Developing and assessing tutorials for quantum mechanics:
Time dependence and measurements

Paul J. Emigh,[1] Gina Passante,[2] and Peter S. Shaffer[3]

[1]*Department of Physics, Oregon State University, Corvallis, Oregon 97330, USA*
[2]*Department of Physics, California State University Fullerton, Fullerton, California 92831, USA*
[3]*Department of Physics, University of Washington, Seattle, Washington 98195, USA*

Tutorials in Physics: Quantum Mechanics is a set of curricular materials for supplementing upper-division quantum mechanics instruction, developed by the physics education group at the University of Washington. We describe the development of a sequence of three tutorials that focused on quantum measurements and time dependence. This sequence of tutorials is based on prominent findings in the research literature related to student understanding, as well as classroom techniques and strategies that have proved effective in other instructional contexts in physics. We also present data from a set of four exam questions used to evaluate how effectively these tutorials improve student ability to use correct mathematical and conceptual reasoning to predict the time evolution of quantum probabilities. We find that student performance on each question is significantly better than on corresponding questions given after only traditional instruction, with effect sizes that range from small to moderate. This improvement, along with a decrease in the prevalence of common incorrect lines of reasoning, supports our claim that this set of three tutorials from *Tutorials in Physics: Quantum Mechanics* assists advanced undergraduate students' in developing a conceptual understanding of some of the fundamental quantum concepts about time evolution and measurements when given after traditional lecture instruction. However, our results also show that there are some topics with which many students still struggle and, thus, that there remains room for further research and development of materials for these topics.

I. INTRODUCTION

A growing body of research on the teaching and learning of quantum mechanics has shown that many physics students have serious difficulty with a broad range of key quantum concepts, even after instruction [1–8]. Two fundamental and interrelated topics with which students struggle are measurements and time dependence [9–15]. These ideas underlie the quantum model used to describe and investigate phenomena for which classical physics is insufficient. It is thus crucial that students be able to apply them in theoretical and lab settings. The research has shown that the most prevalent difficulties with these two concepts often transcend context [15,16] and persist across multiple years of instruction, even extending to students in graduate programs [15,17–19].

These findings suggest that lecture instruction in quantum mechanics, especially at the undergraduate level, fails to help students develop a complete understanding of time dependence and measurements. Similar observations and conclusions have been made regarding the effectiveness of lecture instruction for introductory physics [20]. At the introductory level, techniques based on active engagement, such as clicker questions, lecture demonstrations, flipped classrooms, and tutorials, have been effective at improving student understanding [21–23]. Such methods are now also being used in upper-division courses, including quantum mechanics, where they are also being proved effective [7,15,24–26].

The Physics Education Group at the University of Washington (UW) is developing a set of research-based and research-validated instructional materials for supplementing the lectures and textbooks used to teach undergraduate quantum mechanics [27]. This curriculum, *Tutorials in Physics: Quantum Mechanics*, is similar to *Tutorials in Introductory Physics* [23]. The entirety of a typical upper-division quantum course is covered.

This article discusses how an interrelated sequence of three tutorials on time evolution and quantum measurements was developed and assessed. We have chosen these topics for three reasons: (i) the concepts involved are particularly challenging for students, (ii) they are fundamental to quantum mechanics, and (iii) they have been the

focus of a substantial body of research on student under-standing. This particular sequence has been in development at UW over a period of approximately fifteen years.

Each section in this article describes part of the process of developing and assessing *Tutorials in Physics: Quantum Mechanics*. Section II presents the background of the courses and methods involved in the project. Section III gives the theoretical perspective on research-based curriculum development. Section IV presents the relevant literature on student understanding of time evolution and measurements. Section V describes the tutorial curriculum and how its development was informed by prior research. Section VI then discusses results from our assessment of the tutorials. In Sec. VII we reflect on the development of the tutorials, their effectiveness, and directions for future research.

II. BACKGROUND AND METHODOLOGY

The curriculum development and assessment described in this article was primarily conducted in the core quantum mechanics course at UW. The course consists of two ten-week quarters that cover the entirety of Griffiths' widely used quantum mechanics textbook [28]. Almost all physics majors at UW are required to complete the first quarter (usually in the junior year), which introduces time dependence and measurement theory. The typical enrollment was 70–90 students during this study.

The course meets for 3 h of traditional lecture each week, conducted by a senior faculty member. Students also attend a weekly 50 min small-group section (15–30 students) led by a graduate student teaching assistant (TA). *Tutorials in Physics: Quantum Mechanics* is used in these small-group sections. The tutorials are designed to be flexible, so that they are easy to adopt both for courses with large enroll-ment (as at UW) and with fewer students. There are more than twenty tutorials on topics spanning the content of upper-division quantum mechanics.

The tutorials consist of structured worksheets that help students develop an understanding of the most challenging and fundamental concepts in quantum mechanics and recognize the underlying quantum model. They often make use of well-tested instructional strategies, such as elicit-confront-resolve [29]. Students work in small groups of 3–5 to complete the worksheets. The groups are guided by TAs trained not to answer student questions directly, but instead to ask guiding and probing questions that help students build their own knowledge.

Shown in Fig. 1 is the instructional sequence associated with each weekly tutorial. Before each in-class worksheet
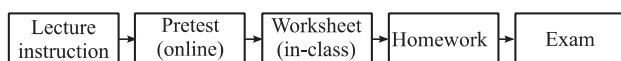

FIG. 1. Timeline of instruction associated with each tutorial topic.

(but after lecture instruction on the same topic) students are given a tutorial "pretest" that has multiple purposes. First, it prompts students to think about relevant ideas prior to the tutorial. Second, instructors and TAs can examine pretest responses to identify student ideas about key concepts for themselves. Third, the results provide data for research-ers investigating student thinking and reasoning. After each tutorial, students complete a hand-written homework assignment that has a structure similar to the in-class worksheets and that helps students extend their under-standing beyond the context used in class. Students are given formative feedback from TAs within one week of submitting the homework.

Each component of a tutorial contributes to students' final grades. Students are given credit for completing the pretests (regardless of whether or not their answers are correct) and for participating in the small-group sections. The tutorial homework is graded, and forms the bulk of the tutorial score. In addition, course exams include questions based on tutorial content, which signals to students that the tutorials are a valuable and integrated part of the course. The exams also provide researchers with a means for assessing the tutorials.

III. PERSPECTIVE ON RESEARCH-BASED CURRICULUM DEVELOPMENT

In this article, we focus on data drawn from two parts of the tutorial curriculum: (i) the tutorial pretests and (ii) the tutorial exam questions. As discussed above, the pretests are given after lecture instruction on a particular topic, and before tutorial instruction. They are given as online surveys, and students are asked to explain their reasoning to each question. The questions are typically focused on a single concept and a correct answer requires a relatively straightforward qualitative application of the relevant concept(s). Quantitative answers are usually not required, although they can be and often are still productive. The exam questions serve as post-tests for the tutorials. The post-test questions retain a qualitative focus, but tend to be more challenging than the pretests.

Development of the tutorial curriculum begins with research on student thinking. To this end, questions are developed that can be answered qualitatively or semi-qualitatively and that ask for explanations of reasoning. Written responses to pre- and post-test questions are then analyzed to identify the underlying line(s) of reasoning used by each student. The answer and explanation are considered together when analyzing each response. Prior articles have focused on this component of the tutorial development process, in which we have identified common incorrect lines of reasoning involving time dependence [13] and energy measurements [14] in different contexts. In this article, we describe the development process after research on student ideas has been conducted. We also assess the impact of the tutorials on student understanding. Thus, we

focus on comparing student responses to pre- and post-test results for questions that test the same underlying concept (although the questions may not be identical). Since the pretests are administered *after* lecture instruction, changes in student performance may be attributed to the impact of the relevant tutorial(s).

We have used the framework for curriculum development that is characteristic of the physics education group at UW [30]. From prior research on student thinking, one can identify patterns in student responses that suggest ways in which students are thinking about the concept. They may be loosely held or more stable. Patterns that are loosely held tend to be sensitive to the framing of the question or to the underlying physical context. Patterns of responses that are more stable tend to be independent of question wording or context. The current article does not detail how these patterns of reasoning are identified, but rather draws on prior literature, which is summarized in Sec. IV, to describe how we developed curriculum to address the results of this literature.

There are two main theoretical frameworks that we have drawn on to interpret the patterns found in student responses and that underlie the development of the tutorials: "conceptual change" [29] and "knowledge in pieces" [31,32]. We have found the idea of conceptual change best describes the patterns of incorrect ideas that tend to be independent of wording and context, which we interpret to be strongly held. For example, many students will apply the idea that "current is used up" to a wide variety of physics problems, and that they continue to apply this idea even after targeted instruction [33,34]. In contrast, we have found that using the lens of knowledge in pieces better describes the patterns in reasoning that fluctuate with different contexts. For example, confusion between different quantities—such as current, voltage, and resistance—has been found in many different contexts, indicating that students apply different components of knowledge in different ways (sometimes incorrectly), as opposed to holding a single incorrect idea [33,34]. These two viewpoints have led us to use two different techniques to address both stable student ideas and loosely held student ideas.

Patterns of reasoning that are interpreted as strongly held are ones that may benefit from an elicit-confront-resolve approach, which has proved highly effective in other instructional contexts [29]. The goal of this strategy is to help students recognize common errors and incorrect lines of reasoning and then reconcile them with the physics formalism they have been taught. This method will only work if the same incorrect responses are elicited to many similar questions. When using this technique, we often start by asking students to make qualitative predictions. Then, students are guided to recognize which concepts and lines of reasoning can be used to make accurate predictions. Lastly, students revisit their predictions to check for consistency; often, we ask them to consider other incorrect predictions that have been identified in the research literature.

The second technique is a model-building approach in which students develop or identify fundamental connections between concepts that are then implemented to predict quantum behavior. This strategy is used when student responses have been found to vary with context and it is not possible to identify a single underlying idea at the root of the incorrect answers. Instead, the tutorials reinforce the concepts by asking students to reflect on how they apply in new situations and as new elements of the model are introduced. Students encounter contexts that increase in complexity with each subsequent tutorial and homework. Throughout the sequence, students are asked to revisit and build upon their previous answers, both within a given tutorial and across multiple tutorials.

We have found that these two techniques are effective when employed in tandem to address a variety of incorrect responses of both types. The set of tutorials on time dependence and measurements are an excellent example because the tutorials build on each other and provide a variety of contexts for students to engage in model building. Furthermore, we have identified key areas in the tutorials where an elicit-confront-resolve approach has been very effective at reducing the number of strongly held ideas identified in the post-tests.

The tutorials and assessments are modified over a period of many years using an iterative approach of development, implementation, and testing. The materials are only considered effective if students are able to apply concepts and reasoning in new contexts (i.e., contexts other than those directly used in the curriculum). Even when particular materials appear effective, we often continue to investigate new and different methods to improve student learning further.

IV. RESEARCH BACKGROUND

There has been research by both the authors and other researchers on student understanding of time dependence and measurements. We do not provide a comprehensive review of the literature, but instead focus on results most relevant to the tutorial curriculum, which are grouped into six categories.

(C1) **Time independence of probability densities**: Students often incorrectly claim that probability densities do not depend on time, regardless of the initial wave function [7–9,13,15–17]. Students use a variety of different lines of reasoning in support of this idea. Many (10%–30%, depending on the particular question) treat the time dependence of any wave function as being due to a *single* time-dependent phase [9,13,15], which vanishes when the probability density is computed— reasoning that is correct only for energy eigenfunctions. Some of these students explicitly write

a generic wave function as $\Psi(x,t) = \psi(x)e^{-iEt/\hbar}$. However, even those who write a wave function correctly as $\Psi(x,t) = \sum \psi_n(x)e^{-iE_n t/\hbar}$ may claim that all time dependence vanishes when $|\Psi(x,t)|^2$ is calculated [13,35–38]. An alternate line of reasoning given by students is that the probability density is time independent because the *wave function* itself is always time independent [8,13,16].

(C2) **Long time evolution**: Students often incorrectly describe the behavior of a state a *long time* after the initial state is given [7,14,16]. One pervasive line of reasoning consistently given by 5%–10% of students is that the state evolves towards or decays into the system's ground state. Another is that the wave function spreads out over time toward an even distribution over all possible values [7,15]. Each of these ideas is relevant for some physical systems, but not universally applicable across all of quantum mechanics.

(C3) **Probabilities of energy measurements**: Students often fail to use the wave function and its evolution in time to determine energy probabilities and their evolution in time [7–16]. Some students (about 10% as reported in Ref. [13]) claim that energy probabilities depend on time whenever the probability density depends on time (i.e., for superposition states). Often, these students do not seem to distinguish between the methods used to determine the probabilities of these different quantities. Other students instead state that energy probabilities are time independent (a correct answer) based on the incorrect idea that the wave function itself is independent of time [14].

(C4) **The state after measurement**: Anywhere from 10%–40% of students give answers indicating that the state of a system after an energy measurement is the same as the state before the measurement [7,9,13–15]. Some instead claim that the system *returns* to the initial state after some amount of time has passed. It is also not uncommon for students to use reasoning based on either the "decay" or the "spreading out" of the wave function (similar to what was described above), especially a long time after a measurement is made [7].

(C5) **Measurements of position**: Some student ideas about measurements of position are similar to student ideas about measurements of energy [6–12,15]. Many believe that position measurements do not affect the state of the system. Some recognize the effect of a position measurement on the wave function, but answer that other aspects of the system, such as energy, are not changed by the measurement [15]. As many as 45% of students state that position remains fixed after a position measurement has been made. Other students claim

that the position will change classically after a position measurement. Lastly, some of the ideas discussed above—that the wave function will "decay" or "spread out" after a long time, or that the state will return to the state prior to the measurement—have also been documented [15].

(C6) **Measurements of momentum**: Some student ideas about momentum measurements appear to be similar to those that are common for position measurements [6,12], but little research on student understanding of momentum has been published.

V. DEVELOPMENT OF TUTORIAL CURRICULUM

The research summarized in Sec. IV indicates that students have a variety of ideas about quantum measurements and time dependence, many of which are incorrect or are often applied incorrectly or inappropriately. In this section, we describe a sequence of tutorials [39] that attempts to address the results of this research. The sequence consists of three tutorials: (i) *Time dependence in quantum mechanics,* (ii) *Energy measurements,* and (iii) *Position, momentum, and energy measurements*. The sequence focuses on fundamental topics that serve as a foundation for much of quantum mechanics, and thus for later tutorials. It is intended to be given after instruction on the time-dependent Schrödinger equation, on basic potentials such as the infinite square well, and on the quantum effects of ideal measurements. This sequence does not use Dirac notation, so it is suitable for a wide variety of courses regardless of structure, textbook, or the order of topics presented.

In the following sections we describe each of the three tutorials in the sequence individually. We first provide an overview of the structure and learning goals for that tutorial before describing the tutorial exercises. Overall, the sequence guides students to develop a strong conceptual model for the fundamental time dependence of wave functions. In this model, a quantum state is written in terms of energy eigenstates, and the time evolution of each energy eigenstate is used to determine the time evolution of the overall state. At the end of the sequence, students should be able to use this model to describe the time evolution of probabilities and probability densities, and should be ready to describe the time evolution of more advanced contexts encountered during later portions of the course.

A. Time dependence in quantum mechanics

The first tutorial in the sequence, Time dependence in quantum mechanics, focuses on helping students develop a functional understanding of the time evolution of wave functions and their corresponding probability densities. Students consider both (i) energy eigenstates and (ii) a

superposition of two energy eigenstates, and then reflect on the differences between these two cases. We expect students to be able to relate the time dependence of the probability density to the time dependence of the wave function. In particular, the tutorial is structured so that students observe a probability density that *depends on time* and construct reasons for *why* it depends on time for themselves. Specific learning goals are that students should be able to

   (1) State whether or not the probability density for a wave function changes in time.

   (2) Reason qualitatively about the changing shapes of wave functions and probability densities.

These learning goals were developed in response to the research on student understanding outlined in C1 and C2.

The tutorial begins by asking students about the time evolution of energy eigenstates (i.e., stationary states). Such a state evolves in time with the phase factor, $e^{-iE_n t/\hbar}$ where E_n is the energy eigenvalue [40]. We give students a hands-on tool for visually representing wave functions to help facilitate the development of a robust understanding of this time dependence. The tool consists of a set of four transparencies showing (i) the real axis, (ii) the imaginary axis, (iii) the ground state of the infinite square well, and (iv) the first excited state of the infinite square well. The transparencies can be combined to form a three-dimensional coordinate system, which students use to represent the time evolution of an energy eigenfunction as rotation in the complex plane. In-class observations suggest that students benefit from the hands-on nature of the activity, from observing the motion of wave functions in three-dimensional space (rather than on the two-dimensional screen of a simulation), and from the opportunity to construct the representation on their own. Researchers have found similar benefits from other tangible representations [41]. This activity also helps students recognize the importance of imaginary numbers to quantum time dependence. The other tutorials in the sequence refer back to this representation frequently.

Students are directed to demonstrate the time evolution of each energy eigenfunction using the transparencies. We explicitly ask them to describe how the frequencies of the two states *differ*. Then, students are guided to recognize that the probability density for each energy eigenfunction is time independent. The tutorial asks students to reflect on how the probability density can be independent of time when the wave function *does* depend on time, and to identify the unique conditions that lead to this case.

The second part of the tutorial directly addresses time independence of probabilities (C1). Students are shown a wave function that represents a *superposition* of energy eigenfunctions. First, they are asked to predict whether or not this wave function depends on time, and whether or not its associated probability density depends on time. Then, students sketch the probability density at different instants in time by identifying the phase of the contributing eigenfunctions at each instant. Lastly, students compare their sketches to a handout showing the time evolution as a sequence of graphs and summarize their results. Instructors typically ask students to extend their reasoning from finite times to consider the long-time limit in order to address the findings in C2.

B. Energy measurements

The second tutorial in the sequence, Energy measurements, focuses on the outcomes and probabilities of energy measurements (especially their time dependence) and on how measuring energy affects the state of a system. Naturally, this tutorial revisits and extends the model that students began to develop in the prior tutorial. After this tutorial, students should be able to

   (1) Use the initial wave function to show that energy probabilities do not depend on time, for any state [42].

   (2) Write the state *immediately* after an energy measurement.

   (3) Find and reason about probabilities at any time after an energy measurement.

These learning goals are in response to the research outlined in C1, C3, and C4.

The tutorial starts by giving a superposition of two energy eigenfunctions of the harmonic oscillator. Students are asked to write an expression for the time dependence of this wave function, which relies on skills practiced in the first tutorial. We then ask students to use their expression to determine whether or not energy probabilities depend on time. We expect students to perform an inner product and use orthonormality, leaving a single nonzero term that evolves with only one phase factor, $e^{-iE_n t/\hbar}$, which vanishes when the probability is found by taking the modulus square (C3). Students are explicitly asked to consider whether or not their answer changes as t goes to infinity (C2). Lastly, students are asked to contrast their answer with the nontrivial time evolution of the state's probability density.

Next, students are told that an energy measurement results in the value E_1 (the first excited state energy) and asked to determine the new state of the system (C4). They consider the outcome of a second energy measurement made either a short or a long time later. Students consider three incorrect statements about what happens to the state after a long time has elapsed, each corresponding to one of the three common incorrect lines of reasoning (return to the initial state, decay to the ground state, and spread over all possible values). They are asked to identify why each statement is inconsistent with the quantum model for time dependence (C2).

A final exercise helps strengthen the model introduced in the first tutorial. Students consider the wave function corresponding to the ground state of the infinite square well potential as a possible wave function to describe a system with the *harmonic oscillator* potential. This wave function

is not easily (or obviously) represented as a superposition of energy eigenfunctions, and students often believe it can be treated as an energy eigenfunction, and that it will evolve in time with a single time-dependent phase factor. The tutorial guides students to recognize that the wave function must be represented as a (nontrivial) superposition of the appropriate energy eigenstates for the system (in this case, the harmonic oscillator) and that each evolves at a different frequency. Students then observe that the probability density for such a superposition depends on time, and reconcile this fact with their prediction (C1 and C3).

C. Position, momentum, and energy measurements

The final tutorial in the sequence is Position, momentum, and energy measurements. Its primary goals are to explore the effect of *position* measurements on quantum systems and to synthesize the previous tutorials. The tutorial achieves these goals by returning to a simple system—the infinite square well—and using elicit-confront-resolve to help students consider and overcome common incorrect lines of reasoning. After the entire sequence, we expect students to build a solid conceptual model for determining the time dependence of any quantity for any state, including those that result from prior measurements. Learning goals specific to this tutorial are that students should be able to
  (1) Write the state *immediately* after a position measurement.
  (2) Find and reason about probabilities at any time after a position measurement.
These learning goals are in response to the research articulated in C5. Additionally, this tutorial supports the model-building effort begun in the previous two tutorials in this sequence.

The tutorial uses elicit-confront-resolve to address student ideas [29]. Students first predict the possible outcomes of a position measurement made on a general quantum system. Then, they predict the possible outcomes of a second position measurement made on a system *immediately after* the first measurement. Last, they predict how (if at all) their answers would differ if instead the second measurement were made after some amount of time has elapsed. They are asked to reflect on three different possible lines of reasoning, but they are not told which is correct.

After students have made predictions and discussed multiple lines of reasoning, they consider a particle initially described by a specific energy eigenstate. They are asked about the time evolution of energy and position probabilities for this state. This exercise is intended to help students recall important elements of the model and the reasoning skills they have developed over the prior two tutorials. Then, students are told the outcome of a position measurement on the state. They sketch the wave function, determine whether or not it can be represented as a superposition of energy eigenfunctions, and use their answers to describe how the probability density evolves in time (C5). Students check

their work and resolve any inconsistencies with their initial prediction by checking with TAs, who may show students a handout or a computer simulation of time evolution after a measurement of position. Finally, students fill out tables to summarize, compare, and contrast their results for the different cases they have considered (C1, C3, and C5).

Preliminary versions of this tutorial also asked students to consider measurements of momentum. However, we found that a majority of groups were unable to finish those versions in the allotted class time. Thus, we shifted the focus of the in-class worksheet away from momentum, and instead structured the homework to incorporate measurements of momentum (and other quantities) into students' model. The homework exercises also guide students to connect the mathematics of quantum operators (i.e., commutation) with the time evolution of probabilities. As discussed above, further research on student reasoning about momentum might prove useful in future development of instructional material.

VI. ASSESSMENT OF TUTORIAL CURRICULUM

In this section, we present some of the results from our efforts to evaluate the impact of the tutorials on student performance. Although the tutorials are based on a substantial body of research literature, we consider them validated only when students demonstrate a more robust understanding of the relevant physical concepts after completing the tutorials than beforehand. Often, assessments of early curriculum reveal possibilities for modification and improvement, and many cycles of development, implementation, and assessment are necessary.

There are several examples of research-based curriculum aimed at improving student understanding of quantum mechanics, especially for the topics of time dependence and measurements [24,43–50]. However, very few articles have discussed the *effectiveness* of such curricula, although preliminary efforts suggest that they can prove effective at the upper-division level [24,49,50].

Our primary assessment method is to compare pre- and post-test results. As discussed in Sec. II, the pretests are given online before each tutorial (but after lecture instruction) and the post-tests are given as part of course exams. The exam questions were written by the authors in consultation with the course instructor. All the questions are qualitative, so calculations are not required to answer correctly. The learning objectives for the tutorials are for students to be able to use fundamental physics concepts to reason correctly about the behavior of quantum systems. Since prior research has shown that such reasoning can be challenging even for graduate students [7,13–15,17], this objective is not necessarily easy for all students to achieve. We therefore consider the tutorials effective if more students are able to demonstrate productive conceptual

reasoning in exam settings (after the tutorials) than after lecture instruction alone.

During this investigation, we have used many different questions as pre- and post-tutorial assessments. Where appropriate, we aggregate pretest results from multiple courses when the same (or nearly the same) questions were administered at the same point in instruction. Typically, the pretests are relatively static, as they are designed not only for research but also as warm-up exercises for the tutorial. It has been our finding that the pretest responses tend not to vary substantially from year to year or from instructor to instructor [51].

In general, we aim for the post-tests to be at least as difficult as and typically more difficult than the corresponding pretests, but to require the same underlying reasoning. However, there are greater constraints on the post-tests than on the pretests because they are part of the course exams. It has therefore been necessary to modify the post-tests substantially from year to year. It has also not been possible to assess the entire sequence of tutorials on a single exam during each year the tutorials were used, due to the timing of the exams. In this article, we have therefore chosen to present results from only a single exam, chosen because (i) it assesses a broad set of our learning goals, (ii) it was given after the entire sequence of tutorials, and (iii) it is the best match for our pretest questions. However, as with the pretests, we have tended to obtain similar results on alternative exam questions. (A description of our results for other exam questions given at different instructional points can be found in Chap. 5 of Ref. [15].)

We recognize that student performance on written questions is not a perfect measure of student understanding. However, we believe that improved performance on the conceptual questions in this section does imply improved understanding of the material. We do not look only at the percentage of students giving the correct answer, but we also determine the percentage of the students giving completely correct reasoning, as well as the percentage giving common incorrect (or incomplete) answers. We compare the pre- and post-test performance

by this measure and report a p value (using a chi-squared test) and an effect size (using Cramer's V) for each question [52,53]. We follow up on this result by discussing common incorrect lines of reasoning on both the pre- and post-tutorial questions so that we can better understand the impact of the tutorials on the lines of reasoning favored by students.

Our presentation of the results is divided into subsections for each of three concepts: (i) the time dependence of position probabilities, (ii) the time dependence of energy probabilities, and (iii) the effects of quantum measurements. In each subsection, we first describe the relevant questions and the criteria we used to analyze student responses. Then, we provide results and analysis for students in the core junior-level quantum mechanics course at UW.

A. Time dependence of position probabilities

The questions shown in Fig. 2 focus on the time evolution of position probabilities (i.e., of the position-space probability density). The pretest was given just before the tutorial Time dependence in quantum mechanics. The post-test was given on a midterm after the full sequence of tutorials. Each question shows students an initial wave function for a particle in the infinite square well potential. On the pretest, the wave function was expressed as a superposition of two energy eigenfunctions, and was given graphically (Ref. [13] discusses this pretest in detail). On the post-test, the given wave function is not explicitly written in terms of energy eigenfunctions. The post-test assesses similar ideas but requires more steps because the wave function must be rewritten in terms of energy eigenstates to find the time dependence. In each case, students are asked whether or not the probability of finding the particle in a given region depends on time, and to explain their answer.
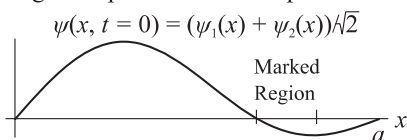
The correct answer to each question is the same: the desired position probability depends on the time at which the measurement is made. A response was considered to have correct reasoning if the answer was supported by

Question 1

A system consists of a particle in an infinite square well of width a.

Pretest 1	**Post-test 1**
The wave function at $t = 0$ is shown below. It is entirely real. Does the probability of finding the particle within the marked region depend on time? Explain.	At $t = 0$, the normalized wave function is given below. Does the probability that the position is given by $x < a/3$ depend on the time at which the measurement is made? Explain.

$$\psi(x, t = 0) = (\psi_1(x) + \psi_2(x))/\sqrt{2}$$

Marked Region

$$\psi(x, t = 0) = \begin{cases} N(x^2 - ax) & 0 < x < a \\ 0 & \text{elsewhere} \end{cases}$$

FIG. 2. A pair of questions about the time dependence of position probabilities. The pretest was given after lecture instruction on time dependence but before the tutorials. The post-test was given after the full sequence of three tutorials focusing on quantum measurements and time dependence.
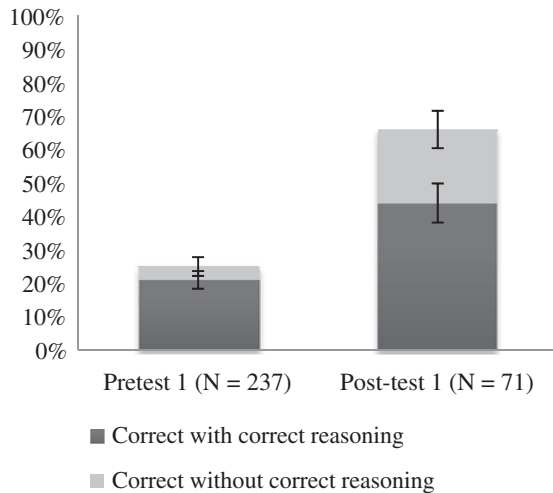
FIG. 3.   Student performance on question 1, before and after completion of the tutorials. Error bars represent the standard error.

either a conceptual or a mathematical explanation. For example, one possible line of reasoning is that the given wave function is a superposition of two or more energy eigenfunctions, each evolving in time at a different (energy-dependent) rate. When computing the probability density (the modulus square of the wave function), there is interference between the time-dependent phase factors. This time dependence can also be identified by direct computation of the probability density.

Figure 3 shows the results for question 1. The post-test results are from a single academic quarter; the results from similar questions given in other quarters have tended to be similar. The pretest results are aggregated across six quarters and across several variants of the question shown in Fig. 2. The pretest performance did not vary considerably from quarter to quarter [13].

Performance on the pretest is poor: only 25% of the students ($N = 237$) answer correctly, with 21% giving correct reasoning. There is substantial improvement on the post-test ($N = 71$), for which 66% answer correctly and 44% give correct reasoning. The difference between the percentages who answer correctly with correct reasoning is

statistically significant ($p < 0.001$) with a moderate effect size (Cramer's $V = 0.55$).

The most common errors on the pretest are those described in Sec. IV (C1). In particular, 10%–20% of students at UW base their pretest answer on the incorrect assertion that the wave function evolves in time via only a single overall phase. After completion of the tutorial sequence, which specifically targets common incorrect student lines of reasoning, only 4% of the students give such answers. This suggests that the tutorial helps reduce the prevalence of this common and persistent error identified by the research literature.

B. Time dependence of energy probabilities

Figure 4 shows a second pair of pre- and post-test questions that probe student understanding of the time dependence of energy probabilities. As with question 1, the pretest was given after lecture instruction and before the three-tutorial sequence. Post-test 2 was part of the same exam as post-test 1. On the pre- and post-test, students are given the same initial wave function for the infinite square well. They are asked to determine whether or not the probability of E_1, the ground state energy, depends on time.

The correct answer is that the probability of measuring E_1 (or any other energy eigenvalue) is independent of time. This probability is given by $|\langle E_1|\psi(t)\rangle|^2$. Since the energy eigenstates are mutually orthogonal, $\langle E_1|\psi(t)\rangle$ has only one nonzero term that evolves in time with a single phase, which cancels when the modulus square is calculated. Student explanations were only considered correct if they explained why all phases vanish in this case.

The results from question 2 are shown in Fig. 5. The pre- and post-test were each given in a single (different) junior-level course at UW. The percentage who answered question 2 correctly is high: 72% before the tutorials ($N = 74$) and 84% after the tutorials ($N = 71$). However, correct reasoning was given by only 11% on the pretest and 32% on the post-test, a difference that is statistically significant ($p < 0.005$) with a small effect size (Cramer's $V = 0.26$).

A substantial fraction of the students who stated that the energy probability is independent of time (a correct answer) did not give correct reasoning. This was true both before

Question 2

A system consists of a particle in an infinite square well of width a. At $t = 0$, the normalized wave function is:

$$\psi(x, t = 0) = \begin{cases} N(x^2 - ax) & 0 < x < a \\ 0 & \text{elsewhere} \end{cases}$$

Pretest 2	Post-test 2
Let P_1 represent the probability that a measurement of this system's energy results in E_1, the ground state energy. Does P_1 change with time? Explain your reasoning.	Does the probability that the energy is equal to E_1 depend on the time at which the measurement is made? Explain.

FIG. 4.   A pair of questions about the time dependence of energy probabilities. The pretest was given after lecture instruction on time dependence but before the tutorials. The post-test was given after the full sequence of tutorials, as part of the same exam as post-test question 1.
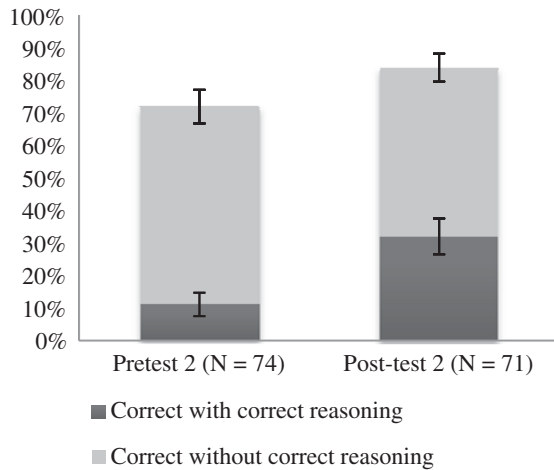
FIG. 5. Student performance on question 2, before and after completion of the tutorials. Error bars represent the standard error.

and after the tutorials. It was very common for students who gave this answer on the pretest to assert that the complex phase(s) would vanish due to the modulus square. While some students explicitly described the given wave function incorrectly as a stationary state and/or as being associated with a single phase, many did not clearly articulate why the phases vanish or even whether the wave function has one phase or many. On the pretest, 27% of the students gave some explanation of this type; on the post-test, only 8% did so. On the exam, it was instead common (27%) for students to claim that energy probabilities are always time independent—accurate for any system with a time-independent Hamiltonian—which we did not consider to be a correct explanation because it is not a complete line of reasoning. This response was not common (4%) on the pretest.

Among students who answered incorrectly, the most common explanation was that the probability of an energy measurement depends on time because the wave function is a superposition of energy eigenstates. Often, these students explicitly noted that such a state is not "stationary," and thus the probability depends on time. The prevalence of this response was essentially the same on the post-test (11%) as it was on the pretest (9%).

Overall, the results for the time dependence of energy measurements are promising but underwhelming. Although more students give correct reasoning and fewer students give incorrect reasoning, a large fraction are not yet able to articulate a fully correct explanation for why energy probabilities do not depend on time, even after the complete sequence of tutorials. A focus on promoting students' reasoning ability may lead to further improvement of the tutorials.

C. The effects of quantum measurements

The set of questions shown in Fig. 6 was designed to assess the impact of the last tutorial in the sequence, Position, energy, and momentum measurements. The focus of this tutorial is on how measurements (in particular, position measurements) affect quantum systems. The pre- and post-tests each ask about a quantum system after a measurement of position has been made, although they do so in different ways. On the pretest, students are told to consider a particle in an energy eigenstate of the infinite square well whose position is measured. (Note that pretests 3a and 3b use different initial states and outcomes of the first measurement.) Students are then given three possible outcomes of a subsequent measurement (energies for pretest 3a and positions for pretest 3b) and asked to identify whether each outcome is *definite, possible but not certain*,

---

Question 3

Pretest 3a

A particle is in the square well from $x = 0$ to $x = a$. At $t = 0$, the energy is measured to be E_2. At $t_1 > 0$ you measure position to be $x \approx a/3$. A short time later at t_2, you measure the energy (assume no other measurements are made between t_1 and t_2). Which of the statements best describes the probability that the measurement will result in each of the following values: E_1, E_2, and E_3? Explain.

  i)   I will definitely measure this value
  ii) It is possible (but not certain) that I will measure this value
  iii) I will definitely not measure this value
  iv) I do not have enough information to tell

Pretest 3b

A particle is in the square well from $x = 0$ to $x = a$. At $t = 0$, the energy is measured to be E_1. At $t_1 > 0$ you measure position to be $x \approx 2a/3$. A short time later at t_2, you measure position again (assume no other measurements are made between t_1 and t_2). Which statement best describes the probability that the result will be each of the following values: $a/3$, $a/2$, and $2a/3$? Explain.

Post-test 3a

A system consists of a particle in an infinite square well of width a. The wave function at $t = 0$ is:

$$\psi(x, \ t = 0) = \begin{cases} N(x^2 - ax) & 0 < x < a \\ 0 & \text{elsewhere} \end{cases}$$

Now suppose that the position is measured to be at $x = a/3$ at time $t_1 > 0$. Suppose you measure the energy at $t_2 > t_1$. Does the probability that the energy is E_1 depend on the time at which the measurement is made? Explain.

Post-test 3b

Suppose instead you measure the position at $t_2 > t_1$ (assume no measurements have been made between t_1 and t_2). Does the probability that the position is given by $x < a/3$ depend on the time at which the measurement is made? Explain.

FIG. 6. Several questions about the effects of measuring position. The pretest was given prior to the final tutorial in the sequence, but after lecture instruction and the first two tutorials. The post-test was given after completion of the full sequence, as part of the same exam as post-tests 1 and 2. The text in the figure has been edited for length.

or *not possible*. Each question asks students to explain their reasoning.

It should be noted that the language associated with quantum mechanics, and especially the language related to measurement, is often easy to misinterpret. To ensure that our results are not arising solely from such misinterpretations, we have varied the language used in pretest 3 from quarter to quarter. For example, we have asked students to consider t_2 as occurring "a few minutes later" instead of "a short time later." In pretest 3b, we have specifically described the measurement either as being "at" or as being "near" $x = 2a/3$. Neither change resulted in substantial variation in student responses, and in general students did not appear to be misinterpreting this language in a way that impacted our assessment of student ideas.

On the post-test, students are given an initial wave function (the same as in post-tests 1 and 2) and then are told the outcome of a position measurement. They are asked whether or not a subsequent measurement of position (post-test 3a) or of energy (post-test 3b) depends on time.

Each question can be answered by approximating the state after the measurement as a position eigenstate (i.e., a delta function) or as a narrow Gaussian. Such a function is a superposition of all energy eigenfunctions (except those equal to zero at the measurement position), and thus the outcome of a subsequent measurement can be any energy eigenvalue. Similarly, all values of position are possible outcomes, except where the wave function is zero at any instant in time.

Since the wave function cannot be represented as a single energy eigenfunction, the reasoning from questions 1 and 2 remains valid. Thus, position probabilities depend on time while energy probabilities do not. For students who use such lines of reasoning, the same analysis criteria were applied. Lines of reasoning based on the uncertainty principle or operator algebra were considered correct as long as they completely described the underlying physics. Since the post-test asks students to reason about time dependence, which requires them to identify that the state is now a superposition of energy eigenstates in order to be considered as correct reasoning, we consider it to be more difficult than the pretest.

The results are shown in Fig. 7. For the questions that ask about a subsequent energy measurement (3a), many students give the correct answer but most do not give correct reasoning. There is a small improvement from before to after the administration of the tutorial ($p < 0.001$, Cramer's $V = 0.18$). For the questions that ask about a subsequent position measurement (3b), the improvement appears to be larger ($p < 0.0001$, Cramer's $V = 0.42$), in part due to the fact that the pretest performance is relatively low [53].

A few incorrect lines of reasoning are very common in the responses to this pair of pretest questions, some of which have been discussed in prior publications [9,10,13–15]. For example, 35% of the students indicated
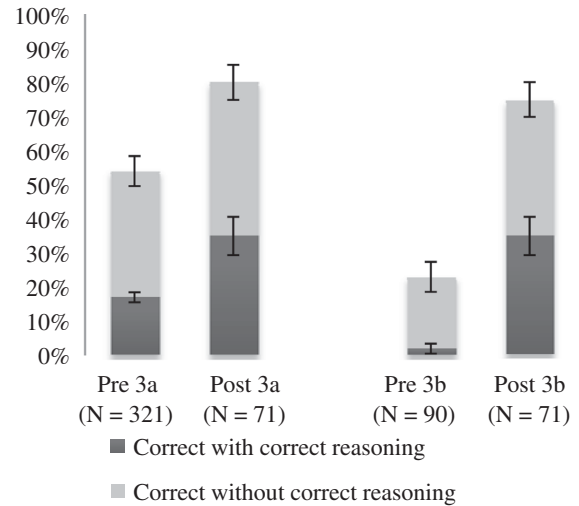

FIG. 7. Student performance on question 3, before and after completion of the tutorials. Error bars represent the standard error.

that an energy measurement (question 3a) would not affect the quantum system, and 10% gave a similar answer for question 3b. This line of reasoning was rare in response to both post-test questions. On pretest 3b, almost half of the students (44%) claimed that the outcome of the original position measurement was the *only* possible outcome of a subsequent position measurement. Many of these students explicitly indicated that the position of the particle is *fixed* following the position measurement (these students did not, for example, appear to be focusing on very short times after the initial measurement). On post-test 3b, 15% of the students gave this answer, a substantial improvement. Furthermore, many students gave incorrect explanations to this pair of questions similar to the incorrect explanations given in response to questions 1 and 2.

VII. DISCUSSION AND CONCLUSION

In this paper we have described the development of a set of three tutorials on time evolution and measurements in quantum systems. We have also discussed the results to several pre- and post-test questions that contribute to our assessment of the effectiveness of these materials. We found that student performance on our post-test questions improves after working through the tutorials, although there is still room for improvement.

The sequence of tutorials uses two main instructional approaches: elicit-confront-resolve and model building. The approach chosen for a particular concept depends on the nature of student ideas revealed by prior research. If student ideas are independent of context and question wording, elicit-confront-resolve was used, whereas if they are context dependent and sensitive to question wording, a model-building approach was chosen.

We presented three sets of pre- and post-test questions, each of which targets a specific concept addressed by the tutorials. On each set of questions, students' performance improved from pre- to post-test with small to moderate effect sizes. It is important to recognize that this article does not present all pre- and post-test questions analyzed as part of our investigation. Due to the nature of the assessments, it was not possible to repeat the same post-test question(s) over several years, and it would be too lengthy to describe each assessment in this paper. The questions presented above were selected because they directly assess several important concepts covered by the tutorials and because they were given after the full sequence of tutorials had been administered. Analysis of other questions we have asked can be found in Refs. [7,15], and the results tend to show similar improvement from pretest to post-test.

In addition to an increase in the number of students answering these questions with correct reasoning, we found that particular patterns of incorrect answers and/or explanations were substantially reduced from pretest to post-test. For example, in response to pretest 2, 27% of students explicitly wrote that the superposition state had a single time-dependent phase and that number decreased to 8% on the post-test. Additionally, some responses to the pre- and post-test questions suggest directions for further study. In particular, we notice that students' use of certain terminology is not always consistent with expert understanding of quantum mechanics. For example, we find that students struggle with the term "stationary state" and the meaning of the phrase "time dependence." We believe that examining how students interpret common language in more detail might help further the improvement of quantum instruction.

It is worthwhile to note that the tutorials do not address all common student ideas. Quantum measurements and time evolution are very challenging for students, and while the tutorials provide much-needed supplemental instruction, they are not intended as a standalone curriculum. We find that working through this sequence of tutorials prepares students to advance their understanding of these topics to new and more complex scenarios, such as systems with degenerate energies or systems with time-dependent Hamiltonians. In particular, the model-building framework of these tutorials lays the groundwork for students to reflect on their understanding and then apply it to more advanced contexts in a coherent manner.

While the results shown are very promising, it is important to note the limitations of this type of work. Because of the course constraints we were not able to ask matching pre- and post-test questions. Additionally, in this type of study we are unable to control for all student activities between the pre- and post-tests. We know that the students attended tutorial sessions, but we do not know how much they studied the relevant material for the exam. Despite these challenges, we feel confident that, based on both the improvement in the number of students that give
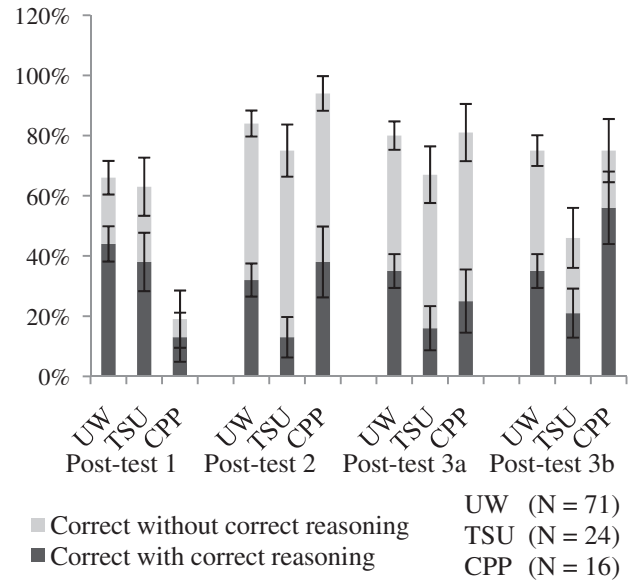

FIG. 8. Student performance on all four post-test questions at three different universities. Error bars represent the standard error.

correct answers with correct reasoning and a decrease in the number of students displaying well-researched incorrect lines of reasoning on the post-test, the tutorial curriculum does lead to improved understanding of quantum mechanical time dependence and measurements.

It is worthwhile to note that most of the students at UW have experience with tutorials from their introductory classes and very quickly embraced the advanced tutorial materials. There may be an increased time from introduction of the tutorial materials to student buy-in at other institutions. Furthermore, UW is a selective institution, and also an institution with relatively low diversity. To examine whether or not our successful implementation of the tutorials can be duplicated at other universities with different instructional environments, we have collected and analyzed data from two institutions with different student populations: Texas State University (TSU) and California Polytechnic University Pomona (CPP). TSU gave the entire sequence of tutorials in a junior-level quantum course similar to the one at UW. CPP did not give the first tutorial in the sequence (Time dependence in quantum mechanics) but did give preliminary versions of the remaining tutorials [54]. Both sites administered the set of four post-test questions described in Sec. VI after students completed the tutorials. Figure 8 shows the results alongside results at UW.

The number of students at each pilot site was much smaller than at UW (24 at TSU and 16 at CPP), and the pilot sites did not administer the pretest questions, so we cannot compare the pre- and post-test performance of these students. Past results suggest that pretest performance at other universities is unlikely to differ considerably from that at UW [13]. We note a general pattern: student performance on individual questions differs somewhat, but the overall performance appears similar. Chi-squared tests [52] on each

of the four questions resulted in p values between 0.04 and 0.62, with effect sizes (using Cramer's V) between 0.11 and 0.21: differences that are neither significant nor substantial. These results suggest that the tutorials may be implemented and contribute to student learning of quantum mechanics not only at UW, where they were developed and tested, but also at additional institutions that may have different student populations and instructional environments.

ACKNOWLEDGMENTS

[1] D. F. Styer, Common misconceptions regarding quantum mechanics, Am. J. Phys. **64,** 31 (1996).

[2] Research on teaching and learning quantum mechanics, *Proceedings of the Annual Meeting of National Association for Research in Science Teaching*, edited by D. Zollman (1999), https://web.phys.ksu.edu/papers/narst/QM_papers.pdf.

[3] S. Vokos, P. Shaffer, L. McDermott, and B. Ambrose, Student understanding of the wave nature of matter: Diffraction and interference of particles, Am. J. Phys. **68,** S42 (2000).

[4] L. Bao and E. F. Redish, Understanding probabilistic interpretations of physical systems: A prerequisite to learning quantum physics, Am. J. Phys. **70,** 210 (2002).

[5] J. Morgan and M. Wittmann, Examining the evolution of student ideas about quantum tunneling, AIP Conf. Proc. **818,** 73 (2006).

[6] H. Sadaghiani, Ph.D. thesis, Ohio State University, 2005, https://etd.ohiolink.edu/pg_10?0::NO:10:P10_ACCESSION_NUM:osu1123878116#abstract-files.

[7] A. Crouse, Ph.D. thesis, University of Washington, 2007, https://digital.lib.washington.edu/researchworks/handle/1773/9781.

[8] C. Singh, Student understanding of quantum mechanics, Am. J. Phys. **69,** 885 (2001).

[9] G. Zhu and C. Singh, Surveying students' understanding of quantum mechanics in one spatial dimension, Am. J. Phys. **80,** 252 (2012).

[10] G. Zhu and C. Singh, Improving students' understanding of quantum measurement I: Investigation of difficulties, Phys. Rev. ST Phys. Educ. Res. **8,** 010117 (2012).

[11] C. Singh, Transfer of learning in quantum mechanics, presented at the Physics Education Research Conference 2004, Sacramento, CA (2004).

[12] E. Cataloglu and R. W. Robinett, Testing the development of student conceptual and visualization understanding in quantum mechanics throughout the undergraduate career, Am. J. Phys. **70,** 238 (2002).

[13] P. J. Emigh, Gina Passante, and P. S. Shaffer, Student understanding of time dependence in quantum mechanics, Phys. Rev. ST Phys. Educ. Res. **11,** 020112 (2015).

[14] G. Passante, P. J. Emigh, and P. S. Shaffer, Examining student ideas about energy measurements on quantum states across the sophomore, junior, and graduate levels, Phys. Rev. ST Phys. Educ. Res. **11,** 020111 (2015).

[15] P. J. Emigh, Ph.D. thesis, University of Washington, 2015, https://digital.lib.washington.edu/researchworks/handle/1773/35301.

[16] G. Passante, P. J. Emigh, and P. S. Shaffer, Investigating student understanding of basic quantum mechanics in the context of time-dependent perturbation theory, in *Proceedings of the Physics Education Research Conference 2013, Portland, OR* (AIP, New York, 2013), pp. 269–272.

[17] C. Singh, Student understanding of quantum mechanics at the beginning of graduate instruction, Am. J. Phys. **76,** 277 (2008).

[18] L. D. Carr and S. B. McKagan, Graduate quantum mechanics reform, Am. J. Phys. **77,** 308 (2009).

[19] C. Porter, A. Bogdan, and A. Heckler, Student understanding of potential, wave functions and the Jacobian in hydrogen in graduate-level quantum mechanics, in *Proceedings of the Physics Education Research Conference 2016, Sacramento, CA* (AIP, New York, 2016), pp. 244–247.

[20] L. C. McDermott and E. F. Redish, Resource letter PER-1: Physics education research, Am. J. Phys. **67,** 755 (1999).

[21] 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).

[22] D. R. Sokoloff and R. K. Thornton, Using interactive lecture demonstrations to create an active learning environment, Phys. Teach. **35,** 340 (1997).

[23] L. C. McDermott and P. S. Shaffer (Physics Education Group), *Tutorials in Introductory Physics*, (Prentice Hall College Div., Englewood Cliffs, NJ, 2001).

[24] C. Singh, Interactive learning tutorials on quantum mechanics, Am. J. Phys. **76,** 400 (2008).

[25] E. Marshman, Ph.D. thesis, University of Pittsburgh, 2015, http://d-scholarship.pitt.edu/25547/.

[26] G. Passante, P. J. Emigh, and P. S. Shaffer, Testing tutorials in upper-division: An example from quantum mechanics,

Proceedings of the Physics Education Research Conference 2014, Minneapolis, MN (AIP, New York, 2014), pp. 199–202.

[27] L. C. McDermott, P. S. Shaffer, and P. R. L. Heron (Physics Education Group), Tutorials in Physics: Quantum Mechanics (to be published).

[28] D. J. Griffiths, *Introduction to Quantum Mechanics* (Pearson Prentice Hall, Englewood Cliffs, NJ, 2005).

[29] See, for example, G. J. Posner, K. A. Strike, P. W. Hewson, and W. A. Gertzog, Accommodation of scientific conception: Toward a theory of conceptual change, Sci. Educ. **66,** 211 (1982).

[30] P. R. L. Heron, Empirical investigations of student understanding, AIP Conf. Proc. **720,** 15 (2004).

[31] A. A. DiSessa, Knowledge in Pieces, *Constructivism in the Computer Age* (Lawrence Erlbaum Associates, Hillsdale NJ, 1988), pp. 49–70.

[32] A. A. DiSessa, Toward an epistemology of physics, Cognit. Instr. **10,** 105 (1993).

[33] L. C. McDermott and P. S. Shaffer, Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding, Am. J. Phys. **60,** 994 (1992).

[34] P. S. Shaffer and L. C. McDermott, Research as a guide for curriculum development: An example from introductory electricity. Part II: Design of instructional strategies, Am. J. Phys. **60,** 1003 (1992).

[35] Results of this type have spurred research on student understanding of complex numbers in quantum contexts, such as Refs. [34–36].

[36] E. Smith, J. Zwolak, and C. A. Manogue, Student difficulties with complex numbers, in *Proceedings of the Physics Education Research Conference 2015, College Park, MD* (AIP, New York, 2015), 311–314.

[37] T. Wan, P. J. Emigh, G. Passante, and P. S. Shaffer, Student understanding of period in introductory and quantum physics courses, in *Proceedings of the Physics Education Research Conference 2016, Sacramento, CA* (AIP, New York, 2016), pp. 380–383.

[38] T. Wan, P. J. Emigh, and P. S. Shaffer, Student understanding of the measurable effects of relative phases in superposition states, in *Proceedings of the Physics Education Research Conference 2017, Cincinnati, OH* (AIP, New York, 2017), pp. 420–423.

[39] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.14.020128 for sample copies of the tutorial worksheets and tutorial homeworks discussed in this article.

[40] This result can be achieved by using separation of variables to solve the time-dependence Schrödinger equation; for more detail, see Sec. 2.1 of Ref. [28].

[41] C. Schiber, H. Close, E. Close, and D. Donnelly, Student use of a material anchor for quantum wave functions, in *Proceedings of the Physics Education Research Conference 2013, Portland, OR* (AIP, New York, 2013), pp. 325–328.

[42] This is true for systems with time-independent Hamiltonians such as those used throughout most instruction at the junior-level course. For a discussion of student difficulties in contexts that involve Hamiltonians that are not time independent, see Ref. [16].

[43] S. B. McKagan, K. K. Perkins, M. Dubson, C. Malley, S. Reid, R. LeMaster, and C. E. Wieman, Developing and researching PhET simulations for teaching quantum mechanics, Am. J. Phys. **76,** 406 (2008).

[44] Several PhET simulations (University of Colorado—Boulder) deal with quantum phenomena: https://phet.colorado.edu/en/simulations/category/physics/quantum-phenomena.

[45] A. Kohnle, D. Cassettari, T. J. Edwards, C. Ferguson, A. D. Gillies, C. A. Hooley, N. Korolkova, J. Llama, and B. D. Sinclair, A new multimedia resource for teaching quantum mechanics concepts, Am. J. Phys. **80,** 148 (2012).

[46] The QuVIS (University of St. Andrews) includes numerous online modules and visualizations: http://www.st-andrews.ac.uk/physics/quvis/.

[47] The Paradigms in Physics project (Oregon State University) takes an innovative approach to teaching quantum mechanics: http://physics.oregonstate.edu/portfolioswiki/.

[48] For the Open Source Physics simulations related to time dependence and quantum measurement, see http://www.compadre.org/OSP/items/detail.cfm?ID=6814.

[49] G. Zhu and C. Singh, Improving students' understanding of quantum measurement II: Development of research-based learning tools, Phys. Rev. ST Phys. Educ. Res. **8,** 010118 (2012).

[50] A. Kohnle, C. Baily, A. Campbell, N. Korolkova, and M. J. Paetkau, Enhancing student learning of two-level quantum systems with interactive simulations, Am. J. Phys. **83,** 560 (2015).

[51] For example, we found that the percentage of students who responded correctly with correct reasoning to questions like pretest 1 was about 13%. The only courses in our analysis that differed from this average by more than 5% were very small ($N \sim 15$). Our results from analysis of pretest responses are discussed in more detail in Refs. [13–15].

[52] The Yates correction to the Chi-squared test was used in cases where one or more categories included 5 or fewer responses.

[53] The chi-squared test was chosen because it is appropriate for categorical data that is not matched. Since the chi-squared test requires unmatched samples, students who completed the post-test were excluded from the pretest samples reported in Refs. [13,14]. While this introduces the possibility of adding a systematic error, the pretest results for the quarter corresponding to the exam reported in this article were in fact between 2%–8% smaller than the pretest results reported here.

[54] The tutorials used at CPP included some small modifications made by the University of Colorado–Boulder.