

Contrasting grading approaches in introductory physics and quantum mechanics: The case of graduate teaching assistants

Emily Marshman,¹ Ryan Sayer,² Charles Henderson,³ and Chandralekha Singh¹

¹*Department of Physics and Astronomy, University of Pittsburgh,
3941 O'Hara Street, Pittsburgh, Pennsylvania 15260, USA*

²*Department of Physics, Bemidji State University, Bemidji, Minnesota 56601, USA*

³*Department of Physics, Western Michigan University,
1903 W. Michigan Avenue, Kalamazoo, Michigan 49008, USA*

(Received 24 October 2016; published 1 May 2017)

At large research universities, physics graduate teaching assistants (TAs) are often responsible for grading in courses at all levels. However, few studies have focused on TAs' grading practices in introductory and advanced physics courses. This study was designed to investigate whether physics graduate TAs grade students in introductory physics and quantum mechanics using different criteria and if so, why they may be inclined to do so. To investigate possible discrepancies in TAs' grading approaches in courses at different levels, we implemented a sequence of instructional activities in a TA professional development course that asked TAs to grade student solutions of introductory physics and upper-level quantum mechanics problems and explain why, if at all, their grading approaches were different or similar in the two contexts. We analyzed the differences in TAs' grading approaches in the two contexts and discuss the reasons they provided for the differences in their grading approaches in introductory physics and quantum mechanics in individual interviews, class discussions, and written responses. We find that a majority of the TAs graded solutions to quantum mechanics problems differently than solutions to introductory physics problems. In quantum mechanics, the TAs focused more on physics concepts and reasoning and penalized students for not showing evidence of understanding. The findings of the study have implications for TA professional development programs, e.g., the importance of helping TAs think about the difficulty of a problem from an introductory students' perspective and reflecting on the benefits of formative assessment.

DOI: [10.1103/PhysRevPhysEducRes.13.010120](https://doi.org/10.1103/PhysRevPhysEducRes.13.010120)

I. INTRODUCTION

At large research institutions in the U.S., graduate students often play an important role in the education of undergraduate students in physics courses at all levels. It is quite common for physics graduate teaching assistants (TAs) to teach introductory physics recitations or labs and grade student work in introductory and advanced courses. However, most TAs receive very little training or guidance about grading [1] and there have been few studies focusing on TAs' grading practices in introductory and advanced courses [2].

Therefore, we investigated TAs' grading practices when grading solutions to introductory physics and quantum mechanics problems. This research study investigates the types of solution features that physics graduate TAs value when preparing introductory physics and quantum mechanics (QM) problem solutions for their students. Furthermore,

we examined the types of solution features that TAs grade on when grading solutions to introductory physics and quantum mechanics problems, whether they use different criteria when grading in the two contexts, and the reasons for the differences in grading in the two contexts (if there are differences). TAs were given an opportunity to compare their grading practices in the two contexts, reflect on their grading goals and practices, and resolve possible conflicts between their goals and practices. The findings of the study can inform professional development leaders interested in helping TAs improve their grading practices.

This study involved 15 graduate TAs participating in a semester-long professional development program at a research university in the U.S. The data collection tool was designed to probe implicit and potentially conflicting perceptions regarding the goals of grading and grading practices. TAs were given an introductory physics problem and a QM problem and were asked to create solutions to the problems that they believed would help their students learn. Then, TAs were given a set of introductory student solutions that were used in prior studies to investigate faculty grading practices [3] as well as a set of QM student solutions that have solution features similar to the introductory student solutions. Some of the solutions are

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.

elaborate and show evidence of understanding while others are brief and do not provide evidence of understanding. All of the steps in the brief solutions to each problem were included in the elaborated solutions. The TAs were asked to grade the student solutions for the introductory physics and QM problems. They were also asked to explain whether they used different criteria when grading student solutions in the two different contexts, and if so, why.

In particular, the study was designed to investigate the following research questions:

- (1) What features do TAs include in their own problem solutions when creating solutions for students in introductory physics and QM?
- (2) Do TAs grade students' solutions to an upper-level QM problem differently than students' solutions to an introductory physics problem?
- (3) What are the TAs' stated reasons for whether (or not) their grading is different for an introductory problem versus a QM problem?

We begin with a literature review before discussing the methodology. Then, we present the findings and follow up with a summary and discussion.

II. BACKGROUND

At the Graduate Education in Physics Conference jointly sponsored by the American Physical Society and the American Association of Physics Teachers, discussions with faculty about teaching assistantships suggest that the majority of physics departments at research institutions in the U.S. employ physics graduate students as TAs for introductory physics courses and for grading in courses at all levels [1]. The TAs are generally expected to do most of the grading, including grading exams in introductory courses and homework and quizzes in both introductory and upper-level courses. Many of the physics departments provide very brief training to the TAs (half day or less) to help them learn how to carry out their teaching responsibilities [1]. However, a handful of departments have provided a semester-long TA professional development program similar to the associated with the present study. Other than the training provided by the department, most conference participants noted that the TAs usually carry out the tasks without significant guidance from their supervising instructor except for a general discussion about how to carry out recitations or how to grade [1].

TAs receive limited training to carry out their teaching duties, and their grading beliefs and practices [4–8] are often based upon their own experiences as students, the expectations of their supervising instructor, and their workload. Research suggests that TAs perceive the difficulty of a problem they are grading from their own perspective instead of the perspective of their students and they struggle to understand the value of thinking about the difficulty of a problem from an introductory student's perspective [6]. Also, graduate TAs often do not

notice features in example solutions that are intended to help students develop effective problem-solving approaches [7], e.g., performing an initial problem analysis (clarifying assumptions, identifying relevant information, and representing the problem in a diagram), planning the solution (identifying intermediate goals and breaking the problem into sub-problems), explicating the solution (explaining and justifying the relevant principles used to solve the problem), and evaluating progress (assessing their understanding and checking their final answer) [9–20]. Moreover, prior research suggests that they do not always engage in grading practices that are intended to help introductory physics students learn desired problem-solving approaches and develop a coherent understanding of physics [2]. Furthermore, some physics faculty members do not necessarily require their students to show evidence of understanding via explication of the problem-solving approach (i.e., they are often hesitant to take off points if the final answer is correct but the problem-solving process is not shown) [8].

The factors mentioned in the preceding paragraphs can impact TAs' beliefs about grading in different courses. Thus, we investigated TAs' beliefs about grading in introductory physics and QM to uncover possible discrepancies in the two contexts. The findings of the study can be useful for professional development of TAs and can be used to help TAs reflect on their grading practices in introductory physics and QM.

III. METHODOLOGY AND DATA COLLECTION

A. Description of TA professional development course

In this investigation, we collected grading data from a mandatory, semester-long TA professional development course led by one of the authors. The course met for 2 hours each week for the entire semester and was meant to prepare the TAs for their teaching responsibilities. The TAs in general were asked to do 1 hour of homework each week pertaining to the professional development course, e.g., related to grading, that was graded for completeness. During class meetings, TAs generally discussed their homework assignment from the previous week in small groups. At the end of the class, they shared what they had discussed in groups while the instructor gave input. The TAs had also attended a one-day new teaching assistant workshop facilitated by the university, but this workshop was general and did not focus on discipline-specific issues in teaching and learning physics. There were 15 first-year graduate students enrolled in the course. The majority of the first-year graduate students were TAs. Most of the TAs were teaching recitations for introductory physics courses for the first time. A few other TAs were also assigned to facilitate a laboratory section or grade students' work in various physics courses for the first time. In the same semester, a majority of the TAs were also tutors in a physics

resource room where introductory students can receive help on assignments such as homework and laboratory reports. The participants consisted of a mix of domestic and international students from nations such as China, India, Turkey, etc. There were 4 female TAs and 11 male TAs. The demographics of the TAs in this course are somewhat similar to national norms among physics graduate students [21].

B. Data collection

1. Data collection tools and artifacts

The data on TAs' beliefs about grading and grading practices were collected using a group administered interactive questionnaire (GAIQ) previously developed and validated by three of the authors in collaboration with two graduate student researchers in physics education for use with TAs or instructors [3]. The GAIQ consists of a series of activities involving worksheets which are designed to clarify a TA or instructor's ideas about helping students learn physics content and effective problem-solving approaches, e.g., performing an initial problem analysis, planning the solution, explicating the solution, and evaluating progress [9–20]. In particular, the grading activities in the GAIQ were based upon the grading framework of Heller, Yerushalmi, Henderson, and Docktor [3,8,9], which posits that grading practices can encourage students to use an effective problem-solving approach. The grading activities in the GAIQ were designed to encourage graders to compare and contrast various types of problem-solving approaches (that may include, e.g., effective problem-solving approaches [9–20]).

The GAIQ involves specific artifacts including an introductory physics problem (see Fig. 1) and two student solutions (see Fig. 2). The introductory physics problem was designed, validated, and approved by four physics instructors who taught introductory physics courses at the University of Minnesota and were used on final exams. The problem involves synthesis of important physics concepts and principles, is context rich, and is difficult enough to require an average student to choose to use an exploratory decision making process as opposed to an algorithmic

procedure [9–20]. There were two student solutions to each problem. The student solutions were designed to mimic actual students' common answers in the final exam and reflect differences between expert and novice problem solving from the research literature such as including a diagram describing the problem, explication of subproblems, justification of steps, evaluation of the final answer, explication of the principles used, evidence of reflective practices, etc. [9–20]. Figure 2 shows an introductory student solution that is elaborated (ISS-E) and an introductory student solution that is brief (ISS-B). ISS-E includes a diagram, articulation of the principles used to find intermediate variables, and clear justification for the final result. However, while there are different ways to interpret the nature of the mistake in the elaborated solution ISS-E, one interpretation is that there are two canceling mistakes that lead to the correct final answer. ISS-B is brief with no explication of reasoning and does not give away any evidence for mistaken ideas. However, the three lines of work in the brief solution ISS-B are also present in the elaborated solution ISS-E.

To investigate TAs' grading practices when grading student solutions to QM problems and compare them with their introductory physics grading, we incorporated a QM problem (see Fig. 3), two student solutions to this problem (see Fig. 4), and a grading worksheet in addition to the grading activities for introductory physics in the present study. The QM problem was developed and iterated over a period of more than ten years and had been used on midterms and exams in several advanced QM courses at a large research university. The difficulty of the QM problem is such that an average student in a quantum mechanics course would choose to use an exploratory decision-making process as opposed to an algorithmic procedure. An initial qualitative analysis of the problem and planning can greatly facilitate the problem-solving process. Two student solutions to this problem were developed and iterated several times between three researchers and were designed to mimic actual student responses to the QM problem from previous years with common conceptual difficulties. The student solutions include a quantum mechanics student solution that is brief (QSS-B) and a

Introductory Physics Problem

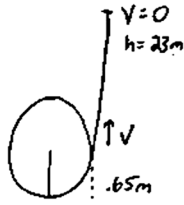
You are whirling a stone tied to the end of a string around in a vertical circle having a radius of 65 cm. You wish to whirl the stone fast enough so that when it is released at the point where the stone is moving directly upward it will rise to a maximum of 23 meters above the lowest point in the circle. In order to do this, what force will you have to exert on the string when the stone passes through its lowest point one-quarter turn before its release? Assume that by the time you have gotten the stone going and it makes its final turn around the circle, you are holding the end of the string at a fixed position. Assume also that air resistance can be neglected. The stone weighs 18 N.

The correct answer is 1292 N

FIG. 1. The introductory physics problem for which the TAs created a solution and graded several student solutions.

ISS-E

ISS-B



$$V^2 = 2gh$$

$$F - mg = \frac{m 2gh}{R}$$

$$F = 18 + \frac{2 \cdot 18 \cdot 23}{.65} = 1292 \text{ N}$$

Energy conservation between top and release

$$\frac{1}{2}mv^2 = mgh$$

$$v^2 = 2gh$$

$$v = \sqrt{2(-9.8)23}$$

$$v = 21.2$$

Uses h instead of h-R

makes sign error

changes sign

Between release and bottom $T \perp v$ so no work done

\therefore Energy is conserved and velocity is the same

$$\sum \vec{F} = m\vec{a}$$

$$T - mg = \frac{mv^2}{R}$$

$$T = 18 + \frac{18}{.65} \cdot 21.2^2$$

$$= 1292 \text{ N}$$

uses v_{release} instead of v_{bottom}

FIG. 2. For the introductory physics problem, an introductory student solution that is elaborated (ISS-E) and an introductory student solution that is brief (ISS-B).

quantum mechanics student solution that is elaborated (QSS-E) (see Fig. 4).

To make comparisons in the grading approaches of the TAs for the introductory physics solutions and QM solutions, the QM solutions developed were made analogous to the two introductory physics solutions, i.e., QSS-B for QM is similar to ISS-B for introductory physics and QSS-E for QM is similar to ISS-E for introductory physics. Both QSS-B and QSS-E include the correct answer. QSS-E, like ISS-E, includes articulation of the principles

used to find intermediate variables, and clear justification for the final result. Similar to ISS-E, the elaborated reasoning in QSS-E reveals a mistake involving writing the wave function immediately after measurement as $\psi(x) = a/2$ rather than $\psi(x) = A\delta[x - (a/2)]$, though the delta function is mentioned later in the solution. Like the brief introductory solution ISS-B, the quantum solution QSS-B is brief with no explication of reasoning, and it does not give any evidence for mistaken ideas on the part of the student. However, the three lines of work in

Quantum Mechanics Problem

For an electron in a one-dimensional infinite square well with well boundaries at $x = 0$ and $x = a$, measurement of position yields the value $x = a/2$. Write down the wave function immediately after the position measurement and without normalizing it show that if energy is measured immediately after the position measurement, it is equally probable to find the electron in any odd-numbered energy stationary state.

FIG. 3. The upper-level quantum mechanics problem.

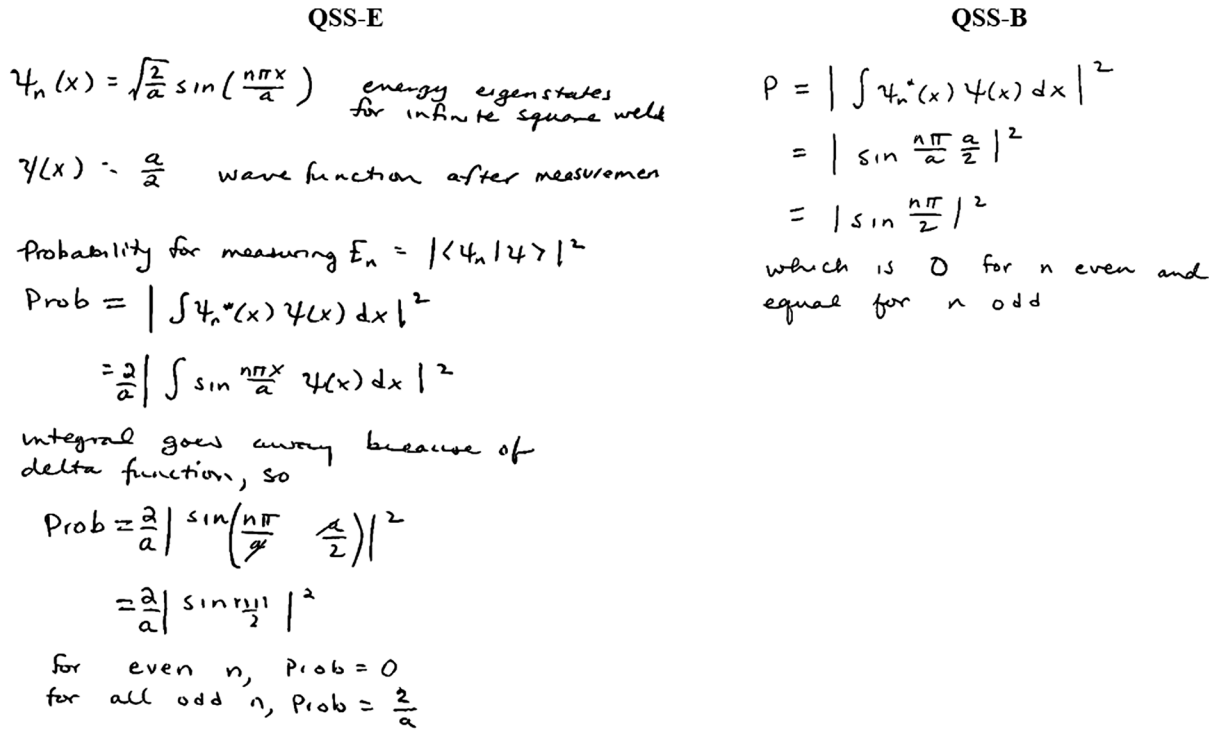


FIG. 4. For the quantum mechanics physics problem, a quantum mechanics student solution that is elaborated (QSS-E), and a quantum mechanics student solution that is brief (QSS-B).

QSS-B are also present in QSS-E. We note that clearly incorrect aspects in the introductory student solution are indicated by boxed notes to help TAs identify where mistakes were made.

2. Implementation of the data collection tool within the professional development course

The sequence of grading activities in the TA professional development course included several stages (see Table I). The activities served as a data collection tool in order to study TAs’ grading decisions and considerations in a simulated environment as well as a learning experience within the professional development program [3].

At the beginning of the semester, TAs were asked to create a solution to the introductory physics problem (see Fig. 1) and the QM problem (see Fig. 3) that they would give to their students to help them learn. The TAs were also asked to individually grade introductory physics solutions ISS-E and ISS-B for both homework (HW) and quiz contexts out of a total score of ten points, list characteristic solution features, and explain their choice of weights for the different features to obtain a final score [see Fig. 5(a)]. The TAs were told to assume that (i) they were the instructors of the class and could structure their grading approaches to improve learning (ii) they had the authority to make grading decisions and (iii) they had told their students how they would be graded. An example response (transcribed) is shown in Fig. 5(a).

During the in-class stage at the beginning of the semester (see Table I), the TAs worked in groups of 3 to 4 in which they were asked to discuss and try to reach an agreement regarding grading the student solutions ISS-E and ISS-B to the introductory physics problem. After they had graded the solutions, a representative from each group shared their grading approaches with the entire class. Two of the authors were present in the class. One researcher coordinated the class work and the discussion at the end of the class which highlighted “best practices” of grading, i.e., grading approaches that promote desired problem solving based upon the grading framework in Refs. [3,8,9]. The instructor of the professional development program also discussed with TAs the disadvantages of grading which focused exclusively on correctness. One researcher observed and documented the TAs’ comments during the group and whole-class discussions.

Right after the introductory grading activities were over, the TAs participated in grading activities in the quantum mechanics context (see Table I). The TAs were provided with a correct instructor’s solution to the QM problem shown in Fig. 3 because many of the TAs struggled to solve the QM problem correctly. Then, the TAs were asked to grade the two student solutions to the QM problem QSS-E and QSS-B (see Fig. 4) for both the homework (HW) and quiz contexts out of a total score of ten points. See Fig. 5(b) for an example response. The TAs were also asked to list characteristic solution features of QSS-E and QSS-B and explain their choice of weights for the different features to

TABLE I. Sequence of TA grading activities.

Time	Activity	Data collection
Beginning of semester	Homework	<ul style="list-style-type: none"> TAs' solutions to the introductory physics problem and QM problem were collected.
	In class	
Immediately after the introductory student grading activities	Homework	<ul style="list-style-type: none"> TAs' worksheets on grading the solutions to the introductory physics problem were collected, scanned, and returned to the TAs. TAs' group worksheets on grading the solutions to the introductory physics problem were collected. TAs' worksheets on grading solutions to the quantum mechanics problem were collected.
	In class	

obtain a final score. After the TAs had graded the solutions to the quantum mechanics problem, they were asked the following questions regarding their grading practices:

(1) Was your grading approach different when grading introductory physics student solutions versus

upper-level quantum mechanics student solutions? If so, why? If not, why not?

(2) How did your grading considerations change when grading introductory physics student solutions versus upper-level quantum mechanics student solutions? What are the reasons for these differences?

Features: ISS-B	Score		Reasons: explain your reasoning for weighing the different features to result with the score you arrived at.
	HW	Q	
The answer is correct, the approach is correct. The steps for getting $v^2=2gh$ are not written	8	10	I gave this student a lower grade on HW because I think that students have enough time to write down all steps, and they should. This answer looks like it has been written just to get a grade, not that the student was learning something while doing the HW. I think that since the approach and the answer are right, this answer gets a full grade on a quiz.

(a)

Features: QSS-E	Score		Reasons: explain your reasoning for weighing the different features to result with the score you arrived at.
	HW	Q	
-Organizing/setting up -One mistake $\Psi(x) \neq \frac{a}{2}$ -Missing $ A $ -Explaining himself adequately -Knowns and unknowns	9	9.5	This student understands the problem and using the correct approach. However, his statement $\Psi(x) = \frac{a}{2}$ is not correct and he omitted $ A $. For homework, he will lose one point, but for quizzes $\frac{1}{2}$ point is enough.

(b)

FIG. 5. (a) One component of a sample TA's worksheet (transcribed) related to introductory student solution ISS-B which was part of the pregrading activity. (b) One component of a sample TA's worksheet (transcribed) related to advanced QM student solution QSS-E which was part of the pregrading activity.

The TAs were not aware that they would be comparing their grading practices in the introductory physics and quantum mechanics contexts before they graded the solutions to the quantum mechanics problem.

One can hypothesize that some of the TAs' beliefs about grading in the quantum mechanics context may have been influenced by the introductory physics grading activities, in particular, the in-class small group and whole class discussions of grading in introductory physics. However, we note that, in this study, we asked TAs to regrade the same solutions to the introductory physics problem in the last week of the professional development program. We found that, on average, there was no noticeable change in the TAs' grading approaches for the introductory physics problem after one semester of professional development and teaching experience (except for a few TAs whose scores went up or down). Thus, it is reasonable to assume that the in-class discussions about introductory physics grading were not likely to change how the TAs graded the solutions to the quantum mechanics problem.

3. Postcourse interviews

After an initial analysis of the collected data, approximately one month after the TA professional development course was over, seven of the TAs in the study volunteered to be interviewed to provide further clarification of their grading beliefs and practices and to investigate whether the grading activities carried out in the TA training class impacted their beliefs about their grading in some manner not captured in their written responses. The researchers hypothesized that it would be valuable to interview the TAs after the semester was over in order to provide further clarification of their grading beliefs and practices and to investigate whether the grading activities carried out in the TA training class impacted their beliefs about their grading in some manner not captured in their written responses. The interviewer had some predetermined questions to ask the TAs (e.g., Can you elaborate on the differences in grading solutions to introductory physics problems compared to grading solutions to QM problems? Did your approach to grading students' solutions to introductory physics problems change after the grading activity involving QM solutions? Have your experiences as a TA in introductory physics caused you to reflect on your grading approach? Have your beliefs about grading changed due to the interventions in the TA professional development course? What caused the change in beliefs?). However, the interviewer also asked additional follow-up questions on the spot to examine TAs' reasoning and also showed them their original written responses to the grading activity to give them an opportunity to clarify their responses on the worksheets if there were any ambiguities in their responses.

IV. RESULTS

A. TAs include effective problem-solving approaches in their own solutions to the introductory physics and QM problem

To investigate research question 1 related to the features the TAs included in their own solutions they would give to their students to help them learn, we examined TAs' own solutions to the introductory physics problem and QM problem. For the QM problem, many TAs stated in class that the homework assignment to create a solution to it to help their students learn was difficult for them because they struggled to solve the problem. However, we find that although many of the TAs struggled to solve the QM problem, thirteen of the TAs included an explanation and justification of their reasoning while solving the problem. In addition, eleven of the TAs broke the problem down into intermediate steps in order to solve it. We also examined the TAs' solutions to the introductory physics problem and found that 13 of the TAs included a diagram and 14 of the TAs explicated and justified their reasoning for using physics concepts. Eleven of the TAs broke the introductory physics problem into subproblems.

Thus, the problem solutions created by the TAs for their students show that, in both introductory physics and QM contexts, they thought it was valuable to explicate their problem-solving approach in the solutions they created for their students. The features included in their solutions suggest that the TAs thought it was important to solve problems using a systematic approach. However, in the following section, we discuss findings that suggest that although TAs created solutions that included effective problem-solving approaches in both the QM and introductory physics contexts, they often did not penalize solutions they were asked to grade in which these features were missing in the introductory physics context. In contrast, in the QM context, TAs were more likely to grade on explication of problem solving and explicit demonstration of conceptual understanding.

B. TAs grade solutions to introductory physics problems differently than solutions to QM problems

To investigate research question 2 (*Do TAs grade students' solutions to an advanced QM problem differently than students' solutions to an introductory physics problem?*), TAs' assigned scores on the QM solution and the introductory physics solution were analyzed. Table II shows the average scores and standard deviations when TAs graded the introductory solutions and the QM solutions in both the homework and quiz contexts. TAs tended to score elaborated solutions higher and brief solutions lower in both the introductory and QM contexts, but the difference was more pronounced for the QM solutions than for the introductory physics solutions. The highest

TABLE II. Average scores assigned to the brief and elaborated solutions to the introductory and QM physics problems, with corresponding standard deviations (S.D.) for each score, with p values for comparison between brief and elaborated solution scores as well as between introductory physics and QM solution scores.

		Introductory physics solutions			QM solutions		
		Brief (ISS-B)	Elaborated (ISS-E)	p	Brief (QSS-B)	Elaborated (QSS-E)	p
HW	Average	6.00	7.40	0.130	4.93	7.67	<0.001
	S.D.	3.16	1.30				
Quiz	Average	7.07	7.93	0.274	6.57	8.47	0.008
	S.D.	2.71	1.24				

disagreement among TAs was about what scores to assign the brief solution to the introductory problem ISS-B (S.D. = 3.16 for the HW context and 2.71 for the quiz context). We performed t tests for comparison, and found that the differences in averages were statistically significant between the QM solutions QSS-B and QSS-E in both the HW context ($p < 0.001$) and quiz context ($p = 0.008$) but not statistically significant for the introductory solutions (see Table II).

Figure 6 shows the distribution of TAs' assigned scores for the elaborated solutions to the QM problem versus the introductory problem in the quiz (left) and homework (right) contexts. The smallest bubbles represent one TA, and a larger bubble shows that many TAs are clustered at that point (the number of TAs at a particular point is proportional to the relative size of the bubble). TAs who are above the diagonal line in the graphs scored the QM solution higher than introductory physics solution. While the scores were mostly grouped near the upper right corner, the scores were somewhat higher for the QM problem than for the introductory problem, though a t test shows that the difference between the means was not statistically significant for either the quiz ($p = 0.307$) or the HW ($p = 0.625$) contexts.

Figure 7 shows the distribution of TAs' assigned scores for the brief solutions to the QM problem v. the

introductory physics problem in the quiz (left) and homework (right) contexts. TAs tended to grade the brief QM solution somewhat lower than the brief introductory solution (fewer TAs are above the diagonal line), though the difference in the means was not statistically significant for either the quiz ($p = 0.574$) or the HW ($p = 0.273$) contexts. Figure 7(a) shows that brief quiz solutions (both introductory and QM) are graded in an inconsistent manner by the TAs. On the other hand, Fig. 7(b) shows that TAs were rather consistent when grading brief QM homework but gave widely varying scores when grading the brief introductory homework.

In addition to examining the scores TAs gave to the brief and elaborated solutions in the introductory physics and QM contexts, we investigated the solution features TAs graded on. In the grading activities, TAs were asked to grade introductory physics solutions and the QM solutions in a HW and a quiz context, list features of each solution, and explain their choice of weights for the features to arrive at a final score. Data analysis involved coding the features listed by TAs in the worksheets into a combination of theory-driven and emergent categories. Twenty-one solution features were identified. The coding was done by two of the researchers. In cases where disagreement occurred, this was usually due to vagueness in the wording of TAs'

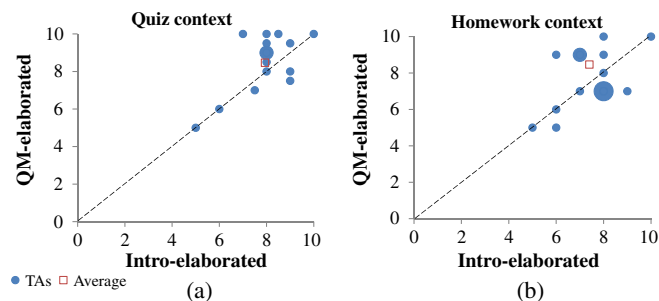


FIG. 6. (a) Distribution of individual scores assigned to the elaborated solutions QSS-E (QM) versus ISS-E (Intro) in the quiz context ($p = 0.307$). (b) Individual scores assigned to the elaborated solutions QSS-E (QM) versus ISS-E (Intro) in the HW context ($p = 0.625$). The relative size of the bubble represents the number of TAs at a particular point.

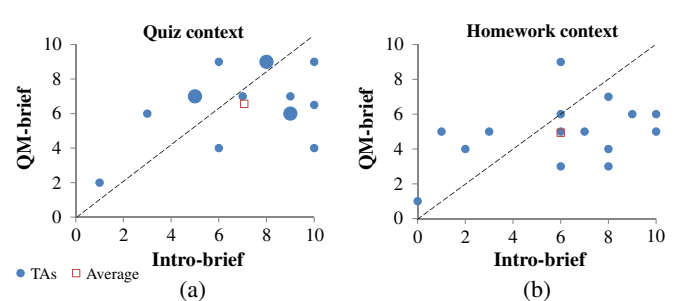


FIG. 7. (a) Distribution of individual scores assigned to the brief solutions QSS-B (QM) versus ISS-B (Intro) versus QSS-B (QM) in the quiz context ($p = 0.574$). (b) Individual scores assigned to QSS-B (QM) versus ISS-B (Intro) in the HW context ($p = 0.273$). The size of the bubble represents the number of TAs ($N = 15$).

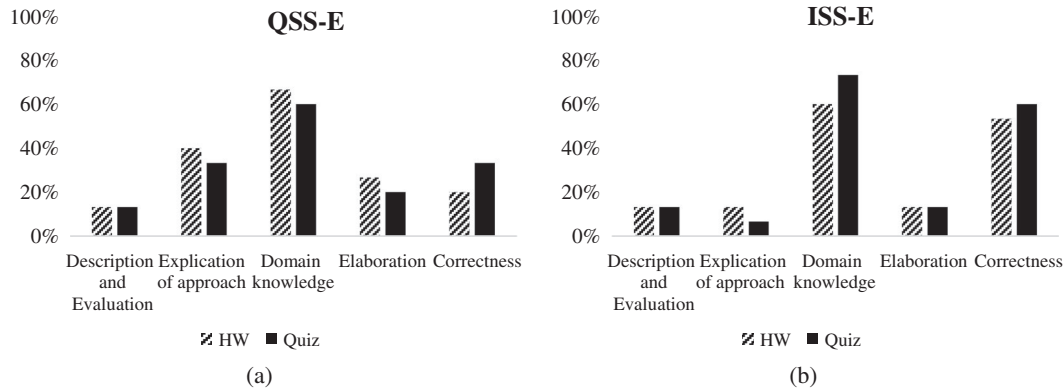


FIG. 8. (a) Percentage of TAs who graded on solution features in the five clusters on the elaborated QM solution QSS-E in the homework and quiz context. (b) Percentage of TAs who graded on solution features in the five clusters on the elaborated introductory physics solution ISS-E in the homework context ($N = 15$ TAs).

written statements. After comparing codes, the researchers discussed any disagreements during multiple meetings until agreement better than 90% was reached.

To facilitate interpretation of the data, the features were analyzed by grouping them into 5 clusters, as shown in Table III. Each solution feature listed by a TA was entered into only one cluster. Cluster 1 (C1) includes features related to effective problem-solving approaches [9–20] (i.e., initial problem analysis as well as evaluation of the final result). Cluster 2 also involves features related to effective problem-solving approaches such as explication of reasoning (i.e., articulation and justification of principles). Cluster 3 (C3) includes domain-specific features, such as invoking relevant physics principles and applying them properly. Cluster 4 (C4) includes features related to elaboration which emerged during the coding process, e.g., “written statements,” “good presentation,” “solution in steps,” and “conciseness.” These features were not assigned to the “explication” category C2 because they were imprecise. Cluster C2 is focused on the explication and justification of the physics principles, whereas C4 is more

about general communication of the solution. For example, we could not differentiate whether a TA who wrote “written statements” meant that the student solution includes an explicit statement of a principle in writing, explicit justification of a principle in writing, or simply a written statement. Thus, we coded “written statements” as belonging in the general category C4 (elaboration). Finally, Cluster 5 (C5) focuses on correctness of algebra and the final answer.

Figure 8 shows the percentages of TAs who graded on solution features in the five clusters in the elaborated QM solution QSS-E and the elaborated introductory physics solution ISS-E when treating the student solutions in a homework and quiz context. When grading the elaborated solutions, many TAs focused on domain knowledge in both introductory physics and quantum mechanics. However, TAs were more likely to grade on cluster C2 (explication) in QM as opposed to introductory physics. In addition, TAs were less likely to grade on cluster C5 (correctness) in QM as opposed to introductory physics.

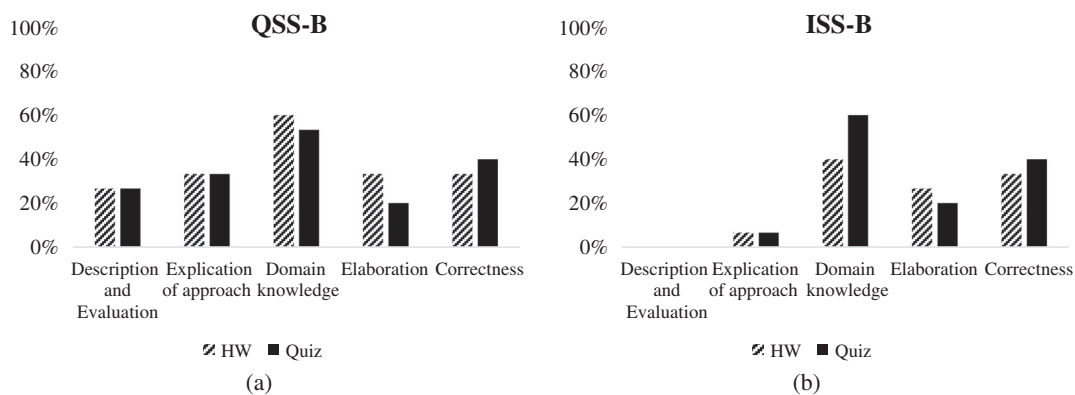


FIG. 9. (a) Percentage of TAs who graded on solution features in the five clusters on the brief QM solution QSS-B in the homework and quiz context. (b) Percentage of TAs who graded on solution features in the five clusters on the brief introductory physics solution ISS-B in the homework and quiz context ($N = 15$ TAs).

Figure 9 shows the percentages of TAs who graded on solution features in the five clusters on the brief QM solution QSS-B and the brief introductory physics solution ISS-B in the homework and quiz context. Again, many of the TAs were focused on correct domain knowledge in both introductory physics in QM. However, the TAs were more likely to grade on C1 (problem description and evaluation) and C2 (explication) in QM as opposed to introductory physics.

These findings suggest that TAs expect students in QM to show evidence of understanding via problem description, evaluation, and explication of their problem-solving approach. However, in grading introductory students, the TAs were mainly focused on domain knowledge and correctness. Below, we discuss TAs' stated reasons for why they graded differently in the two contexts in their written responses and interviews.

C. TAs grade differently in the QM and introductory physics context because of differences in the perceived complexity of the problem and differing expectations of students

To investigate research question 3 [*What are the TAs' stated reasons for whether (or not) their grading is different for introductory problems versus QM problems?*], TAs were asked to write responses to the following two questions, which were part of the QM grading activity: 1. "Was your grading approach different when grading introductory physics student solutions versus upper-level quantum mechanics student solutions? If so, why? If not, why not?" 2. "How did your grading considerations change when grading introductory physics student solutions versus upper-level quantum mechanics student solutions? What are the reasons for these differences?" In addition, a subset of the TAs were interviewed approximately one month after the professional development course to further clarify their views about grading solutions to QM and introductory physics problems. In their written responses, 10 of the 15 TAs (67%) noted that they would grade the QM and introductory physics problems differently, while 5 TAs (33%) noted that they would not grade differently in the two contexts. TAs' written responses about the reasons why they would grade differently or not in the two contexts were analyzed using open coding to generate initial categories grounded in the actual data [22]. Once initial categories emerged from the data, the coding was completed by two of the researchers separately. After comparing codes, any disagreements were discussed and the categories were refined until better than 90% agreement was reached. Table IV shows the categories of TAs' written responses for why (or why not) they would grade differently in the introductory physics and QM contexts, example citations, and the percentages of TAs who mentioned each category. We note that TAs could have written more than one reason

for why they graded differently in the QM and introductory solution contexts.

Out of the 15 TAs in the course, 26% of them expected that students should explicitly demonstrate their understanding when solving QM problems as opposed to introductory physics problems because QM problems are more complex. For example, one TA gave this explanation for why he would use different criteria: "For the upper level courses, the concepts are more complex, need more explanation." Interviews suggest that the TAs with these types of responses were able to gauge the difficulty of a QM problem from the perspective of an advanced student because they were themselves at a similar expertise level. In interviews, some of the TAs explicitly mentioned that QM problems are significantly more difficult than introductory physics problems. One interviewed TA explained that she had thought about how much more difficult a QM problem is and, therefore, focused significantly more on the proof of understanding when grading solutions to QM problems, stating: "In QM I don't expect people to be able to do things in their mind, so if they're not writing it down I kind of feel they don't know it." The TAs who felt that the QM problem is significantly more difficult than the introductory physics problem often stated that they expected the students in QM courses to explicate the problem-solving approach and they would grade them on the use of a systematic approach to problem solving that includes a conceptual analysis of the physics problem (but they did not have the same high expectations when grading introductory solutions). Some interviewed TAs who claimed that QM is more difficult than introductory physics mentioned that introductory physics is easier than QM because introductory physics is more concrete. For example, one TA mentioned "For introductory physics we can make an example to understand the questions more clearly." Another TA compared herself to her students (instead of putting herself in introductory students' shoes and thinking from introductory students' perspective) and stated: "In introductory physics, since I can do it in my mind, I think that introductory students can do it [in their minds] too."

In their written responses, 26% of the TAs claimed that advanced students should demonstrate their understanding when solving QM problems because they are already expected to have learned physics concepts as well as effective problem-solving approaches. For example, one interviewed TA stated that "high-level students have gone through many years of training, what they need is [to] interpret the problem [to get credit]." This TA felt that after many years of training, students should be able to articulate their thought processes explicitly in their solution in order to receive credit. In addition, 40% of the TAs noted that they would grade advanced students on the explication of concepts, but that they would grade introductory students on whether they used correct formulas or mathematical

TABLE III. Sample features sorted into clusters and sample citations.

C1 problem description and evaluation	Visual representation (e.g., “diagram,” “figure,” “graph”); articulating the target variables and known quantities (e.g., “knowns or unknowns,” “list of variables,” “nothing labeled”); evaluation of the reasonability of the final answer (e.g., “check,” “double check what they did”)
C2 explication of problem-solving approach	articulation of principles (e.g., “labels energy conservation use,” “text showing knowledge of concepts”); justifying principles (e.g., “explained the reason he used the formulas,” “explanation for constant velocity,” “no demonstration for why the first equation holds”)
C3 domain knowledge	Essential principle invoked (e.g., “sums forces, energy conservation,” “has not written [the stationary state for an infinite square well] explicitly,” “knows how to calculate the probability of an event” “does not write wave function after measurement,”); essential principle is applied adequately [e.g., “mistake $\psi(x) \neq a/2$,” “wrong $\psi(x)$, correct probability”]
C4 elaboration	Explanation; written statements (e.g., “verbal explanations,” “narration”, “no text,” “doesn’t explain anything,” “no words,” “no statements”)
C5 correctness	Algebraic errors (e.g., “makes sign error”); correct final answer (e.g., “final result right”)

steps and got the correct answer. For example, one interviewed TA stated, “If a student is majoring in physics, they should be able to understand all the concepts perfectly to be able to solve complicated problems. In upper-level courses, I think the student should understand everything they are doing, they are not allowed to just use an equation because they have seen people use [that equation] before.” This TA emphasized that a formula-fitting approach was

acceptable in introductory physics courses but not in advanced physics courses for physics majors. Another TA claimed that her grading focused more on concepts in QM and that “in introductory physics (assuming the students are not majoring in physics) it’s okay if they only learn how to use equations and how to solve problems because they might have not seen physics problems before in their life...” Yet, another TA who only valued

TABLE IV. Explanation of categories used for coding 15 TAs’ stated differences or similarities when grading student solutions for introductory versus QM physics problems and percentages of TAs mentioning each category. TAs could mention more than one category so the percentages do not add up to 100%.

Category	Examples	% of TAs
Demonstrating understanding is more important in QM because the subject is more complex.	–“Expect more explanations (in QM) because it’s a more difficult course.” –“For the upper level courses, the concepts are more complex, need more explanation.”	26
Demonstrating understanding is more important in QM because it is expected of advanced students	–“high-level students have gone through many years of training, what they need is interpret the problem [to get credit].”	26
Grading should focus more on conceptual understanding in QM and more on procedures (use of equations, calculations, solving steps, correct math) in introductory physics	–“If a student is majoring in physics, they should be able to understand all the concepts perfectly to be able to solve complicated problems.” –“I will consider (grading) more on the interpretation of problems when grading upper level quantum mechanics students. As for the introductory level students, I will consider more on their calculation, solving steps”.	40
Problem features such as diagrams and lists of unknown quantities are more important for introductory physics problems than for QM problems.	–“Upper level student should not waste time on drawing graphs that they are familiar with, they can decide if they need a diagram to help themselves.” –“Focus more on concept understanding than diagram or list (in QM).”	20
The grading standards should be the same for introductory and QM physics problems.	–“I think whether a student majors in the field or not, they should be held up to the same standard in grading, because the difference already exists in how hard the questions are, and to reach the objective of the course, students should be expected to do things right even in introductory courses.” –“I would put equal weight on different criteria and look for whether they are present/absent and correct/incorrect. That means an equal framework for both seniors (QM) and freshmen (introductory physics).”	33

clarification and interpretation in problem solution in the context of QM stated, “I will consider (grading) more on the interpretation of problems when grading upper level quantum mechanics students. For introductory students, I will consider more on their calculation...” These TAs, in general, required more of advanced students than introductory students in terms of whether they needed to clearly explain why they were using some concepts to solve a problem. However, some of these TAs who cared about students demonstrating their conceptual understanding in QM were not as critical of mathematical mistakes in QM. For example, an interviewed TA noted the differences in mathematical complexity between introductory physics and QM stating, “the introductory physics involves more fundamental mathematics while the upper-level quantum mechanics always requires integrals or other upper-level mathematics. So I will be more tolerant to the mathematical mistakes in the upper-level physics course [but he would not tolerate if QM solutions did not clearly explain why some concepts were applied].” The TA stated that “in [advanced physics] exams and quizzes, skipping some steps are tolerable and minor issues compared with introductory students.”

Also, while a majority of TAs expected QM solutions to demonstrate conceptual understanding and explication of the problem-solving approach, about 20% of the TAs stated that drawing a diagram and listing given values and unknown variables are not important when solving QM problems (although they are important when solving introductory physics problems). However, for the particular QM problem the TAs were asked to grade in this study, the answer could have been checked by drawing a diagram of the wave function after the measurement of position (with a delta function in the middle of the well) and also drawing the stationary state wave functions for an infinite square well. By drawing these diagrams, one can rationalize that since the even-numbered stationary state wave functions are zero at the center of the well, the probability of finding the particle in an even-numbered stationary state would be zero after the measurement of energy.

Some TAs held contradictory beliefs regarding the importance of conceptual understanding in QM. For example, in an interview, one TA stated, “in introductory physics we expect that the student is still learning, but when you are doing something like QM we expect that you understand the basic physics and you can easily implement it in your advanced work, so we expect somewhat more understanding.” In the latter part of the interview, when explicitly asked if students learning QM face similar challenges to students learning introductory physics, the same TA continued, “QM is in itself a difficult thing to understand...so this is a factor...the problem-solving pattern will be the same for both but the concept may be different. Conceptually, I will be lenient [when grading QM]...when it comes to getting the answer perfectly and

reaching a good result...” This TA first claimed that he expected more understanding in QM. However, when this TA was asked an explicit question about the similarities and differences between introductory and advanced students, he then stated that he would grade more leniently on conceptual understanding in QM since QM is difficult. This TA appeared to change his initial views somewhat after the interviewer asked probing questions [23].

Five out of the 15 TAs noted that both introductory physics and QM should have the same grading standards. One TA explained why she graded QM and introductory student solutions similarly stating, “they should be held up to the same standard in grading, because the difference already exists in how hard the questions are, and to reach the objective of the course, students should be expected to do things right even in introductory courses.” In the interview, this same TA stated, “the difference should be embodied in the difference of the questions, not the grading. The way you do things should be held up to the same standard for all kinds of students.” This TA felt that even though the topic may be different in introductory physics and QM, the grading should focus on similar standards for all students. Another TA, who had graded the brief solutions lower than the elaborated solutions in both the QM and introductory physics contexts, noted that he did not think that grading should change based upon the level of the student: “it is the same physics, different concepts. If she or he is at this [advanced] level, no need to grade different.” Furthermore, another TA noted that he would use similar standards for grading in both the QM and introductory physics contexts, stating “I would put equal weight on different criteria and look for whether they are present, absent, correct, or incorrect. That means an equal “framework” for both seniors (advanced students) and freshmen (introductory physics students).” Although five TAs stated that their grading practices would be similar in the QM and introductory physics contexts, there was a contradiction in what some of these TAs claimed they would do and what they actually did. A comparison of their scores on the QM and introductory physics solutions shows that 3 out of these 5 TAs scored the brief introductory physics solution higher than the elaborated introductory physics solution on a quiz. On the other hand, none of these 5 TAs scored the brief QM solution higher than the elaborated QM solution in a quiz context.

Some interviews suggest that a few TAs’ beliefs about grading in QM and introductory physics may have been somewhat positively impacted by the intervention in the professional development course. For example, one interviewed TA stated, “In QM I don’t expect people to be able to do things in their mind, so if they’re not writing it down I kind of feel they don’t know it, but in introductory physics since I can do it in my mind, I (used to) think that intro students can do it too. But then I learned that that’s not the case, if I can’t do quantum in my mind then they can’t do

(introductory physics) in their mind.” Further discussions with the TA suggests that she was learning to put herself in her students’ shoes and was beginning to recognize that advanced students learning QM are similar to introductory students learning introductory mechanics. At the beginning of the semester this TA gave ISS-B a score of 8/10 in the HW context and noted in her explanation of her score that “the final answer is correct.” However, when asked at the end of the semester to grade ISS-B once again, she gave ISS-B in the HW context a score of 6/10 and wrote the following: “As a homework problem, the student has to show me that they understand what is going on and write down the steps.” The interview and grading data suggest that she was starting to realize that she needed to require evidence of understanding even in introductory physics student solutions. It is possible that this shift in her opinion was the combined effect of the grading activities in the professional development course and her own experiences in teaching and learning. However, based upon their re-grading of the problem in the last class of the professional development course, we note that, on average, there was no noticeable change in the TAs’ grading approaches for the introductory physics problem after one semester of professional development and teaching experience (except for a few TAs whose scores went up or down).

V. SUMMARY AND IMPLICATIONS

In this study, we found that a majority of the TAs thought that solution features such as explicating the problem-solving approach, drawing diagrams, breaking the problem into subproblems, evaluating the answer, etc., can help students learn in both introductory physics and QM and they included those features in their prepared solutions for their students. Most of the solutions that the TAs created for the introductory physics problem and the QM problem included effective problem-solving approaches. However, a majority of the TAs expected elaborated solutions that explicated the problem-solving approach and explicitly demonstrated conceptual understanding from students in a QM course but not from students in introductory physics courses. Only one-third of the TAs stated that they would grade introductory student solutions and QM solutions in a similar manner. An analysis of their actual grading shows that even among those TAs, some were stricter in grading QM solutions than the introductory solutions. TAs stated that they graded differently in the introductory physics and QM context because they thought that QM is more complex than introductory physics and that grading in QM should focus more on concepts and grading in introductory physics should focus more on steps and procedures. Many TAs also mentioned that they graded advanced students’ solutions more strictly than introductory students’ solutions because they had higher expectations of advanced physics students.

Several TAs explicitly noted that the introductory physics problem was easier than the QM problem, suggesting

that they had not considered the difficulty of the introductory problem from the perspective of an introductory student [6]. Prior research has shown that TAs often view solutions to introductory physics problems as obvious [6] and TAs who think that introductory problems are easy often inferred correct understanding in a student’s solution when there was no evidence of it [2]. In the study presented here, many TAs claimed that since the introductory physics problem was easier than the QM problem, they did not require introductory students to explicate the problem-solving approach and show conceptual understanding. This viewpoint was a factor in why they graded introductory solution heavily on correctness as opposed to on the explication of the problem-solving approach. On the other hand, many TAs claimed that the QM problem was more difficult than an introductory physics problem, suggesting that they did consider the difficulty of the QM problem from the perspective of an advanced student. They claimed that since the QM problem was more difficult than the introductory physics problem, they expected students in QM to explicate their reasoning and demonstrate effective problem-solving approaches in order to obtain a high grade. However, to an introductory student, introductory physics is also highly abstract and conceptually difficult. The discrepancy in TAs’ considerations of introductory and advanced students’ perspectives is one possible reason for why TAs graded solutions to introductory physics problem and QM problem differently.

Professional development programs for TAs can include early discussions of the difficulties that introductory physics students face when solving introductory-level problems and the importance of helping students learn and develop better problem-solving approaches. In particular, the professional development of TAs may be improved by “framing” for TAs via activities and explicit discussions that the challenges encountered by introductory students when solving introductory physics problems are analogous to those that advanced students face in solving QM problems. It also may be helpful to encourage the TAs to explicitly think about their own problem-solving approaches when they solve the QM problem and why those approaches would also help introductory students when they solve problems in introductory physics. TAs might be asked to list the ways in which they are similar to introductory physics students, possibly helping them realize that students learning introductory physics face similar challenges to advanced students learning QM.

Furthermore, many TAs included effective problem-solving approaches (e.g., explicated the problem-solving approach) in their own worked-out solutions for students in introductory physics and QM. However, they did not grade on the explication of the problem-solving approach in introductory physics. In fact, many TAs claimed that introductory students are novices and it is fine for them to focus only on the correct formulas, mathematical steps,

and final answer during problem solving. On the other hand, these same TAs claimed that students in QM should be required to demonstrate conceptual understanding because they have learned more physics and are expected to have learned to use effective problem-solving approaches. TAs' stated belief that they should be stricter in grading in QM than in introductory physics because introductory students had not learned problem-solving skills and were not experts in physics (so a lenient grading standard should be appropriate for them) suggests that they had not considered whether grading practices can serve as a formative assessment tool [3,8,9,24–32]. TAs' stated beliefs about grading in introductory physics and QM suggest that they considered grading solely as a summative assessment of student learning (i.e., to evaluate what students had learned so far). Their view of grading as summative assessment may have been one reason why TAs graded differently in introductory physics and QM (i.e., TAs' stated beliefs that since introductory students are novices, they are not expected to use effective problem-solving approaches, whereas advanced students have learned more physics and are expected to use effective problem-solving approaches). We hypothesize that TAs' grading practices in an exam setting may be similar to or more lenient than a

quiz setting due to time constraints and stress in an exam setting, although future work can focus on the factors that influence TAs' exam grading practices. Similar to the importance of investigating introductory students' epistemological beliefs about physics [33], it is also valuable to investigate TAs' beliefs about the purposes of grading in homework, quiz, and exam contexts and use their beliefs as resources.

Professional development courses or programs can allow more time and support for the TAs to reflect on the benefits of formative assessment and ways to support students in developing better problem-solving approaches and learning physics [3,8,9,24–32]. TAs may be asked to contemplate how and why their grading practices would differ in a homework, quiz, or exam setting and reflect on whether grading in these different settings can be used as a formative assessment for both the student and the instructor.

ACKNOWLEDGMENTS

We thank the United States National Science Foundation for Grant No. PHY-1505460.

-
- [1] For example, see <http://www.aps.org/programs/education/graduate/conf2008/index.cfm>.
- [2] E. Yerushalmi, E. Marshman, A. Maries, C. Henderson, and C. Singh, Grading practices and considerations of graduate students at the beginning of their teaching assignment, *Proceedings of the Physics Education Research Conference, 2014, Minneapolis, MN*, edited by P. Engelhardt, A. Churukian, and D. Jones (AIP, New York, 2015), p. 287, doi: [10.1119/perc.2014.pr.068](https://doi.org/10.1119/perc.2014.pr.068); C. Henderson, E. Marshman, A. Maries, E. Yerushalmi, and C. Singh, Instructional goals and grading practices of graduate students after one semester of teaching experience, *Proceedings of the Physics Education Research Conference, 2014, Minneapolis, MN*, edited by P. Engelhardt, A. Churukian, and D. Jones (AIP, New York, 2015), p. 111, doi: [10.1119/perc.2014.pr.024](https://doi.org/10.1119/perc.2014.pr.024); E. Yerushalmi, R. Sayer, E. Marshman, C. Henderson, and C. Singh, Physics graduate teaching assistants' beliefs about a grading rubric: Lessons learned, *Proceedings of the Physics Education Research Conference, 2016, Sacramento, CA*, edited by D. Jones, L. Ding, and A. Traxler (2016), doi: [10.1119/perc.2016.pr.097](https://doi.org/10.1119/perc.2016.pr.097); C. Henderson, E. Marshman, R. Sayer, C. Singh, and E. Yerushalmi, Graduate teaching assistants use different criteria when grading introductory physics vs. quantum mechanics, *Proceedings of the Physics Education Research Conference, 2016, Sacramento, CA* edited by D. Jones, L. Ding, and A. Traxler (2016), doi: [10.1119/perc.2016.pr.030](https://doi.org/10.1119/perc.2016.pr.030).
- [3] E. Yerushalmi, C. Henderson, K. Heller, P. Heller, and V. Kuo, Physics faculty beliefs and values about the teaching and learning of problem solving. I. Mapping the common core, *Phys. Rev. ST Phys. Educ. Res.* **3**, 020109 (2007); E. Yerushalmi, C. Henderson, W. Mamudi, C. Singh, and S. Lin, The group administered interactive questionnaire: An alternative to individual interviews, *AIP Conf. Proc.* **1413**, 97 (2012).
- [4] F. Lawrenz, P. Heller, and R. Keith, Training the teaching assistant, *J. Coll. Sci. Teach.* **22**, 106 (1992); *Recruiting and Educating Future Physics Teachers: Case Studies and Effective Practices*, edited by C. Sandifer and E. Brewe (American Physical Society, PhysTEC, 2015); A. H. Schoenfeld, Toward a theory of teaching-in-context, *Issues Educ.* **4**, 1 (1998).
- [5] A. Maries and C. Singh, Improving one aspect of pedagogical content knowledge of teaching assistants using the TUG-K, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020120 (2013); Performance of graduate students at identifying introductory students' difficulties with kinematics graph, *Proceedings of the Physics Education Research Conference, 2014, Minneapolis, MN*, edited by P. Engelhardt, A. Churukian, and D. Jones (AIP, New York, 2015), p. 171 doi: [10.1119/perc.2014.pr.039](https://doi.org/10.1119/perc.2014.pr.039); Teaching assistants' performance at identifying common introductory student difficulties in mechanics revealed by the Force Concept Inventory, *Phys. Rev. Phys. Educ. Res.* **12**, 010131 (2016).

- [6] C. Singh, Categorization of problems to assess and improve proficiency as teacher and learner, *Am. J. Phys.* **77**, 73 (2009); Rethinking tools for training teaching assistants, *AIP Conf. Proc.* **1179**, 59 (2009); A. Mason and C. Singh, Surveying graduate students' attitudes and approaches to problem solving, *Phys. Rev. ST Phys. Educ. Res.* **6**, 020124 (2010); Helping students learn effective problem solving strategies by working with peers, *Am. J. Phys.* **78**, 748 (2010); Impact of guided reflection with peers on the development of effective problem solving strategies and physics learning, *Phys. Teach.* **54**, 295 (2016); Surveying college introductory physics students' attitudes and approaches to problem solving, *Eur. J. Phys.* **37**, 055704 (2016).
- [7] S. Lin, C. Henderson, W. Mamudi, E. Yerushalmi, and C. Singh, Teaching assistants' beliefs regarding example solutions in introductory physics, *Phys. Rev. ST Phys. Educ. Res.* **9**, 010120 (2013); S. Lin, C. Singh, W. Mamudi, C. Henderson, and E. Yerushalmi, TA-designed vs. research-oriented problem solutions, *AIP Conf. Proc.* **1413**, 255 (2012).
- [8] C. Henderson, E. Yerushalmi, V. Kuo, P. Heller, and K. Heller, Grading student problem solutions: The challenge of sending a consistent message, *Am. J. Phys.* **72**, 164 (2004).
- [9] J. Docktor and K. Heller, Robust assessment instrument for student problem solving, in *Proceedings of 82nd NARST Annual International Conference, 2009*; see http://groups.physics.umn.edu/physed/People/Docktor/talks_papers/Docktor_NARST09_paper.pdf; K. Heller and P. Heller, *The Competent Problem Solver for Introductory Physics* (McGraw-Hill, New York, 2000); P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving, *Am. J. Phys.* **60**, 627 (1992); P. Heller and M. Hollabaugh, Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, *Am. J. Phys.* **60**, 637 (1992).
- [10] J. Larkin, J. McDermott, D. Simon, and H. Simon, Expert and novice performance in solving physics problems, *Science* **208**, 1335 (1980); A. Schoenfeld, What's all the fuss about metacognition?, in *Cognitive Science and Mathematics Instruction*, edited by A. Schoenfeld (Lawrence Erlbaum, Hillsdale, New Jersey, 1987); D. Schwartz, J. Bransford, and D. Sears, Efficiency and innovation in transfer, in *Transfer of Learning From a Modern Multidisciplinary Perspective*, edited by J. Mestre (Information Age, Greenwich, CT, 2005), pp 1–51; G. Hatano and K. Inagaki, Two courses of expertise, in *Child Development and Education in Japan* (Freeman, New York, NY, 1986), pp. 262–272.
- [11] F. Reif, Systematic problem solving, in *Applying Cognitive Science to Education: Thinking and Learning in Scientific and other Complex Domains* (MIT Press, Cambridge, England, 2008), pp. 201–227; J. Heller and F. Reif, Prescribing effective human problem solving processes: Problem description in physics, *Cognit. Instr.* **1**, 177 (1984); F. Reif, Millikan Lecture 1994: Understanding and teaching important scientific thought processes, *Am. J. Phys.* **63**, 17 (1995); A. Van Heuvelen, Learning to think like a physicist: A review of research based instructional strategies, *Am. J. Phys.* **59**, 891 (1991).
- [12] J. Mestre, J. Docktor, N. Strand, and B. Ross, Conceptual problem solving in physics, in *Psychology of Learning and Motivation*, edited by J. Mestre and B. Ross (Academic Press, New York, 2011), Vol. 55, pp. 269–298.
- [13] C. Singh, When physical intuition fails, *Am. J. Phys.* **70**, 1103 (2002).
- [14] D. Maloney, Research on problem solving: Physics, in *Handbook of Research on Science Teaching and Learning*, edited by D. Gable (MacMillan, New York, 1994).
- [15] W. Leonard, R. Dufresne, and J. Mestre, Using qualitative problem solving strategies to highlight the role of conceptual knowledge in solving problems, *Am. J. Phys.* **64**, 1495 (1996).
- [16] T. Nokes-Malach, K. VanLehn, D. Belenky, M. Lichtenstein, and G. Cox, Coordinating principles and examples through analogy and self-explanation, *Eur. J. Psychol. Educ.* **28**, 1237 (2013).
- [17] R. Atkinson, A. Renkl, and M. Merrill, Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps, *J. Educ. Psychol.* **95**, 774 (2003).
- [18] M. Chi, M. Lewis, P. Reimann, and R. Glaser, Self-explanations: How students study and use examples in learning to solve problems, *Cogn. Sci.* **13**, 145 (1989).
- [19] M. Chi and K. VanLehn, The content of physics self-explanations, *J. Learn. Sci.* **1**, 69 (1991).
- [20] M. Chi, Self-explaining expository texts: The dual processes of generating inferences and repairing mental models, in *Advances in Instructional Psychology*, edited by R. Glaser (Erlbaum Associates, Hillsdale, NJ, 2000), pp. 161–238.
- [21] See <https://www.aip.org/statistics/data-graphics/demographic-profile-physics-phds-classes-2010-2011-2012-combined>.
- [22] A. Strauss, *Qualitative Research for Social Scientists* (Cambridge University Press, Cambridge, England, 1987).
- [23] N. S. Rebello, D. A. Zollman, A. R. Allbaugh, P. V. Engelhardt, K. E. Gray, Z. Hrepic, and S. F. Itza-Ortiz, Dynamic Transfer: A Perspective from Physics Education Research, in *Transfer of Learning from a Modern Multidisciplinary Perspective*, edited by J. P. Mestre (Information Age Publishing, Greenwich, CT, 2005), pp. 217–250.
- [24] National Research Council 2001, Knowing What Students Know: The Science, and Design of Educational Assessment, *Committee on the Foundations of Assessment*, edited by J. Pellegrino, N. Chudwosky, and R. Glaser (National Academy Press, Washington, DC, 2001).
- [25] P. Black and D. Wiliam, Assessment and classroom learning, *Assess. Educ.* **5**, 7 (1998); Inside the black box: Raising standards through classroom assessment, *Phi Delta Kappan* **92**, 81 (2010); Developing the theory of formative assessment, *Educational Assessment*, *Eval. Accountability* **21**, 5 (2009); P. Black, C. Harrison, C. Lee, B. Marshall, and D. William, *Assessment for Learning: Putting It Into Practice* (Open University Press, Buckingham, 2003).
- [26] E. Yerushalmi, E. Cohen, A. Mason, and C. Singh, What do students do when asked to diagnose their mistakes?

- Does it help them? I. An atypical quiz context, *Phys. Rev. ST Phys. Educ. Res.* **8**, 020109 (2012); E. Yerushalmi, E. Cohen, A. Mason, and C. Singh, What do students do when asked to diagnose their mistakes? Does it help them? II. A more typical quiz context, *Phys. Rev. ST Phys. Educ. Res.* **8**, 020110 (2012).
- [27] A. Mason, E. Yerushalmi, E. Cohen, and C. Singh, Learning from mistakes: The effect of students' written self-diagnoses on subsequent problem solving, *Phys. Teach.* **54**, 87 (2016); B. Brown, A. Mason, and C. Singh, Improving performance in quantum mechanics with explicit incentives to correct mistakes, *Phys. Rev. Phys. Educ. Res.* **12**, 010121 (2016); A. J. Mason and C. Singh, Do advanced students learn from their mistakes without explicit intervention?, *Am. J. Phys.* **78**, 760 (2010).
- [28] C. Singh, Assessing student expertise in introductory physics with isomorphic problems. I. Performance on a nonintuitive problem pair from introductory physics, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010104 (2008); Assessing student expertise in introductory physics with isomorphic problems. II. Effect of some potential factors on problem solving and transfer, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010105 (2008); A. Mason and C. Singh, Using categorization of problems as an instructional tool to help introductory students learn physics, *Phys. Educ.* **51**, 025009 (2016); A. J. Mason and C. Singh, Assessing expertise in introductory physics using categorization task, *Phys. Rev. ST Phys. Educ. Res.* **7**, 020110 (2011).
- [29] S. Y. Lin and C. Singh, Categorization of quantum mechanics problems by professors and students, *Eur. J. Phys.* **31**, 57 (2010); Using isomorphic problems to learn introductory physics, *Phys. Rev. ST Phys. Educ. Res.* **7**, 020104 (2011); Using an isomorphic problem pair to learn introductory physics: Transferring from a two-step problem to a three-step problem, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020114 (2013); Challenges in using analogies, *Phys. Teach.* **49**, 512 (2011); S. Siddiqui and C. Singh, Surveying instructors' attitudes and approaches to teaching quantum mechanics, *AIP Conf. Proc.* **1289**, 297 (2010); How diverse are physics instructors' attitudes and approaches to teaching undergraduate level quantum mechanics?, *Eur. J. Phys.* **38**, 035703 (2017).
- [30] B. White and J. Frederiksen, Inquiry, modeling, and metacognition: Making science accessible to all students, *Cognit. Instr.* **16**, 3 (1998).
- [31] A. Schoenfeld, When good teaching leads to bad results: The disasters of "well-taught" mathematics courses, *Educ. Psych.* **23**, 145 (1988).
- [32] T. Angelo and K. Cross, *Classroom Assessment Techniques: A Handbook for Faculty* (National Center for Research to Improve Postsecondary Teaching and Learning, Ann Arbor, MI, 1998); T. Crooks, The impact of classroom evaluation practices on students, *Rev. Educ. Res.* **58**, 438 (1988).
- [33] A. Elby, Another reason that physics students learn by rote, *Am. J. Phys.* **67**, S52 (1999); D. Hammer and A. Elby, Tapping students' epistemological resources, *J. Learn. Sci.* **12**, 53 (2003); L. Jaber and D. Hammer, Learning to feel like a scientist, *Sci. Educ.* **100**, 189 (2016); D. Hammer, Two approaches to learning physics, *Phys. Teach.* **27**, 664 (1989); Epistemological beliefs in introductory physics, *Cognit. Instr.* **12**, 151 (1994).