

Testing Tutorials in Upper-Division: An Example from Quantum Mechanics

Gina Passante, Paul J. Emigh, and Peter S. Shaffer

Department of Physics, University of Washington, Seattle, WA 98195-1560

Abstract: The Physics Education Group at the University of Washington has been developing and testing *Tutorials in Physics* for use in upper-division courses on quantum mechanics. Like *Tutorials in Introductory Physics*, these materials are intended to supplement instruction in lecture-based courses, with a focus on improving student conceptual understanding. We describe the overall impact of the tutorials on student learning in a junior-level quantum mechanics sequence at the University of Washington. Our findings indicate that students at all levels (by academic performance in the course) benefit from the use of tutorials. Moreover, on course evaluations, students have overwhelmingly ranked the tutorials as being helpful to their learning, with the highest ratings coming from the highest performing students.

Keywords: physics education research, quantum mechanics, upper-division, tutorials, student understanding

PACS: 01.40.Fk, 01.40.G-, 03.65.-w

INTRODUCTION

Quantum mechanics is often considered one of the capstone ideas in a physics degree. The underlying concepts are becoming increasingly relevant across a broad range of fields. There is mounting evidence, however, that even majors often complete their study unable to reason correctly about certain foundational ideas [1-3]. As has been found in studies at the introductory level, many difficulties seem to be independent of lecturer, textbook, or university [4,5].

To help address this problem, the Physics Education Group at the University of Washington (UW) is developing a set of tutorials (*Tutorials in Physics: Quantum Mechanics* [6]) for use in upper-division courses on quantum mechanics. The approach we have used is one that has proved effective in introductory physics courses. Like *Tutorials in Introductory Physics* [7], the tutorials on quantum mechanics are a supplement to lecture instruction that aim to improve student learning by focusing on specific conceptual and reasoning difficulties. They are being created through an iterative process of research, curriculum development, and assessment.

In upper-division courses, however, the number of students and the opportunities for data collection are much more limited than at the introductory level. In addition, the different course structures and expectations of students, teaching assistants, and instructors can impose additional constraints. In this paper, we discuss some of the differences we have encountered and describe methods that we have been using to assess the overall impact of the tutorials.

CONTEXT FOR RESEARCH AND CURRICULUM DEVELOPMENT

The structure of the tutorials on quantum mechanics is similar to what has been discussed in previous articles describing *Tutorials in Introductory Physics* [8]. Each tutorial sequence consists of a pretest (typically taken by students online); a tutorial section in which students work together in small groups through a tutorial worksheet; and a short tutorial homework that is graded by teaching assistants. The sequence is followed by post-test questions that are given as a regular part of each course examination.

Each tutorial pretest is administered before the relevant tutorial, but after lecture instruction on the corresponding material. They typically take between 15 and 25 minutes to complete and are multiple-choice with explanations of reasoning. The post-test questions constitute about one-third of each of the two midterm exams and one-fifth of the final exam. The pretest and post-test questions are never identical. We refer to questions written to assess the tutorials as “tutorial exam questions” and the remaining questions (written by the course instructor) as “lecture exam questions.”

At UW, the quantum mechanics course has traditionally had three 50-minute lectures (or two 1.5 hour lectures) and one 50-minute recitation each week. The recitations are led by graduate student teaching assistants (TAs). Over the course of this study, tutorials have gradually replaced TA-led discussions during certain weeks. Students receive participation credit for attendance; however, at times as many as half of the students may not attend.

Since the students who attend a given tutorial may not be representative of the entire class, we have needed methods for assessment that differ from those we have used to examine the impact of the introductory tutorials. (For those tutorials, we have typically compared pre- and post-test results for all students.) In this paper, we discuss two methods. The first is an analysis of pre- and post-test data for only those students who attended the relevant tutorial(s). The second method is an aggregated comparison of all the post-test responses (across all questions in a given quarter) for students who did and did not attend the relevant tutorial(s). We also present results from a survey asking students about the extent to which they regarded the tutorials as beneficial to their learning. (Note that it is not the goal of this paper to describe specific student difficulties or to discuss the development of the curriculum.)

The data presented in this paper is from the two-quarter junior-level quantum mechanics sequence at UW. There were 69 (58) students who completed the first (second) quarter. Although tutorials have been used for many years, we report results for only one year since both the post-tests and the tutorials are modified annually. The pretest results that are reported are consistent with those obtained in previous years ($N \approx 300$). The post-test results have risen steadily.

DESCRIPTION OF PRETESTS AND POST-TESTS

Five tutorials were used during the first quarter of instruction discussed in this paper: *Time dependence in quantum mechanics*; *Energy measurements*; *Position, momentum, and energy measurements*; *Angular momentum in quantum mechanics*; and *Addition of angular momentum*. We also discuss *Time-independent perturbation theory*, used in the second quarter. From these topics, three matched pre- and post-test questions were given to probe the impact of the tutorials. In our experience the pre- and post-test questions present a similar level of difficulty to students and require the same chain of reasoning to answer.

Energy measurements: The first set of pre- and post-test questions are for the tutorials on time dependence and energy measurements. On the pretest, students are given the state $\Psi(x, 0) = i\sqrt{1/3}\psi_0 - \sqrt{2/3}\psi_1$, where ψ_n is the n^{th} energy eigenstate of the harmonic oscillator potential, with energy E_n . One question asks whether there is ever a time when the probability of measuring E_1 is equal to zero. Four choices are given: *Yes, there are times when this probability is zero*; *Yes, but only after waiting a long time*; *No, there is no time when the probability is zero*; and *There is not enough information to answer*. The

correct answer is that there is no such time, since the probability of an energy measurement is the absolute value square of the corresponding coefficient. On the post-test, students are told that the state of a particle in an infinite square well of width a is $\psi(x, 0) = \sqrt{30/a^5}(x^2 - ax)$. They are asked if the probability that the energy is measured to be E_1 depends on the time of the measurement. The question is similar to the pre-test, although open-ended and the state is not written as a superposition of energy eigenstates.

Angular momentum: The second pre- and post-test set cover the pair of tutorials on angular momentum and spin. The pretest that preceded both tutorials gave students the wave function for a 3-dimensional system as: $\psi(r, \theta, \phi) = f(r)Y_2^1(\theta, \phi)$, where $Y_2^1(\theta, \phi)$ is a spherical harmonic. Students were asked to identify the possible results of a measurement of L_z . The post-test question was identical except that students were given the initial state in Dirac notation, specifying the quantum numbers l , m_l , s , and m_s . To answer both questions, students needed to recognize that the possible results of the measurement are $m_l\hbar$ for each value of m_l present in the wave function.

Perturbation theory: The third set of questions is from time-independent perturbation theory. Both the pretest and post-test questions provided students with a graphical description of a perturbed infinite square well potential and asked if the perturbation to the ground state energy was *greater than*, *less than*, or *equal to* the perturbation to the first excited state energy. On the pretest, the perturbed potential was $+V_0$ ($-V_0$) on the left (right) halves, respectively. On the post-test, the perturbed potential was $+V_0$ at the center and decreased linearly to zero on the left and right sides. A correct answer required students to compare the overlap of the wave function squared with each perturbation.

ASSESSMENT THROUGH PRETESTS AND POST-TESTS

In this first assessment of student performance, we include only students who attended the relevant tutorial(s). Table 1 shows the number and percentage of students who answered correctly out of those who

TABLE 1. Percent correct (independent of reasoning) on pretests and post-tests of those students who worked through the relevant tutorial. The difference is statistically significant with $P < 0.001$ for all three questions.

| | Pretest | Post-test |
|---------------------|-------------|-------------|
| Energy measurements | 60% (15/25) | 92% (36/39) |
| Angular momentum | 30% (10/33) | 85% (29/34) |
| Perturbation theory | 35% (12/34) | 90% (35/39) |

worked through the relevant tutorial. Although the numbers are small, the results suggest good improvement in each case. The difference between each set of pre- and post-test percentages are statistically significant with $P < 0.001$. This level of improvement was not obtained in early versions of the tutorials; many cycles of research and development have been required [9].

On the *energy measurements* post-test question, the percentage of students who gave the correct answer is very high. However, on both the pre- and post-test not all students provided complete explanations.

Angular momentum is a difficult topic for students. We found that a single tutorial on this topic was not sufficient. It was recently expanded to a set of two, which allowed for a more thorough treatment in order to address common difficulties. The results in Table 1 represent an improvement over previous versions of the tutorial.

In perturbation theory, most students are able to give a correct equation for the correction to the energy. However, on the pre- and post-test questions, students are provided with a graphical description of the functional form. Thus, they are required to reason qualitatively to compare the effect of the perturbation on two different energy levels. We have found this to be very difficult for students. On recent versions of the tutorial, student performance has increased from 35% on the pretest to 90% after tutorial instruction.

ASSESSMENT THROUGH AGGREGATE EXAM PERFORMANCE

The previous section focused only on pre- and post-test results for students who attended the relevant tutorial and took the exam. It excluded students who did not attend the tutorial(s), many of whom also did not take the pretest. The number of these students was not negligible, so we tried to find a way to compare the post-test performance of students who did and did not attend the tutorials.

Since the class was relatively small, we decided to aggregate the data from all (17) tutorial exam questions. In total, 1167 post-test responses to individual questions were analyzed: 707 from students who attended the relevant tutorial and 431 from students who did not. The percentage of tutorial exam questions answered correctly by students who attended

the tutorial was 61% in contrast to 38% from students who did not attend tutorial, a difference of 23%. Responses were considered correct if they contained the right answer and some portion of correct reasoning.

It is not possible, however, to attribute the difference only to the tutorials. A similar comparison for the lecture exam questions yields 61% and 46% respectively for the averages for students who did and did not attend the tutorials, a 15% difference. This result might indicate that working through the tutorials improves performance on all exam questions (both tutorial and non-tutorial), but it might simply indicate that stronger students attended the tutorials.

To determine the extent to which performance on the tutorial exam questions is related to working through the tutorial, we needed to do the analysis for students of similar ability. We divided the class into three populations based on total score on the *lecture* exam questions. Within these top, middle, and bottom performing populations, we then compared the lecture exam scores of students who did and did not attend tutorial. (See the right half of Table 2.) Note that the difference in scores within each third is small.

Although the average difference in lecture exam scores is small for each third of the class, it is large when the groups are aggregated (15%). This apparent paradox is due to unequal percentages of students who attended the tutorial in each third (87%, 59%, and 44% in the top, middle, and bottom thirds, respectively). In other words, students who attended the tutorial had a higher average lecture exam score than those who did not because more top students are in that population. This result is known as Simpson's paradox [10].

Because student performance on the *lecture* exam questions within each third is similar, we tested for the impact of attending tutorials within each third. (See the left half of Table 2.) The results indicate that students who attended tutorial performed better on the *tutorial* exam questions than their peers. In fact, the average performance of students in the bottom third who attended the tutorials is similar to that of students in the middle third who did not attend the tutorial.

There is a statistically significant difference, at the 99% confidence level, in the middle and bottom thirds of the class. The p-value for the upper third of students is 0.14. However, the uncertainty in this group is large due to the small number of students in this category who did not attend the tutorials.

TABLE 2. Average performance on exam questions. Asterisks indicate a difference significant at the 99% confidence level [11].

| | Tutorial exam questions (% correct) | | | Lecture exam questions (% score) | | |
|--------------|-------------------------------------|-------------|------------|----------------------------------|-------------|------------|
| | Tutorial | No tutorial | Difference | Tutorial | No tutorial | Difference |
| All students | 61% | 38% | 23% | 61% | 46% | 15% |
| Top 1/3 | 70% | 64% | 7% | 76% | 75% | 1% |
| Middle 1/3 | 48% | 35% | 13%* | 56% | 53% | 3% |
| Bottom 1/3 | 34% | 19% | 15%* | 39% | 34% | 5% |

STUDENT PERCEPTION OF TUTORIALS

At the end of the quarter we asked students for feedback on the tutorial component of the course. They were asked to rate their agreement with the statements below on a 6-point scale: strongly agree, agree, somewhat agree, somewhat disagree, disagree, and strongly disagree.

- Q1. The tutorials were helpful to my learning.
Q2. I would like to see tutorials in other advanced courses (e.g., E&M or classical mechanics).

Figure 1 shows the average for each third of the class, where strongly disagree was given a value of zero and strongly agree a value of five. The highest performing students, those in the top third, rated the tutorials higher than the others. In fact, every student in the top third agreed with both statements to some degree.

The average score on the survey is consistent with our observations from each year that we have used the tutorials on quantum mechanics. The majority of students have consistently requested that more time be devoted to tutorials in the small group sections. In part, because of student experiences with the tutorials in quantum mechanics, upper-division electro-magnetism tutorials [12] have been introduced into the recitation sections at the request of students.

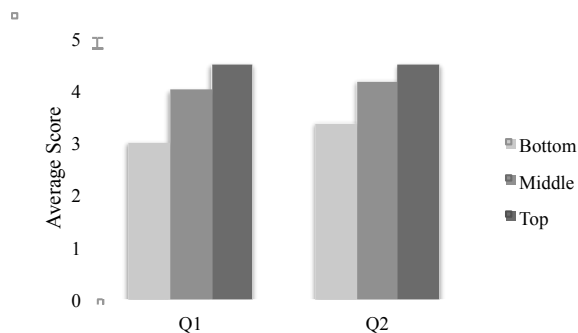


FIGURE 1. Responses to the survey questions for the top, middle, and bottom thirds of the course as defined by their scores on the lecture exam questions.

CONCLUSION

This article describes methods for assessing curriculum that we are finding useful for upper-level courses. In cases where the populations are smaller, tutorial attendance can become a significant factor in the analysis. The findings are encouraging and suggest that there can be significant benefits from the use of *Tutorials in Physics: Quantum Mechanics* in a junior-level course on quantum mechanics. The results indicate an overall improvement in conceptual

understanding on a variety of topics for students who attended the tutorials. In addition, we have shown that these students are better able to answer tutorial exam questions than their peers who did not attend the relevant tutorial(s). This difference in performance was greatest for students in the middle and bottom thirds of the class. Although we did not see a statistically significant difference for the top third of students, this may have been due to the small number of students in that population who do not attend tutorials. However, we have found that these students overwhelmingly believe that the tutorials are of benefit to their learning.

In the future we plan to extend this analysis as well as to report on the development and assessment of individual tutorials. We will also document instructional strategies that have proved effective at addressing student conceptual and reasoning difficulties. In addition, we will document findings from faculty who are using these materials at other institutions.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contributions of current and past members of the Physics Education Group at UW. Andrew Crouse was primarily responsible for the initial design of many of the tutorials. We would also like to thank UW faculty member John Rehr for welcoming this research in his courses. This research has been supported in part by the NSF under Grant No. DUE-1022449.

REFERENCES

- [1] D.F. Styer, *Am. J. Phys.* **64**, 31 (1996).
- [2] C. Singh, *Am. J. Phys.* **69**, 885 (2001).
- [3] A. Crouse, "Student Understanding of Quantum Mechanics", Ph.D. Dissertation, University of Washington (2007).
- [4] C. Singh, *American Journal of Physics* **76**, 277 (2008).
- [5] G. Zhu and C. Singh, *Am. J. Phys.* **80**, 252 (2012).
- [6] L.C. McDermott, P.S. Shaffer, P.R.L. Heron, and the Physics Education Group, *Tutorials in Physics: Quantum Mechanics*, Preliminary edition (2014).
- [7] L.C. McDermott, P.S. Shaffer, the Physics Education Group at the University of Washington, *Tutorials in Introductory Physics*, First Edition (Prentice-Hall, Upper Saddle River, NJ, 2002).
- [8] L.C. McDermott and P.S. Shaffer, *Am. J. Phys.* **60**, 994 (1992); P.S. Shaffer and L.C. McDermott, *Am. J. Phys.* **60**, 1003 (1992).
- [9] For information on the iterative cycle of research and curriculum development, see Ref. [8].
- [10] C.R. Blyth, *J. Am. Stat. Assoc.*, 67:364-366 (1973).
- [11] Results for 'all students' were not tested for significance due to unequal representation from each third of the class.
- [12] Based on EM tutorials from CU-Boulder: colorado.edu/physics/EducationIssues/cts/index.htm