Teaching problem solving in undergraduate physics courses: An endorsement for deliberate practice

Kelly Miller, Olivia Miller, and Georgia Lawrence

Department of Physics and Division of Engineering and Applied Sciences, Harvard University, 9 Oxford Street, Cambridge, Massachusetts 02138, USA

(Received 11 August 2025; accepted 7 July 2026; published 27 August 2026)

Developing expertlike problem-solving skills is a central goal of undergraduate physics education. In this study, we investigate the impact of teaching explicit problem-solving frameworks, combined with deliberate practice, on students' problem-solving approaches. Using multidimensional scaling to analyze students' decision-making patterns, we compare the similarity of students taught with these methods to physics experts and to students taught with traditional repeated practice. Our results show that students who received structured frameworks and targeted feedback through deliberate practice exhibited problem-solving behaviors significantly more aligned with those of experts. These findings suggest that pedagogies emphasizing explicit strategy instruction and feedback are more effective than rote repetition for fostering expertise. We recommend the integration of these approaches into physics curricula to better support the development of skilled and adaptive problem solvers.

DOI: 10.1103/7pqt-gd9c

I. INTRODUCTION

The importance of developing problem-solving skills during an undergraduate education in science, technology, engineering, and mathematics (STEM) is widely accepted [1–3]. Recent engineering graduates claim problem solving as one of the most valuable skills they use in their careers [4]. In a 2011 study of physics PhDs 10 years into their careers, the American Institute of Physics (AIP) reported that problem solving was cited as the second most important factor related to their career success (second only to hard work) [5]. While many physics education research studies have been conducted on students' conceptual difficulties and pedagogies that effectively promote content mastery, fewer studies have focused on how to develop and especially measure students' ability to solve physics problems [6]. Several physics education researchers have raised the issue of the community placing too much emphasis on conceptual understanding at the expense of problem-solving skill development [7]. Over the past decade, physics education research has substantially expanded its focus on problem solving—both to better characterize what expertlike problem solving entails and to develop instruction and assessments that target those skills. In particular, Docktor and Mestre synthesized the discipline-based education research in physics and summarized a large body of problem-solving work through 2014, highlighting persistent challenges in helping students coordinate conceptual knowledge, representations, and metacognition during problem solving [8]. Building on this foundation, more recent work has produced validated tools for evaluating students' written problem solutions (e.g., rubric-based approaches) [9] as well as classroom-ready scaffolds and templates to support planning, monitoring, and reflection during problem solving [10]. Complementary studies in recent years have also examined targeted interventions to strengthen strategic components of problem solving such as metacognitive reflection [11] and self-explanation [12]. Despite this growing literature, there remains a need for studies that directly test whether teaching an explicit, generalizable problem-solving framework shifts students' problem-solving approach toward more expertlike patterns *and* to measure those shifts using a stand-alone instrument designed to emphasize process over content mastery.

The purpose of this study is to investigate the following three research questions pertaining to teaching and assessing problem-solving skills in the undergraduate physics classroom:

1. How can students' problem-solving skills be effectively measured and evaluated using a stand-alone diagnostic instrument or survey?
2. What could such an instrument tell us about the difference between novice (student) versus expert problem solvers?

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

3. How do novice problem-solving skills differ between those explicitly taught a problem-solving framework and those not taught a framework?

II. BACKGROUND

A. Problem solving and cognition

The Program for International Student Assessment (PISA) defines the problem-solving competency as "an individual's capacity to engage in cognitive processing to understand and resolve problem situations where a method of solution is not immediately obvious." Problem solving has also been defined as "the cognitive process of moving toward a goal when the path is uncertain" [13]. Similarly, Carl Wieman and Wendy Adams define it as "cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver" [14]. All three definitions emphasize that for true problem solving to occur, the solver does not know, at least initially, how to solve the problem. This leads to an important distinction between "authentic problems" and "exercises." Typical textbook problems are exercises because to solve them, students are required to apply a known procedure. The designation of a task as an authentic problem is dependent on the solver—if they know how to solve it, it is, by definition, not an authentic problem.

All the above definitions emphasize that problem solving is a cognitive process. Other definitions have stressed that cognitively, the problem solver must construct knowledge and organize behavioral processes to cope with difficulties, and as a result, problem solving is closely linked to creativity [15]. Cognitively, problem solving involves behaviors associated with educational goals that are high up on Bloom's taxonomy. "Authentic" problems require cognitive processes that appear high on Bloom's taxonomy because they require students to design, construct, develop, and formulate solutions that they have never seen before. Metacognition (awareness and understanding of one's own thought processes) is also considered to be an important cognitive skill needed to be an effective problem solver [13]. Managing time and direction, determining the next step, monitoring understanding, asking skeptical questions, and reflecting on one's own learning process are all important behaviors associated with successful problem solving [13].

B. Teaching problem-solving skills

Given its importance in postcollege career success, there is a trend in STEM pedagogy toward emphasizing problem-solving skills over content knowledge [16]. Several teaching strategies have been developed to help students become competent problem solvers. Many introductory physics courses teach students to follow a prescriptive problem-solving procedure based on frameworks of expert problem solving. George Pólya, a Hungarian mathematician, developed the first framework, a four-step problem-solving procedure [17]. Heller's competent problem solver is a similar, five-step strategy based on Heller and Reif's prescriptive model of physics problem solving [18]. Mazur developed a four-step problem-solving strategy to help frame the process for introductory physics students [19]. The steps involved in each of these frameworks are outlined in Table I and organized in such a way that similar steps from the three frameworks are lined up with one another. All three strategies involve an initial step of understanding the problem, making assumptions, and thinking about how best to represent the physics. They all also involve a final, metacognitive step that prompts students to reflect on the reasonableness of their answer. The pedagogical value of this final step is well founded in the literature, which shows that metacognition is needed to be an effective problem solver [13]. This is, by no means, a comprehensive list of frameworks but provides an overview of those most commonly used in the context of physics problem solving.

C. Deliberate practice versus repeated practice

Deliberate practice and repeated practice are both pedagogical practices used to build skills, but they differ significantly in structure, feedback, and effectiveness. Deliberate practice is more targeted, feedback-driven, and focuses on breaking skills down into smaller "subskills." Repeated practice, on the other hand, is often unstructured, focuses on repetition, and has infrequent or nonexistent feedback. Deliberate practice has been shown to be more effective for developing expertise than simple repeated practice [20]. Repeated practice has been shown to be effective for skill maintenance or basic improvement, whereas deliberate practice is more effective for skill mastery and expertise, particularly in high-level skills such as problem solving [21,22].

D. Novices versus experts

Much of the work on physics problem solving has studied the differences between novice and expert problem solvers [13,14,23–27]. Many studies have shown that students often learn how to solve quantitative problems by plugging values into algorithmic equations and pattern matching. Consequently, they are not developing the necessary skills to transfer their understanding to unseen situations [28–33]. Other studies have found differences in both the knowledge structure and the problem-solving strategy typically used by experts compared to novices. Reif and Heller [33] found that the main difference between novices and experts was how they organized and used their knowledge in the context of solving a problem. Experts organize their knowledge in a structured, cohesive way and can activate these knowledge structures when they are needed. Novices typically do not have these knowledge structures; rather, their knowledge consists of

TABLE I. Problem-solving frameworks used to scaffold the process for students.

Pólya's framework (1945)	Heller's framework (1984)	Mazur's framework (2014)
Understand the problem -translate the situation and goals into fundamental concepts and decide reasonable idealizations and approximations needed	**Focus the problem** -develop a qualitative description of the problem -first visualize the problem by creating a sketch and then write down a statement of what you want to find **Describe the physics** -use qualitative understanding of the problem to prepare for the quantitative solution	**Getting started** -organize information given and understand what is asked in the problem -think about which principles apply and which simplifying assumptions are helpful
Devise a plan -apply specialized techniques and use specific concepts to connect the situation with the goal	**Plan the solution** -translate the physics description into a set of equations that represent the problem	**Devise Plan** -ascertain what information is needed to solve the problem -what physical relationships (or equations) can be applied?
Carry out the plan -follow through with the specialized plan developed in "devise a plan"	**Execute the plan** -execute the planned solution -use equations planned to find a numerical solution	**Execute plan** -follow the steps outlined in the previous sections -carry out any necessary mathematical operations
Look back -ask the questions: Can I check the result? Can I derive the solution differently?	**Evaluate the plan** -check work to see if it is properly stated, reasonable, and has answered the question asked	**Evaluate result** -reflect on your work, examine the result, and determine whether your answer is reasonable

random facts and equations that are context-specific and lacking in conceptual meaning [34,35]. Experts redescribe the problem and use qualitative arguments to plan solutions before describing the details of the problem from a mathematical perspective. Novices tend to rush to quickly string together various miscellaneous mathematical equations [33].

Wieman and Adams [14] found that when solving authentic problems, experts spend more time analyzing, planning, and managing their own behavior than novices and generally demonstrate a more holistic and systematic approach. Heller [13] found that experts usually follow specific steps when solving a problem: understanding the problem, determining the concepts, making the plan, solving the problem, and evaluating the outcome. Novices, on the other hand, first try to solve problems by using mathematical expressions. Expert problem solvers take more time to understand the problem and the concepts involved, as well as to explore the relationships between concepts. Novice problem solvers cannot establish these relationships, especially when the problem is complex [13].

E. Measuring problem solving

A range of methodologies are used to measure problem-solving skill level. They range from interviews [36] to think-aloud protocols [37] to expert self-reflections [17] to observing participants solving problems in a lab setting [38] to computer-delivered assessment [39].

1. Computer-delivered assessments

Computer-delivered assessment is emerging as a common methodology for studying problem solving. The PISA 2012 assessment of problem solving was computer-delivered for the first time. These assessments involve problems that do not require any specific disciplinary knowledge for their solution, so the focus can be placed on measuring the cognitive processes fundamental to problem solving. Students are presented with a scenario, and a series of test items about the scenario are posed to measure students' performance in four cognitive processes involved in problem solving (exploring and understanding, representing and formulating, planning and executing, and monitoring and reflecting) [39]. Each test item occupies a single computer screen, and students proceed from item to item without being able to skip items or go back to previous items. Test items consist of a variety of response formats, which can be automatically coded (multiple-choice, drag-and-drop, and constructed-response), and free text response, which requires coding by experts.

2. Interviews and think-aloud protocols and rubrics

Many studies that measure and assess problem solving use interviews and/or think-aloud protocols. Students (and sometimes experts) are interviewed by researchers who pose questions designed to understand the approach being taken to solving a problem. Think-aloud is a semistructured cognitive interviewing method in which a person is asked to verbalize their thought process as they do a specific task, during which they are recorded (on paper, audio, or video) for further analysis. In the study published by Alii *et al.* [37], 21 students were asked to narrate their thinking while solving physics problems, and all data were recorded. Afterwards, qualitative interviews were held with the students. Interviews (or think-aloud recordings) are typically transcribed, and the text is analyzed and categorized (or coded) by several experts in the subject-matter domain. Interrater reliability is monitored by measuring the level of agreement between the expert coders to ensure validity. The interviews and generation of codes are often iterative. Rubrics are a common measurement tool used to evaluate students' written work and are often used as part of the methodology in problem-solving studies. Halim *et al.* studied students' ability to apply problem-solving strategies in physics [40]. Students were asked to solve routine problems on paper, and rubrics were used as the measurement tool. Burkholder *et al*. developed a solution template for students to use as they solve problems [10]. Responses to students' problem solving using this template were collected, and then the template was used as a rubric for measuring how expertlike students' problem-solving thinking was.

F. Characterization of expert problem-solving process

Recently, important research has been done using think-aloud interviews to characterize the subskills of problem solving to understand how experts make decisions when they solve problems. Price *et al*. [36] interviewed 52 science and engineering experts and asked them to explain how they go about solving problems in their fields. The study found a surprising degree of overlap in the decisions made by experts across various fields. This consistency suggests that there are common decision elements shared across disciplines in science and engineering. Table II summarizes this framework and provides a list of the decisions that Price *et al.* found experts commonly made when solving problems in their fields.

The framework of decisions defined by Price *et al.* (outlined in Table II) provides a structured approach to assessing problem-solving skills, and it is this framework that we base our analysis for this study. The rows shaded in gray, in Table II, represent the decisions that we focus on. Due to the nature of the questions on our problem-solving diagnostic instrument (described below in the methods section), we only coded for decisions 3 through 15. The other decisions were outside the scope of the questions on the instrument.

III. METHODS

We developed a computer-delivered physics problem-solving diagnostic instrument designed to measure an individual's ability to solve a complex physics problem. This instrument consists of a series of questions that ask respondents to consider a problem and, without explicitly solving it, devise a procedure to solve it. The instrument was designed to decouple (as much as possible) respondents' problem-solving skill levels from their mastery of the underlying physics. Respondents are evaluated on their ability to devise a plan for solving the problem as opposed to solving it. We used Price *et al.*'s [36] problem-solving decision framework (outlined in Table II) to code respondents' devised plans for solving the problem presented in the diagnostic instrument. We administered this survey to experts in physics and physics students (novices). We tested two different populations of students from two different introductory physics courses at Harvard University—each with a different pedagogical approach to problem solving. One course explicitly taught a problem-solving framework, and the other course did not.

A. Problem-solving diagnostic instrument

The problem-solving diagnostic instrument was designed to elicit the problem-solving process a respondent would take to solve a physics problem without having them solve it. The problem presented is in the context of mechanics, and the stem is shown in Fig. 1.

The context of the problem is that you are working for a shipbuilding company and have been assigned to determine whether you can attach a monitor to the wall of the ship using a set of four magnets. The diagnostic consists of 14 questions and, like the PISA (2012) [39] problem-solving assessment, uses a variety of response formats (constructed-response, multiple-choice, many-choice, and drag-and-drop). The diagnostic is designed to start with broad (unscaffolded) questions about the process the respondent would use to solve the problem. The first question on the diagnostic is:

> Imagine you are working for a physics consulting firm, and you are recommending the most secure way to hang the monitor on the wall of the ship (assuming you only have 4 magnets). Write a detailed plan of your recommendation. Include as many details as you think are necessary for someone else to be able to implement it. Your response should be 1–2 paragraphs long.

After committing to a detailed plan in the first question, respondents are then guided through 13 more questions probing key features of the problem and potential solution

TABLE II. Problem-solving decision framework presented in Price *et al.* [36].

Category name	Decision	Description
A: Selection and goals of the problem	1. What is important in the field?	Identifying key issues or topics of significance in the field
	2. Opportunity fits solver's expertise?	Evaluating if the problem aligns with the researcher's knowledge and skills
	3. Goals, criteria, and constraints?	Defining objectives, criteria for success, and limitations
B: Important features and information	4. Important features and info	Determining available information relevant to problem solving
	5. What predictive framework?	Deciding among possible predictive frameworks
	6. How to narrow down the problem?	Often involves formulating specific questions and hypotheses
	7. Related problems?	Identifying similar problems and useful solutions from other contexts
	8. Potential solutions?	What are potential solutions? (This is based on experience and fitting some criteria for a solution they have for a problem having general key features identified)
	9. Is the problem solvable?	Deciding if the problem is solvable and if the solution is worth pursuing, given the difficulties, constraints, and uncertainties
C: Predictions and frameworks	10. Approximations and simplifications to make?	Simplifying the problem to make it easier to solve
	11. How to decompose into subproblems?	Breaking the problem up into more tractable subproblems
	12. Most difficult or uncertain areas?	Identifying areas of particular difficulty and/or uncertainty
	13. What information is needed?	Identifying the information needed to solve the problem
	14. Priorities	Prioritizing the most important considerations
	15. Specific plan for getting information	Identifying general requirements for the problem-solving approach they will pursue
D: Narrowing down the problem	16. Which calculations and data analysis?	Identifying calculations and data analysis needed
	17. How to represent and organize information?	Identifying the best way to represent and organize available information to provide clarity and insight
	18. How believable is the information?	Is the information valid, reliable, and believable?
	19. How does information compare to predictions?	Comparing new information to expected results
	20. Any significant anomalies?	Figuring out how to follow up on anomalous results
	21. Appropriate conclusions?	Identifying appropriate conclusions based on the data
	22. What is the best solution?	Deciding on the best solution(s)
E: Related problems	23. Assumptions and simplifications appropriate?	Deciding if previous simplifications and predictive frameworks are still appropriate
	24. Additional knowledge needed?	Deciding if more information is needed
	25. How well is the problem-solving approach working?	Deciding if the problem-solving approach needs to be modified
	26. How good is the solution?	Deciding if the chosen solution is adequate
F: Implications and communication	27. Broader implications?	Considering the wider impact and significance of the findings
	28. Audience for communication?	Identifying who needs to be informed about the findings
	29. Best way to present work?	Deciding how to effectively share the research outcomes
G: Ongoing skill and knowledge development	30. Stay up-to-date in field	Continuously learning about the latest developments in the field
	31. Intuition and experience	Using accumulated intuition and experience in problem solving
	32. Interpersonal and teamwork	Developing skills necessary for effective collaboration
	33. Efficiency	Enhancing productivity and the ability to make timely decisions
	34. Attitude	Maintaining a positive and open-minded approach to challenges

paths; these subsequent questions are more targeted (e.g., what assumptions would you make? Draw a diagram). The instrument starts with the broad, unscaffolded devise a plan question so that it does not influence respondents' process-oriented thinking with the more scaffolded questions.

This progression from the open-ended to the scaffolded questions is an intentional structure designed to gather as much information as possible from respondents with a wide range of problem-solving skills. Increasing the scaffolding as the respondent progresses from one section to the next allows us to gather information about

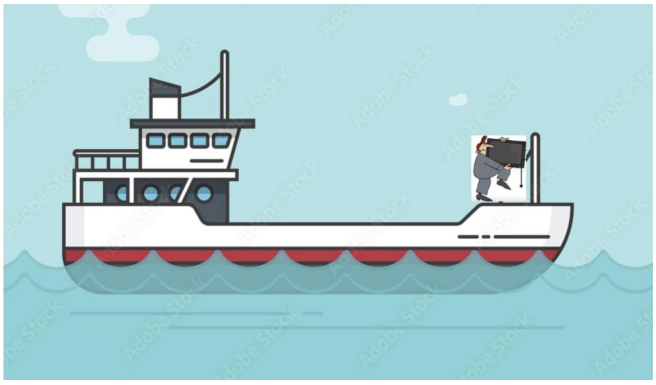


FIG. 1. Problem stem for the problem-solving diagnostic instrument. Cargo ship illustration. Image © Adobe Stock/vladwel (Asset ID: 167186621), used under standard license.

respondents' problem-solving processes even if they are not able to devise a cohesive plan in the first section.

The instrument is administered through the Harvard Qualtrics interface. Qualtrics is a web-based survey software for creating surveys. It also collects and stores data from survey respondents. The full survey as it appears in Qualtrics is included in Appendix A in Supplemental Material [41]. Upon logging on to each survey on Qualtrics, respondents are told that they will be presented with an introductory physics problem with questions designed to probe their thinking about how to solve the problem. Respondents are also told that the questions will be presented in sections, and they will not be able to go back to previous sections once they advance to the next one. They are also instructed to only use the information that is provided in the problem and to not use any outside resources. Finally, they are told that they should complete the diagnostic in one sitting and that while there is no time limit, it should take around 45 min to complete.

B. Application of problem-solving decision framework

To align our coding scheme with the construct measured by the diagnostic instrument, we restricted our analysis to decisions corresponding to the *planning phase* of problem solving. Specifically, we coded for decisions 3–15 from the Price *et al.*'s framework (categories A–C), which involve defining goals, identifying relevant features, selecting predictive frameworks, and outlining an approach to solution. We did not code for decisions in categories D and E, as these primarily reflect processes that occur during or after execution of a solution strategy. Category D includes decisions related to carrying out the plan (e.g., selecting calculations, organizing representations, and analyzing results), while category E involves evaluative and iterative decisions (e.g., assessing the validity of assumptions or determining whether additional information is needed). These decisions typically require intermediate results, data, or feedback from attempted solution steps. In contrast, the diagnostic instrument, particularly the initial devise a plan prompt, was designed to elicit respondents' ability to formulate a solution strategy *prior to execution*. As such, it does not provide the context necessary to reliably observe decisions that depend on performing calculations, analyzing results, or revising assumptions. For example, while decision 10 (approximations and simplifications) reflects an initial modeling choice, decision 23 (whether those assumptions remain appropriate) requires subsequent evaluation of results and was therefore outside the scope of this instrument. This alignment ensures that the coding framework reflects the specific phase of problem solving being measured and avoids conflating planning behaviors with execution- and evaluation-dependent processes. This phase-specific alignment between instrument design and coding framework is consistent with prior work that distinguishes planning, execution, and evaluation as separable components of problem solving (e.g., Pólya, Heller, and Mazur).

C. Instrument validation

We validated the instrument using the methodology described in "Development and Validation of Instruments to Measure Learning of Expert-Like Thinking" (Adams and Wieman) [42], which stresses the use of student interviews as the most important part of developing and validating formative assessments designed to measure the expertness of students' thinking. The steps taken to develop and validate the problem-solving diagnostic instrument are outlined below. The process was iterative and involved student and expert interviews at several stages.

(1) Consulted several experts to solicit advice and feedback on the initial problem stem. These experts were all college instructors of introductory physics who had been teaching for at least 10 years. Several candidate problems were considered and iterated on. The following selection criteria were considered:
   - Authenticity—problems where the solution method is not immediately obvious to even expert problem solvers
   - Complexity—problems complex enough that, to solve them, the solver must state their assumptions and have a detailed method
   - Open-ended—problems with multiple solution paths that depend on initial assumptions

(2) Conducted interviews of both students and experts. With the problem stem developed, we conducted six interviews of introductory physics students at Harvard University and four introductory physics instructors. We showed them the problem and asked them to talk through the ways in which they would solve it. These 10 interviews allowed us to identify

student thinking about the problem and the various ways this was different from expert thinking. We revised the problem stem for clarity. At this stage, the two revisions we made to the problem stem based on feedback were specifying the number of magnets (the first iteration did not give a specific number of magnets) and including the image of the boat, the man, and the monitor. Students indicated that having a picture of a boat and waves would help signal that the rocking of the boat should be taken into consideration.

(3) Created open-ended survey questions to probe student thinking more broadly. This first version of the survey was administered to 18 introductory physics students at Harvard. We subsequently interviewed a subset of six of these students (chosen based on the diversity of their responses to the first devise a plan question).

(4) Iterated the instrument based on students' responses and interview findings. We did one round of iterations (described below).

In the first version of the survey, we had one open-ended question on the survey designed to probe respondents' problem-solving process. This question was as follows:

What would you do to determine if this method of hanging the monitor would work? Write a detailed plan.

We were hoping that, in their response to this question, students would include their assumptions and a list of things that they would need to know if their method for hanging the monitor would work.

After interviewing the six students, we changed this one open-ended question to three more targeted questions. We added Qs 2 and 3 because some of the students we interviewed said that they did not think to include assumptions or known quantities. We analyzed the responses to Q1 separately because we believe that knowing to include assumptions in a detailed plan is, in itself, an indicator of more expert problem solving.

Q1 Write a detailed plan of your recommendation.
Q2 What information would you need to determine if your method of hanging the monitor in this way would work?
Q3 What assumptions or approximations would you need to make to determine if your method of hanging the monitor in this way would work?

(5) Conducted validation interviews on the new version of the survey with five (new) novices and four experts (instructors of introductory physics but different from those interviewed in step 2). Based on the feedback at this stage, we modified Q1 to the following:

Q1 Write a detailed plan of your recommendation. Include as many details as you think are necessary for someone else to be able to implement it. Your response should be 1–2 paragraphs long.

These changes were based on the feedback from students that they had no idea how much to write and how much detail to include. In this iteration of the question, we wanted it to be clear that students should include all the details that they thought were important, and the expectation was that their responses would be quite long.

(6) Administered the instrument to students and experts on a wider scale.

D. Coding the devise a plan responses

After the final version of the instrument was developed, we conducted an experiment to determine if physics students' responses to the first survey item (devise a plan) differed depending on whether they were explicitly taught a problem-solving approach in their introductory physics class. We administered the survey to two large introductory classes at Harvard: the deliberate practice course (DP; $N = 72$), which explicitly taught students a problem-solving process, and the repeated practice course (RP; $N = 116$), which did not, as well as to 13 experts. The experts were all college physics instructors who had been teaching introductory physics for at least 10 years. Three content experts studied each of the responses to the first survey item and coded it based on Price *et al.*'s [36] list of decisions made by experts when solving problems in their fields (outlined in Table II). Each of the three coders made a list of decisions from Table II reflected in each of the devise a plan responses. The decision codes were compared across the three coders, and all cases where there was not perfect agreement were discussed and resolved.

Figures 2(a) and 2(b) show two examples of novice responses to item 1 on the survey and the corresponding decision codes.

Figure 2 shows representative responses to the devise a plan prompt from two students [Figs. 2(a) and 2(b)] and one expert [Fig. 2(c)], along with their corresponding decision codes. We identified portions of each response that corresponded to one or more decisions from Price *et al.*'s framework and inserted the relevant decision number in parentheses immediately after each coded portion. The response in Fig. 2(a) represents a common novice pattern in which the respondent quickly proposes a specific solution. The student suggests using a "four-point suspension system" and provides a qualitative justification based on load distribution. This response includes several decision codes, including proposing a potential solution (decision 8), referencing a related problem (decision 7), identifying important information from the problem (decision 4), invoking a predictive framework (decision 5),

Assuming that the monitor is not too heavy and the magnets are strong enough, (10) the most secure way to hang the monitor on the wall of the ship would be to use all four magnets in a configuration known as a "four-point suspension system." (8) In a four-point suspension system, two magnets are attached to the top corners of the monitor, and two more magnets are attached to the corresponding points on the wall. (7) The magnets should be arranged so that the north pole of one magnet faces the south pole of the other magnet. (4) By using this configuration, the weight of the monitor will be evenly distributed across all four magnets, (5) ensuring that the monitor stays securely attached to the wall of the ship. Additionally, the four-point suspension system provides redundancy, meaning that if one magnet fails, the other three will still be able to support the weight of the monitor. (3)	Approximations & simplifications? (10) Potential solutions? (8) Related problems? (7) Important features & info? (4) Predictive framework? (5) Goals, criteria, constraints? (3)

I would want to find the strength of the magnets, especially through the distance of the walls to determine if they are strong enough to hold the weight of the monitor. (13) Additionally, there would need to be some kind of stress test to see how the magnets would hold if the ship encounters rough waters. (15) If the magnetic field is strong enough to hold the monitor in place past the limits of what is expected to be encountered, then it would seem like a viable option to hang the monitor. (3) However, we would also need to take into consideration our ability to remove the magnets if we want to move the monitor, which is the entire reason we are looking at magnets in the first place, as well as any potential damage that could come to the monitor from being in close proximity to the magnet. (12)	What info needed? (13) Specific plan for getting info? (15) Goals, criteria, constraints? (3) Most difficult or uncertain areas? (12)

The first thing I would do is make sure that the magnets do not impact the functioning of the television by bringing them close to the television. (14) In reality, I might just simply run the experiment of trying it, and seeing if it holds, (15) but you may be wanting to know more about how I would conceptually solve this, right? So, I would first think about the forces, and Torques involved on the television if you had it being held up by the four corners by whatever means. (4) That is, I would imagine the motion of the boat and other bumps, the television might encounter, say, from people. (12) Then I would use statics (8) to calculate what the perpendicular and parallel forces on whatever may be holding the corners of the TV. (13) Once I do that, I would look into whether or not, they are magnets that would not move if those forces were put on them. (5) This would include perpendicular, magnetic forces, as well as parallel friction forces. (3)	What to prioritize among many competing considerations. (14) Specific plan for getting info? (15) Important features & info? (4) Most difficult or uncertain areas?? (12) What info needed? (13) Potential solutions? (8) Predictive framework? (5) Goals, criteria, constraints? (3)

FIG. 2. (a)–(c) Examples of (a) and (b) novice and (c) expert responses to the devise a plan question on the problem-solving diagnostic instrument. Numbers in parentheses within each response identify the corresponding decision codes from Table II. Response (a) is from an RP student, response (b) is from a DP student, and response (c) is from an expert.

and identifying specific goals, criteria, or constraints within the problem (decision 3). The response also includes an assumption about magnet strength and monitor weight (decision 10). These statements are organized around a single proposed configuration and are not supported by an explicit plan for testing the approach or addressing uncertainty in the problem. In contrast, the response shown in Fig. 2(b) reflects a different and more expertlike planning

pattern. Rather than committing immediately to a specific solution, the respondent focuses on identifying what must be known to determine whether the proposed approach would work. The response explicitly identifies information needed, such as magnet strength (decision 13), and proposes a method for obtaining that information through testing (decision 15). This response also references criteria and constraints that must be satisfied, such as holding the monitor securely under expected conditions while maintaining the ability to reposition it (decision 3). In addition, the respondent highlights key uncertainties, such as performance under rough water and possible damage to the monitor (decision 12). This emphasis on identifying unknowns, evaluating feasibility, and proposing ways to obtain relevant information aligns more closely with expert problem-solving behavior, in which considerable effort is devoted to defining what must be determined before selecting a solution. While the response in Fig. 2(b) does not articulate a predictive framework or propose a concrete solution, it reflects a more strategic and evaluative approach to planning than the solution-driven response in Fig. 2(a). The expert response in Fig. 2(c) represents a more cohesive, complete, and "physics"-based plan than both the novice responses. Like the response in Fig. 2(b), it identifies important uncertainties, specifies needed information, and proposes methods for obtaining that information (decisions 12, 13, and 15). It also explicitly frames the problem using a predictive framework grounded in physics by mentioning forces, torques, statics, magnetic forces, and friction (decisions 4 and 5) and identifies potential solution paths (decision 8) while outlining relevant criteria and constraints (decision 3). Additionally, the expert prioritizes among competing considerations (decision 14) by stating that "the first thing" to do would be to determine whether the magnets would interfere with the operation of the television.

We selected these examples because they represent common response patterns observed across the dataset. They illustrate two distinct novice approaches: one that is solution-driven [Fig. 2(a)] and one that is information- and uncertainty-driven [Fig. 2(b)]. The latter shares important features with expert problem solving, especially in its focus on evaluating how feasible the solution is and identifying what must be known before proceeding. The expert response in Fig. 2(c) shows these decisions combined with a well-articulated predictive framework and coordinated into a comprehensive problem-solving strategy. All three of these examples show a qualitative interpretation of the differences summarized quantitatively in Table III.

E. Testing novices from two classes with different pedagogical approaches to problem solving

We administered the problem-solving diagnostic survey to two groups of novices in two different calculus-based introductory physics classes at Harvard University. Both courses are comprised of students with similar physics backgrounds (nonmajors), and historically, precourse FCI scores from both courses have indicated that both populations come in with a similar baseline physics knowledge. While both courses covered the same content, one course (DP) explicitly taught students to use a four-step problem-solving procedure (Mazur's framework from Table I) for all problems encountered throughout the course. The other course (RP) did not present or model a framework for problem solving. Both courses required students to solve the same number and types of problems (during class and in the homework); however, their approach to teaching problem solving was different. At the beginning of DP, the instructor outlined each of the four steps in Mazur's framework [19] (getting Started, devise a plan, execute the plan, and evaluate the result) and

TABLE III. Frequency of decisions [from Price *et al.*'s problem-solving framework (Table II)] represented in the devise a plan responses.

Decision	Frequency of decision in DP responses ($N = 72$)	Frequency of decision in RP responses ($N = 116$)	Frequency of decision in expert responses ($N = 13$)
3. Goals, criteria, and constraints?	63	82	12
4. Important features and info	55	70	9
5. What predictive framework?	33	46	6
6. How to narrow down the problem?	0	1	1
7. Related problems?	3	2	0
8. Potential solutions?	42	97	3
9. Is the problem solvable?	8	9	2
10. Approximations and simplifications to make?	12	13	2
11. How to decompose into subproblems?	3	3	1
12. Most difficult or uncertain areas?	25	22	8
13. What information is needed?	27	27	12
14. Priorities	10	19	6
15. Specific plan for getting information	20	16	11

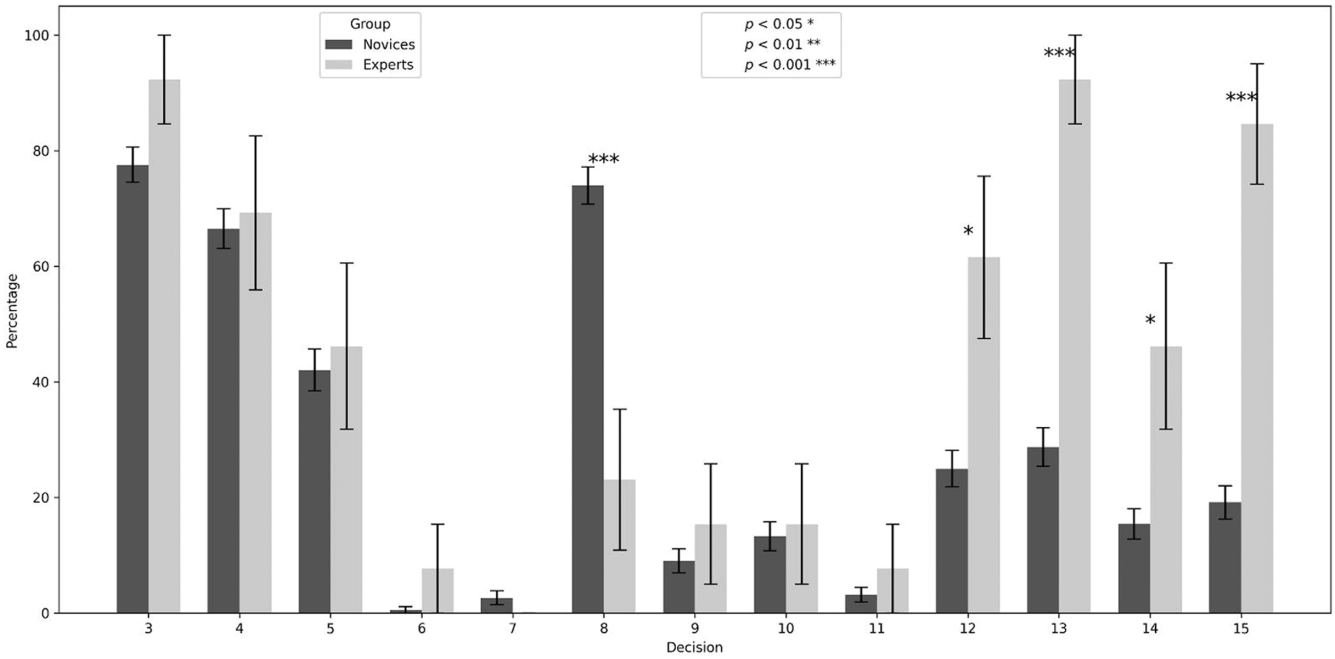


FIG. 3. Percentage of novices (both DP $N = 72$ and RP $N = 116$) [$N = 188$ (dark gray)] and experts [($N = 13$) (light gray)], which showed evidence of problem-solving decisions 3 through 15 in their devise a plan response. Error bars represent the standard error of the mean; asterisks denote the p values for a chi-squared test between the two groups.

walked through several examples of problems solved using this procedure. This framework is also modeled throughout the required textbook used in every aspect of the course throughout the course wherever students are required to solve a problem. Students in DP are required to explicitly use this framework for every problem they encounter in the course, including any submitted homework problem. DP students receive consistent, targeted feedback on their problem-solving process within the prescribed framework. To receive credit for submitted homework, DP students must have each of the four steps enumerated for all problems. DP students are assessed on following the four-step problem-solving process instead of the correctness of the final solution. In RP, on the other hand, there is no problem-solving framework presented, and students are not provided feedback on their problem-solving process. The assessment for RP is based solely on correctness and not on following a specific problem-solving procedure. From a pedagogical point of view, DP follows deliberate practice, with an emphasis on breaking the problem-solving process into subskills and providing students with targeted feedback. RP, on the other hand, uses a repeated practice (unstructured) approach, which focuses on repetition.

IV. RESULTS

Table III summarizes the frequency of the decisions found in all the responses to the devise a plan question on the problem-solving diagnostic instrument. Devising a plan only involves the first three categories of decisions from Table II (defining goals of the problem, identifying important features and information, andcreating predictions and frameworks), and therefore, only decisions 3 through 15 are represented in the responses. The frequency with which these decisions appear across each respondent category ($N = 201$) is found in Table III. Each decision was counted at most once for each respondent, even when evidence of that decision appeared in multiple sentences of the response.

A. Novices versus experts

To determine what we can learn from this problem-solving diagnostic instrument about the differences between novice and expert problem solvers, we compared the frequency with which novices made the decisions listed in Table III to that of experts. This is represented in Fig. 3, which shows the percentage of each group (novices/experts) that demonstrated problem-solving decisions 3 through 15 in their response to the devise a plan question on the problem-solving diagnostic instrument. The novices in this plot represent both groups of novices (DP $N = 72$ and RP $N = 116$). The experts plotted represent all the experts ($N = 13$). We performed a chi-squared analysis to determine which decisions showed statistically significant differences between novices and experts.

Figure 3 shows that experts consider the most difficult or uncertain areas of the problem (decision 12), statistically significantly more than novices. Experts are also

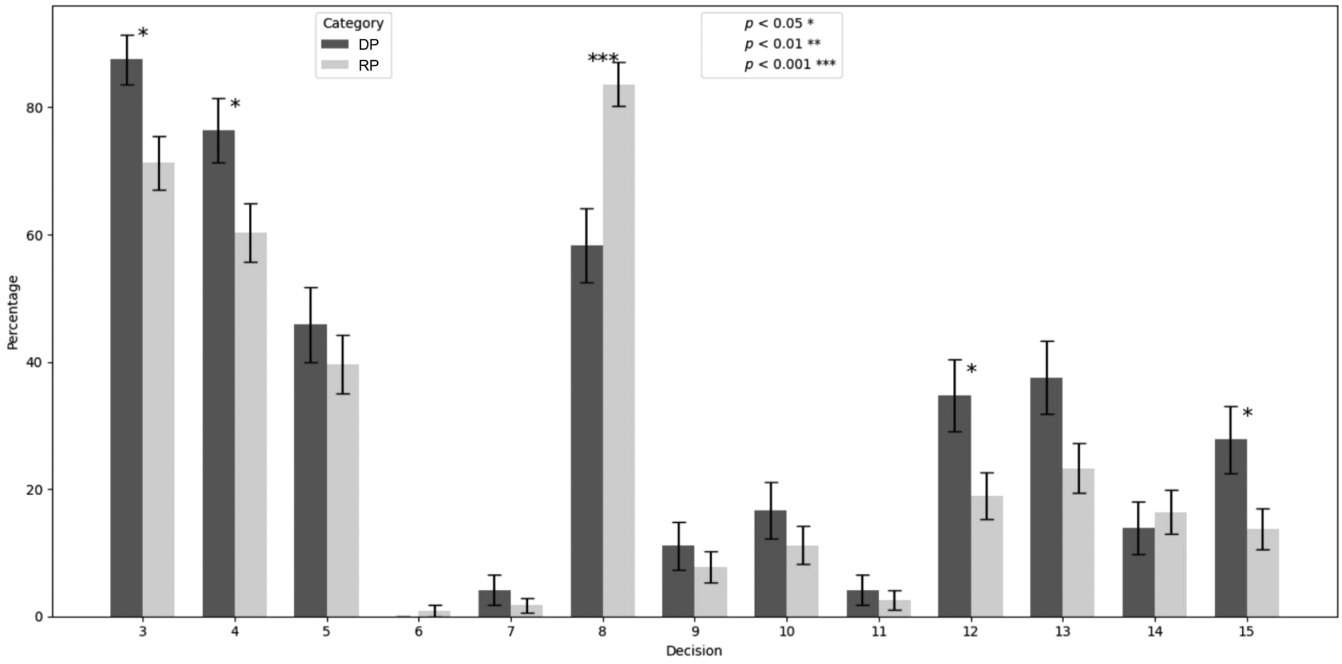


FIG. 4. Percentage of DP students (dark gray) and RP students (light gray), which showed evidence of problem-solving decisions 3 through 15 in their devise a plan response. Error bars represent the standard error of the mean; asterisks denote the p values for a chi-squared test between the two groups.

better at focusing on the information they will need to solve the problem (decision 13), the specific plan for getting information (decision 15), and how to prioritize the most important considerations (decision 14). Novices, on the other hand, are more likely than experts to focus on the potential solution (decision 8).

B. Comparing two different pedagogical approaches to problem solving

To determine if novice problem-solving skills differ between those explicitly taught a problem-solving framework and those not taught a framework, we compared the frequency with which DP students made the decisions listed in Table III to that of RP students. DP explicitly taught students how to apply a framework to solving problems, whereas RP did not. This comparison is represented in Fig. 4, which shows the percentage of each group (DP and RP) that demonstrated problem-solving decisions 3 through 15 in their response to the devise a plan question on the problem-solving diagnostic instrument. We performed a chi-squared analysis to determine which decisions showed statistically significant differences between novices and experts.

Figure 4 shows students taught a problem-solving procedure (i.e., DP students) define goals, criteria, and constraints (decision 3) determine core features or ideas related to the problem (decision 4), identify areas of particular difficulty and/or uncertainty (decision 12), and

TABLE IV. Summary of decisions with statistically significant differences between novice and expert groups and/or DP and RP groups.

Decision	Novice (DP and RP) ($N = 179$)(%)	Expert ($N = 13$) (%)	p value (chi-squared)	DP ($N = 72$) (%)	RP ($N = 116$) (%)	p value (chi-squared)
3) Goals, criteria, and constraints?	77.1	92.3	0.3	87.5	70.7	0.01
4) Important features and info?	66.5	69.2	1.0	76.4	60.3	0.03
8) Potential solution?	**73.9**	23.1	0.0003	58.3	**83.6**	0.0002
12) Most difficult or uncertain areas?	25.0	**61.5**	0.01	**34.7**	19.0	0.02
13) What information is needed?	28.7	**92.3**	0.00001	**34.5**	23.3	0.05
14) Priorities?	15.4	46.1	0.01	13.8	16.4	0.8
15) Specific plan for getting information	19.1	**84.6**	0.00001	**27.8**	13.8	0.02

outline a specific plan for getting information (decision 15) statistically significantly more than students not taught a problem-solving procedure (i.e., RP students) who are more likely to focus on a potential solution (decision 8).

Table IV provides a summary of the statistically significant results from Figs. 3 and 4. To determine which group of students was more expertlike in their decisions, we bolded the larger percentage for decisions that showed significant differences between both experts and novices and between DP and RP students. For example, in decision 8, the results from both chi-squared tests are significant. A higher percentage of novices (73.9%, compared to 23.1% of experts) referenced decision 8 in their response. Decision 8 (focusing on a potential solution) was also referenced more by RP students (compared to DP students) (83.6%, compared to 58.3%), implying that RP students are less aligned with experts on this decision. The other three decisions, which show statistically significant differences between both sets of groups (decisions 12, 13, and 15), show a higher percentage of experts and a higher percentage of DP students. From this, it appears that the decisions made by experts are also more commonly made by DP students. Overall, our data show that novices are more likely to focus on generating a potential solution (i.e., where to place the magnets on the back of the monitor), whereas more expert problem solvers focus on what information is needed to solve a problem and a specific plan for *how* to solve it. Experts are also more likely to identify the difficult or uncertain areas of a problem (for example, the fact that the monitor is being mounted on a ship and the ship will rock back and forth).

V. DISCUSSION

We find that teaching students a specific problem-solving framework, with the use of deliberate practice, does make them more expertlike in their problem-solving decisions compared to the use of repeated practice (with no problem-solving framework and lots of repetition). This study contributes to the literature that emphasizes the importance of explicit instruction in problem-solving frameworks within undergraduate physics education. As established in previous research [17–19], experts approach problems not merely by applying formulas but by strategically structuring their solution process, incorporating metacognitive reflection, and adopting coherent frameworks. Our results align with these findings and extend them by demonstrating quantitatively that students who were taught to apply a structured problem-solving framework exhibited decision patterns more closely resembling those of experts than students who were not given explicit framework instruction.

This distinction underscores the pedagogical value of deliberate practice as articulated in the literature [20]. Deliberate practice—characterized by breaking down complex skills into subskills, providing focused and ongoing feedback, and encouraging incremental mastery—was central to the instructional approach given to the DP cohort. The alignment of their problem-solving decisions with those of experts highlights the importance of structured, feedback-rich practice in cultivating higher-order problem-solving skills. In contrast, repeated practice, which emphasizes rote repetition with no targeted feedback, appears less effective for achieving expertlike problem-solving skills.

Our findings are consistent with prior work highlighting the importance of metacognitive processes in expert problem solving, particularly those related to evaluating the feasibility and reasonableness of solutions [13]. While our instrument does not directly assess students' evaluation of final answers, it does capture earlier planning behaviors such as identifying constraints, uncertainties, and information needs that may support more effective monitoring and evaluation during later stages of problem solving. The disparity between "experts" and students who did not engage with a structured problem-solving framework (RP) suggests that content knowledge alone is inadequate for expertise; without scaffolding, students tend to rely on surface-level cues rather than engaging deeply with underlying physical principles. Our study provides empirical support for the targeted use of problem-solving frameworks as effective pedagogical tools. Not only do these frameworks scaffold novice performance, but—as indicated by the similarities between DP and experts—they also help undergraduate students develop expertlike problem solving. This suggests that integrating deliberate practice into physics instruction could help bridge the gap between novice and expert problem-solving skills.

While the findings provide strong evidence for the effectiveness of explicit problem-solving frameworks and deliberate practice pedagogies, several limitations should be acknowledged. First, the study was conducted within the context of introductory physics courses at a single institution, and therefore, the generalizability of the results to other academic settings, disciplines, or educational levels is limited.

Second, the analysis relies on responses to a diagnostic instrument and self-reported strategies, which are subject to interpretation biases and might not fully capture the complex cognitive processes underlying expert problem solving. While efforts were made to ensure consistency in coding and interpretation, some subjectivity in coding problem-solving decisions is inevitable, despite consensus procedures among expert coders.

Third, the two student comparison groups (DP and RP) were comprised of different students, albeit with similar academic backgrounds, and both were drawn from the same population of Harvard undergraduates, most commonly sophomores and juniors fulfilling premedical requirements. While previous research at our institution

(including baseline assessments such as the force concept inventory) has repeatedly shown that these populations are nearly indistinguishable in terms of prior physics knowledge and coursework, we cannot entirely rule out the influence of unmeasured differences such as motivation, study habits, or prior informal exposure to problem-solving strategies. The absence of direct measurement of initial problem-solving skills in this study further limits our ability to attribute observed differences wholly to instructional method. While we believe it is reasonable to assume both cohorts entered the courses with very limited and similar problem-solving abilities, future studies would benefit from including a preinstruction assessment of problem-solving approaches to further control for potential confounding variables.

Additionally, the duration over which the interventions were implemented was relatively short. Longitudinal studies would be necessary to determine the persistence of the observed differences and whether students continue to exhibit more expertlike behaviors over time. Additionally, while the study highlights overall trends in problem-solving decisions, it does not disentangle the relative contributions of individual instructional techniques or feedback components within the broader deliberate practice approach. Further investigation, employing both quantitative and qualitative methodologies, would help clarify which aspects of the instructional interventions are most critical for developing expertise. Despite these limitations, the findings offer valuable insights into physics instruction and provide a foundation for future research aimed at optimizing problem-solving pedagogy.

VI. CONCLUSION

In this study, we demonstrate that teaching students to apply an explicit problem-solving framework, coupled with deliberate practice pedagogy, contributes to the development of expertlike problem-solving behaviors in undergraduate physics courses. We show that students exposed to these targeted instructional strategies become more similar to experts in their decision-making processes than those engaging in unstructured or repetitive practice alone. These results reinforce prior literature advocating for the integration of metacognitive scaffolding and targeted feedback into science education, highlighting that deliberate practice—not mere repetition—drives meaningful skill acquisition. Given these findings, we strongly recommend that physics instructors adopt explicit problem-solving frameworks as a core component of their pedagogy and prioritize deliberate practice techniques. This approach appears to facilitate the transition from novice to expertlike thinking. Future research should expand on these findings by exploring the long-term impact and transferability of these instructional techniques across diverse educational contexts.

ACKNOWLEDGMENTS

Several people contributed to the work described in this paper. K. M. conceived of the basic idea for this work. K. M., O. M., and G. L. designed and carried out the experiments and analyzed the results. K. M. wrote the first draft of the manuscript; all authors subsequently took part in the revision process and approved the final copy of the manuscript. Carl Wieman, Eric Burkholder, Argenta Price, Shima Salehi, Greg Kestin, Logan McCarty, and Eric Mazur provided feedback on the research throughout its development. The authors acknowledge the use of ChatGPT, an artificial intelligence (AI) language model developed by OpenAI, to assist with language editing, formatting consistency, and manuscript revision. All substantive content, analysis, interpretation, and conclusions were generated and verified by the authors, who take full responsibility for the final manuscript.

DATA AVAILABILITY

The data that support the findings of this article are openly available [1]; embargo periods may apply.

[1] ABET, *Criteria for Accrediting Engineering Programs, 2019–2020* (Engineering Accreditation Commission, ABET, Baltimore, MD, 2019), https://www.abet.org/wp-content/uploads/2018/11/E001-19-20-EAC-Criteria-11-24-18.pdf.

[2] C. L. Dym, Design, systems, and engineering education, Int. J. Eng. Educ. **20**, 305 (2004), https://www.ijee.ie/articles/Vol20-3/IJEE2517.pdf.

[3] C. L. Dym, A. M. Agogino, O. Eris, D. D. Frey, and L. J. Leifer, Engineering design thinking, teaching and learning, J. Eng. Educ. **94**, 103 (2005).

[4] D. Jonassen, J. Strobel, and C. B. Lee, Everyday problem-solving in engineering: Lessons for engineering educators, J. Eng. Educ. **95**, 139 (2006).

[5] AIP, Data on Employment for PhD Recipients (accessed 2024).

[6] P. R. L. Heron and D. E. Meltzer, The future of physics education research: Intellectual challenges and practical concerns, Am. J. Phys. **73**, 390 (2005).

[7] C. Hoellwarth, M. J. Moelter, and R. D. Knight, A direct comparison of conceptual learning and problem-solving

ability in traditional and studio style classrooms, Am. J. Phys. **73**, 459 (2005).
[8] J. L. Docktor and J. P. Mestre, Synthesis of discipline-based education research in physics, Phys. Rev. ST Phys. Educ. Res. **10**, 020119 (2014).
[9] J. L. Docktor, J. Dornfeld, E. Frodermann, K. Heller, L. Hsu, K. A. Jackson, A. Mason, Q. X. Ryan, and J. Yang, Assessing student written problem solutions: A problem-solving rubric with application to introductory physics, Phys. Rev. Phys. Educ. Res. **12**, 010130 (2016).
[10] E. W. Burkholder, J. K. Miles, T. J. Layden, K. D. Wang, A. V. Fritz, and C. E. Wieman, Template for teaching and assessment of problem solving in introductory physics, Phys. Rev. Phys. Educ. Res. **16**, 010123 (2020).
[11] A. Reinhard, A. Felleson, P. C. Turner, and M. Green, Assessing the impact of metacognitive postreflection exercises on students' problem-solving skillfulness, Phys. Rev. Phys. Educ. Res. **18**, 010109 (2022).
[12] V. Gjerde, V. Havre Paulsen, B. Holst, and S. Kolstø, Problem solving in basic physics: Effective self-explanations based on four elements with support from retrieval practice, Phys. Rev. Phys. Educ. Res. **18**, 010136 (2022).
[13] K. Heller and P. Heller, Cooperative Problem Solving in Physics: A User's Manual (2010), https://www.aapt.org/Conferences/newfaculty/upload/Coop-Problem-Solving-Guide.pdf.
[14] W. K. Adams and C. E. Wieman, Analyzing the many skills involved in solving complex physics problems, Am. J. Phys. **83**, 459 (2015).
[15] E. Ince, An overview of problem-solving studies in physics education, J. Educ. Learn. **7**, 191 (2018).
[16] J. Krajcik, Three-dimensional instruction: Using a new type of teaching in the science classroom, Sci. Scope **39**, 16 (2015).
[17] G. Pólya, *How to Solve It: A New Aspect of Mathematical Method* (Princeton University Press, Princeton, NJ, 1945), https://press.princeton.edu/books/paperback/9780691164076/how-to-solve-it.
[18] J. I. Heller and F. Reif, Prescribing effective human problem-solving processes: Problem description in physics, Cognition Instr. **1**, 177 (1984).
[19] E. Mazur, *Principles & Practice of Physics* (Pearson, 2014), https://www.pearson.com/en-us/subject-catalog/p/principles--practice-of-physics/P200000006983/9780135610862?utm_source=google&utm_source=googleutm_medium=cpcutm_medium=paid_searchutm_campaign=pla_shopping_allutm_campaign=pla_shopping_allutm_term=utm_ad=utm_ad_campaign_id=22839090472adgroup=186046697914cq_cmp=22839090472gad_source=1gad_campaignid=22839090472gbraid=0AAAADnfL3uZrMMGkls_GGdMTsEbl9RH9gclid=CjwKCAjw4dDTBhAqEiwAkHYmSiApbAdXww7yv1r0wecTTYY7RIa2h4SJR1h6ZD14gRSYPq0jt6gSpxoC8FYQAvD_BwE.
[20] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, The role of deliberate practice in the acquisition of expert performance, Psychol. Rev. **100**, 363 (1993).
[21] C. H. Marcus, L. R. Newman, A. S. Winn, K. Antanovich, Z. Audi, A. Cohen, A. W. Hirsch, H. K. Harris, K. A. Miller, and C. D. Michelson, TEACH and repeat: Deliberate practice for teaching, Clin. Teach. **17**, 688 (2020).
[22] P. Ward, Y.-J. Chen, K. Higginson, and X. Xie, Teaching rehearsals and repeated teaching: Practice-based physical education teacher education pedagogies, J. Phys. Educ. Recreat. Dance **89**, 20 (2018).
[23] A. H. Schoenfeld, Toward a theory of teaching-in-context, Issues Educ. **4**, 1 (1998).
[24] P. B. Kohl and N. D. Finkelstein, Patterns of multiple representation use by experts and novices during physics problem solving, Phys. Rev. ST Phys. Educ. Res. **4**, 010111 (2008).
[25] P. T. Hardiman, R. Dufresne, and J. P. Mestre, The relation between problem categorization and problem solving among experts and novices, Mem. Cognit. **17**, 627 (1989).
[26] T. de Jong and M. G. Ferguson-Hessler, Cognitive structures of good and poor novice problem solvers in physics, J. Educ. Psychol. **78**, 279 (1986).
[27] J. Larkin, J. McDermott, D. P. Simon, and H. A. Simon, Expert and novice performance in solving physics problems, Science **208**, 1335 (1980).
[28] E. Mazur, Peer instruction: Getting students to think in class, AIP Conf. Proc. **399**, 981 (1997).
[29] W. J. Leonard, R. J. Dufresne, and J. P. Mestre, Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems, Am. J. Phys. **64**, 1495 (1996).
[30] E. Mazur, Qualitative versus quantitative thinking: Are we teaching the right thing, Opt. Photon. News **3**, 38 (1992).
[31] E. Kim and S.-J. Pak, Students do not overcome conceptual difficulties after solving 1000 traditional problems, Am. J. Phys. **70**, 759 (2001).
[32] E. F. Redish, Changing student ways of knowing: What should our students learn in a physics class, in *Proc. World View Phys. Educ* (2005), Vol. 1, https://www.physics.umd.edu/perg/papers/redish/IndiaPlen.pdf.
[33] F. Reif and J. I. Heller, Knowledge structure and problem solving in physics, Educ. Psychol. **17**, 102 (1982).
[34] A. Van Heuvelen, Learning to think like a physicist: A review of research-based instructional strategies, Am. J. Phys. **59**, 891 (1991).
[35] M. S. Sabella and E. F. Redish, Knowledge organization and activation in physics problem solving, Am. J. Phys. **75**, 1017 (2007).
[36] A. M. Price, C. J. Kim, E. W. Burkholder, A. V. Fritz, C. E. Wieman, and J. Momsen, A detailed characterization of the expert problem-solving process in science and engineering: Guidance for teaching and assessment, CBE–Life Sci. Educ. **20**, 36 (2021).
[37] M. Alii, C. A. Talib, N. H. Ibrahim, J. Surif, and A. H. Abdullah, The importance of monitoring skills in physics problem solving, Eur. J. Educ. Stud. **1**, 1 (2016), https://oapub.org/edu/index.php/ejes/article/view/54.
[38] K. A. Ericsson, R. R. Hoffman, A. Kozbelt, and A. M. Williams, *The Cambridge Handbook of Expertise and Expert Performance*, edited by K. A. Ericsson, R. R. Hoffman, A. Kozbelt, and A. M. Williams, 2nd ed. (Cambridge University Press, Cambridge, 2018), p. 165,

[39] D. Ramalingam, R. Philpot, and B. McCrae, The PISA 2012 assessment of problem solving (2017).

[40] A. Halim, Y. Yusrizal, S. Susanna, and T. Tarmizi, An analysis of students' skill in applying the problem-solving strategy to the physics problem settlement in facing AEC as global competition, J. Pendidik. IPA Indones. **5**, 1 (2016), https://scispace.com/pdf/an-analysis-of-students-skill-in-applying-the-problem-4bv1s0d2j0.pdf.

[41] See Supplemental Material at http://link.aps.org/supplemental/10.1103/7pqt-gd9c for Appendix A, which contains the full problem-solving diagnostic instrument as it appears in Qualtrics.

[42] W. K. Adams and C. E. Wieman, Development and validation of instruments to measure learning of expert-like thinking, Int. J. Sci. Educ. **33**, 1289 (2011).