Sophisticated epistemologies of physics versus high-stakes tests: How do elite high school students respond to competing influences about how to learn physics?

Sevda Yerdelen-Damar¹ and Andrew Elby²

¹Secondary Science and Mathematics Education, Yüzüncü Yıl University, 65080, Van, Turkey

²Department of Teaching & Learning, Policy & Leadership, University of Maryland, College Park, Maryland 20742, USA

(Received 9 October 2015; published 26 February 2016)

This study investigates how elite Turkish high school physics students claim to approach learning physics when they are simultaneously (i) engaged in a curriculum that led to significant gains in their epistemological sophistication and (ii) subject to a high-stakes college entrance exam. Students reported taking surface (rote) approaches to learning physics, largely driven by college entrance exam preparation and therefore focused on algorithmic problem solving at the expense of exploring concepts and real-life examples more deeply. By contrast, in recommending study strategies to “Arzu,” a hypothetical student who doesn’t need to take a college entrance exam and just wants to understand physics deeply, the students focused more on linking concepts and real-life examples and on making sense of the formulas and concepts—deep approaches to learning that reflect somewhat sophisticated epistemologies. These results illustrate how students can epistemically compartmentalize, consciously taking different epistemic stances—different views of what counts as knowing and learning—in different contexts even within the same discipline.

DOI: 10.1103/PhysRevPhysEducRes.12.010118

I. INTRODUCTION

This study explores how Turkish high-school physics students respond to two opposing influences, namely, their sophisticated epistemologies and the high-stakes test that determines college admissions. Specifically, we probe students’ perceptions of their own approaches to learning and how those approaches differ (or not) from pursuing “deep learning.”

Students’ approaches to learning, in particular whether they employ *deep* or *surface approaches* [1,2], affect their learning outcomes [1,3–8]. Two prominent factors affecting students’ approaches to learning are (i) *personal epistemologies*, their views about what counts as knowledge and knowing [9–11], and (ii) *modes of assessment*, the ways in which students are graded or otherwise rewarded [8,12–19]. Previous work shows that physics (and science) students with sophisticated epistemologies—those who view knowledge as richly interconnected concepts instead of isolated facts and formulas and view learning as constructing their own understanding rather than absorbing information from authority—tend to take *deep approaches to learning* [8,11,20–25]. Conversely, many high-stakes testing regimens tend to narrow curricula and lead to more *surface approaches to learning* [26–37]. Most previous

work, however, has foregrounded *either* personal epistemologies or high-stakes testing and has not explored tensions and interactions between them.

In this paper, we begin exploring these interactions by looking at a population who have unusually sophisticated personal epistemologies of physics *and* who are subject to a particularly high-stakes test—a population for whom tensions or interactions between their sophisticated epistemologies and their high-stakes testing preparation are likely to be salient and visible. Our participants were 10th grade students at an elite secondary school in Turkey. The students were engaged in a physics curriculum that led to dramatic increases in students’ epistemological sophistication (within the context of their physics course, at least), as documented elsewhere and summarized here [38,39]. However, the students were all planning to take the Turkish College Entrance Exam, which is very high stakes; only half of students taking the exam gain admission to *any* college in Turkey. The new contribution of this paper is exploring the study behaviors students claimed to adopt in light of these competing influences. Furthermore, unlike most previous discussions of the effects of high stakes testing on students [26,28,34], this paper includes the students’ own views rather than teachers’ recounting of students’ views. In brief, this paper addresses the following research question:

When physics students subject to high-stakes testing have engaged in a curriculum that helped them develop more sophisticated epistemologies of physics, how do they claim to approach the learning of physics in light of those potentially competing influences?

Published by the American Physical Society under the terms of the *Creative Commons Attribution 3.0 License*. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

Our results we hope are relevant to reform efforts immersed in a background culture of high-stakes testing. Several different results were possible. One, because even the most traditional, “rote” physics problems can also be approached through sense making [11], the students might have decided to take a deep approach to learning in preparation for the College Entrance Exam (CEE), thereby preparing for the test and achieving deep understanding. Alternatively, disconcerted by the mismatch between the usual approaches to studying for the CEE and the unusual physics curriculum in which they were engaged, students might have decided that the physics curriculum and associated *epistemic stances*—views about what counts as knowing and learning in a given context—were a course-specific aberration and that “real” physics learning aligns with what the CEE rewards. Interestingly, neither of these trajectories describes the majority of students in this study. Instead, they *epistemically compartmentalized*, by which we mean they consciously took different epistemic stances toward learning physics for the CEE and learning physics for deep understanding.

In this paper, we first situate our work in the literature on students’ personal epistemologies in science (particularly physics) and the effects of high-stakes testing on students, tying these pockets of research together through the construct of *deep vs surface approaches to learning*. Then, we describe the Turkish high-stakes testing regimen and briefly describe the participants and the curriculum in which they were engaged. Next, we describe our survey-based methods, and our results. Finally, in our discussion, we relate the epistemic compartmentalization we observed to previous studies and suggest directions for future research.

A. Personal epistemologies and high-stakes testing affect approaches to learning

In this review of the effects of (i) personal epistemologies and (ii) high-stakes testing on students’ approaches to learning, we use the construct of *deep vs surface approaches to learning* to bridge these two pockets of literature. Discussing the bridge first will better enable us to chart the landscape on each side.

1. Deep vs surface approaches to learning

Following Biggs [4], we define approaches to learning as “the ways a particular student has of going about selecting and learning from” curriculum content or other related teaching tasks [4] (p. 381). According to Biggs [2], an approach to learning combines a motive and congruent strategies to deal with a particular task. Kember [40], relying on a review of studies addressing students’ approaches to learning, also concluded that a learning approach can be decomposed into an intention and a strategy.

An early study by Marton and Saljo [1] found that university students displayed different types of learning processes while reading passages and answering questions about them. Marton and Saljo called those types of processing deep level and surface level:

In the case of surface-level processing the students directs his attention towards learning the text itself (the sign)...he is more or less forced to keep to a rote-learning strategy. In the case of deep-level processing, on the other hand, the students direct toward the intentional content of the learning material (what is signified), i.e., he is directed towards comprehending what the author wants to say about [1].

Biggs’ [2] (p. 129) classification scheme also includes *surface approaches* and *deep approaches*, largely aligned with Marton and Saljo’s scheme:

...*surface*, where the motive is to meet institutional requirements minimally, and the congruent strategy is limiting the target to essentials that may be reproduced through rote learning.

...*deep*, where the motive is intrinsic interest in the content learned, and the congruent strategy is discovering meaning and acquiring competence by reading widely, interrelating with existing knowledge, etc.

However, Biggs also introduces a category called *surface achieving*, defined by a goal of high achievement in school and strategies equivalent to those of surface level. Similarly, he introduces *deep achieving*, where the strategies mirror those of deep level learning but in the service of both achievement and intrinsic interest. Introducing these additional categories allows Biggs to tease apart goals (motives) from strategies. In most subsequent studies, as discussed below, researchers do not make this distinction. Furthermore, in later work Biggs and colleagues [41] revised Biggs’ widely used Study Processes Questionnaire [42] and verified that it displayed a two-factor structure corresponding to deep vs surface, in which the two factors are negatively correlated. Therefore we will use “surface” to refer to both surface and surface achievement and similarly for “deep.” In interpreting our own data, by contrast, we will resurrect surface achieving as a category.

(a) *Approaches to learning in science*.—Only a limited number of studies of science students have labeled themselves as exploring “students’ approaches to learning.” The classification schemes in these papers, however, align well with the deep vs surface distinction just discussed. For instance, Case and Gunstone [5] analyzed 11 students’ approaches to learning in a second-year college chemical engineering course, by conducting interviews and examining student journals. They identified three distinct approaches to learning: (i) “A conceptual approach, where the intention is to understand concepts”; (ii) “An algorithmic approach, where the intention is to remember calculation methods for solving problems”; and (iii) “An

information-based approach, where the intention is to remember information that can be supplied in response to assessment questions” [5] (p. 806). Case and Gunstone pointed out that a conceptual approach can be considered deep while the algorithmic and information-based approaches can be considered two forms of the surface-level learning, since the intentions and strategies do not focus on understanding.

Chin and Brown [6] by contrast adopted the deep vs surface framework from the outset but “aimed to explore in detail qualitative differences between a *deep* and *surface approach to learning* science among 8th graders.” They observed six students’ laboratory activities in a chemistry unit and also interviewed students individually before and after instruction about the related science concepts. They then used students’ scores on a learning approach questionnaire and the teacher’s evaluation to categorize each student as taking a *deep* or *surface approach to learning*. Chin and Brown [6] (p. 130) found that

When students used a deep approach, they generated their ideas more spontaneously... They also gave microscopic theorylike explanations which described nonobservable entities and cause-effect relationships or referred to relevant daily life experiences in trying to understand a phenomenon. Explanations associated with a surface learning approach tended to be reformulations of the question, a black box variety which did not refer to a mechanism, or macroscopic descriptions which referred only to what was visible.

Furthermore, “When students used a deep approach, they also displayed more cognitive appraisal and regulatory control of the learning process through ongoing reflective thinking.” They were also more likely to generate mental images, analogies, hypotheses, predictions, and thought experiments [6].

Other pockets of literature reach conclusions that align with Chin and Brown [6]. For example, studying high school students engaged in a genetics unit, Jimenez-Alexandre, Rodriguez, and Duschl [43] documented a “distinction in argumentation discourse between ‘doing science’ vs ‘doing school.’” When doing science, students displayed many of the practices documented by Chin and Brown [6], such as generating analogies and causal explanations. Doing school corresponded closely to surface approaches. Indeed, other researchers have also documented how traditional classroom discourse patterns and ways of framing “school science” can favor surface approaches to learning science [44–49].

In summary, the notion of deep vs surface *approaches to learning* from Marton and Saljo and from Biggs has proven useful for researchers in multiple disciplines; and science education researchers have documented the salience of that distinction and of other distinctions that align closely with deep vs surface. For these reasons, we adopt this distinction as an analytical lens through which to discuss and connect

the effects of epistemologies and of high-stakes testing on students’ approaches to learning.

2. Relating personal epistemology to students’ approaches to learning

In science and in other disciplines, a range of studies have documented that students with more sophisticated personal epistemologies tend to take deeper approaches to learning, pursuing conceptual as opposed to rote understandings. For example, Schommer, Cruse, and Rhodes [10] administered an epistemological questionnaire, a study strategy inventory, and a comprehension task in which students read a passage about statistics and then took a test probing their conceptual understanding of the material. Their regression analysis, controlling for previous mathematics coursework and other factors, revealed that students who viewed knowledge as “complex” (conceptual and interconnected) took a deeper approach to learning and scored higher on the comprehension task. Similarly, Watters and Watters [8] investigated the relationship between college students’ epistemological beliefs about learning and knowledge, approaches to learning, and achievement in first-year biological chemistry and biochemistry courses. They found a significant relationship between students’ epistemological beliefs about learning as identified in student discussions, approaches to learning expressed on a survey, and performance in complex problem-solving tasks. And Chiou, Liang, and Tsai [50], using an epistemology survey (the Conceptions of Learning Biology questionnaire) and their Approaches to Learning Biology questionnaire, found that students’ degree of epistemological sophistication about learning biology correlated with the degree to which they professed a *deep approach to learning*. In physics, Hammer [11] found that university students’ epistemological sophistication aligned with in-the-moment decisions they made while learning new material from the textbook or solving a problem, e.g., about whether to reconcile equations with everyday intuitions.

Similar results have been documented at the secondary level. For instance, Lee, Johanson, and Tsai [21] found that Taiwanese high school students who viewed the learning of science as absorbing problem-solving procedures by heart tended to take *surface approaches to learning*, while those who viewed learning science as understanding and applying ideas and seeing the world in a new way tended to take *deep approaches to learning*.

Note that, although Hofer and Pintrich [9] define personal epistemology as views about the nature of knowledge and knowing but not views about the nature of learning, the studies just cited and many others include views about learning under the umbrella of personal epistemology. To connect to this literature, and for theoretical reasons touched upon below in our Conceptual

Framework, we define personal epistemologies as views about the nature of knowledge, knowing, and learning.

Although the notion of self-regulated learning does not correspond precisely to deep approaches to learning, the overlap is substantial: both involve monitoring one's own understanding and going beyond mere memorization. For this reason, the literature documenting and modeling the connection between epistemological beliefs and self-regulated learning [51–55], partly synthesized by Muis [56], counts as further evidence that sophisticated epistemologies correlate with deep approaches to learning.

In short, a substantial literature supports our assumption that a physics curriculum which successfully fosters more sophisticated epistemological views should help students see the value of deep approaches to learning. Our data support this assumption.

B. The effects of high-stakes testing on students

In this section, we first review literature about the effects of high stakes testing generally. Then we zero in on science. We show that, for multiple reasons, high-stakes testing regimens tend to favor surface approaches to learning.

Following Madaus (as cited in Au [57], p. 258), we define a “test [as] high-stakes when its results are used to make important decisions that affect students, teachers, administrators, communities, schools, and districts.” High-stakes testing has received extensive criticism for its negative impact on curriculum, instruction, and students. For example, many researchers argue that high-stakes testing leads to surface (rote) learning [26–37]. Driven by exam pressures, teachers downplay or ignore content not tested on the exam and emphasize factual and rote procedural knowledge at the expense of higher-order conceptual and procedural knowledge [37,58–60].

Valli *et al.*'s [37] longitudinal study of 25 elementary schools began the year before the No Child Left Behind Act (NCLB) was implemented and continued through the first few years of NCLB, enabling a stark documentation of this phenomenon. They observed that the level of cognitive demand of mathematics and reading lessons decreased every year in all schools. As the study progressed, teachers were more likely to ask lower-order questions and pose lower-level tasks because of the increased press of test accountability. Other studies also show that high-stakes testing leads instruction to become more teacher centered, emphasizing drill activities and test-taking skills [26–28,30,34–36,57,58,61,62]. Teachers spend much instructional time on practice for the high-stakes tests [26,30,34,37] and design their classroom assessments to mirror the content and format of those tests [26,37].

For these and other reasons, high-stakes testing can affect students, often in negative ways [26–28,34,35,63]. Such testing decreases students' motivation, self-efficacy, and confidence [26,34,35]. And according to teachers, students often feel anxiety and stress about the tests

[26,28,34,35]. For example, responding to a survey about the effects of high-stakes tests, only 28% of teachers felt their students were more prepared for learning and 15% indicated that their students had more confidence [34]. By contrast, 61% felt that their students suffered from more anxiety, 24% felt their students were less confident, and 48% indicated that testing decreased students' love of learning.

Similar results have come from studies focused on science education. The most striking negative effect is that high-stakes testing decreases the time given for experimentation and hands-on or active-learning practices aligned with research on how best to foster deep learning [64–70]. It also narrows curriculum [65,67], limits freedom and creativity in teaching [69,70], and leads to memorization-based learning [67,68,70].

Indeed, Anderson [71] systematically reviewed 35 studies investigating impacts of test-based accountability on science instruction. According to this review, 26% of reviewed studies reported positive effects of test-based accountability on science education while 97% of reviewed studies reported negative aspects of test-based accountability. These percentages sum to greater than 100% because some studies reported both positive and negative effects. In these studies, teachers reported that high-stakes testing impedes the use of research-based reforms. They claimed that testing influences their instructional activities, often preventing meaningful science learning.

Other work suggests that Valli *et al.*'s [37] findings about the cognitive demand of classroom tasks apply to science as well. For instance, Wideen *et al.* [70] interviewed teachers in Grades 8, 10, and 12 drawn from 10 districts in British Columbia, to explore the relation between high-stakes testing and the teaching of science. They observed that the range of instructional strategies, and in particular inquiry-oriented teaching, decreased from Grade 8 to Grade 12. Based on teachers' comments on high-stakes testing at Grade 12—in particular, that the testing regimen eliminated opportunities for spontaneity and depth—the researchers claimed that testing might cause these instructional differences.

Case studies of individual teachers support researchers' interpretation of their data in the large-*N* studies discussed above. For instance, Smith and Southerland [68] explored how two elementary teachers perceived national science standards, state-mandated science curricula, and associated criterion-referenced testing and the impacts of these reforms on their instruction. Both teachers felt pressured to cover all content on the high-stakes tests. One of the teachers said she preferred lecture and discussion to hands-on or inquiry-based instruction to cover the necessary content, since she viewed her job as transmitting factual knowledge to help students pass the tests. The other teacher responded to high-stakes testing by integrating a few

multiple-choice questions like those in state-mandated test into classroom assessments.

In summary, the literature suggests two mechanisms by which high-stakes testing leads students to adopt surface approaches to learning. First, teachers adjust their instruction in ways that support surface approaches at the expense of deep approaches to learning. Second, the tests themselves reward and therefore encourage surface learning. Still, what's underrepresented in this body of research is students' own views, as Anderson [71] points out. Studies about the impacts of high-stakes testing on students have relied mainly on teachers', administrators', or parents' perceptions. Our study helps to fill this gap by asking students both closed- and open-ended questions about their responses to the Turkish CEE, a very high-stakes test.

C. Literature review summary

In general and in science, sophisticated personal epistemologies push students toward deep approaches to learning, while high-stakes testing pushes students toward surface approaches to learning. These generalizations motivate this study's focus on students with comparatively sophisticated epistemologies who are subject to high-stakes testing. How do they respond to the competing influences?

D. Theoretical framework for personal epistemologies

As discussed in Methods below, deep vs surface learning distinction serves as our primary interpretive lens. However, in our Discussion, we go a step further and interpret our results in terms of epistemic stances and epistemic compartmentalization. To define those terms carefully and situate them within a broader cognitive model of personal epistemologies (or what some researchers prefer to call epistemic thinking), we now briefly discuss the framework on which we rely.

Most research on personal epistemologies attributes epistemological stages or belief systems to students, as Hofer and Pintrich [9] discuss. According to these "unitary" accounts, students' epistemological views—their views about the nature of knowledge, knowing and (in some studies) learning—are fairly stable and robust, at least within a given academic domain or discipline. By contrast, Hammer and colleagues [72–74] introduce a knowledge-in-pieces [75] framework for modeling personal epistemologies. According to this *epistemological resources* perspective, nonexpert epistemologies consist of networks of finer-grained knowledge elements—epistemological resources—whose patterns of activation depend on context. In this framework, a student doesn't necessarily have a single "true" epistemology. Instead, the framework allows that different contexts trigger different networks of epistemological (and other) resources, corresponding to different views about what counts as knowing and learning in that

context—different *epistemic stances*. We adopt the epistemological resources framework when interpreting our results, though we also discuss how our results might look from a unitary perspective.

Within the resources framework, it is possible for students to slip from one epistemic stance to another beneath conscious awareness, in response to contextual cues. We define *epistemic compartmentalization*, by contrast, as *deliberately* taking different epistemic stances toward the same discipline in different contexts. We argue below that the students in our study epistemically compartmentalize.

II. SUBJECT SELECTION AND CHARACTERISTICS

Instead of including subject selection as a subsection of Methods, we break it out as a separate section for two reasons. First, to help readers understand our subjects' responses to our survey questions, we must discuss the physics component of the Turkish CEE, which almost all our subjects were planning to take. Second, in order to establish that our participants can help us address our research question,

When physics students subject to high-stakes testing have engaged in a curriculum that helped them develop more sophisticated epistemologies of physics, how do they claim to approach the learning of physics in light of those potentially competing influences?

we need to show that the curriculum did indeed foster epistemological change. Doing so will require us to summarize the results of a previous study [38,39].

A. Student population from which participants were selected

The study included 46 (24 female, 22 male) 10th grade students aged 15 to 17. They were drawn from two "Anatolian teacher training high schools" in Ankara, the capital of Turkey. These elite schools accept students who score well in a competitive government entrance exam similar in many ways to the CEE discussed below. All elementary students take this high school entrance exam in 8th grade. Their scores determine which high schools they may choose to attend. High school in Turkey lasts four years, 9th through 12th grade. After high school, students wishing to go on to university take another national exam, the CEE. Thus, our subjects already had experience successfully studying for a competitive national exam, and most were already studying for the CEE. Although Anatolian teacher training schools focus on preparing students for jobs as faculty of education in colleges, students can apply to other majors based on their CEE scores.

B. High-stakes Turkish College Entrance Examination

In this section, we discuss the CEE and how students typically prepare for it. Then, we briefly summarize prior work on how the CEE affects Turkish high school physics curriculum, instruction, and students, as a window into the cultural context in which this study occurs.

1. College Entrance Examination in Turkey

The two-stage CEE, administered in 12th grade by the Student Selection and Placement Center (OSYM), determines which students may attend college. The first stage, covering material taught through 9th grade, screens students for the second stage. First-stage physics items are mostly conceptual, with a few numerical questions. The second stage, covering all high school curricula, is used to place students in undergraduate programs (other than distance education programs). The physics items are mostly quantitative. Figure 1 presents a question from the 2010 second stage [76].

In 2010, only 55% of students taking the CEE were accepted into *any* university [77]. Since most desirable careers in Turkey require a university education, success on this exam is important to students and their parents.

2. Exam preparation

CEE preparation begins in elementary school and continues through high school and beyond if students do not pass the CEE on their first try in 12th grade. Exam preparation activities include self-directed home study, private tutoring, and tutoring centers.

In individual home study, students typically use CEE guide books to study summaries of topics, exam tactics, and practice questions. Some families hire private tutors to guide and supplement this process. Because private tutors are expensive, however, most families enroll their

college-bound children in *dershane*, for-profit tutoring centers where professional teachers tutor in a classroom setting [78], similar in some ways to Japanese *juku* (“cram schools”). High school students attend these tutoring centers on weekends and after school on weekdays. High school graduates who failed previous CEEs also attend. These centers focus on training students to solve problems as quickly as possible.

3. Effect of the CEE on curriculum, instruction and students in Turkey

The Turkish high school physics curriculum was recently changed to reflect research in science education and educational psychology. Changed features include contextualized real-life examples, inquiry-based approaches to learning, “spiraling” to revisit concepts at more advanced levels, and more emphasis on conceptual learning, and higher order thinking skills [79–80]. These features, however, are in conflict with the CEE and with the *dershane*. Most CEE physics questions are not contextualized to real-life scenarios. *Dershane* treat a given topic all at once instead of spiraling back to increasing levels of abstraction and difficulty. The deeper inquiry and problem-solving skills and habits of mind targeted by the new curriculum are not measured on the CEE. Because of this disconnect, students may pressure teachers to return to “traditional” teaching, and if teachers do not, students may feel the need to rely on *dershane* even more (personal communication with two high school physics teachers, March 18, 2010). Similar tensions affect assessment. The new physics curriculum encourages the use of both traditional assessments (multiple choice, etc.) and alternative techniques such as concept maps and performance-based assessments [80]. Alternative assessment techniques are rarely used, however, because they take more time and differ in format from CEE items (personal communication with two high school physics teachers, March 18, 2010).

On a frictionless inclined plane, Object K having a mass of 2 kg is attached to Spring X in a state of equilibrium as shown in the figure. The elasticity coefficient of the spring is equal to 60 N/m. How much does the spring compress? ($g = 10 \text{ m/s}^2$; $\sin 37^\circ = 0.6$; $\cos 37^\circ = 0.8$; ignore the mass of the spring.)

(A) 0.2 m, (B) 0.3 m, (C) 0.4 m, (D) 0.5 m, (E) 0.6 m.

FIG. 1. Item from the second stage of the 2010 CEE physics portion.

The effects of the CEE on the enactment of the reform-based physics curriculum has been investigated by interviewing and surveying Turkish teachers about the curriculum [81–83]. In those studies, teachers pointed out the misalignment between the structure of the CEE and the curriculum. For example, Mercan [83] interviewed 39 physics teachers from 27 different high schools in Istanbul. Ninety percent of the teachers claimed that the main aim of their instruction is to prepare students for the CEE. Seventy-seven percent reported that they did not change their teaching approaches and 90% acknowledged that they did not include experiments in their instruction. In terms of assessment, 64% of teachers argued that assessment tasks and response formats suggested by the curriculum are incompatible with those used in the CEE. Based on teachers' explanations, Mercan concluded that the primary reason for the misalignment between the new curriculum and its enactment in schools is the CEE.

The CEE affects Turkish students in other ways, too. Kelecioğlu [63] surveyed 1215 seniors in high school about the CEE. They reported that studying for the exam decreased their learning in regular school. Similarly, according to a Turkish Education Association report [84], 37% of 2854 high school students reported that studying for the CEE makes them tired and bored. 62% reported that if they didn't need to study for the CEE, they would spend some of their extra time studying more for their school courses. 53% would read fiction or poetry, 43% would participate in art activities, 54% would play sports, 58% would spend more time with their friends, and 57% would spend more time with their families.

Studying for the CEE even cuts into students' time in school. In elite high schools, almost all seniors skip for ~45 days to study for the CEE, often at *dershane*. School administrations must accept this because 17- and 18-year olds have the legal right to skip school for this duration. Indeed, because college acceptance rates are a measure of a school's success, some school administrators support these absences. This indicates that not only students but also some school personnel think that attending school is not the way to maximize CEE scores.

In summary, the CEE plays a huge role in the life of Turkish high school students who wish to attend college. Essentially all participants in this study planned to take the CEE.

C. Curriculum and classroom context

Now we switch gears to briefly discuss the curriculum in which our participants were engaged and the effects of the curriculum on their epistemological views.

The inquiry-based unit on forces and motion was designed by the first author to improve both conceptual and epistemological understandings in physics. She chose

this topic because students have a great deal of everyday experiences and intuitive knowledge to draw on. The 10th grade curricular unit was taught by two regular classroom teachers, not the curriculum designer, and it consisted of two 45-minute class sessions per week lasting for four months [38].

1. Epistemological agenda in an inquiry-oriented physics unit

Specifically, the curriculum was structured using the 7 E learning cycle model developed by Eisenkraft [85]. In each cycle, students make predictions about physical phenomena, collect data to test their predictions, support, refine, and evaluate their explanations using the evidence, provide coherent causal explanations, and discuss their findings and conclusions with their classmates. However, explicit and implicit epistemological prompts and activities were integrated into these learning cycles, with the goal of helping students view physics as “the refinement of everyday thinking” [86]. In other words, the curriculum aimed to help students refine and build on—rather than ignore or abandon—their common-sense thinking, and to view the learning of physics as reconciling common sense and formal physics concepts, to the extent possible [38]. The first author adapted strategies from “Tutorials in Physics Sense-Making” [87] to achieve this goal.

The development and final form of the curriculum are presented elsewhere [38].

2. Epistemological shifts

We now briefly summarize results presented in more detail elsewhere [38,39] about the effects of this curriculum on students' epistemologies. The control group was students taught the same course by the same teachers but not using the first author's curricular unit to teach forces and motion. At the end of the forces and motion unit, the treatment and control groups both took the Turkish version of the Maryland Physics Expectations Survey version 2 [88,89], a revision of the original MPEX [90]. MPEX2, consisting of Likert scale items, probes to what extent students view physics knowledge as formulas vs concepts (expressible as formulas), as weakly connected pieces vs a coherent web of ideas, and as something absorbed from authority vs constructed for oneself. (By contrast, the original MPEX focused more heavily than MPEX2 does on students' expectations about how to succeed in a particular course.) Sample MPEX2 items include

Tamara just read something in her physics textbook that seems to disagree with her own experiences. But to learn physics well, Tamara shouldn't think about her own experiences; she should just focus on what the book says.

To understand physics, the formulas (equations) are really the main thing; the other material is mostly to help you decide which equations to use in which situations.

The treatment group, who are the participants in this study, scored significantly higher than the control group [one-way ANCOVA, $F(1, 100) = 19.97$, $p = .000$], with an effect size of $d = 0.90$. The mean post-test scores of treatment and control group were 99.9 ($SD = 14.7$) and 87.2 ($SD = 13.1$), respectively. Considering the maximum possible score, 145, and the minimum possible score, 29, on the survey [91], we can conclude that the treatment-group students espoused more sophisticated views—that physics is a coherent web of concepts rather than piecemeal facts and formulas, best learned by making sense of ideas rather than just absorbing information [38].

Given the physics-class context in which students completed this survey and given that the questions refer to course-related experiences such as reading a physics textbook, it is likely (and indeed, intended by MPEX and MPEX2 creators) that students thought about their physics classroom experiences when completing the survey. So, from these results, we infer that the treatment group was epistemologically sophisticated compared to the control group in the sense of taking a more sophisticated epistemic stance in the context of their physics course. For the argument we are making in this paper, this is sufficient; our argument does not depend on that stance corresponding to a generalized physics epistemology.

In summary, the participants in our study were subject to a high-stakes testing environment that pushes students toward *surface approaches to learning*, but they were immersed in a physics curriculum that helped them develop comparatively sophisticated epistemic stances toward learning physics. This establishes that our participants are well suited for our research question about the competing influences of epistemological sophistication and high-stakes testing.

III. METHODS

A. Survey instrument

Our survey was adapted from and is nearly equivalent to the one validated and used by Elby [12]. It was translated into Turkish by the first author. The translated version was edited by a university-level English instructor by matching it to the original form. The Turkish version's questions about assessments refer specifically to the CEE, while the original version refers to midterm and final exams in students' courses. See Appendix A for the complete survey, in English. Here, we describe the overall gist of the survey, which consists of three parts.

In part 1, students indicated how they typically study physics:

Assume that you divide your study time for physics into percentages. Specify the percentage you give to studying

(a) concepts and relationships among the concepts (for example, velocity, acceleration, and force are examples of concepts; an example of the relationship among them may be the tendency of objects to continue moving in a straight line at a constant speed unless a push or pull changes the motion), (b) formulas, (c) practice problems in the test books, and (d) real-life examples.

Open-ended questions then ask students to explain the reasons behind their choices.

In part 2, we asked students to imagine a hypothetical student:

Arzu is a student just like you, with the same age and abilities. But, since Arzu and her family will move to another country where no high-stakes entrance exam is used after she graduates from her high school, she does not need to worry about the College Entrance Examination. So, *her goal is simply to understand physics more deeply.*

The survey then asks students how they would advise Arzu to study to reach her goal, using the same questions as were asked in part 1. This allows us to compare how a student's study strategy *for herself* compares to her suggested study strategy *for Arzu*.

Finally, in part 3, we asked questions that included how they think Arzu would perform on the CEE.

Students took the survey after completing three of the four months of the inquiry-based physics unit described above, and they would take the CEE two years later. To create time for students to respond in detail, we asked them to complete the surveys at home, on paper. Forty six of the 52 students in the treatment group completed the survey. Six students did not return the survey.

Following Elby [12], we assumed that students' study-strategy suggestions for Arzu reflect the epistemic stance they imagine taking toward understanding and learning physics deeply. On the other hand, their self-reports of their own approaches to learning likely indicate a mix of their actual behaviors and their expectations about how to succeed [12], particularly in light of the CEE. Students' responses to open-ended questions, discussed below, confirm that the different parts of the survey tapped into these different aspects of students' views, enabling us to explore to what extent the high-stakes testing environment (and other factors) lead students to approach learning in ways that misalign with their epistemic stance about what counts as deep understanding and how to achieve it.

B. Analysis

In the Results section itself, we explain the straightforward calculations used to generate our quantitative results about differences between students' study-time allocations and those they recommend for Arzu. Here, we describe how we processed the open-ended survey responses.

1. Coding and interpretation of open-ended responses

To categorize students' written responses, the first author used a bottom-up approach that borrowed tools from multiple traditions in qualitative research. Although not working within a grounded theory paradigm, since we expected to ultimately interpret our data through a lens of deep vs surface approaches to learning, we started with open coding as described by Glaser and Strauss [92]. We categorized students' responses into fairly low-inference categories. For instance, here is "Esra's" response to the question "How do you study physics?"

I think it is more important to study the logic of the topic when you study physics. We can relate the topic to our life to not forget the logic of the topic. For example, in motion topic, consider it is asked how observers in Car A and Car B see each other. If Car A is faster than Car B, I consider myself in a moving car. I remember the car which travels in the positive direction relative to me is faster than the car where I am. This leads me to understand the logic of the question and the topic. After that, I try to come to know the formulas and use them correctly on questions. If I have too many incorrect answers, I ask help from my friends. This generally helps me.

The first author initially assigned the following four codes: *Relate to everyday life to make sense of the topic*, *learn formulas*, *answer many questions*, and *ask friends for help*. Again following aspects of Glaser and Strauss [92] and of phenomenographic methods [93], for each question, the first author sorted the coded responses into categories based on commonalities, iteratively refining the sort in light of new responses. Refined versions of the initial four codes ended up as five categories in the final coding scheme, namely *relate to everyday life*, *make sense of what is learned*, *memorize formulas*, *solve many questions/problems*, and *ask friends/teachers for help*. This example illustrates how our final codes are fairly low inference but combine participant-specific codes into broader categories. Our final coding scheme for "How do you study physics?" included additional categories not reflected in Esra's response; see Table II.

We acknowledge that some of our codes are higher inference than others. With Esra's response, our highest inference comes in coding "After that, I try to come to know formulas and use them correctly on questions" as *memorize formulas*. After all, "come to know" could refer to sense-making, not memorization. We used contextual cues and emerging patterns in the data to interpret "come to know" as "memorize" in this case. First, this strategy is introduced as "after that," where "that" refers to sense-making strategies Esra had just discussed, indicating that she may view this "coming to know" as something she does *after* sense making, not as part of sense making. Supporting this interpretation is the fact that Esra did not use time-sequence

words like "after" or "then" previously in her response. Second, the strategy of studying the formulas and then applying them to numerous practice problems was described by many students, often using a Turkish word that corresponds more closely to "commit to memory" than to "come to know." Because Esra's response fell into this pattern, we made the inference that "come to know formulas" likely meant "memorize formulas."

To provide a sense of how frequently we assigned lower-inference vs higher-inference codes, in Appendix B, we provide five quotes, from different students that were all assigned the code *solving as many questions as possible* (our shorthand for solving lots of problems though not necessarily at the expense of other study strategies), and we dissect what assumptions were needed to assign the code in each case.

Finally, as elaborated in the Results and Discussion section, we interpreted our final coding categories in light of previous work on deep vs surface approaches to learning. Prior to interpreting our codes, we had not grouped our codes into themes; for instance, we had not combined *repeating what is learnt in class* and *studying topics from a book* into a theme such as "Study physics topics from authoritative sources."

2. Interrater agreement

For each open-ended survey prompt, 30% of students' responses were randomly selected and independently coded by the second author. We chose 30% because we estimated this would generate 50+ codes per prompt for the prompts most likely to yield interrater disagreements, namely, "How do you study Physics?" and "How should Arzu study?" We calculated percentage agreements as follows: With respect to a given survey prompt, we divided the number of codes on which both coders agreed by the total number of codes. For example, to calculate the interrater agreement for "How do you study physics," 15 students' responses were coded by the second author, using the categories that the first author found in her analysis. The total number of codes obtained from the 15 selected students' responses, not double-counting cases where both authors assigned the same code to the same student, was 63. The coders agreed on 53 of those 63 codes. So, the interrater agreement for that category was 53/63 or 84%. For the second open-ended prompt, "How should Arzu study," the interrater agreement was 91%. For all other prompts, it was 100% [94].

In checking interrater reliability and presenting our results below (except for the final open-ended prompt), we only paid attention to coding categories the first author assigned to at least 20% of the responding students, since interrater disagreement about idiosyncratic responses does not threaten the arguments we are making with this data.

TABLE I. Students' study-time allocations for themselves and for Arzu ($N = 46$).

Category	Mean for self	SD	Mean for Arzu	SD	Difference
Concepts	24	13	37	14	-13
Formulas	22	14	11	9	11
Practice Problems	38	18	11	8	27
Real-life Examples	17	13	42	20	-25

IV. RESULTS

First, we present students' self-reported study-time allocations for themselves vs their recommendations for Arzu. Then, we present results from the coding-based analysis of free responses, which shed light on the reasons behind the first set of results.

A. Study time allocations: Self vs suggestions for Arzu

Recall that students indicated how they divide their study time and then indicated how Arzu should divide *her* study time. Table I shows the mean scores for students' self-reported study-time percentages and the percentages of time they advised Arzu to spend. The last column shows the differences between those two means. Paired-samples t tests showed that the differences are all statistically significant ($p < 0.05$).

Students reported spending the largest percentage of their time doing practice problems (mostly from test books, as students report in their free responses) and only a small amount of time thinking about real-life examples. By contrast, students recommended that Arzu spend most of her time on concepts and real-life examples and minimal time on formulas and practice problems from test books. So, the students think they study in ways not optimal for someone (Arzu) whose goal is to learn physics deeply. Put another way, students *distort* their study habits away from those they think would be productive for pursuing deep understanding [12].

To quantify this distortion, we calculated the *distortion percentage* for each student, defined as the percentage of study time that the student spends differently than she would have Arzu spend. An example will clarify what we mean. Suppose a student spends 25% of her time on formulas, 25% on concepts, 25% on practice problems, and 25% on real-life examples. Suppose, however, that this student wants Arzu to spend 15% of her time on formulas, 35% on concepts, 25% on practice problems, and 25% on real-life examples. The student's distortion percentage is 10%; that's the percentage of time she spends differently from how she would have Arzu allocate her time. Specifically, the student differs from Arzu in spending an "extra" 10% of her total time on formulas rather than

FIG. 2. Distribution of distortion percentages ($N = 46$).

concepts. By playing around with examples, the reader can confirm that distortion percentages can range from 0% to 100% and are calculated as follows: (i) find the concepts distortion percentage by taking the absolute value of [(percentage of time student focuses on concepts) — (percentage of time she would have Arzu focus on concepts)], and similarly for the formulas, practice problems, and real-life examples; (ii) sum those four distortion percentages; and finally (iii) divide by 2, to avoid double counting [12].

Figure 2 shows the distribution of students' distortion percentages. The mean and median are 42% and 38%, respectively. For the middle half (25th to 75th percentile) of students, the distortion percentages range from 24% to 55%. These results underscore those of Table I; students claim to study differently from how they advise studying in pursuit of deep understanding.

B. Students' reasons for their study strategies

We now turn to students' open-ended responses, starting with how they say they study physics. Table II shows the percentages of students giving the most common responses (those expressed by >20% of students), some examples of which we now provide. Percentages sum to >100% because most students were coded into multiple categories, as illustrated by Esra's response above (see Methods).

Most students claimed to study physics by solving many multiple-choice questions from test books, questions that mostly require correct use of formulas, so that they could learn different types of questions and practice solving questions quickly. Reviewing what they learned in physics class before diving into practice problems was another common approach:

I generally review what we learned in class that day for 5–10 min when I return to my dormitory. Then, I look at formulas and solve all tests in my test book. Moreover, I solve questions in different resource books to be familiar with different types of questions. Indeed, I study similarly for mathematics, cchemistry, and biology, which are tested in the CEE.

46% of students reported that, before starting to solve practice problems, they reviewed the topics from books. These books were generally test-preparation books which include a summary of topics and example questions together with solution tactics and practice problems.

Students rarely mentioned school physics textbooks in their responses. For instance,

When I study physics, I try to understand the topic completely. For this, I study the topic and examples from at least 2 or 3 test books.

In summary, from Table II and from these examples and others like them, we see that four of the seven most commonly used study strategies—studying topics from test-preparation books, studying examples from the test-preparation books (with an eye toward rote, quick problem solving), memorizing formulas, and solving as many problems as possible—are surface approaches to learning, focused on rote reproduction rather than deeper understanding. Two other common approaches, going over what was covered in class and asking teachers or friends for help, could be conducted in either a deep or a surface way, and students' responses did not provide us with warrants to classify those approaches one way or the other. Only one of the seven commonly expressed strategies, “making sense of what is learned,” aligns with a deep approach to learning. And only 28% of students claimed to use that strategy.

TABLE II. Coding categories for students' free responses to “how do you study physics?”.

Study approaches	Percentages
Solving as many questions as possible	85
Repeating what is learnt in class	50
Studying examples in books or notes (to learn problem-solving tactics)	46
Studying topics from a book (generally test book, rarely textbook)	46
When unable to solve a problem on the first try, asking questions to a teacher or a friend, or asking a teacher or friend to explain topics that still are not understood	43
Making sense of what is learnt (concepts, formulas, topics)	28
Memorizing formulas	24

TABLE III. Coding categories for students' responses to “how should Arzu study physics?”.

Study approaches	Percentages
Doing experiments and/or making observations related to the topic	59
Relating physics knowledge to everyday life	35
Making sense of what she learns in school (concepts, formulas, topics)	30
Searching for more information on topics	26
Understanding the logic behind formulas	26
Studying concepts and the relationships among the concepts	22

C. Students' suggested study strategies for Arzu

By contrast, most of the common strategies suggested for Arzu differ from students' own study strategies and align with a deep approach to learning. Table III shows the most common categories of responses, those expressed by >20% of participants.

For example, the second-most common suggestion was relating what is learned to everyday life. As one student wrote,

...Especially, she [Arzu] should pay particular attention to the concepts she learns in her class and she should connect these concepts to daily-life examples to increase the durability of her knowledge. She should analyze everything with her physics knowledge even when she walks.

Moreover, 30% of students viewed making sense of what is learned in school as an important approach for Arzu:

I think, instead of memorizing formulas, she should make sense of them. Studying this way enables her to not forget her knowledge and to prevent her from memorizing physics knowledge. Since she makes sense of the topics, she can easily find any information she needs...

Indeed, 26% of students specifically mentioned learning the logic behind formulas instead of memorizing them. And 22% of students mentioned the need for Arzu to study concepts and the relations among them. As one student wrote,

She does not spend so much time learning formulas and solving problems as I do. She should allocate her time to learn concepts and the relationships among concepts and their connection with everyday life.

In brief, four of the six most common categories of students' suggestions for Arzu align with deep approaches to learning: relating physics knowledge to everyday life, making sense of what is learned (instead of just

memorizing it), understanding the logic behind formulas, and studying the concepts and the relations among them. The other two of the common suggestions for Arzu, doing experiments and/or making observations related to the topic and searching for more information on topics, could align with either surface or deep approaches to learning, depending on how they are implemented. However, students generally offered these as suggestions for deepening rather than just broadening Arzu's knowledge. For example, one student wrote,

If Arzu is not concerned about the CEE, she should not spend her time on solving tests and memorizing formulas. Instead, she should do research about what she learns in her class from the internet or other resources to learn the topic very deeply.

This suggestion is to use the internet to help “learn the topic very deeply,” not just to gather more information. Similarly, explaining why Arzu should do experiments, a student writes,

...At the same time, Arzu is interested in experiments from which physics formulas are derived and she tries some experiments which she is able to conduct by herself. She tries to understand experiments. For example, the velocity versus time graph for a motion with constant velocity is like that [drew a straight horizontal line]. The students who learn physics for the CEE like me learn this graph from books. However, Arzu learns this graph by making experiments. She learns for herself by using her experiences while I learn or memorize textbook information.

The student here notes that Arzu's experiments should help Arzu understand where formulas (and associated graphs) come from. The student underscores this point by flagging his own memorization-based approach to learning velocity graphs and contrasting that approach to how Arzu could use experiments to “[learn] for herself.” Along these same lines, a few students (<20%) proposed that Arzu should solve some practice problems—but with the aim of evaluating her own learning or increasing the durability of her learning rather than increasing her problem-solving speed.

Of course, not every student gave Arzu epistemologically sophisticated advice aligned with a deep approach to learning. A small percentage (<10%, 3 students) suggested “rote” approaches. For example, one student suggested strategies reminiscent of *dershane*:

If Arzu wants to learn physics deeply, she should not be content with what is taught in school, since the physics taught in the school is limited to topics suggested by the Ministry of Education. If she wants to learn

physics deeply, she should study the topics from different resource books and solve lots of questions. She must study by following a plan. If she does not want to be content with these, she should take lessons from experts.

These students' post MPEX2 scores ranged from 61 to 74. Considering the treatment group's mean score of 99.9 (with $SD = 14.7$), we can infer that those three students were epistemologically less sophisticated than the other students in the group.

For the most part, however, students' suggestions for Arzu aligned with a deep approach to learning, showing that the students had taken an sophisticated epistemic stance in suggesting advice for Arzu—a stance that physics knowledge is conceptual (as opposed to just formulas), connected to everyday experience, and something that can be made sense of instead of just memorized.

D. Reasons for students' study-time allocations for themselves and Arzu

We now discuss the reasons students indicated in their open-ended responses for their study-time allocations and why Arzu's should differ.

1. Practice problems

72% of students explicitly mentioned the CEE in discussing their focus on practice problems. Since the practice problems in test books are developed in part from previous CEEs, students view these problems as particularly good preparation. For example, one student wrote,

I study physics to attend to a good college. So my studying is exam-oriented. Because I think the students who solve a lot of practice problems will be more successful in the CEE, I allocate most of my study time to practice problems.

Specifically, students wrote that they solved numerous questions to learn all types of exam-like questions, to learn solution tactics, and to become faster at solving the problems. Familiarity with question types was considered a key to CEE success. And since the CEE is timed, students do not want to get slowed down by unfamiliar question types or inefficient solution strategies. As two students wrote,

Solving lots of questions means that I see lots of the question types. The more questions I solve, the more successful I will be in the CEE. Solving lots of questions increases our question-solving speed.

My aim is to succeed on the CEE. Physics will play an important role in my success in the exam. To be

prepared for physics questions in the exam and to reinforce what I learn, it is essential to solve questions in test books and understand their solutions.

2. Real-life examples and concepts

28% of students explicitly stated that they ignored or underemphasized real-life examples because doing so would not help them answer CEE questions. For example,

To be honest, I cannot allocate too much time to studying daily-life examples. If I did, it might be better for me. But I can solve the questions for the exam by only knowing topics and formulas.

Along these same lines, 48% students wrote that relating what is learned in school to everyday life is necessary for Arzu to learn physics deeply. They suggested that studying the physics of real-life examples could help her make sense of what she learns and make her learning more permanent.

She should spend most of her time on this to learn physics deeply. This increases the durability of her knowledge. Most of physics is related to life. If she makes connections between the two, she understands the topic.

Similarly, in advising Arzu, many students noted that studying concepts and relating them to real life is essential for learning physics deeply.

3. Formulas

54% of students reported that most CEE and test-book questions required knowing formulas:

I try to memorize formulas on physics topics for the CEE because the time is important in the CEE and some questions cannot be solved without formulas. So studying formulas is important for me.

The response just quoted illustrates a related issue: Although 15% of students noted that they could solve CEE questions without formulas, they thought knowing the formulas would increase their speed. As another student put it,

We can solve problems quickly and easily by knowing formulas. Of course, we can solve problems by using our logic but it takes time.

In brief, the results from this section reinforce our interpretation of earlier results: Students focus on practice problems and formulas largely because they need to solve problems quickly on the CEE, but they realize that a deeper understanding requires a different approach, centered around connecting concepts to real life.

E. How did students feel about needing to prepare for the CEE?

Our research question and survey instrument did not address students' emotions. Nonetheless, many students did express their feelings. Although these results are not systematic, they are worth presenting to motivate future research focused on students' emotions when faced with tensions between learning deeply and studying for exams.

Many students' responses hinted at frustration or at least ruefulness about the test-cramming lifestyle into which they were forced:

If I were not concerned about [CEE phase 1] and [CEE phase 2]: I would do a lot of experiments and make connections among the results of these experiments. Instead of memorizing formulas without understanding, I would try to understand the concepts that underlie formulas. By relating the topics to each other, I would find the connections between the topics I do not understand and the topic I understand; then, I would use this connection to understand. I would find examples in my everyday life. I would apply what I learn to my life in order to not forget. However, I am concerned about [CEE].

I can't make so many experiments in school. After understanding the topic, I study examples. Then, I solve at least 100 questions...

Another student was more explicit in labeling herself "unlucky" compared to Arzu:

Because we worry about the CEE, we study formulas more. We only study for the exam. Arzu is luckier than us in terms of this. If I were her, I would study topics with experiments. And I would do searches about topics more.

When these and other students explained why they allocated little time to studying real-life examples, they often seemed regretful. They claimed to value studying real-life examples but to lack the time. In the following example, the student is angry with educational policy makers for obliging students to take an exam to proceed to higher education. She blamed them for preventing her from learning physics:

Indeed, the best learning method for physics is giving everyday examples and applying what we learn to our life. However, I gave little time to this because of our education policy. Education policy says I cannot do so much [thinking about real-life examples], but must solve test questions.

Another student expressed a similar feeling:

To me, this [studying real life examples] is the most useful method for studying physics. But I cannot allocate so much time because I am studying for the CEE.

Of course, inferring emotions from written responses is imprecise at best; but in this paper we didn't want to ignore the frustration that students seemed to be expressing.

V. DISCUSSION

A. Conscious distortions of study habits

The point of this study was to explore how students think they approach learning physics when subjected to the competing influences of epistemological sophistication (e.g., ability to take a sophisticated epistemic stance toward learning physics) pushing them toward deep approaches to learning and high-stakes testing pulling them toward surface approaches to learning. Different outcomes were plausible. Because rote physics problems can also be approached through sense making [11], the students might have decided to take a deep approach to learning, thereby achieving a durable understanding and preparing for the CEE. No participants claimed to take this approach. Some noted that, while it's possible for someone with a deep understanding to "use logic" to solve CEE problems, doing so would take too long, given the CEE's time constraints.

Another plausible outcome, in light of the tension between the culture of CEE test preparation and the unusual physics curricular unit in which students were engaged, would have been for students to decide that the epistemic climate [56] of the curricular unit was a course-specific aberration and that "real" physics aligns more closely with what the CEE rewards. During early weeks of the unit, the class threatened to head in that direction. Many students noticed the tension and expressed confusion to the first author, who was often present as an observer. As one student put it,

...After the curriculum change, I got confused like other students: Should I solve questions full time to be good at test tactics or should I relate topics to daily life and give examples. I am now indecisive.

The teachers felt the tension, too. They worried if they were adequately preparing students for exams. However, the teachers persisted, as evidenced by the four-month duration of the unit; and the students remained sufficiently intellectually engaged to learn, as evidenced by the MPEX2 epistemological survey results mentioned above, and pre-post conceptual gains discussed elsewhere [38]. So, most students did not simply dismiss the epistemic views encouraged by the curricular unit.

Instead, most students displayed epistemic compartmentalization, by which we mean consciously taking different epistemic stances toward what counts as knowing and

learning physics in different contexts. Specifically, students drew a distinction between what counts as physics learning and knowing for the CEE and what counts as physics knowing and learning for deep understanding.

Before continuing, we need to defend our claim that students' own approaches to learning and their advice to Arzu corresponded to different epistemic stances as opposed to just different strategies undertaken within the same epistemic stance. To see why such an argument is needed, consider a hypothetical student who thinks that what counts as knowing physics for the CEE is the same as (or perhaps a watered down version of) what counts as knowing physics for Arzu, e.g., connected understanding of concepts and how those concepts are expressed by equations. But suppose the student adopts a *surface approach* to CEE preparation due to time constraints and the culture of *dershane*. This hypothetical student's different *strategies* for CEE preparation vs deep learning do not correspond to different epistemic stances. If this hypothetical student is typical of the actual students in this study, then we have not demonstrated epistemic compartmentalization.

However, students' written responses (discussed above) show that they were indeed taking different epistemic stances toward CEE preparation vs advising Arzu how to learn physics deeply. Their advice to Arzu explicitly mentioned the nature of physics knowledge as having a "connection with everyday life," as involving "relationships among concepts," and as something one can "[make] sense of." Within the epistemological resources framework, this stance might be toy modeled as including epistemological resources such as *Knowledge as constructed stuff* and *Rule systems* (a resource for understanding a form of knowledge in which rules are interconnected into a coherent whole). By contrast, in discussing their CEE study strategies, the students make clear that the physics knowledge they must learn does not include connections to everyday life. Instead, it consists of formulas to be memorized (as opposed to made sense of) and problem-solving procedures to drill. Those different procedures correspond to "lots of the question types," with no suggestion of coherence among them. A toy model of the network of epistemological resources corresponding to this stance might include *Knowledge as propagated stuff* and *Rule* (a resource for understanding the notion of a fixed rule) but not *Rule systems*. Our argument doesn't depend on the correctness of these toy models of the students' stances in terms of epistemological resources. Our point is that the students' vision of the physics knowledge they need to learn for the CEE differs from the vision of physics knowledge underlying their advice to Arzu. For this reason, we attribute distinct epistemic stances, and hence epistemic compartmentalization, across these two contexts.

Below, we relate this compartmentalization to other results in the literature on students' epistemologies. For now, however, we discuss the results of our study in more detail.

It's not surprising that academically elite students mold their study habits to prepare for the CEE, a test that, in the words of one student, determines their future careers. What's striking is the awareness and articulateness with which the students describe how and why their study habits and underlying epistemic stances differ substantially from those they recommend for pursuing deep understanding. When advising Arzu, most of the students know the importance of thinking about real-life observations and/or experiments and connecting them to physics concepts; they have a sense of deep approaches to learning, corresponding in part by a sophisticated epistemic stance. But that's not how they study physics for the CEE, for which they adopt surface approaches to learning based on a different epistemic stance.

High-stakes testing in Turkey, the United States, and elsewhere is intended in part to encourage and reward students' learning. So, we find irony and heartbreak in the result that the high-stakes testing regimen in Turkey systematically impedes deep learning of physics for college-bound students.

Defenders of high-stakes testing could argue that the problem isn't high-stakes tests *per se*, but rather, the rote, time-pressured nature of the CEE. Our study does not help settle this debate. However, our data point to characteristics that high-stakes tests should incorporate in order to encourage deep approaches to learning. In brief, students and teachers must believe that the most reliable and efficient way to succeed on the test is to actually understand the material deeply. If students think that test scores resulting from memorization and repetitive rote problem solving would be greater than or equal to the scores resulting from attaining a deep understanding, students will take the shortcuts. From students' responses in this study, we speculate that test creators wishing to encourage pursuit of deep understanding must avoid the reality and perception the following:

- *Time limitations* that do not allow students to figure things out.
- Problems for which memorization or rote approaches lead to success, even if approaches based on deep understanding also work.
- *Predictable problems or problem types* that teachers or test-prep programs think students can be drilled on.

B. Related results on epistemic compartmentalization, and what our study adds

The epistemic compartmentalization displayed in this study is a special case of the more general phenomenon of epistemological variability, which is when people manifest different epistemological stances, beliefs, or levels in different contexts. Research on personal epistemologies shows that students display different epistemological views about different disciplines, e.g., in thinking that knowledge in introductory science is more certain than knowledge in psychology [95,96]. Within physics, a smaller body of

literature documents epistemic variability stemming from differences in the physics task or activity, such as collaboratively learning concepts during a physics discussion section vs answering questions in a clinical interview framed as exploring "how you think" [97–100]. Those studies, however, did not focus on students' conscious awareness of the variation in their epistemic stances. Our study shows that a student's epistemological variability within physics can be epistemic compartmentalization—a conscious choice to take different epistemic stances toward different learning-related tasks within the discipline.

This study builds on and we think, improves upon Elby's [12] study using a similar survey. Elby's participants were community college students in several different introductory physics courses taught by different professors implementing a mix of different reforms aimed at fostering conceptual (but not necessarily epistemological) development. Elby, unlike the first author of this study, did not independently probe participants' epistemologies of physics using a separate instrument such as MPEX2. Hence, he could not establish that his participants had developed atypically sophisticated epistemologies. By contrast, the first author of this study documented a $d = 0.90$ effect size difference in epistemological sophistication (in the context of their physics courses, at least) between our participants and other physics students in the same grade in the same school taught by the same physics teachers. In addition, the CEE is a higher-stakes test than the course-specific midterms and exams considered by Elby's [12] participants. For these reasons, our study more cleanly documents epistemic compartmentalization between students' views about what constitutes knowing and learning for a high-stakes test and what constitutes knowing and learning for deep understanding. And our study, unlike Elby's, can address our research question, "When physics students subject to high-stakes testing have engaged in a curriculum that helped them develop more sophisticated epistemologies of physics, how do they claim to approach the learning of physics in light of those potentially competing influences?"

Another previously documented example of epistemic compartmentalization in physics comes from Gray and collaborators [101]. In some studies using the Likert-scale Colorado Learning Attitudes about Science Survey (CLASS), the researchers posed all items twice, asking (i) the student's own opinion and (ii) what answer the student thinks a physicist would give. So, these studies ask students to take on the epistemic stance they attribute to experts (physicists). Gray *et al.* document large "splits" between students' own epistemological views about physics and those they ascribe to experts, splits that resemble those we documented between the study strategy students attribute to themselves and recommend to Arzu. But the epistemic compartmentalization found by Gray *et al.* differs substantially from the compartmentalization documented in our study. We show that students draw a distinction

between how they learn physics for a test and how they think “a student just like you” should learn physics in pursuit of deep understanding. In other words, we document that students draw a distinction between how they actually study and how they could imagine students *like themselves* learning in pursuit of deep understanding—as opposed to how they imagine experts know and learn.

In brief, our study goes beyond prior work on epistemic variability and compartmentalization by showing that students capable of taking sophisticated epistemic stances toward physics draw a dramatic distinction between what counts as learning for a high-stakes test and what counts—for students—as learning for deep understanding, and by showing that students respond to this distinction by “distorting” their study habits toward surface approaches to learning.

C. What this study does and does not show

A possible critique of this study is that we relied on students’ self-reports of their study habits, not on observations of their actual behaviors. We agree that if our study were about students’ study habits, then our reliance on self-reports would be a fatal flaw. But this study isn’t about students’ study habits. It’s about their *perceptions* of differences between what they do to prepare for a high-stakes test and what they recommend doing to pursue deep understanding. Our conclusions about epistemic compartmentalization, how students consciously distinguish between what counts as learning for the CEE and what counts as learning for deep understanding, could not emerge from even the most fine-grained look at their study strategies. Of course, an ethnographic exploration of students’ actual learning behaviors, following them into their homes and *dershane* (cram schools), would yield fascinating data about how the CEE affects their day-to-day lives and how tensions between “CEE culture” and aspects of their personal epistemologies play out. Nonetheless, for addressing our research question, self-reports of behavior—and crucially, students’ explanations for their (perceived) choices—were what we needed.

We must acknowledge, however, that the *sizes* of the distortion effects in Table II and Fig. 1 may be larger than would be found by experiments using a different design. That is because the order of our survey questions—first asking students about their own study habits, then introducing a hypothetical student who does *not* need to worry about a college entrance exam—could prime students to accentuate differences between their own study strategies and their advice to Arzu. Still, we see evidence that those differences are psychologically real, not just a priming effect. Students were articulate and sometimes even emotional about the differences between Arzu’s situation and their own predicaments; they described not just how but also why they study differently from Arzu. Calling our results a “priming effect” is saying that students have

access to an epistemic stance associated with their advice to Arzu, a stance they might not take in everyday circumstances but that nonetheless exists in the sense that it can be tapped into when primed strongly enough. That students *can* take such a stance, but do not do so when studying for the CEE, is worth documenting, we think. Furthermore, the MPEX2 results discussed above suggest that students took a sophisticated epistemic stance toward physics even when priming wasn’t an issue.

Another feature of our study is both a strength and a limitation: the focus on elite high school students. The limitation is we cannot conclude that typical Turkish high school physics students, or physics students in other countries, respond as our participants did to the tensions between their personal epistemologies and a high-stakes testing environment. The advantage of our participant pool, however, is that the participants—because of their overall academic abilities, comparatively sophisticated epistemologies, and experiences in the first author’s epistemologically focused physics unit—were unusually well prepared to adopt and sustain deep approaches to learning physics. The fact that these students chose not to do so points to the power of the CEE to disrupt deep learning. Students less well prepared academically or less sophisticated epistemologically might be even more liable to succumb to the culture of high-stakes testing and its affordances toward surface approaches to learning. Of course, only future studies of a cross section of college-bound Turkish students could confirm this conjecture.

Interestingly, despite our documentation of epistemic compartmentalization, our study does not count as evidence against “unitary” (stage based or belief-systems based) theories of personal epistemology. To see why not, imagine that the students in our study have robust, context-independent, relatively sophisticated epistemological beliefs about physics, including the centrality of formulas and associated problem-solving procedures that express conceptual meaning. By this account, the epistemic stance students take when studying for the CEE draws on some but not all aspects of their (coherent, context-independent) epistemological belief system, e.g., the centrality of formulas and problem-solving procedures but not their connection to conceptual meaning. By contrast, researchers working from an epistemological resources perspective would model students as activating different networks of epistemological resources when studying for the CEE vs advising Arzu, and these two networks of resources would not necessarily be conjoined into a coherent whole that could itself be activated as a unit. Under either model, however, epistemic compartmentalization is occurring in the sense that students are relying upon different aspects of their personal epistemologies when advising Arzu vs describing how they study for the CEE.

In our view, the biggest limitation of this study is our reliance solely on written responses, with no interviews.

Interviews would have enabled us to pursue in more detail the tensions students felt between pursuing deep learning and preparing for the CEE. Still, we think this study is a worthwhile early exploration of test-induced epistemic compartmentalization.

D. Instructional implications

From an instructional standpoint, our results contain both good news and bad news. The good news: even students immersed in a culture of high-stakes testing can engage with a reform-oriented curriculum deeply enough to achieve some epistemological sophistication, as indicated by our participants' advice to Arzu and by the first author's earlier study. This sophistication may contribute to their ability to distinguish test preparation from deeper learning. In the future, the students could presumably leverage their understanding of what counts as learning physics deeply, perhaps in a reform-oriented physics course in college.

The bad news is that most students say they will not use their insights about deep approaches to learning when studying for the CEE. Given the hundreds of hours students spend on CEE preparation, this lost opportunity is heart-breaking to us as physics education researchers and practitioners. It also frustrates many students, as shown above.

Should Turkish high school physics teachers address this problem by emphasizing to students that gaining a deep understanding of physics can help them on the CEE? We question this advice for two reasons. First, we think our participants are correct that they must practice solving problems quickly (and hence by rote) in order to finish the CEE on time. Second, the students' epistemic compartmentalization of test preparation and deep learning into two distinct epistemic activities is potentially productive, both for CEE preparation and for future deep learning.

E. Implications for research

This study has methodological implications for researchers studying students' epistemologies and beliefs about learning. As demonstrated here and in Elby [12], students' views about what constitutes deep knowing and learning can differ from their views about what counts as knowing and learning in a specific context. This insight should affect both the construction and interpretation of surveys probing epistemologies and other beliefs. For instance, consider this Likert-scale item from the CLASS:

Knowledge in physics consists of many disconnected topics.

Our participants would likely agree if they were thinking of the "physics" needed for the CEE but would likely disagree if they were thinking of the physics needed for Arzu. As a result, researchers must be cautious in interpreting "beliefs" from survey items such as this. Our critique applies equally strongly to a similar item on MPEX [90], MPEX2, and epistemology surveys more generally.

Our study also suggests that at least some students have articulate views about the effects of high-stakes testing on their approaches to learning. Our results only scratch the surface of Turkish students' emotional reactions to the CEE and its broader effect on their academic and nonacademic lives. As mentioned above, interview-based and ethnographic studies could probe students' experiences more deeply.

VI. CONCLUSION

In this study, we explored how Turkish students reconcile—or do not reconcile—their experiences in a conceptually and epistemologically oriented physics curricular unit with their experiences preparing for the high-stakes CEE. Our subjects displayed epistemological sophistication, both in this study and in a previous study [38,39]. They understood aspects of a deep approach to learning. But the students perceived that CEE success hinges on their ability to quickly solve problems by rote. So, students' approaches to studying physics, largely driven by CEE preparation, focused on absorbing formulas and problem-solving algorithms at the expense of exploring concepts and real-life examples. By contrast, in recommending study strategies to "Arzu," a hypothetical student who doesn't need to take a college entrance exam and just wants to understand physics deeply, the students focused on linking physics concepts to each other and to real-life experiences and experiments. Overall, the students epistemically compartmentalize "learning physics deeply" and "learning physics for the CEE" into two distinct activities. In this way, the CEE hinders rather than scaffolds the learning of physics.

As discussed above, most studies about the effects of high-stakes testing report on data from teachers, administrators, or parents, but not students. Using students' written responses to open-ended questions extends the literature by (i) documenting how *students themselves* perceive the effects of high-stakes testing on their approaches to learning, and (ii) showing that high-stakes testing can induce surface approaches to learning even when students have the epistemological sophistication—and in many cases, the desire—to take a deep approach to learning. This study further highlights the insidiousness of high-stakes tests; even when the teacher didn't *teach* to the test, students still felt compelled to *learn* to the test.

ACKNOWLEDGMENTS

This study was supported by The Scientific and Technological Research Council of Turkey.

APPENDIX A: THE COMPLETE SURVEY USED IN THIS STUDY

1. How do you study physics?

2. Arzu is a student just like you, with the same age and abilities. But, since Arzu and her family will move to another country where no high-stakes entrance exam is used after she graduates from her high school, she does not need to worry about the College Entrance Examination. So, her goal is simply to understand physics more deeply. How should she study to reach her goal?

3. Assume that you divide your study time for physics into percentages. Specify the percentage you give to studying (a) concepts and relationships among the concepts (for example, velocity, acceleration and force are examples of concepts; an example of the relationship among them may be the tendency of objects to continue moving in a straight line at a constant speed unless a push or pull changes the motion), (b) formulas, (c) practice problems in the test books, and (d) real-life examples.

a. Studying the concepts and the relationships among the concepts ...%

Why do you give that percentage?

b. Studying formulas ...%

Why do you give that percentage?

c. Studying practice problems in the test books ...%

Why do you give that percentage?

d. Studying real-life examples ...%

Why do you give that percentage?

4. How would Arzu's percentages be?

a. Studying the concepts and the relationships among the concepts ...%

Why would she give that percentage?

b. Studying formulas ...%

Why would she give that percentage?

c. Studying the practice problems in the test books ...%

Why would she give that percentage?

d. Studying real-life examples ...%

Why would she give that percentage?

APPENDIX B: EXAMPLES OF LOWER-INFERENCE VS HIGHER-INFERENCE CODES

- (1) I solve all tests in my test book. Moreover, I solve questions in different resource books to be familiar with different types of questions.
- (2) During solving tests, first, I study examples to learn question solving styles. Then, I try to solve as many questions as possible.
- (3) I repeat what I learn and study exercises and solve problems from test books at my leisure to consolidate them.
- (4) I examine examples related to the topic and then, I solve questions related to the topic from my books.
- (5) I start to solve test questions. If there are any questions I cannot solve, I put a question mark next to questions and I continue. Then I ask these questions to my teacher. If I cannot solve most of the questions, I try to find similar example questions from my notebook and I try to practice by solving similar questions over and over. Or I ask my teacher from my training center to explain the topic again.

Quotes 1 and 2 clearly fit this code (low inference). With medium-low inference, we assign this code to quote 3 (the student refers to multiple test books) and quote 4 (the student refers to multiple books and lists this strategy as the second of just two strategies). With medium inference, we also place quote 5 in this category, on the grounds that the student lays out a conditional loop of working through problem after problem until reaching one or more for which he needs help. One form of help is solving yet more problems "over and over" from other sources. But even when the help takes another form, we infer that, after receiving help, the student either resumes solving practice problems about the topic at hand or else moves on to the next topic and starts the loop again, solving lots of problems.

-
- [1] F. Marton and R. Säljö, On Qualitative Differences in Learning: I-Outcome and process, *Br. J. Educ. Psychol.* **46**, 4 (1976).
 - [2] J. Biggs, The role of metacognition in enhancing learning, *Aust. J. Educ.* **32**, 127 (1988).
 - [3] J. B. Biggs, Individual and group differences in study processes, *Br. J. Educ. Psychol.* **48**, 266 (1978).
 - [4] J. Biggs, Individual differences in study processes and the quality of learning outcomes, *Higher Educ.* **8**, 381 (1979).
 - [5] J. M. Case and R. F. Gunstone, Approaches to learning in a second year chemical engineering course, *Int. J. Sci. Educ.* **25**, 801 (2003).
 - [6] C. Chin and D. E. Brown, Learning in science: A comparison of deep and surface approaches, *J. Res. Sci. Teach.* **37**, 109 (2000).
 - [7] E. Hazel, M. Prosser, and K. Trigwell, Variation in learning orchestration in university biology courses, *Int. J. Sci. Educ.* **24**, 737 (2002).
 - [8] D. J. Watters and J. J. Watters, Approaches to learning by students in the biological sciences: Implications for teaching, *Int. J. Sci. Educ.* **29**, 19 (2007).
 - [9] B. K. Hofer and P. R. Pintrich, The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning, *Rev. Educ. Res.* **67**, 88 (1997).

- [10] M. Schommer, A. Cruse, and N. Rhodes, Epistemological Beliefs and Mathematical Text Comprehension: Believing it is simple does not make it so, *J. Educ. Psychol.* **84**, 435 (1992).
- [11] D. Hammer, Epistemological beliefs in introductory physics, *Cognit. Instr.* **12**, 151 (1994).
- [12] A. Elby, Another reason that physics students learn by rote, *Am. J. Phys.* **67**, S52 (1999).
- [13] N. J. Entwistle and A. Entwistle, Contrasting forms of understanding for degree examinations: the student experience and its implications, *Higher Educ.* **22**, 205 (1991).
- [14] A. Entwistle and N. Entwistle, Experiences of understanding in revising for degree examinations, *Learn. Instr.* **2**, 1 (1992).
- [15] J. Gulikers, T. J. Bastiaens, P. A. Kirschner, and L. Kester, Relations between student perceptions of assessment authenticity, study approaches and learning outcome, *Studies in educational evaluation* **32**, 381 (2006).
- [16] F. Marton and R. Säljö, On qualitative differences in learning-II Outcome as a function of the learner's conception of the task, *Br. J. Educ. Psychol.* **46**, 115 (1976).
- [17] K. Scouller, The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay, *Higher Educ.* **35**, 453 (1998).
- [18] C. Tang, Assessment and Student Learning: Effects of Modes of Assessment on Students' Preparation Strategies, in *Improving Student Learning: Theory and Practice*, edited by G. Gibbs (Oxford Centre for Staff Development, Oxford, 1994), pp. 151–170.
- [19] D. Watkins, Factors influencing the study methods of Australian tertiary students, *Higher Educ.* **11**, 369 (1982).
- [20] K. M. Edmondson and J. D. Novak, The interplay of scientific epistemological views, learning strategies, and attitudes of college students, *J. Res. Sci. Teach.* **30**, 547 (1993).
- [21] M. H. Lee, R. E. Johanson, and C. C. Tsai, Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis, *Sci. Educ.* **92**, 191 (2008).
- [22] K. Hogan, Relating students' personal frameworks for science learning to their cognition in collaborative contexts, *Sci. Educ.* **83**, 1 (1999).
- [23] S. A. Rosenberg, D. Hammer, and J. Phelan, Multiple epistemological coherences in an eighth-grade discussion of the rock cycle, *J. Learn. Sci.* **15**, 261 (2006).
- [24] W. A. Sandoval, Understanding students' practical epistemologies and their influence on learning through inquiry, *Sci. Educ.* **89**, 634 (2005).
- [25] C. C. Tsai, An analysis of scientific epistemological beliefs and learning orientations of Taiwanese eighth graders, *Sci. Educ.* **82**, 473 (1998).
- [26] L. M. Abrams, J. J. Pedulla, and G. F. Madaus, Views from the classroom: Teachers' opinions of statewide testing programs, *Theory Into Practice* **42**, 18 (2003).
- [27] A. L. Amrein and D. C. Berliner, The effects of high-stakes testing on student motivation and learning, *Educ. Leadership* **60**, 32 (2003).
- [28] M. A. Barksdale-Ladd and K. F. Thomas, What's at stake in high-stakes testing teachers and parents speak out, *J. Teach. Educ.* **51**, 384 (2000).
- [29] Y. J. Dori, From nationwide standardized testing to school-based alternative embedded assessment in Israel: Students' performance in the matriculation 2000 project, *J. Res. Sci. Teach.* **40**, 34 (2003).
- [30] J. Dounay, High-stakes testing is high-stress, too, *Educ. Digest.* **65**, 9 (2000).
- [31] C. A. Grant, Oppression, privilege, and high-stakes testing, *Multicult. Perspect.* **6**, 3 (2004).
- [32] F. M. Hess and F. Brigham, How kids pay for high-stakes testing, *Educ. Digest.* **66**, 11 (2000).
- [33] D. Hursh, The growth of high-stakes testing in the USA: accountability, markets and the decline in educational equality, *Br. Educ. Res. J.* **31**, 605 (2005).
- [34] M. G. Jones, B. D. Jones, B. Hardin, L. Chapman, T. Yarbrough, and M. Davis, The impact of high-stakes testing on teachers and students in North Carolina, *Phi Delta Kappan* **81**, 199 (1999).
- [35] G. F. Madaus, The effects of important tests on students: Implications for a national examination system, *Phi Delta Kappan* **73**, 226 (1991).
- [36] L. A. Shepard, Will national tests improve student learning?, *Phi Delta Kappan* **73**, 232 (1991).
- [37] L. Valli, R. Croninger, M. Chambliss, A. Graeber, and D. Buese, *Test Driven: High-Stakes Accountability in Elementary Schools* (Teachers College Press, New York, 2008).
- [38] S. Yerdelen-Damar, Ph.D. thesis, Middle East Technical University, Ankara, Turkey, 2013.
- [39] S. Yerdelen-Damar and A. Eryılmaz, The impact of metacognitive 7 E learning cycle on students' epistemological understandings, *Kastamonu Educ. J.* (to be published).
- [40] D. Kember, The intention to both memorise and understand: Another approach to learning?, *Higher Educ.* **31**, 341 (1996).
- [41] J. Biggs, D. Kember, and D. Y. Leung, The revised two-factor study process questionnaire: R-SPQ-2 F, *Br. J. Educ. Psychol.* **71**, 133 (2001).
- [42] J. B. Biggs, *Study Process Questionnaire Manual. Student Approaches to Learning and Studying* (Australian Council for Educational Research Ltd., Radford House, Frederick St., Hawthorn, Australia, 1987).
- [43] M. P. Jimenez-Aleixandre, A. B. Rodriguez, and R. A. Duschl, "Doing the lesson" or "doing science": Argument in high school genetics, *Sci. Educ.* **84**, 757 (2000).
- [44] L. K. Berland and D. Hammer, Framing for scientific argumentation, *J. Res. Sci. Teach.* **49**, 68 (2012).
- [45] L. K. Berland and B. J. Reiser, Making sense of argumentation and explanation. *Sci. Educ.* **93**, 26 (2009).
- [46] A. Cavaghetto, B. M. Hand, and L. Norton-Meier, The nature of elementary student science discourse in the context of the science writing heuristic approach, *Int. J. Sci. Educ.* **32**, 427 (2010).
- [47] P. Hutchison and D. Hammer, Attending to student epistemological framing in a science classroom, *Sci. Educ.* **94**, 506 (2010).

- [48] A. W. Oliveira and T. D. Sadler, Interactive patterns and conceptual convergence during student collaborations in science, *J. Res. Sci. Teach.* **45**, 634 (2008).
- [49] I. Tabak and E. Baumgartner, The teacher as partner: Exploring participant structures, symmetry, and identity work in scaffolding, *Cognit. Instr.* **22**, 393 (2004).
- [50] G. L. Chiou, J. C. Liang, and C. C. Tsai, Undergraduate students' conceptions of and approaches to learning in biology: A study of their structural models and gender differences, *Int. J. Sci. Educ.* **34**, 167 (2012).
- [51] J. A. Greene, K. R. Muis, and S. Pieschl, The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments: Conceptual and methodological issues, *Educ. Psychol.* **45**, 245 (2010).
- [52] I. Bråten and H. I. Strømsø, The relationship between epistemological beliefs, implicit theories of intelligence, and self-regulated learning among Norwegian postsecondary students, *Br. J. Educ. Psychol.* **75**, 539 (2005).
- [53] R. Bromme, S. Pieschl, and E. Stahl, Epistemological beliefs are standards for adaptive learning: a functional theory about epistemological beliefs and metacognition, *Metacogn. Learn.* **5**, 7 (2010).
- [54] J. L. Burnette, E. H. O'Boyle, E. M. VanEpps, J. M. Pollack, and E. J. Finkel, Mind-sets matter: A meta-analytic review of implicit theories and self-regulation, *Psychol. Bull.* **139**, 655 (2013).
- [55] M. B. Paulsen and K. A. Feldman, The conditional and interaction effects of epistemological beliefs on the self-regulated learning of college students: Motivational strategies, *Res. High. Educ.* **46**, 731 (2005).
- [56] K. R. Muis, The role of epistemic beliefs in self-regulated learning, *Educ. Psychol.* **42**, 173 (2007).
- [57] W. Au, High-stakes testing, and curricular control: A qualitative metasynthesis, *Educ. Res.* **36**, 258 (2007).
- [58] D. Gerwin and F. Visone, The freedom to teach: contrasting history teaching in elective and state-tested courses, *Theor. Res. Soc. Educ.* **34**, 259 (2006).
- [59] M. G. Gunzenhauser, High-stakes testing and the default philosophy of education, *Theory Into Practice* **42**, 51 (2003).
- [60] R. G. Lomax, M. M. West, M. C. Harmon, K. A. Viator, and G. F. Madaus, The impact of mandated standardized testing on minority students, *J. Negro Educ.* **64**, 171 (1995).
- [61] M. Goldberg, Losing students to high-stakes testing, *Educ. Digest* **70**, 10 (2005).
- [62] K. E. Vogler, Impact of a high school graduation examination on Tennessee science teachers' instructional practices, *Am. Secondary Educ.* **35**, 33 (2006).
- [63] H. Kelecioğlu, Ortaöğretim öğrencilerinin üniversiteye giriş sınavları ve sınavın öğrenimlerine etkisi hakkındaki görüşleri, *Hacet. Unit. Egit. Fak.* **23**, 135 (2002).
- [64] M. Galton, Continuity and progression in science teaching at key stages 2 and 3, *Cambridge J. Educ.* **32**, 249 (2002).
- [65] E. W. Jenkins, The impact of the national curriculum on secondary school science teaching in England and Wales, *Int. J. Sci. Educ.* **22**, 325 (2000).
- [66] K. S. Louis, K. Febey, and R. Schroeder, State-mandated accountability in high schools: Teachers' interpretations of a new era, *Educ. Eval. Policy Anal.* **27**, 177 (2005).
- [67] R. M. Pringle and S. C. Martin, The potential impacts of upcoming high-stakes testing on the teaching of science in elementary classrooms, *Res. Sci. Educ.* **35**, 347 (2005).
- [68] L. K. Smith and S. A. Southerland, Reforming practice or modifying reforms?: Elementary teachers' response to the tools of reform, *J. Res. Sci. Teach.* **44**, 396 (2007).
- [69] A. R. Taylor, M. G. Jones, B. Broadwell, and T. Oppewal, Creativity, inquiry, or accountability? Scientists' and teachers' perceptions of science education, *Sci. Educ.* **92**, 1058 (2008).
- [70] M. F. Wideen, T. O'Shea, I. Pye, and G. Ivany, High-stakes testing and the teaching of science, *Can. J. Educ.* **22**, 428 (1997).
- [71] K. J. Anderson, Science education and test-based accountability: Reviewing their relationship and exploring implications for future policy, *Sci. Educ.* **96**, 104 (2012).
- [72] D. Hammer and A. Elby, On the form of a personal epistemology, in *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, edited by B. K. Hofer and P. R. Pintrich (Lawrence Erlbaum, Mahwah, NJ, 2002), pp. 169–190.
- [73] D. Hammer and A. Elby, Tapping epistemological resources for learning physics, *J. Learn. Sci.* **12**, 53 (2003).
- [74] L. Louca, A. Elby, D. Hammer, and T. Kagey, Epistemological resources: Applying a new epistemological framework to science instruction, *Educ. Psychol.* **39**, 57 (2004).
- [75] A. A. diSessa, Towards an epistemology of physics, *Cognit. Instr.* **10**, 105 (1993).
- [76] OSYM Exam Archive, <http://osym.gov.tr/belge/1-11489/2010-yili.html>.
- [77] OSYM 2010 publications, <http://www.osym.gov.tr/belge/1-12381/yuksekgoretime-giriste-okul-turu-ve-ogrenim-durumuna-go-.html>.
- [78] A. Tansel and F. Bircan, Demand for education in Turkey: A tobit analysis of private tutoring expenditures, *Econ. Educ. Rev.* **25**, 303 (2006).
- [79] Turkish physics curriculum development commission, <http://w3.gazi.edu.tr/~bgunes/fizik/>.
- [80] A. Eryılmaz and A. Arslan, Proposed teaching and assessment approaches in the curriculum, presented at GIREP- International Research Group on Physics Teaching, Physics Curriculum Design, Development and Validation, Nicosia, Cyprus, 2008 (unpublished).
- [81] H. Ş. Ayvaci, Views of physics teachers about context based approach, *Dicle University J. Ziya Gökalp Faculty of Educ.* **15**, 42 (2010).
- [82] S. Demir and A. Demir, New High School Instructional Programs in Turkey: Problems, Expectations and Suggestions, *Elem. Educ. Online* **11**, 35 (2012).
- [83] F. Ç. Mercan, Turkish physics teachers' views about the 2007 physics teaching program, and its implementation, *Educ. Res. Rev.* **8**, 1559 (2013).
- [84] Turkish Education Association, <http://portal.ted.org.tr/yayinlar/ortaogretimeveyuksekgoretimegecissistemi.pdf>.
- [85] A. Eisenkraft, Expanding the 5 E model, *Sci. Teach.* **70**, 56 (2003).

- [86] A. Einstein, Physics and Reality, *J. Franklin Inst.* **221**, 349 (1936).
- [87] A. Elby, R. E. Scherr, T. McCaskey, R. Hodges, E. F. Redish, D. Hammer, and T. Bing, *Maryland Tutorials in Physics Sense-making*, DVD (National Science Foundation, Washington, DC, 2007).
- [88] E. F. Redish and D. Hammer, Reinventing college physics for biologists: Explicating an epistemological curriculum, *Am. J. Phys.* **77**, 629 (2009).
- [89] S. Yerdelen-Damar, A. Elby, and A. Eryilmaz, Applying beliefs and resources frameworks to the psychometric analyses of an epistemology survey, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010104 (2012).
- [90] E. F. Redish, J. M. Saul, and R. N. Steinberg, Student expectations in introductory physics, *Am. J. Phys.* **66**, 212 (1998).
- [91] Yerdelen-Damar's study applied the abridged version of the Turkish MPEX2 in which Items 1, 26 and 32 were not used for reducing students' reading burden. Thus, the possible maximum and minimum scores were less than those of the original MPEX2. The treatment and control group scores are not easily compared to those of other populations because researchers often report MPEX2 scores in terms of the percentage of favorable and unfavorable (and neutral) responses, which effectively collapses the 5-point Likert scale to 3 points. We kept the 5-point scale, and for each item, each student was awarded 1 point for the worst answer (e.g., strongly disagreeing with a favorable statement), 2 points for the second worst answer, and so on, up to 5 points for the best answer, e.g., strongly agreeing with a favorable statement. So, on the 29-item survey, the lowest possible score is $29 \times 1 = 29$, while the highest possible score is $29 \times 5 = 145$. As a back-of-the-envelope approximation, however, we can characterize how MPEX2 results reported in "standard form" could correspond to the results reported in this study. For a generic population of MPEX2 respondents, if we assume that "favorable" answers are evenly split between the two best answers (e.g., agreeing or strongly agreeing with a favorable statement) and unfavorable answers are evenly split between the two worst answers, then the treatment group's average MPEX2 score of 99.9 could correspond to 55% favorable, 20% neutral, and 25% unfavorable responses, almost exactly the results reported by Redish and Hammer for an epistemology focused introductory college physics class. By contrast, the control group's average MPEX score of 87.2 corresponds to equal percentages of favorable and unfavorable responses, such as 40% favorable, 20% neutral, and 40% unfavorable.
- [92] B. G. Glaser and A. L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research* (Aldine de Gruyter, Chicago, 1967).
- [93] G. S. Åkerlind, Variation, and commonality in phenomenographic research methods, *Higher Educ. Res. Dev.* **31**, 115 (2012).
- [94] To decide whether the second coder should code more responses for "How do you study physics," the only prompt for which interrater agreement was $<90\%$ (namely, 53 out of 63), we set 75% agreement as the absolute lowest acceptable and considered the question, If the "true" interrater reliability is 75%, what is the probability that a set of 63 codes would yield ≥ 53 coding agreements. Assuming a "weighted coin flip" (binomial) distribution with a true agreement probability of 0.75, the probability of ≥ 53 agreements is 0.058. So, if the true agreement percentage were only 75%, it's very unlikely we would have obtained 84% agreement. This gave us sufficient confidence that our interrater reliability is above 75%. We should acknowledge that the use of percentage agreement might overestimate the agreement rate compared to Cohen's kappa and other conceptually similar interrater reliability coefficients, since percentage agreement does not adjust for agreements by chance.
- [95] B. K. Hofer, Dimensionality and disciplinary differences in personal epistemology, *Contemp. Educ. Psychol.* **25**, 378 (2000).
- [96] K. R. Muis, L. D. Bendixen, and F. C. Haerle, Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework, *Educ. Psychol. Rev.* **18**, 3 (2006).
- [97] L. Lising and A. Elby, The impact of epistemology on learning: A case study from introductory physics, *Am. J. Phys.* **73**, 372 (2005).
- [98] J. Leach, R. Millar, J. Ryder, and M. G. Sere, Epistemological understanding in science learning: The consistency of representations across contexts, *Learn. Instr.* **10**, 497 (2000).
- [99] W. M. Roth and A. Roychoudhury, Physics students' epistemologies and views about knowing and learning, *J. Res. Sci. Teach.* **31**, 5 (1994).
- [100] K. R. Muis and B. Gierus, Beliefs about knowledge, knowing, and learning: differences across knowledge types in physics, *J. Exp. Educ.* **82**, 408 (2014).
- [101] K. E. Gray, W. K. Adams, C. E. Wieman, and K. K. Perkins, Students know what physicists believe, but they don't agree: A study using the CLASS survey, *Phys. Rev. ST Phys. Educ. Res.* **4**, 020106 (2008).