difference in our approaches to in-lab instructions indicates a difference in theoretical perspective—or perhaps a difference in how we understand the instructional implications of theory.

For Hammer, students' framing the activity as “doing school” [29] is a primary concern that drives many of the decisions affecting the structure of in-lab guidance. For Scherr and Holmes, while this framing certainly presents a concern, they take the approach of using students' expectations to follow instructions as a way to nudge students toward productive arenas of experimental work and scientific reasoning. That is, students themselves may not initially see the point of making a histogram and so do so because the instructions tell them to, but thinking about how to best make a histogram and comparing and interpreting that histogram provide a rich context for reasoning about and with representations.

There is an additional, affective dimension to Scherr's and Holmes's decisions around in-lab guidance: the detailed instructions seek to ensure that students do not feel lost or overwhelmed. To be clear, while the instructions at Cornell and UWB are more detailed *relative to Tufts*, they are much less detailed than those of a “cookbook” lab. Indeed, much of their design work here involves writing open-ended prompts that provide guidance to students while also shifting decision making onto students.

Across the curricula, the designs reflect different choices about an agentive trade-off in the level of detail to provide. There is also an “affective trade-off,” with respect to

students' experiences of uncertainty, greater at Tufts with less written guidance. Students learning to manage and regulate potential feelings of discomfort—meta-affective learning [25] is itself a desired outcome. While the in-lab instructions at Cornell and UWB provide support for students with respect to agency and affect, at Tufts that falls much more to the instructor, a graduate or undergraduate TA. In this way, TAs at Tufts are responsible and empowered to be more flexible and adaptable in how they respond to specific students and emergent idiosyncrasies.

Of course, the lab instructors at Cornell and UWB are also empowered and able to respond to students, but the onus is reduced. Part of the reasoning for more granular instructions at Cornell and UWB is that they minimize the potential negative influence of inexperienced TAs or instructors with discordant pedagogical beliefs. Indeed, the Tufts lab design places a heavy demand on TAs and experience has shown the importance of supporting their learning as instructors. This is another form of trade-off: Maintaining that support in the face of varying enrollments (and thus numbers of TAs) and levels of engagement and skill of both the TAs and the faculty assigned to that role has proved to be challenging. There has been noticeable variation in the effectiveness of the labs across different semesters.

Finally, we note a key difference regarding curricular objectives with respect to statistical analysis. At UWB and Cornell, much of the guidance within the materials concerns the definition and use of the t' statistic. At Tufts, there is no “content” objective in this lab with respect to canonical statistics. Rather, the objective is entirely epistemological: that students recognize and take on the challenge of quantifying uncertainty. The difference in course structure (stand-alone lab course versus component of a lecture course) plays a role in this decision making: the stand-alone structure affords opportunities to engage in invention activities about statistics.

C. Models of students

Moving forward, one area in particular where this exercise in comparison has provided inspiration for each of us is the degree to which our design decisions are informed by often unstated “working models” of the students. At Cornell and Tufts, experience has taught Holmes and Hammer that students in their labs will have heard of both Galileo and the simple harmonic period equation, while that is not Scherr's experience at UWB. These considerations were consequential in how we operationalized our theoretical commitments. Scherr and Holmes both have a sense of the affective consequences for their students of a lack of in-lab instructions: in part, the guiding nature of their in-lab instructions is designed to meet students where they are.

We refer to these expectations as working models to reflect both their unfinished and partially tacit nature and to highlight these expectations as connected to our theoretical

commitments. Implicit in our desire to shift students' framings is an expectation of how students will think and act coming into our labs. Research has shaped our expectations [2–5,9,14,92–95], along with our experiences in and out of the classroom. Still, as we engaged in our comparisons, we often found ourselves needing to explicate and clarify our understandings of the students. In reflecting on the ways in which the Cornell and UWB labs structure student engagement with statistical and representational tools, the Tufts instructional team realized that the lab design was shaped by unstated expectations about students' abilities to generate and use novel representational forms. This motivated a redesign of the third lab session to include more explicit attention to data representation and analysis.

Additionally, in discussing our various expectations of students—as individuals—we also recognized how structuring collaborative groups remains an open question. Collaboration is built into our curriculum on a fundamental level, and at each school, we engage in a handful of strategies to promote the emergence of equitable and supportive small groups. Still, we are all uncertain how to best to support groups' working effectively, as well as how to assess the “effectiveness” of collaboration.

VII. CLOSING THOUGHTS

The differences among our lab activities demonstrate that a theoretical framework of epistemic agency and framing can engender multiple, distinct forms of pedagogical structures. Our labs provide a sample, but they certainly do not span the entire design space. There is plenty of opportunity, for curriculum designers as for instructors, to innovate, to explore, to try things out, as well as to conduct scholarship that deepens our understanding of what works. Given the heterogeneity of national and global lab contexts [96,97], there is not only opportunity but also necessity, for local inquiry by instructors, as well as for the development of curricula that are flexible and adaptable to local circumstances.

In reflecting on how curriculum developers and the broader PER ecosystem can work toward more useful research outputs, Elby and Yerdelen-Damar introduced the notion of *instructionally generative fodder*: descriptions of curriculum, related research of student activity and learning in designed contexts, and more general research that provides actionable insight into learning and how to support it [42]. To optimize the generativity of curricular products, they suggest that developers make the rationale for design decisions available and clear to instructors, including points of uncertainty.

In recent years, there has been a push for PER scholars to include more detailed and more useful descriptions of theory in curriculum development. In doing this work, researchers have commonly focused on design decisions and choices—making more decisions explicit and connecting them to theoretical frameworks. We suggest a slight shift in focus, to emphasize instead *design reasoning*. That

is, descriptions of curriculum should bring to the fore the reasoning involved in how design decisions are made and how a curriculum is expected to mediate learning. This suggestion is largely a rephrasing and synthesis of Elby and Yerdelen-Damar's [42] discussion of generativity in curricular products; it also echoes other arguments that it is important for teachers to understand the purposes and aims of curricular materials [37–41,43].

Consider how the experimental question is posed in our three pendulum lab activities. Focusing on design decisions, one might say that our understanding of student agency and framing guided our decisions about how to phrase the experimental question. This explains the decision, but provides no insight into how we arrived at different phrasings. Focusing on the design reasoning requires us to discuss expectations of students (at Cornell and Tufts) to be aware of the pendulum period equation, to describe how these questions lead into an epistemologically productive failure and an elicit-confront-resolve approach, and to articulate how each of these phrasings works to emphasize the substance of students' data and experimental work.

A shift to design reasoning also echoes the argument that Brookes *et al.* [56] outline in their paper on the motivations for ISLE. Building from educational theorists MacMillan and Garrison [82], they “suggest that if we are to make better connections between education theory and classroom implementation, both researchers and implementers need to articulate their underlying intentionalities more explicitly.”

Robertson *et al.* [43] illustrate the utility of knowledge of what we are calling design reasoning in a study of learning assistants in a university-level physics course. They specifically provide evidence that “knowledge of the purposes of questions or sequences of questions within [research-based instructional materials] [...] can support (or serve as a resource in) the enactment of responsive instruction” (p. 151). Robertson *et al.* build from a long tradition of education scholars who maintain that making visible the “pedagogical judgments” [40] and “hidden decisions”

[41] of curricular materials can support teachers in enacting theoretically oriented and research-based curricula [37–39,98]. We extend this claim to recommend that curricular materials and research on curriculum development in PER ought to foreground design reasoning.

Our examination of our own curricula helped us recognize and articulate aspects of the reasoning that play significant roles in our developing and refining activities and materials. We contend that a focus on design reasoning can help make descriptions of curricula more generative for practitioners. Design reasoning can also be generative for researchers: expectations about how a curriculum will support learning should inform studies of what takes place. Given that research-based pedagogies are taken up in noticeably different ways (if they are taken up at all) [73–75], future research should explicate the consequential aspects of curricula and crucial interactions.

ACKNOWLEDGMENTS

We would like to thank the faculty and staff at Cornell, UWB, and Tufts involved in developing and implementing these labs. We acknowledge the Cornell University College of Arts and Sciences Active Learning Initiative for providing support for part of the development of the *Thinking Critically in Physics* curriculum. We also thank members of the Tufts STEM Education program, in particular Sophia Jeon, Miguel Vasquez-Vega, Ira Caspari, and the students in the Qualitative Research in STEM Education course for feedback on initial comparative analyses. Finally, we appreciate the thorough feedback of our anonymous reviewers, in particular, the connections to Robertson *et al.* [43].

DATA AVAILABILITY

The data that support the findings of this article are openly available [32,33,86].

-
- [1] G. Galilei, *Dialogues Concerning Two New Sciences* (Macmillan, New York, 1638).
 - [2] E. M. Smith, M. M. Stein, and N. G. Holmes, Surprise! Shifting students away from model-verifying frames in physics labs, *presented at PER Conf. 2018*, [10.1119/perc.2018.pr.Smith](https://doi.org/10.1119/perc.2018.pr.Smith).
 - [3] E. M. Smith, M. M. Stein, and N. G. Holmes, How expectations of confirmation influence students' experimentation decisions in introductory labs, *Phys. Rev. Phys. Educ. Res.* **16**, 010113 (2020).
 - [4] E. M. Smith and N. G. Holmes, Let's just pretend": Students' shifts in frames during a content-reinforcement lab, *presented at PER Conf. 2020*, pp. 484–489, [10.1119/perc.2020.pr.Smith_M](https://doi.org/10.1119/perc.2020.pr.Smith_M).
 - [5] M. Sundstrom, A. M. Phillips, and N. G. Holmes, Problematizing in inquiry-based labs: How students respond to unexpected results, *presented at PER Conf. 2020*, pp. 539–544, [10.1119/perc.2020.pr.Sundstrom](https://doi.org/10.1119/perc.2020.pr.Sundstrom).
 - [6] A. M. L. Phillips, M. Sundstrom, D. G. Wu, and N. G. Holmes, Not engaging with problems in the lab: Students' navigation of conflicting data and models, *Phys. Rev. Phys. Educ. Res.* **17**, 020112 (2021).
 - [7] S. Jeon, I. Descamps, and D. Hammer, “I guess my question is like”: Problematizing in an introductory college

- physics lab, in *Proceedings of the 17th International Conference of the Learning Sciences—ICLS 2023* (International Society of the Learning Sciences, 2023), <https://repository.isls.org/handle/1/9848>.
- [8] M. Sundstrom, R. K. Fussell, A. M. L. Phillips, M. Akubo, S. E. Allen, D. Hammer, R. E. Scherr, and N. G. Holmes, Instructing nontraditional physics labs: Toward responsiveness to student epistemic framing, *Phys. Rev. Phys. Educ. Res.* **19**, 020140 (2023).
 - [9] I. Descamps, S. Jeon, N. G. Holmes, R. E. Scherr, and D. Hammer, Dynamics of productive confirmation framing in an introductory lab, *Phys. Rev. Phys. Educ. Res.* **20**, 020111 (2024).
 - [10] A. Boudreaux and A. Elby, How curriculum developers' cognitive theories influence curriculum development, *Phys. Rev. Phys. Educ. Res.* **16**, 020144 (2020).
 - [11] 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, Greenwich, CT, 2005).
 - [12] A. A. diSessa, Toward an epistemology of physics, *Cognit. Instr.* **10**, 105 (1993).
 - [13] D. Tannen, *Framing in Discourse* (Oxford University Press, New York, 1993).
 - [14] R. E. Scherr and D. Hammer, Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics, *Cognit. Instr.* **27**, 147 (2009).
 - [15] 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).
 - [16] M. J. Ford, The game, the pieces, and the players: Generative resources from two instructional portrayals of experimentation, *J. Learn. Sci.* **14**, 449 (2005).
 - [17] B. W. Frank and R. E. Scherr, Interactional processes for stabilizing conceptual coherences in physics, *Phys. Rev. ST Phys. Educ. Res.* **8**, 020101 (2012).
 - [18] D. G. Wu, A. B. Heim, M. Sundstrom, C. Walsh, and N. G. Holmes, Instructor interactions in traditional and nontraditional labs, *Phys. Rev. Phys. Educ. Res.* **18**, 010121 (2022).
 - [19] R. E. Scherr, R. S. Russ, T. J. Bing, and R. A. Hodges, Initiation of student-TA interactions in tutorials, *Phys. Rev. ST Phys. Educ. Res.* **2**, 020108 (2006).
 - [20] L. D. Conlin and R. E. Scherr, Making space to sensemake: Epistemic distancing in small group physics discussions, *Cognit. Instr.* **36**, 396 (2018).
 - [21] M. Zhang, K. Ansell, and E. Kuo, Navigating moments of uncertainty and socio-emotional risks in small-group work, presented at *PER Conf. 2021*, pp. 468–473, [10.1119/perc.2021.pr.Zhang](https://doi.org/10.1119/perc.2021.pr.Zhang).
 - [22] J. Watkins, D. Hammer, J. Radoff, L. Z. Jaber, and A. M. Phillips, Positioning as not-understanding: The value of showing uncertainty for engaging in science, *J. Res. Sci. Teach.* **55**, 573 (2018).
 - [23] L. Appleby, V. Dini, L. Withington, E. LaMotte, and D. Hammer, Disciplinary significance of social caring in postsecondary science, technology, engineering, and mathematics, *Phys. Rev. Phys. Educ. Res.* **17**, 023106 (2021).
 - [24] J. Gouvea, L. Appleby, L. Fu, and A. Wagh, Motivating and shaping scientific argumentation in lab reports, *CBE Life Sci. Educ.* **21**, ar71 (2022).
 - [25] J. Radoff, L. Z. Jaber, and D. Hammer, “It’s scary but it’s also exciting”: Evidence of meta-affective learning in science, *Cognit. Instr.* **37**, 73 (2019).
 - [26] V. Dini and D. Hammer, Case study of a successful learner’s epistemological framings of quantum mechanics, *Phys. Rev. Phys. Educ. Res.* **13**, 010124 (2017).
 - [27] N. G. Holmes, D. Kumar, and D. A. Bonn, Toolboxes and handing students a hammer: The effects of cueing and instruction on getting students to think critically, *Phys. Rev. Phys. Educ. Res.* **13**, 010116 (2017).
 - [28] K. N. Quinn, M. M. Kelley, K. L. McGill, E. M. Smith, Z. Whipps, and N. G. Holmes, Group roles in unstructured labs show inequitable gender divide, *Phys. Rev. Phys. Educ. Res.* **16**, 010129 (2020).
 - [29] M. P. Jiménez-Aleixandre, A. Bugallo Rodríguez, and R. A. Duschl, “Doing the lesson” or “doing science”: Argument in high school genetics, *Sci. Educ.* **84**, 757 (2000).
 - [30] Z. Y. Kalender, E. Stump, K. Hubenig, and N. G. Holmes, Restructuring physics labs to cultivate sense of student agency, *Phys. Rev. Phys. Educ. Res.* **17**, 020128 (2021).
 - [31] M. Scardamalia and C. Bereiter, Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media, *J. Learn. Sci.* **1**, 37 (1991).
 - [32] R. E. Scherr and N. G. Holmes, Quantifying uncertainty and distinguishing data sets in introductory physics, *CourseSource* **10** (2023).
 - [33] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/6g9j-s2cm> for the instructional materials for the pendulum lab activity at Tufts University.
 - [34] N. G. Holmes, J. Olsen, J. L. Thomas, and C. E. Wieman, Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content, *Phys. Rev. Phys. Educ. Res.* **13**, 010129 (2017).
 - [35] E. M. Smith and N. G. Holmes, Best practice for instructional labs, *Nat. Phys.* **17**, 662 (2021).
 - [36] *Responsive Teaching in Science and Mathematics*, edited by A. Robertson, R. Scherr, and D. Hammer (Routledge, New York City, NY, 2015).
 - [37] L. S. Shulman, Those who understand: Knowledge growth in teaching, *Educ. Res.* **15**, 4 (1986).
 - [38] L. Shulman, Knowledge and teaching: Foundations of the new reform, *Harv. Educ. Rev.* **57**, 1 (1987).
 - [39] D. L. Ball and D. K. Cohen, Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform?, *Educ. Res.* **25**, 6 (1996).
 - [40] E. A. Davis and J. S. Krajcik, Designing educative curriculum materials to promote teacher learning, *Educ. Res.* **34**, 3 (2005).
 - [41] D. B. Harlow, Structures and improvisation for inquiry-based science instruction: A teacher’s adaptation of a model of magnetism activity, *Sci. Educ.* **94**, 142 (2010).

- [42] A. Elby and S. Yerdelen-Damar, Rethinking the relationship between instructors and physics education researchers, *Phys. Rev. Phys. Educ. Res.* **16**, 020151 (2020).
- [43] A. D. Robertson, K. E. Gray, C. E. Lovegren, K. L. Killough, and S. T. Wenzinger, Curricular knowledge as a resource for responsive instruction: A case study, *Cognit. Instr.* **39**, 149 (2021).
- [44] V. K. Otero and D. E. Meltzer, The past and future of physics education reform, *Phys. Today* **70**, No. 5, 50 (2017).
- [45] F. Reif and M. St. John, Teaching physicists' thinking skills in the laboratory, *Am. J. Phys.* **47**, 950 (1979).
- [46] R. K. Thornton, Tools for scientific thinking-microcomputer-based laboratories for physics teaching, *Phys. Educ.* **22**, 230 (1987).
- [47] J. M. May, Historical analysis of innovation and research in physics instructional laboratories: Recurring themes and future directions, *Phys. Rev. Phys. Educ. Res.* **19**, 020168 (2023).
- [48] E. Etkina, A. Karelina, M. Ruibal-Villasenor, D. Rosengrant, R. Jordan, and C. E. Hmelo-Silver, Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories, *J. Learn. Sci.* **19**, 54 (2010).
- [49] E. M. Smith, M. M. Stein, C. Walsh, and N. G. Holmes, Direct measurement of the impact of teaching experimentation in physics labs, *Phys. Rev. X* **10**, 011029 (2020).
- [50] B. R. Wilcox and H. J. Lewandowski, Impact of instructional approach on students' epistemologies about experimental physics, *presented at PER Conf. 2016, Sacramento, CA*, pp. 388–391, [10.1119/perc.2016.pr.092](https://doi.org/10.1119/perc.2016.pr.092).
- [51] B. R. Wilcox and H. J. Lewandowski, Open-ended versus guided laboratory activities: Impact on students' beliefs about experimental physics, *Phys. Rev. Phys. Educ. Res.* **12**, 020132 (2016).
- [52] B. R. Wilcox and H. J. Lewandowski, Students' views about the nature of experimental physics, *Phys. Rev. Phys. Educ. Res.* **13**, 020110 (2017).
- [53] C. Walsh, H. J. Lewandowski, and N. G. Holmes, Skills-focused lab instruction improves critical thinking skills and experimentation views for all students, *Phys. Rev. Phys. Educ. Res.* **18**, 010128 (2022).
- [54] D. R. Sokoloff, P. W. Laws, and R. K. Thornton, RealTime physics: Active learning labs transforming the introductory laboratory, *Eur. J. Phys.* **28**, S83 (2007).
- [55] R. K. Thornton and D. R. Sokoloff, Learning motion concepts using real-time microcomputer-based laboratory tools, *Am. J. Phys.* **58**, 858 (1990).
- [56] D. T. Brookes, E. Etkina, and G. Planinsic, Implementing an epistemologically authentic approach to student-centered inquiry learning, *Phys. Rev. Phys. Educ. Res.* **16**, 020148 (2020).
- [57] N. G. Holmes and D. A. Bonn, Doing science or doing a lab? Engaging students with scientific reasoning during physics lab experiments, *presented at PER Conf. 2013, Portland, OR*, pp. 185–188, [10.1119/perc.2013.pr.034](https://doi.org/10.1119/perc.2013.pr.034).
- [58] N. G. Holmes, J. Day, A. H. K. Park, D. A. Bonn, and I. Roll, Making the failure more productive: Scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities, *Instr. Sci.* **42**, 523 (2014).
- [59] N. G. Holmes, J. Ives, and D. A. Bonn, The impact of targeting scientific reasoning on student attitudes about experimental physics, *presented at PER Conf. 2013, Minneapolis, MS*, pp. 119–122, [10.1119/perc.2014.pr.026](https://doi.org/10.1119/perc.2014.pr.026).
- [60] N. G. Holmes and E. M. Smith, Operationalizing the AAPT learning goals for the lab, *Phys. Teach.* **57**, 296 (2019).
- [61] H. J. Lewandowski, D. R. Bolton, and B. Pollard, Initial impacts of the transformation of a large introductory lab course focused on developing experimental skills and expert epistemology, *presented at PER Conf. 2018, Washington, DC*, [10.1119/perc.2018.pr.Lewandowski](https://doi.org/10.1119/perc.2018.pr.Lewandowski).
- [62] A. Werth, C. G. West, and H. J. Lewandowski, Impacts on student learning, confidence, and affect in a remote, large-enrollment, course-based undergraduate research experience in physics, *Phys. Rev. Phys. Educ. Res.* **18**, 010129 (2022).
- [63] K. Funkhouser, W. M. Martinez, R. Henderson, and M. D. Caballero, Design, analysis, tools, and apprenticeship (DATA) lab, *Eur. J. Phys.* **40**, 065701 (2019).
- [64] N. G. Holmes, B. Keep, and C. E. Wieman, Developing scientific decision making by structuring and supporting student agency, *Phys. Rev. Phys. Educ. Res.* **16**, 010109 (2020).
- [65] E. Brewster, L. Kramer, and G. O'Brien, Modeling instruction: Positive attitudinal shifts in introductory physics measured with CLASS, *Phys. Rev. ST Phys. Educ. Res.* **5**, 013102 (2009).
- [66] J. Biggs, Enhancing teaching through constructive alignment, *Higher Educ.* **32**, 347 (1996).
- [67] H. J. Lewandowski, R. Hobbs, J. T. Stanley, D. R. Dounas-Frazier, and B. Pollard, Student reasoning about measurement uncertainty in an introductory lab course, *presented at PER Conf. 2017, Cincinnati, OH*, pp. 244–247, [10.1119/perc.2017.pr.056](https://doi.org/10.1119/perc.2017.pr.056).
- [68] B. Pollard, R. Hobbs, J. T. Stanley, D. R. Dounas-Frazier, and H. J. Lewandowski, Impact of an introductory lab course on students' understanding of measurement uncertainty, *presented at PER Conf. 2017, Cincinnati, OH*, pp. 312–315, [10.1119/perc.2017.pr.073](https://doi.org/10.1119/perc.2017.pr.073).
- [69] B. Pollard and H. J. Lewandowski, Transforming a large introductory lab course: Impacts on views about experimental physics, *presented at PER Conf. 2018, Washington, DC*, [10.1119/perc.2018.pr.Pollard](https://doi.org/10.1119/perc.2018.pr.Pollard).
- [70] B. Pollard, A. Werth, R. Hobbs, and H. J. Lewandowski, Impact of a course transformation on students' reasoning about measurement uncertainty, *Phys. Rev. Phys. Educ. Res.* **16**, 020160 (2020).
- [71] W. Sandoval, Conjecture mapping: An approach to systematic educational design research, *J. Learn. Sci.* **23**, 18 (2014).
- [72] P. Cobb, Putting philosophy to work: Coping with multiple theoretical perspectives, in *Second Handbook of Research on Mathematics Teaching and Learning* (Information Age Publishing, Charlotte, NC, 2007).
- [73] M. Dancy, C. Henderson, and C. Turpen, How faculty learn about and implement research-based instructional strategies: The case of peer instruction, *Phys. Rev. Phys. Educ. Res.* **12**, 010110 (2016).

- [74] C. Turpen and N. D. Finkelstein, The construction of different classroom norms during peer instruction: Students perceive differences, *Phys. Rev. ST Phys. Educ. Res.* **6**, 020123 (2010).
- [75] Y. Affriyenni, H. Georgiou, and N. Finkelstein, Navigating the adoption of research-based instructional strategies within the complex nature of higher education, *Phys. Rev. Phys. Educ. Res.* **21**, 020124 (2025).
- [76] B. M. Zwickl, V. Ikoro, and S. Allie, Characterizing lab environments using activity theory, in *The International Handbook of Physics Education Research: Teaching Physics*, edited by M. F. Taşar and P. R. L. Heron (AIP Publishing LLC, Melville, New York, 2023), pp. 0–1–10–30, [10.1063/9780735425712_010](https://doi.org/10.1063/9780735425712_010).
- [77] D. B. Harlow, V. K. Otero, A. E. Leak, S. Robinson, E. Price, and F. Goldberg, Learning about teaching and learning while learning physics: An analysis of 15 years of responsive curriculum development, *Phys. Rev. Phys. Educ. Res.* **16**, 020155 (2020).
- [78] L. C. McDermott, M. L. Rosenquist, and E. H. van Zee, Student difficulties in connecting graphs and physics: Examples from kinematics, *Am. J. Phys.* **55**, 503 (1987).
- [79] I. A. Halloun and D. Hestenes, Common sense concepts about motion, *Am. J. Phys.* **53**, 1056 (1985).
- [80] A. B. Arons, Achieving wider scientific literacy, *Daedalus* (Boston) **112**, 91 (1983), <http://www.jstor.org/stable/20024855>.
- [81] American Association of Physics Teachers, Goals of the introductory physics laboratory, *Am. J. Phys.* **66**, 483 (1998).
- [82] C. J. B. Macmillan and J. W. Garrison, *A Logical Theory of Teaching: Erotetics and Intentionality* (Kluwer Academic Publishers, Dordrecht; Boston, 1988).
- [83] K. A. Oliver, A. Werth, and H. Lewandowski, Student experiences with authentic research in a remote, introductory course-based undergraduate research experience in physics, *Phys. Rev. Phys. Educ. Res.* **19**, 010124 (2023).
- [84] A. Werth, K. Oliver, C. G. West, and H. J. Lewandowski, Assessing student engagement with teamwork in an online, large-enrollment course-based undergraduate research experience in physics, *Phys. Rev. Phys. Educ. Res.* **18**, 020128 (2022).
- [85] A. A. diSessa and P. Cobb, Ontological innovation and the role of theory in design experiments, *J. Learn. Sci.* **13**, 77 (2004).
- [86] <https://www.physport.org/curricula/thinkingcritically/>.
- [87] N. G. Holmes and D. A. Bonn, Quantitative comparisons to promote inquiry in the introductory physics lab, *Phys. Teach.* **53**, 352 (2015).
- [88] A. I. Sabra, Ibn al-haytham: Brief life of an Arab mathematician, *Harvard Magazine*, 2003, <https://www.harvardmagazine.com/2003/09/ibn-al-haytham-html>.
- [89] M. Kapur and K. Bielaczyc, Designing for productive failure, *J. Learn. Sci.* **21**, 45 (2012).
- [90] S. Jeon, I. Descamps, and D. Hammer, Affect, meta-affect, and epistemology in an introductory physics lab, *presented at PER Conf. 2024, Boston, MA*, [10.1119/perc.2024.pr-Jeon](https://doi.org/10.1119/perc.2024.pr-Jeon).
- [91] V. Otero, S. Pollock, and N. Finkelstein, A physics department's role in preparing physics teachers: The Colorado learning assistant model, *Am. J. Phys.* **78**, 1218 (2010).
- [92] D. Hu and B. M. Zwickl, Examining students' views about validity of experiments: From introductory to Ph.D. students, *Phys. Rev. Phys. Educ. Res.* **14**, 010121 (2018).
- [93] D. Hu and B. M. Zwickl, Examining students' personal epistemology: The role of physics experiments and relation with theory, *presented at PER Conf. 2018, Cincinnati, OH*, pp. 11–14, [10.1119/perc.2017.juried.003](https://doi.org/10.1119/perc.2017.juried.003).
- [94] D. Hu, B. M. Zwickl, B. R. Wilcox, and H. J. Lewandowski, Qualitative investigation of students' views about experimental physics, *Phys. Rev. Phys. Educ. Res.* **13**, 020134 (2017).
- [95] D. Hu and B. Zwickl, Examining students' personal epistemology: The role of physics experiments and relation with theory, *presented at PER Conf. 2017, Cincinnati, OH*, pp. 11–14, [10.1119/perc.2017.juried.003](https://doi.org/10.1119/perc.2017.juried.003).
- [96] N. G. Holmes and H. J. Lewandowski, Investigating the landscape of physics laboratory instruction across North America, *Phys. Rev. Phys. Educ. Res.* **16**, 020162 (2020).
- [97] G. Geschwind, M. Alemani, M. F. J. Fox, P. S. W. M. Logman, E. Tufino, and H. J. Lewandowski, Development of a global landscape of undergraduate physics laboratory courses, *Phys. Rev. Phys. Educ. Res.* **20**, 020117 (2024).
- [98] J. T. Remillard, Examining key concepts in research on teachers' use of mathematics curricula, *Rev. Educ. Res.* **75**, 211 (2005).