Testing the development of student conceptual and visualization understanding in quantum mechanics through the undergraduate career

E. Cataloglu[a)]
Science Education Program, College of Education, The Pennsylvania State University, University Park, Pennsylvania 16802

R. W. Robinett[b)]
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 5 March 2001; accepted 8 June 2001)

In order to probe various aspects of student understanding of some of the core ideas of quantum mechanics, and especially how they develop over the undergraduate curriculum, we have developed an assessment instrument designed to test conceptual and visualization understanding in quantum theory. We report data obtained from students ranging from sophomore-level modern physics courses, through junior–senior level quantum theory classes, to first year graduate quantum mechanics courses in what may be the first such study of the development of student understanding in this important core subject of physics through the undergraduate career. We discuss the results and their possible relevance to the standard curriculum as well as to the development of new curricular materials.
[DOI: 10.1119/1.1405509]

I. INTRODUCTION

Have you ever heard anyone in physics, perhaps even yourself, say something like this? ''*You know, I never really understood that until the third time I saw it, probably in grad school*.'' Even if one does not specify the exact topic involved in such a statement, many students who have progressed through a typical undergraduate curriculum and gone on to graduate study (and perhaps even beyond) in physics can easily think of many subjects in any of four core areas, classical mechanics, electricity and magnetism, statistical mechanics and thermodynamics, and quantum mechanics, which might be the subject of such a wistful comment. One reason this is true is that there are any number of standard topics in these areas that are treated, with varying degrees of sophistication and/or approached with varying levels of mathematical machinery, first at the introductory level (sometimes in large lecture courses or otherwise), once more in specialized, junior–senior courses for physics majors, and yet again in the first year of graduate study.

While there are an increasingly large number of educational studies which focus on student conceptions (or alternative conceptions or misconceptions) of a specific set of topics at a given curricular level, there are far fewer attempts at probing how students' understanding of common or core topics in a given discipline evolve over the course of a typical undergraduate career. As part of an NSF-funded project[1] to develop modern web-based instructional materials related to undergraduate quantum mechanics, we have also been exploring the development of student understanding in some selected areas of quantum theory and how this evolves over the course of a typical undergraduate physics majors experience. We have not been able to find any similar studies in the physics pedagogical or science education literature, so this may well be the first detailed examination of how an increasingly important component of a physics major's training affects their understanding in a core area over an entire undergraduate career.

In order to obtain benchmark data to use to test the efficacy of any new quantum mechanics pedagogical materials, as well as to better inform our own development of web-based instructional modules, we have developed and tested a conceptual assessment instrument[2] which has now been given to over 160 students at our institution during a recent three semester period (Fall 1999, Spring 2000, and Fall 2000). The test, which we call the Quantum Mechanics Visualization Instrument, or QMVI, focuses on conceptual and visualization understanding and we have used this survey to probe student understanding of a subset of quantum mechanics core ideas at various stages of what we feel is a fairly typical undergraduate career, at least in an American college or university environment.

After briefly describing, in Sec. II, our motivations in selecting the particular areas of interest on which we have focused, we proceed in Sec. III to discuss the development of the QMVI. We then provide, in Sec. IV, a preliminary item analysis using data obtained from students in a sophomore-level modern physics course, a junior–senior level quantum theory course (mostly for physics and astronomy majors), a first-year graduate course in physics, as well as a one-semester graduate level introduction to quantum chemistry. Using these data, we briefly discuss in Sec. V the implications these results might suggest for the standard undergraduate curriculum as well as the impact this study has already had on our own plans for the development of educational materials. Finally, in Sec. VI we discuss our conclusions and prospects for future physics educational research (PER) in this area.

II. MOTIVATION FOR THE QMVI

In studying student understanding in any area of the physics curriculum, it is natural to focus on a rather specific and well-defined core content area and student population, say Newtonian mechanics as taught in calculus-based introductory physics courses. The development of evaluation tools, such as the well-known Force Concept Inventory or others in the area of classical mechanics[3] or similar ones in fields such as electricity and magnetism[4] have benefited from using repeated testing and study of individual test items, given to large numbers of students, most typically in introductory courses, where one expects the curriculum to be relatively well defined.

If we wish to probe student understanding in a specific area over the entire undergraduate career, we would like to focus on a set of topics which appear repeatedly throughout a standard curriculum, even (or perhaps especially) if they recur, couched in increasingly more sophisticated levels of mathematical formalism. We also wished to concentrate on areas which had not been studied so extensively in the PER literature, but which are clearly of importance in the undergraduate and graduate curricula. We have chosen to focus on the development of students conceptual understanding of core topics in quantum mechanics, especially as evidenced by their visualization skills, and we will briefly explain our motivations for choosing this content area and approach.

A. Why quantum mechanics?

While many subjects in classical mechanics or electricity and magnetism are definitely covered several times in a typical undergraduate and first-year graduate curriculum, one can argue that the core material of an introductory course in quantum theory may be seen in more different contexts than any other topic. Starting from the last few chapters of many introductory texts (which may include discussions of matter waves, particle in a box quantization, and even applications to modern devices), students are typically exposed to a relatively common set of ideas in a sophomore level modern physics course, and likely exposed again in a junior–senior level quantum theory course if they are physics or astronomy majors. These courses, in turn, are often the prerequisites for applications courses covering such topics as solid state physics, atomic and molecular physics, and nuclear and particle physics, all of which make extensive use of many aspects of quantum mechanics. Much of the same core material is seen yet again (or at least assumed) in a first-year graduate course or sequence. An undergraduate, therefore, might easily see, for example, the infinite well used as a pedagogical example or a model physical system[5] in up to five courses before he or she graduates. Given the increasing impact that quantum mechanical ideas have on technological applications, it is unlikely that the importance of the concepts and model systems so often covered in the standard undergraduate curriculum can be understated.

Another issue which makes tests of student understanding of quantum mechanical concepts highly relevant is the manner in which quantum ideas are sometimes perceived by and/or described to undergraduates. While there are clearly many lingering student misconceptions about classical mechanics and E&M, these topics are seldom, if ever, described by adjectives such as weird or strange in the same way which is so common for quantum mechanics. Students are occasionally encouraged to approach the subject with the idea that it is almost impossibly difficult to understand and that it is so completely different from other branches of physics that ones intuition is of little or no use. There are, of course many semiclassical connections (WKB methods, wave packet motion, etc.) which can help bridge the gap between the classical and quantum worlds and which can presumably help students focus on the more truly radical aspects of quantum mechanics rather than on such fairly straightforward aspects as, for example, the shape/form of quantum wave functions in a potential well. We choose, to a large extent, to focus on these and related classical connections as not all undergraduate curricula emphasize the most radical aspects of quantum theory, while most standardly used texts do make contact, to some extent, with classical mechanics.

The importance of this relationship to classical ideas can also be justified as it is also a subject of much current experimental research involving effects such as wave packet revivals (which can be experimentally observed, for example, in Rydberg atoms[6]), a topic which has been extensively discussed in the pedagogical literature[7–9] in a manner which is readily accessible to undergraduates at this level. This classical connection is also motivated by earlier PER studies which show that student misconceptions about related classical concepts[10–14] can carry over to similar quantum mechanical model systems. (Student misconceptions which are more specific to quantum[15] or wave[16] mechanics can also be used to inform such studies.) Inspired by all these issues, we have chosen to include a significant component of problems which focus on this classical–quantum connection.

One special instance of an overlap between the quantum and classical descriptions which we include is the momentum–space description of quantum theory. This aspect not only has connections to classical mechanics which are nicely complementary to that of the more standard position–space formulation, but also is of increasing relevance to experimental realizations of quantum systems, especially in situations involving scattering, in such areas as solid state, nuclear, and particle physics.

Finally, there are a wide variety of approaches to the teaching of quantum mechanics, even at the undergraduate level. A quick electronic search of our institution's fine library collection for the keywords classical mechanics, electricity and magnetism, and then quantum mechanics finds 16, 30, and 57 entries, respectively, with similar ratios for other standard descriptions of books on these topics. While many of the classical mechanics and E&M books are found to have relatively similar approaches, the variety of styles and emphases in the textbooks on quantum mechanics is much larger. Texts which focus on very formal aspects of the subject (starting with the formalism of spin systems, Hilbert spaces, and the like) are available, as are many examples which focus on the Schrödinger equation approach. Some texts include many physical examples, including ones making direct, sometimes numerical connection to experimental results, while some provide few, if any, physical insights. So, compared to other subfields, there is seemingly an even wider array of possible topics which one might consider as constituting the core ideas, and so research in as wide a variety of such topics as possible will be useful.

For example, one recent survey[17] has focused on student understanding of topics such as quantum measurement theory and time evolution of quantum states, both discussed in a rather formal manner, and focusing on students at the end of a full-year upper-level course. To complement such research, in our study we have focused on more conceptual and nonformal aspects of quantum mechanics, but have intentionally included some material on the time development of wave packets, in a more physical and less mathematical fashion, even including some semiclassical aspects, in order to make contact with such studies, but addressing these important issues from a different point of view.

B. Why conceptual understanding?

As mentioned above, many of the differences between the presentation of the core ideas of quantum mechanics at various levels throughout the undergraduate and graduate curriculum are related to the increasing level of abstraction, mathematical sophistication, or calculational formalism used. Students in a sophomore-level modern physics course, for example, will almost certainly not have been exposed to such important topics as perturbation theory, variational methods, operator techniques and the like. In a similar way, applications of quantum mechanics ideas to topics such as solid state physics, atomic, or nuclear/particle physics are much more often left to the discretion of the instructor or the format of the text used in such a course. While forming an important part of the undergraduate curriculum, such topics are often considered separate from the core ideas themselves. For example, the GRE® in Physics Study Guide (Ref. 33) lists the percentages of questions asked in two recent tests as *Atomic physics* (10%), *Quantum mechanics* (10%), and *Advanced topics* (*nuclear and particle physics, condensed matter physics*, etc. (9%), so that the core ideas of quantum theory are explicitly separated out from areas of application.

Tests which focus on the development of conceptual understanding have a much better chance at determining how basic ideas change over the undergraduate curriculum, independent of specific applications to which students may or may not have been exposed. The focus on conceptual issues is also a way to focus attention on the core concepts and separate them from the various levels of mathematical sophistication used to study them as students progress. In that same regard, the use of visualization as a method of probing student understanding, while interesting in its own right, can also be used to shift the focus away from the more abstract mathematical methods used later in a students career.

C. Why visualization?

One aspect of the presentation of the core material in many undergraduate courses which has changed dramatically over the last 50 years[18] has been the increased ability to present numerically exact calculations of a wide variety of solutions to quantum mechanical problems, not just a few mathematically tractable and exactly soluble closed-form special cases. Many of the most recent examples of modern course materials,[19–24] including not just textbooks, but especially specialized software, now allow students to visualize the results of more sophisticated examples, including both multidimensional systems and time-dependent phenomena.[25,26] Since students are now much more routinely exposed at a very early stage, even in their introductory math and physics courses, to more sophisticated graphical and visual representations of experimental data, solutions to differential equations, and the like,[27] new pedagogical materials such as these will almost certainly have an impact on how quantum mechanics courses are taught. Therefore, data on visualization skills for testing the effectiveness of such new materials will likely be of increasing importance.

As if to emphasize this, the National Science Foundation[28] has noted that visualization is a form of communication which transcends application and technological boundaries and that visualization can be an important tool for scientific understanding and learning. Educational research has pointed to the advantages of visualization in developing scientific understanding by providing more opportunity for students to establish connections with pre-existing knowledge structures.[29–31] It has also been pointed out[32] that scientists routinely use visualization to translate data into pictures of various kinds, looking for consistent or inconsistent patterns, much as is done in standard "*what's wrong with this picture*" type of graphical questions.

Are students already expected to have a more sophisticated appreciation of the visual representation of data or concepts? One example of such a likely need can be seen in standardized tests of undergraduate physics knowledge, specifically the GRE®. Sample GRE® Physics tests are available (published by the Educational Testing Service itself) for sale and are eagerly studied by many junior and senior physics majors intending on pursuing graduate degrees. A sample of three such tests,[33] from the years 1985, 1991, and 1996, shows a pattern of increasing use of graphs and figures, both in the statement of questions as well as in the multiple choice solutions themselves. For example, for those three years, the number of questions which rely on graphs or figures (either conceptual or with data) ranges from 22 to 33 to 31, while the number of individual questions for which the actual multiple choice answers consist solely of different graphical images increases from 1 to 3 to 5.

Motivated by factors such as these, we have focused on the visualized representation of conceptual problems and their solutions in the development of the QMVI. In the next section, we describe how the topics and approaches we wish to emphasize, namely (i) conceptual understanding of the core (modern physics level) quantum mechanics material, (ii) connections to classical mechanics, (iii) understanding via visualization, and (iv) extensions of standard ideas to more modern areas such as two-dimensional systems and time-dependent phenomena, and the overlaps of these four areas, were implemented in the QMVI.

III. CONSTRUCTION OF THE QUANTUM MECHANICS VISUALIZATION INSTRUMENT (QMVI)

Motivated by the factors described in Sec. II, we began construction of a quantum mechanics conceptual assessment instrument by making an informal survey (using information readily found on the web) about topics covered in standard modern physics and junior–senior level quantum mechanics courses, and popular textbooks used in support of these, at many American colleges and universities. Using recent editions of frequently used standard modern physics textbooks[34–37] and other more innovative examples,[38] we identified topics typically covered in the 4–7 chapters devoted to basic quantum theory in these texts, especially to one-dimensional, nonrelativistic quantum mechanics. These include the probabilistic interpretation of the Schrödinger wave function, properties of solutions of the Schrödinger equation (including correlations between the amplitude and wiggliness of the wave function and the potential energy function, allowed forms for solutions in simple cases, including the infinite well, quantum mechanical tunneling and barrier penetration, etc.), the uncertainty principle and the Pauli exclusion principle. These content areas were also examined by various physics faculty members who have taught relevant courses in the recent past. (Graduate students from education were also asked to review the final version of the exam questions, for language and general readability.)

We then formulated versions of questions, almost all involving visualization, dealing with all of these topics, in some cases including two complementary versions to check for consistency. For example, a rather typical problem, QMVI[15] (in a hopefully obvious notation), is similar to, but more focused than, one of the more challenging (and open-ended) end-of-chapter problems in Refs. 35 and 37, dealing with an asymmetric infinite well; in the same spirit, both QMVI[16] and [19], which ask about a slanted infinite well, are similar to problems in Refs. 35, 36, and 38. We included a few questions which focus on similar issues (probability ideas and motion described using the potential energy function), but described in a classical context, to assess how students approached similar problems in a non-quantum mechanical framework, motivated by earlier PER studies.[11–13] We also included five questions (QMVI[21]–QMVI[25]) which focus on less traditionally seen material, including visualization of multidimensional wave functions and time-dependent phenomena, basically wave packet propagation, to judge how well students could extend their understanding to material often seen in more modern presentations as well as to make contact with earlier studies[17] on student understanding of the time development of quantum states. We note that in each case we have used numerically exact representations of quantum wave functions, not just sketches, to ensure that the phenomena being illustrated accurately reflects real solutions of the Schrödinger equation.

The format chosen for each question was to have a single question on one page, with five multiple choice answers. Ample space was also provided for a required 2–3 line written response and students were also asked to denote their confidence level in their answer by circling a response ranging from *very certain* to *very uncertain*. Students in all of the studies performed so far have been given the same instructions and graded in the same manner, namely they are told that they will receive 2 points for a correctly chosen multiple choice answer and 0, 1, or 2 points depending on the correctness and completeness of their written response. This (possibly more subjective) grading was all done by one person, namely one of the authors (R.W.R.), to assure uniformity, but who also used a grading rubric for assigning these points. In this way, the maximum number of points per question is 4 and with our 25-item test the total possible score is 100. A student who randomly guesses (circling random answers and providing no written feedback) would thus have an average score of 10. With N such students, the estimated standard error of the mean would be $20/\sqrt{N}\approx 5$ for the sample sizes of roughly $N=15$–20 (per class) we have used. In a similar way, we see that combined average scores of approximately 10 ± 5 % on an individual question indicate responses no better than random guessing.

The administration of such assessment instruments is best done in as controlled environment as possible, ideally in an in-class exam situation. Due to the time constraints of the courses involved, and the varying syllabi of the instructors who participated over the course of the study (seven different faculty members in four different courses over a three semester period), however, we were unable to administer the QMVI in this way. For reasons of consistency, as well as to ensure faculty involvement (in some semesters, faculty declined to participate), in all of the tests performed so far the QMVI offerings have been given in the form of an extended take-home exam, most often for a small amount of extra credit in the course, with very specific instructions to (i) work individually, (ii) to spend no more than 1–2 hours on the test, and (iii) to use no other resource than the textbook used for the particular course involved (as each course had an assigned text). Little or no evidence of collaboration amongst students was seen (as evidenced by the written comments) amongst the responses.

The students agreed to have their grade information made available to the authors as part of their investigation and some information on the longitudinal development of individual students scores on the QMVI is therefore possible.[2] Students were not allowed to keep copies of the exam and only their overall score was reported back to them (as part of the allocation of extra credit points) but not their answers on individual items: in this way we feel that the test was rather secure from one semester to the next. (Information on the site where the QMVI can be downloaded has so far only been available to interested physics instructors upon request and was not publicly accessible on the web at any time during the study.) The tests were given out at similar times during each study, namely the last 1–2 weeks of the semester, after the bulk of the instruction on one-dimensional quantum mechanics, but before the final exam period.

The first few versions were shown to various physics faculty colleagues and others who provided feedback and a complete version of a 25-question test (V0.3) was first given to students in the Fall 1999 semester. Based on student responses, especially their written feedback and explanations, and instructor comments, a slightly revised version (the one currently in use, V0.4) was used in both the Spring 2000 and Fall 2000 semesters and the results of that version will be discussed here.

Because of this format, motivated by our desire to obtain written feedback and confidence information, the existing version is rather long and is not reproduced in its entirety here; two sample pages are shown in Appendix A as an example of the format. Both versions of the QMVI, the grading rubric we used to assign the written score points, worked out solutions, and other background material are all readily available at

www.phys.psu.edu/faculty/RobinettR/QM/QMVI/QMVI.html

in a variety of formats (Postscript and PDF files). (The development of the QMVI is discussed in an upcoming Ph.D. thesis.[2]) We now turn our attention to the results obtained from the various offerings of the QMVI over the last three semesters.

IV. QMVI RESULTS FROM THREE SEMESTERS

Since one of our stated goals was to determine the progress of student understanding through the undergraduate curriculum, various versions of the QMVI have been given to four distinct courses in at least one semester during the development phase. These four courses, with some important background information, are described below.

ModPh is a 3-credit, one-semester, sophomore-level modern physics course using a textbook at the level of Serway, Moses, and Moyer.[34] Students are typically exposed to roughly five chapters of introductory quantum mechanics (*Particle Nature of Matter* through *Quantum Mechanics in Three Dimensions* in Ref. 34) and then proceed to applications including atomic, molecular, solid state, and particle physics. The course is offered both fall and spring semesters and requires completion of the second semester of introductory physics (E&M) and the second semester of introductory

Table I. Collected data for three semesters offerings of two versions of the QMVI. Data on the average score (with standard error of the mean included), number of respondents in each class out of the total enrollment (resp/total), and average grade point average (GPA) of the students who participated (obtained with their permission) at the end of the semester during which they took the QMVI are shown. The average score for the combined Sp00/Fa00 ModPh group is 28.5 ± 2.4.

	ModPh	ModPh-H	UgQM1	GrQM1	QChem
Fa99 (V0.3)	26.7 ± 3.0		47.9 ± 3.6	55.2 ± 8.4	
resp./total	45/52		17/23	5/33	
⟨GPA⟩	3.25 ± 0.09		3.55 ± 0.12	3.68 ± 0.17	
Sp00 (V0.4)	28.8 ± 2.8	69 ± 3	58.3 ± 4.4		
resp./total	26/41	2/8	15/26		
⟨GPA⟩	3.46 ± 0.07	3.68 ± 0.05	3.45 ± 0.15		
Fa00 (V0.4)	27.6 ± 5.1	65	45.4 ± 3.8	55.5 ± 3.4	29.7 ± 4.6
resp/total	8/23	1/1	19/29	13/22	14/21
⟨GPA⟩	3.26 ± 0.13	3.86	3.30 ± 0.10	3.60 ± 0.08	3.51 ± 0.09

calculus. Students enrolled in this course are typically science (physics and astronomy) and engineering (electrical engineering, engineering science, etc.) students.

UgQM1 is a 4-credit, one-semester, junior–senior level introduction to quantum mechanics using a textbook at the level of Griffiths.[39] Students typically work through such a book, up to and including the hydrogen atom. The **ModPh** course is a prerequisite for this class as is a course in ordinary and partial differential equations. An optional second semester course (which would be described as **UgQM2**) is occasionally offered if sufficient demand exists and continues with applications, but was not available during the three-semester trial period. The students enrolled in this course are predominantly science undergraduates (physics and astronomy) and some engineering (electrical engineering and engineering science) undergraduate and graduate students.

GrQM1 is a 3-credit, first-semester, first-year graduate course in quantum theory using a textbook at the level of Cohen–Tannoudji *et al.*[40] The audience is almost entirely first year graduate students in physics who are also required to take the second-semester course (**GrQM2**) as well; the sequence is offered in fall (**GrQM1**) and spring (**GrQM2**) only. On occasion, a talented undergraduate takes the course and the Fall 2000 results include one such student. A course at the level of **UgQM1** is a prerequisite for this class.

QChem is a 3-credit, one-semester, first-year graduate course in quantum chemistry using a textbook at the level of Levine.[41] The course is cross listed with a similarly titled undergraduate course, but the enrollment is almost exclusively graduate chemistry students; the Fall 2000 sample includes only one undergraduate. The course typically covers the basics of quantum mechanics in the first six chapters of Ref. 41 before turning to applications to chemistry. Students enrolled in this course are expected to have had a two-semester course in physical chemistry as well as the second semester of introductory physics (E&M) and so have experienced a similar background (both in physics and math) to those students enrolled in **ModPh**, but typically 1–2 years earlier in their careers.

A. Global results

The first version of the QMVI (V0.3) to be offered to students was given during the Fall 1999 semester. Based on the written feedback of the students and comments from faculty colleagues, and an item analysis, a revised test (V0.4) was constructed and that same version has now been given in two semesters, Spring 2000 and Fall 2000. Some of the results are shown in Table I and include the average scores (out of 100, scored as described above) and standard error of the mean (standard deviation divided by $\sqrt{N}$ where N is the number of respondents) as well as the average grade point average (GPA) (calculated immediately after the end of the semester when grades first become available) for those students who participated in the study. Several aspects of how the courses were organized are especially relevant.

(i) The same textbook[34] was used in **ModPh** for all three semesters, and the same faculty member taught both the Sp00 and Fa00 semesters; because of the similarities in average score and GPA, instructor, and textbook, we have combined those two data sets as representative of the performance of a typical **ModPh** class. The average value for the combined Sp00/Fa00 set is then 28.5 ± 2.4 for 34 students. In both Sp00 and Fa00 semesters there were some Honors students who participated in the course, and in the QMVI study, but who were officially registered for a slightly different version of the class itself, which we label **ModPh-H**. Their data (three students in all, with consistently much higher scores) are shown separately and not included in the overall **ModPh** data or discussed in our analyses due to the very small number of students involved.

(ii) The same instructor taught the **UgQM1** course in both Fa99 and Fa00 semesters and each time used the same popular text.[39] Based on our web surveys of sample syllabi and textbook usage, we think that this approach is rather representative of the way the typical undergraduate QM course is taught, both at our institution and elsewhere, so we will use the Fa00 data to represent the junior–senior level of expertise. The Sp00 course was taught by a different instructor (not one of the authors) and used a different textbook,[42] one written by one of the authors (R.W.R.), which focuses on many of the same ideas and approaches emphasized in the QMVI. This unintentional experiment may provide a possible window on the variability of the QMVI score depending on the pedagogical approach used or emphasis placed on different cur-

Table II. The average (and standard error of the mean) scores for each question for the QMVI (V0.4). The Spring and Fall 2000 data for the ModPh course are combined and the other Fa00 scores, as well as the Sp00 UgQM1 scores are included. The average score (and standard error of the mean) for the entire test for each group is also included (in parentheses) at the top of each column for comparison.

No.	ModPh Sp00/Fa00 (28.5 ± 2.4)	UgQM1 Fa00 (45.4 ± 3.8)	GrQM1 Fa00 (55.5 ± 3.4)	QChem Fa00 (29.7 ± 4.6)	UgQM1 Sp00 (58.3 ± 4.4)
1	50 ± 08	88 ± 05	83 ± 09	46 ± 12	77 ± 10
2	26 ± 07	66 ± 10	82 ± 10	36 ± 13	87 ± 08
3	20 ± 06	53 ± 10	52 ± 13	14 ± 09	65 ± 12
4	57 ± 08	86 ± 07	77 ± 12	54 ± 12	87 ± 09
5	83 ± 06	96 ± 04	100 ± 0	89 ± 07	88 ± 08
6	38 ± 08	87 ± 07	88 ± 08	36 ± 13	78 ± 10
7	12 ± 04	54 ± 11	54 ± 12	18 ± 10	53 ± 13
8	14 ± 05	26 ± 08	31 ± 11	04 ± 03	28 ± 11
9	33 ± 06	49 ± 08	54 ± 08	36 ± 11	68 ± 09
10	14 ± 05	39 ± 11	33 ± 13	29 ± 10	27 ± 11
11	09 ± 03	14 ± 07	02 ± 02	18 ± 09	05 ± 05
12	33 ± 07	45 ± 09	75 ± 10	27 ± 10	75 ± 09
13	52 ± 08	79 ± 09	73 ± 12	57 ± 13	83 ± 09
14	16 ± 06	16 ± 08	38 ± 12	07 ± 05	27 ± 11
15	25 ± 06	32 ± 09	52 ± 10	13 ± 07	52 ± 12
16	27 ± 07	63 ± 10	62 ± 12	07 ± 07	65 ± 10
17	10 ± 04	14 ± 07	92 ± 06	14 ± 08	50 ± 12
18	61 ± 08	75 ± 09	46 ± 14	79 ± 11	82 ± 10
19	21 ± 06	54 ± 09	63 ± 10	16 ± 09	88 ± 07
20	09 ± 03	12 ± 07	13 ± 08	05 ± 04	42 ± 11
21	40 ± 08	37 ± 10	63 ± 12	59 ± 12	63 ± 12
22	24 ± 06	28 ± 10	56 ± 12	43 ± 13	65 ± 11
23	18 ± 05	08 ± 06	54 ± 11	23 ± 10	50 ± 12
24	13 ± 05	05 ± 05	29 ± 08	11 ± 05	48 ± 11
25	06 ± 03	11 ± 04	13 ± 09	04 ± 03	03 ± 03

ricular material at this level. We will therefore briefly compare the **UgQM1** Fa00 and Sp00 data in Sec. IV D.

(iii) The Fa00 **GrQM1** and **QChem** data will be used as being representative of the first year graduate physics and chemistry quantum mechanics courses, respectively.

The similarity in scores (overall and question-by-question) between the sophomore-level **ModPh** and the graduate-level **QChem** student scores is perhaps not surprising as the required math and physics background for the two courses are very similar and the amount of introductory material (roughly 5–6 chapters of one- and three-dimensional quantum theory) covered in the two courses are seemingly very comparable, both in scope, topics, and the level of mathematical sophistication assumed and physical insight provided. Small differences in the individual responses will be briefly discussed below.

B. Item analysis

While one must be extremely cautious about drawing detailed information from an item analysis with the relatively small number of students who have taken the QMVI during the development phase, especially when broken down into each course level, it is still worthwhile to examine the question-by-question results for any obvious trends. We also supplement the raw scores (scaled to a maximum of 100) for each test question with information obtained from the written responses. Each question from the V0.4 version which we are analyzing in detail is denoted by QMVI[n] where n runs over the 25 item pool. The combined Fa00/Sp00 ModPh, Fa00 UgQM1, Fa00 GrQM1, and Fa00 QChem data are used as being most representative of a typical undergraduate or graduate curriculum and textbook. A more complete listing of all of the V0.4 results, along with the Sp00 UgQM1 data, and including errors, is shown in Table II for completeness. The data shown here are rounded and no errors are shown. Averages for the test as a whole for each of the four courses are listed (in parentheses) at the top of each column for comparison so one can more easily see which questions were found to be, on average, easier or harder.

As mentioned above, the first four questions focus on aspects of classical mechanics which may be useful for an understanding of quantum theory. The first question, QMVI[1], asks students to recall the connection between the classical force and the potential energy function $[F(x) = -dV(x)/dx$, but asked purely graphically], while QMVI[2]–[4] focus on probability concepts, but in a classical context. QMVI[4], for example, asks students to interpret familiar classical motions (harmonic oscillator and accelerated particle) in terms of computer generated snapshots of many measurements of the particle position.

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
1	50	88	83	43
2	26	66	83	36
3	20	53	52	14
4	57	86	77	54

In these cases, as with a number of others we'll discuss, students do exhibit an increased level of understanding between the sophomore and more advanced levels. In a number of cases, the increases from **ModPh** to **UgQM1** are such that there is little room left at the top for students to do much better at the **GrQM1** level. The two questions which require students to infer information about the particles speed from a potential energy diagram (QMVI[2] and [3]) have consistently lower averages than those (namely, QMVI[1] and [4]) which are described more directly in physical terms; similar problems with classical misconceptions and difficulty in reading $V(x)$ plots have been noted in earlier studies.[11–13]

The next three questions all involve quantum mechanical probability ideas, ranging from interpreting $|\psi(x)|^2$ as the probability density (QMVI[5]), to calculating probabilities numerically (QMVI[6] which requires the proper normalization of a wave function[43]), to the numerical evaluation of expectation values (QMVI[7]).

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
5	83	96	100	89
6	38	87	88	36
7	12	54	54	18

Students at all levels clearly understood the qualitative meaning of $|\psi(x)|^2$ as the probability density, and after an undergraduate course could perform simple normalization calculations, but even at the graduate level students found it difficult to understand visually presented wave function information and use it to evaluate, or even estimate, expectation values numerically. One possible reason for this effect is that few of the more advanced texts ask students to manipulate real numbers.

Only a single question was included which had no visual component, but which asked students to estimate, via scaling arguments, the position–space uncertainty Δx, of a wave function from a very simple functional form; this step was required in order to obtain a dimensionally correct estimate for the corresponding momentum–space spread, Δp.

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
8	14	26	31	4

Students typically understood at each level that they were required to use the uncertainty principle connection, $\Delta x \cdot \Delta p \sim \hbar$, but clearly most students could not properly estimate the position–space spread, either by sketching the given functional form, by the use of ''full width at half max'' arguments, or even by dimensional analysis.

Two, somewhat related questions, are designed to probe student understanding of quantum mechanical wave functions (especially their amplitude and wiggliness and how the two are correlated) given a physical description of the corresponding classical motion. Examples for a uniformly accelerating particle (QMVI[9]) and a bouncing particle (QMVI[12]) gave similar results, near the overall average score.

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
9	33	49	54	36
12	33	45	75	27

Three questions are designed to focus on the mathematical concepts and results involved in the solution of the time-independent Schrödinger equation (TISE), especially at an intermediate stage of a derivation, but phrased in a conceptual manner. Specifically, students are asked about the form of formal solutions of SE as a differential equation in several different physical regimes, as well as the imposition of boundary conditions to obtain physically acceptable solutions.

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
10	16	39	33	29
11	09	14	02	18
17	10	14	92	14

These questions ranked among the lowest in all four groups studied. For both QMVI[10] and [11], the typical problem experienced by students, expressed through the written comments, was a confusion between known standard final answers, as distinct from the process of deriving an ultimately physically meaningful result through intermediate stages of mathematical manipulation. Students found it difficult to go through the steps of a derivation and separate the formal solution of the time-independent Schrödinger equation as a second-order differential equation for an arbitrary value of the energy eigenvalue E and the imposition of boundary conditions (either at infinity or some other boundary) as resulting in the quantization of the energies. The large jump in score for QMVI[17] over the course of the undergraduate career is perhaps understandable as it relies on what might be described as a trick which, if students see once they typically recall. [This problem asks students to recognize that the odd-parity solutions of the Schrödinger equation for a symmetric one-dimensional potential, $V(x)$, will also satisfy the related half-well problem where $V(x)$ is the same for $x>0$, but which is also characterized by an impenetrable wall at $x=0$. We note that students who answered basically the identical question in Fa99 and Sp00 had scores of 66 ± 7 and 50 ± 12, respectively; the average score on a very similar question dealing with a half harmonic oscillator problem, in recent two GRE® offerings, were 31% and 39%. Our understanding is that this aspect was mentioned in the two earlier semesters, but was likely not emphasized in the Fa00 case.]

A set of five complementary questions was included to probe student understanding of the qualitative behavior of wave functions in different potential energy situations, including more familiar bound state examples in position–space (QMVI[15], [16], and [19]), but also examples in one-dimensional scattering geometries (QMVI[14]) and also a version requiring students to confront the problem in momentum–space (QMVI[20]).

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
14	16	16	38	07
15	25	32	52	13
16	27	63	62	07
19	21	54	63	16
20	09	12	13	05

These problems, all of which ask students to account for the qualitative form of a stationary state wave function from very general arguments about the shape of the potential, are of a type pioneered by French and Taylor,[44] but which are now routinely discussed in a variety of texts, both at the modern physics and more advanced levels. More standard problems of this type for position–space bound states, such as QMVI[15], [16], and [19] all show increasing scores. A very similar one (QMVI[14]) involving scattering from a one-dimensional square barrier (one of the most standard 1D scattering geometries) shows significantly lower scores. (We note that somewhat more formal, but similar questions dealing with one-dimensional scattering on the GRE® also show low scores.[45]) Students typically do not use the same intuitive connections between the form of the potential and the shape of the allowed bound state wave functions when dealing with scattering states; they most frequently try instead to implement formalism dealing with transmitted and reflected fluxes, instead of focusing more directly on the wave functions. In a similar vein, QMVI[20], which asks students to also discuss stationary state solutions, *à la* French and Taylor, but in momentum–space, has scores which remain at the level of random guessing. This indicates that they are seemingly unable to generalize their improving competencies in position–space conceptualization of *how likely is it to find the particle in this part of the well* questions to the *how fast is the particle likely to be moving* type arguments involved in understanding probability densities for the momentum variable.

Even though students did increase their understanding of some problems involving interpretation of $V(x)$ diagrams, it is interesting to look at combinations of questions involving physical descriptions of motion and the connection to the resulting probability densities (either classical or quantum mechanical), averaging over three such questions, namely $(4+9+12)/3$, compared to three similar items where the questions are phrased in terms of the potential energy function, such as $(3+15+19)/3$.

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)	Problem description
$(4+9+12)/3$	41	60	69	39	‘physical’
$(3+15+19)/3$	22	46	56	14	$V(x)$

Clearly the physically described questions are always above the mean, while those requiring reading a $V(x)$ plot are consistently lower.

Two relatively standard questions are included which focus on simple properties of solutions of the infinite well problem (QMVI[13]) and on the Pauli exclusion principle and its effect on the filling of energy levels (QMVI[18]).

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
13	55	79	73	57
18	59	75	46	78

A familiar pattern of scores which increase after UgQM1 to a plateau is seemingly evident. The somewhat lower score in GrQM1 for the energy level filling question is likely due to the structure of the course where applications involving spin and the Pauli principle are typically seen only in the second semester.

In the last section of five problems, involving physical visualization beyond the standard curriculum, two questions are included which focus on simple two-dimensional quantum mechanical systems, a 2D infinite square well (QMVI[21]), and a 2D infinite circular well (QMVI[22]).

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
21	40	37	63	59
22	24	28	56	43

Students seem to be able to generalize the 1D square well results fairly well, even at the sophomore level, to obtain the 2D energy eigenvalues and match them to a multidimensional plot of probability densities. A similar question with circular symmetry such as QMVI[22], on the other hand, is not handled so successfully even though only qualitative information on the wave function (wiggliness or patterns of nodes) is required.

Finally, in the area of time-dependent solutions to the Schrödinger equation, focusing on wave packet motion, we include problems involving classical analogs of a collision with an infinite wall (QMVI[23]), a fairly standard spreading wave packet question (QMVI[24]), and a question which requires students to understand, in some detail, the $\exp(-iE_n t/\hbar)$ time dependence of individual bound state wave functions (QMVI[25]).

QMVI[n] (ave)	ModPh (28)	UgQM1 (45)	GrQM1 (55)	QChem (30)
23	18	08	54	23
24	13	05	29	11
25	06	11	13	04

All three problems have individual scores which are far lower than the overall averages, with the detailed time-dependence question averaging no better than random guessing (or worse). These problems reinforce the results of Ref. 17 which found that student understanding of the time development of quantum states was poor, even when phrased in conceptual terms as in these questions.

C. Student confidence data

In addition to answering the specific QMVI questions in both a multiple choice and written answer format, students were asked to rate their confidence on a four-point scale, ranging from *very certain*, *somewhat certain*, *somewhat uncertain*, to *very uncertain* by circling one of those four statements at the bottom of each page. These answers were transferred to a numerical scale, using values of 100, 67, 33, and 0 for the four choices. Thus, each student response had not only a numerical score (0–4, easily scaled to 0–100) but also some quantitative measure of the students confidence in their answer. To examine these combined data, we averaged the numerical score (0–100 scale) and the student confidence (0–100 scale) over all student responses for each individual question to obtain two values, S_n and C_n, for each question, $n=1,25$. In general, one would imagine that the higher the average score, the greater the student confidence, so we did a different least-squares fit to the 25 point data set for each separate course. The best-fit lines of the form $S=\alpha C+\beta$ and the corresponding correlation coefficients are shown in Table III.

For the three physics courses, the correlations were rather large indicating a reasonable degree of congruence between students perceptions and the correctness of their answers.

Table III. The slope and intercept parameters for a best straight-line fit to the individual score (S_n) and confidence (C_n) data points ($n=1,25$) for each of the four offering of the QMVI. The substantial correlation coefficient for the three physics offerings is further evidence of the validity of the test for these groups of students. The similar slopes (α), but varying intercepts (β), for the undergrad (UgQM1) versus grad (GrQM1) students indicates an increased confidence for the more advanced students.

Course	Best fit line $S=\alpha C+\beta$	Correlation coefficient
ModPh	$S=0.68C-3.10$	$\rho=0.70$
UgQM1	$S=1.03C-11.6$	$\rho=0.85$
GrQM1	$S=1.10C-34.0$	$\rho=0.65$
QChem	$S=0.87C-14.4$	$\rho=0.17$

These observations were also useful in helping to establish the overall validation of the instrument. The question which fell the furthest from the best-fit line for both the UgQM1 and GrQM1 cases was that for QMVI[11] indicating that there was a severe misunderstanding of that question. While the best-fit lines for these two populations had very similar values of α, the difference of roughly 20 in the β values imply that the graduate students were rather more confident of their answers.

D. Comparing different courses or approaches

As mentioned above, the students in the **ModPh** and **QChem** groups had very similar overall scores, and the more detailed data in Table II also exhibits a strong question-by-question correlation as well. There were only a few questions which had some perhaps measurable differences in response rate (but only at the one-sigma level). For the **QChem** group, QMVI[21] and [22] were both higher (perhaps due to the chemists familiarity with visualizing bonding, clearly emphasized in the typical textbooks used[41]) as was QMVI[18] (where the filling of energy levels and the Pauli principle are paramount in understanding atomic structure and chemical properties); for the **ModPh** group, only QMVI[15] and [16] had slightly better responses, both of which make use of the physics students presumably more extensive experience with the use of potential energy functions. Such comparisons are perhaps useful as it has been said that ''...*the preparation beginning graduate students have in quantum mechanics is spotty at best*...'' [46] so that the use of material appropriate for undergraduate physics majors, including a strong visual environment[46] can be a ''...*valuable resource in the teaching and learning of quantum mechanics*.''

The other obvious comparison is between the Fa00 (traditional approach, using Ref. 39) and the Sp00 (more visual approach, using Ref. 42) UgQM1 course approaches which give overall QMVI scores which are fairly different (at the level of slightly more than 2σ). The scores for the Sp00 UgQM1 test are, in fact, closer to the Fa00 GrQM1 results, almost across the board, except for four problems. For QMVI[17], where students who have discussed parity in one-dimensional solutions typically do very well, the UgQM1 score is typical of earlier semesters offerings at this level, but not at the mastery level of the graduate students who almost all get this one right; for QMVI[18], the artificially low score of the GrQM1 students has been discussed above. On the other hand, for QMVI[19] and [20], which require more sophisticated use of French–Taylor ideas,[44] including in momentum–space, those students who have been routinely exposed to such ideas seem to do better, even than graduate students in a typical post–undergraduate course. Clearly, there is a need for more detailed studies of individual item responses from a variety of undergraduate courses, especially using different texts, teaching styles, or instructional methods (use of computer software versus more traditional materials), etc.

E. Newer data

During the editorial review process of this paper, we were able to obtain QMVI results from a set of undergraduate and graduate students at another US institution, namely the University of Arizona (thanks to Professor I. Novodvorsky). The QMVI was given there to two groups.

(i) Four graduate students, completing the second semester of a two-semester grad sequence (and hence most similar to the **GrQM2** course mentioned above) using the textbooks by Sakurai and Baym.

(ii) Nine undergraduate students, completing the second semester of a two-semester undergraduate sequence (and hence most similar to the **UgQM2** at our institution) using the popular standard textbook by Liboff.[47] Considering the detailed syllabus for the course, we consider the instructional methods used and topics covered to be of the standard type employed in many such undergraduate courses and hence most directly comparable to the Fa99 and Fa00 UgQM1 data in Table I, as opposed to the Sp00 results.

In each case, the students were given very similar instructions on how to complete the QMVI so that the circumstances under which it was administered were as comparable as possible.

While the sample sizes for these two groups are even smaller than obtained in the data analyzed above, it is worthwhile to quote the global results for comparison. For the graduate course, the average score was 56.0 ± 5.0 (Arizona, Sp00, V0.4) which is to be compared to the 55.2 ± 8.4 (Fa99, V0.3) and 55.5 ± 3.4 (Fa00, V0.4) values in Table I, which show obvious similarities, within the clearly large associated uncertainties. The only obvious difference between the item-by-item responses (clearly visible even with the very large uncertainties due to small sample sizes) is that the Arizona group did far better on question No. 18 (roughly 3.8σ) involving the Pauli Exclusion principle. We indicated in our item analysis discussion in Sec. IV B that the relatively low score of the GrQM1 group on this question might be due to the fact that this material was typically seen in the second semester of the grad course, which the Arizona students did indeed cover.

For the slightly larger undergraduate sample, the average score was 49.9 ± 5.6 (Arizona, Sp00, V0.4) which we argue should be compared to 47.9 ± 3.6 (UgQM1, Fa99, V0.3) and 45.4 ± 3.8 (UgQM1, Fa00, V0.4), once again, equal within the errors. While the sample size is too small to provide a reliable item analysis by itself, it is still interesting to compare the question-by-question responses, which we show in Table IV.

If we combine the standard errors of the mean (in quadrature) for the two averages for each question, the difference between the response rates for the two samples are all equal to within at most 1.2 (combined) standard deviations, except for questions Nos. 21, 23, and 24; each of those were an-

Table IV. The average (and standard error of the mean) scores for each question for the QMVI (V0.4) for the Fa00 (UgQM1, Penn State) and Sp00 (Arizona) undergraduate data. The average scores for the two samples are 45.4±3.8 (UgQM1, Penn State) and 49.9±5.6 (UgQM2, Arizona).

No. Ave	UgQM1 (PSU) 45.4±3.8	UgQM2 (ARIZ) 49.9±5.6	No. Ave	UgQM1 (PSU) 45.4±3.8	UgQM2 (ARIZ) 49.9±5.6	No. Ave	UgQM1 (PSU) 45.4±3.8	UgQM2 (ARIZ) 49.9±5.6
1	88±05	78±10	11	14±07	00±00	21	37±10	86±10
2	66±10	56±17	12	45±09	47±05	22	28±10	50±10
3	53±10	53±14	13	79±09	67±16	23	08±06	50±12
4	86±07	61±13	14	16±08	36±14	24	05±05	36±11
5	96±04	97±03	15	32±09	47±12	25	11±04	17±10
6	87±07	75±14	16	63±10	50±15			
7	54±11	67±14	17	14±07	33±13			
8	26±08	28±14	18	75±09	89±10			
9	49±08	33±12	19	54±09	61±14			
10	39±11	25±12	20	12±07	06±05			

swered significantly better (at the 2.6–3.5σ level) by the second-semester Arizona group. This could be due, for example, to the additional coverage provided by the longer, two-semester course. If we calculate the average score for only the first 20 questions (excluding those questions, 21–25, which we have described above as covering ''...*less traditionally seen material*...'' such as multidimensional wave functions and time-dependent phenomena) the resulting averages are even closer, namely 40.2±4.8 (Arizona, Sp00, V0.4) versus 41.9±3.3 (UgQM1, Fa00, V0.4).

The agreement, overall and even at this level of detail, using these admittedly rather small data sets, is reassuring and further suggests that QMVI scores after a standard undergraduate quantum course may be similar and a reasonable measure of students understanding of the material we wish to test.

V. IMPLICATIONS FOR QUANTUM MECHANICS EDUCATION

While preliminary, some of the results from our individual item analysis do suggest that many students cannot apply or extend standard curricular material to some novel problems. Based on some of our results, we will be developing some new educational materials to address some of these topics.

One focus will be on the introduction of probability distribution ideas into the context of even classical mechanics problems to address student understanding of specific and focussed questions such as QMVI[3], but also to better address the general difficulty students have with the interpretation of potential energy plots. In this context, the use of projection of trajectory techniques[48] may prove useful: one benefit from such an approach is that it is just as easy to discuss classical probability distributions for momenta (the analogs of $|\phi(p)|^2$) as it is for position–space measurements. Many existing materials (both printed and electronic) can be used, as well as more novel approaches such as the ensemble measurement approach illustrated in QMVI[4] which makes direct connections to experimental determinations, binning of data, direct normalization of probability distributions and the numerical evaluation of average values (which could have an impact on such questions such as QMVI[7]). Many of these ideas can be easily incorporated into or tested in the setting of existing courses on classical mechanics as well, providing materials which are typically not covered in standard undergraduate textbooks at this level, which can, in turn, help motivate students at an earlier stage of their undergrad career to consider probabilistic descriptions of motion which can then be built upon in later classes in quantum theory. Such classical probability ideas might also prove useful to related areas such as statistical mechanics.

A second set of materials will focus on the French–Taylor type problems, incorporating many excellent existing ideas from available computer software and printed materials to discuss, both in an intuitive and in a more formal manner, the many connections between classical motions, the potential energy function, and shape (wiggliness and magnitude) of quantum wave functions. Clearly, extensions into more than one dimension in a variety of different geometries (as in QMVI[20] and [21]) will be valuable.

One set of materials, which is already under development, will focus, in some detail, on the relatively simple, and very focused, problem of an asymmetric infinite well,[49] that is, investigating the bound state properties of a quantum mechanical system described by a potential of the form

$$V(x)=\begin{cases}\infty & \text{for } x<-a,\\ 0 & \text{for } -a<x<0,\\ +V_0 & \text{for } 0<x<+b,\\ \infty & \text{for } +b<x.\end{cases}$$

This problem has the benefit that the level of mathematical sophistication required to solve the Schrödinger equation analytically and to match the boundary conditions is no higher than that required for the standard treated problem of the finite well. The variation in potential implies that conceptual analyses of the wave function form (amplitude versus wiggliness) on each side of the well, *à la* French and Taylor can be easily understood (but one should note that there are surprises in a number of special cases[49]). The explicit form of the Schrödinger equation solutions, real or complex exponentials or sine/cosine versus cosh/sinh functions, depending on whether the energy eigenvalue E satisfies $E>+V_0$ or $E<+V_0$, can be understood by most students. The need to have both $e^{-\kappa x}$ and $e^{+\kappa x}$ solutions in the $0<x<+b$ region for tunneling cases where $E<+V_0$ and the fact that the appropriate boundary conditions at $x=-a,+b,0$ are all equally important should have an impact on student difficulties with QMVI[10] and [11]-type question. Momentum–space analyses of the problem are also possible to make contact with problems such as QMVI[20] and one can easily extend this problem (by eliminating the infinite walls at $x = -a, +b$) to the standard scattering from a step-potential probed in QMVI[14]. Phrased in this context, students will be encouraged to make use of more intuitive ideas about wave function properties for scattering problems, but which can just as easily make contact with more familiar ideas such as probability flux.

Finally, a set of materials dealing with many aspects of the time development of quantum systems is underway, motivated especially by the poor results, even for advanced students, on QMVI[23], [24], and [25]. While standard wave function spreading (using a Gaussian as an analytically calculable example) is discussed in many texts, the ability to visualize these effects for a variety of other shapes will clearly be useful. Scattering geometries can easily be included, including such simple changes as bounces from infinite walls,[26,50] which can focus attention on the time development not only of the wave packets themselves, but also of their expectation values which have clear classical connections.

A very important model system which exhibits an incredibly rich array of quasiclassical and purely quantum effects is that of wave packet revivals in the infinite square well.[7–9] The simple question of how quasiclassical periodicity is exhibited,[8,9] how the wave packet spreads within the enclosed area, and especially how it reforms during the revivals, is a fascinating interplay of classical and quantum ideas, simple mathematical methods and numerical calculations, and the springboard to comparison with real measurements of wave packet revivals in more physical systems.[6] Discussions of the simple time dependence of a two-state quantum system which can then be extended to the more delicate interplay between the many stationary state components of a bound state wave packet can be addressed with a wide variety of methods.

VI. CONCLUSIONS, DISCUSSION, AND OUTLOOK

We have developed an assessment instrument which we hope will be able to shed light on the development of students' conceptual and visual understanding of quantum mechanics. Based on the (admittedly limited) data we have obtained (from our own institution and one other) during the test period, we believe that we have identified clear differences in understanding between the sophomore-level modern physics, junior–senior level quantum theory, and first year graduate student level of competencies. For example, using the Fa00 UgQM1 and GrQM1 (and combined Fa00/Sp00 ModPh) data as an example, we find differences in QMVI scores of 17.8 ± 4.5 (sophomore to jr/sr level) and 10.1 ± 5.1 (jr/sr to grad level) between the three levels of instruction. Additional data from the University of Arizona are consistent with the advanced undergraduate and graduate results we have studied, often in great detail, as discussed in Sec. VI E, further suggesting an increase in ability as students progress through the standard physics curriculum. One set of data (the Sp00 UgQM1 results) suggests that instruction using materials designed to focus on questions of conceptual understanding and visualization can increase competencies in these areas. Preliminary item analyses (as in Sec. IV B) can be used to focus attention on areas where student understanding remains weak after traditional instruction, even at the graduate level.

Given the preliminary results reported here, based on limited data from only two institutions, one of our goals is to continue the development and testing of the QMVI questions in a variety of settings, obtaining as much data as possible on the current (or any future) version of the QMVI.[51] We are continuing to solicit faculty colleagues from a variety of other US colleges and universities to offer the QMVI to appropriate classes to gain more data on the variability of results in different curricular settings, using different textbooks, and pedagogical approaches. Any interested reader who would like to participate in some aspect of this study by offering the QMVI at their institution (in any of the courses discussed above, or others) is encouraged to contact the second author via electronic mail. Clearly, given the large content domain of quantum mechanics, even at the level of modern physics, further refinement of the questions based on a wide variety of student responses will be very important to obtain.

We also hope to obtain data in other ways, such as with more detailed analyses of the written responses to the QMVI,[52] with interviews of students who have taken the QMVI, as well as offering the test both in a pre-instruction mode as well as the post-instruction mode which we have had available so far to assess, more quantitatively, possible gains on the QMVI. Variations of the QMVI which might be suitable for delivery on the web, possibly including more interactive questions, or even animated versions of some of the existing questions, are also under development. And finally, one of the main focal points will be the development of instructional materials (web-based) to address some of the student difficulties we have identified, and the testing of these modules.

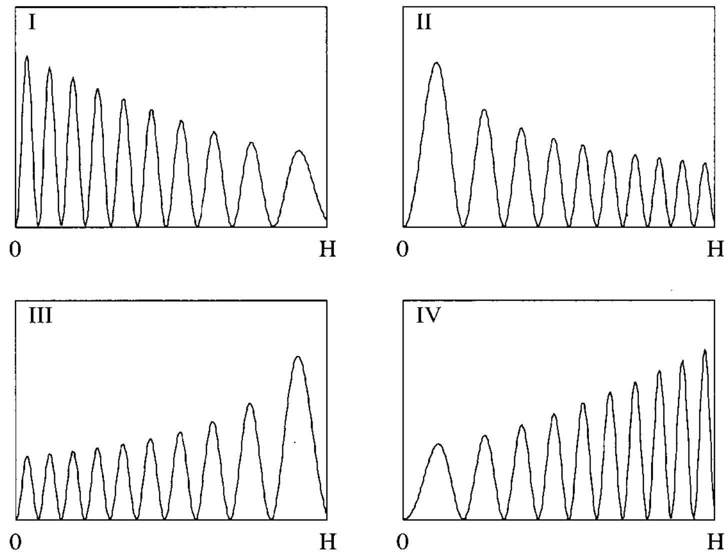
ACKNOWLEDGMENTS

We appreciate the suggestions offered by various physics and chemistry faculty at Penn State University and their cooperation in offering versions of the QMVI to their students; we specifically wish to thank M. Cole, J. Collins, S. Coutu, W. Ernst, S. Hammes-Schiffer, R. Herman, and M. Strikman. We are also grateful for helpful suggestions made by PSU College of Education colleagues, especially V. Lunetta. We have also benefited from conversations with E. Redish and D. Zollman. We especially thank I. Novodvorsky for her assistance in obtaining student data related to the QMVI from the University of Arizona. The work of R.W.R. was supported, in part, by NSF Grant No. DUE-9950702.

APPENDIX A

We reproduce here, as examples, two pages from the latest version (V0.4) of the QMVI indicating the standard format of the questions.

$|\psi(z)|^2$ vs. z



12. A particle is dropped from a height H under the influence of gravity and bounces, without loss of energy, from a flat surface (at $z = 0$) back up to the same height. Which of the plots above would be the best representation of the quantum mechanical position-space probability density, $|\psi(z)|^2$ versus z, of an energy eigenstate for this system.

(a) *I*

(b) *II*

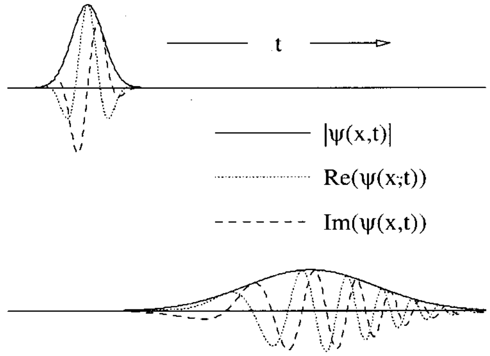
(c) *III*

(d) *IV*

(e) None of them are possible solutions of this problem.

↑ Explain your answer in 1-2 sentences in the space above. ↑
↓ Circle the statement below which describes how you feel about your answer. ↓

very certain	somewhat certain	somewhat uncertain	very uncertain



24. The real (dotted) and imaginary (dashed) parts of a position-space wave function representing a free particle wave packet, along with the modulus or absolute value (solid) of $\psi(x,t)$ is shown in the figure above. As time progresses, and the wave packet spreads, the real and imaginary parts of the wave packet appear to be more 'wiggly' in the leading edge of the wave than in the trailing edge. Which statement below best expresses what is shown in the figure above?

(a) All position-space wave functions spread with time because their real and imaginary parts have different phase velocities.

(b) All position-space wave functions spread with time because their real and imaginary parts get increasingly 'out of phase' with each other.

(c) The different momentum components which are used to construct a free particle wave packet travel at different speeds.

(d) As the spread in the position-space wave function, Δx_t, increases with time, the spread in the corresponding momentum-space wave function, Δp_t, must decrease with time to ensure that the uncertainty principle relation is maintained.

(e) The real and imaginary parts of any wave packet must spread in such a way that the modulus or absolute value $|\psi(x,t)|$ has the same **shape** as $|\psi(x,0)|$, but is just wider.

↑ Explain your answer in 1-2 sentences in the space above. ↑
↓ Circle the statement below which describes how you feel about your answer. ↓

very certain	somewhat certain	somewhat uncertain	very uncertain

APPENDIX B

We collect, in Table II, a question-by-question item analysis of the V0.4 QMVI results for the combined Sp00/Fa00 ModPh, Fa00 UgQM1, Fa00 GrQM1, Fa00 QChem, and Sp00 UgQM1 data. The data presented here are obtained using the multiple choice plus written response combination described in Sec. III (2 points for the correct answer plus 0, 1, or 2 points depending on written response for a possible total of 4 points per question). Similar data using only the correct answer (4 points for a correct multiple choice response) are available for comparison at the web site mentioned in Sec. III.

[a]Electronic mail: exc18@psu.edu
[b]Electronic mail: rick@phys.psu.edu

[1]R. W. Robinett, Web-based quantum mechanics tutorials for undergraduates, NSF Award 9950702. For more information, updated periodically, go to www.ehr.nsf.gof/pirs_prs_web/search/default.asp. See also our own web site, http://www.phys.psu.edu/faculty/RobinettR/QM/QMVI/QMVI.html for downloadable copies of much of the background material generated by this study.

[2]E. Cataloglu, "Development of an achievement test in quantum mechanics: The quantum mechanics visualization instrument (QMVI)," Ph.D. thesis, Penn State University, in preparation.

[3]D. Hestenes, M. Wells, and G. Swackhammer, "Force concept inventory," Phys. Teach. **30**, 141–158 (1992); D. Hestenes and M. Wells, "A mechanics baseline test," *ibid.* **30**, 159–166 (1992); R. K. Thornton and D. Sokoloff, "Assessing student learning of Newton's laws: the force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula," Am. J. Phys. **66**, 338–352 (1998).

[4]T. O'Kuma, C. Hieggelke, D. Maloney, and A. Van Heuvelen, "Developing conceptual surveys in electricity and magnetism," Announcer **28**, 81 (1998); "Preliminary interpretation of the CSE/CSM/CSEM student results," *ibid.* **29**, 82 (1999); "Some results form the conceptual survey of electricity and magnetism," *ibid.* **30**, 77 (2000).

[5]For example, in addition to the standard use of the infinite well as a bound state model, along with the Pauli principle, to help explain aspects of the

structure of solids and even nuclear matter, the infinite well can be used to discuss modern aspects of quantum mechanics ranging from supersymmetry [F. Schwabl, Quantum Mechanics (Springer-Verlag, Berlin, 1990)] to wave packet revivals (Refs. 7–9).

[6]See, e.g., M. Nauenberg, C. Stroud, and J. Yeazell, ''The classical limit of an atom,'' Sci. Am. **270**, 44–49 (1994).

[7]R. Bluhm, V. A. Kostelecký, and J. A. Porter, ''The evolution and revival structure of localized wave packets,'' Am. J. Phys. **64**, 944–953 (1996).

[8]R. W. Robinett, ''Visualizing the collapse and revival of wave packets in the infinite square well using expectation values,'' Am. J. Phys. **68**, 410–420 (2000).

[9]D. F. Styer, ''Quantum revivals versus classical periodicity in the infinite square well,'' Am. J. Phys. **69**, 56–62 (2001).

[10]N. S. Rebello and D. Zollman, ''Conceptual understanding of quantum mechanics after using hands-on and visualization instructional materials,'' *Research on Teaching and Learning Quantum Mechanics* (National Association for Research in Science Teaching, 1999), pp. 2–6.

[11]R. Steinberg, M. Wittmann, L. Bao, and E. F. Redish, ''The influence of student understanding of classical physics when learning quantum mechanics,'' *Research on Teaching and Learning Quantum Mechanics* (National Association for Research in Science Teaching, 1999), pp. 41–44.

[12]E. F. Redish and B. Lei, ''Student difficulties with energy in quantum mechanics,'' www.physics.umd.edu/rgroups/ripe/perg/quantum/aapt97qe.htm

[13]L. Bao, P. Jolly, and E. F. Redish, ''Student difficulties with quantum mechanics,'' www.physics.umd.edu/rgroups/ripe/ph421/bao/talks/qm9608.htm

[14]L. Bao, E. F. Redish, and R. Steinberg, ''Student misunderstandings of the quantum wavefunction,'' Summer AAPT Announcer **28**, 92 (1998).

[15]D. F. Styer, ''Common misconceptions regarding quantum mechanics,'' Am. J. Phys. **64**, 31–34 (1996); **64**, 1202(E) (1996).

[16]S. Vokos, P. S. Schaffer, B. S. Ambrose, and L. C. McDermott, ''Student understanding of the wave nature of matter: Diffraction and interference of particles,'' Phys. Ed. Res. Suppl. **68**, S42–S51 (2000).

[17]C. Singh, ''Student understanding of quantum mechanics,'' Am. J. Phys. **69**, 885–895 (2001).

[18]For example, in the famous textbook by L. Schiff, *Quantum Mechanics*, 1st ed. (McGraw-Hill, New York, 1949) of half a century ago, the figures illustrating the solutions of the simple harmonic oscillator problem are reproduced from an even earlier work, by L. Pauling and E. B. Wilson, *Introduction to Quantum Mechanics* (McGraw-Hill, New York, 1935) which, in turn, were drawn, by a draftsman, on graph paper.

[19]B. Thaller, *Visual Quantum Mechanics: Selected Topics with Computer-Generated Animations of Quantum-Mechanical Phenomena* (Springer-Verlag, New York, 2000).

[20]S. Brandt and S. Dahmen, *The Picture Book of Quantum Mechanics*, 3rd ed. (Springer-Verlag, New York, 2001).

[21]J. Bayfield, *Quantum Evolution: An Introduction to Time-Dependent Quantum Mechanics* (Wiley, New York, 1999).

[22]J. R. Hiller, I. D. Johnston, and D. F. Styer, *Quantum Mechanics Simulations: The Consortium for Upper-Level Physics Software* (Wiley, New York, 1995).

[23]S. M. McMurry, *Quantum Mechanics* (Addison-Wesley, Reading, PA, 1993).

[24]J. Jarecki, *Graphical Schrödingers Equation* (Physics Academic Software, Raleigh, 1998).

[25]For example, one of the first pedagogical papers describing the simplest aspects of wave packet propagation and interaction with potential barriers and wells was written by A. Goldberg, H. M. Schey, and J. L. Schwartz, ''Computer-generated motion pictures of one-dimensional quantum-mechanical transmission and reflection phenomena,'' Am. J. Phys. **35**, 177–186 (1967). The authors used one of the most powerful computers then available at the Livermore Lab (now dwarfed by modern PCs in speed) and had ''...the probability density projected on a cathode-ray tube. From the tube, photographs are made, and in turn are processed into the successive frames of a film.'' Typical software packages now allow students to reproduce these results with arbitrary initial conditions and potential parameters, almost instantaneously. See Ref. 26 for more references on time-dependent wave packet propagation.

[26]A rather comprehensive list of references to discussions of wave packet propagation in many model quantum mechanical systems, both for scattering and for bound states, in the pedagogical literature is contained in M. Doncheski and R. W. Robinett, ''Anatomy of a quantum 'bounce,' '' Eur. J. Phys. **20**, 29–37 (1999).

[27]*Scientific Visualization in Mathematics and Science teaching*, edited by D. A. Thomas (Association for the Advancement of Computing in Education, Virginia, 1995).

[28]B. H. McCormick, T. A. DeFanti, and M. D. Brown, ''Visualization in scientific computing,'' Comput. Graph. **21**, 14–18 (1987).

[29]F. M. Dwyer, ''The effect of over responses in improving visually programming science instruction,'' J. Res. Sci. Teach. **9**, 47–55 (1972).

[30]W. G. Holliday, ''The effects of verbal and adjunct pictorial-verbal information in science instruction,'' J. Res. Sci. Teach. **12**, 77–83 (1975).

[31]J. W. Rigney and K. A. Lutz, ''Effect of graphics analogies of concepts in chemistry on learning and attitude,'' J. Ed. Psych. **68**, 305–311 (1976).

[32]R. R. Gotwals, ''Scientific Visualization in Chemistry: Better Living Through Chemistry, Better Chemistry Through Pictures: Scientific Visualization for Secondary Chemistry Students,'' in Ref. 27, pp. 153–180.

[33]*GRE®: Practicing to take the Physics Test* (Educational Testing Service, Princeton, 1997). This contains three complete Physics GRE examinations as well as data on student responses.

[34]R. A. Serway, C. Moses, and C. Moyer, *Modern Physics*, 2nd ed. (Saunders College, Fort Worth, TX, 1997).

[35]K. S. Krane, *Modern Physics*, 2nd ed. (Wiley, New York, 1996).

[36]A. Beiser, *Concepts of Modern Physics*, 5th ed. (McGraw-Hill, New York, 1995).

[37]P. Tipler and R. A. Llewellyn, *Modern Physics*, 3rd ed. (W. H. Freeman, New York, 1999).

[38]T. A. Moore, *Six Ideas That Shaped Physics*, Unit Q, Particles Behave Like Waves (McGraw-Hill/WCB, Boston, 1998).

[39]D. J. Griffiths, *Introduction to Quantum Mechanics* (Prentice Hall, Englewood Cliffs, NJ, 1995).

[40]C. Cohen-Tannoudji, Bernard Diu, and Franck Laloë, *Quantum Mechanics*, Vols. I and II (Wiley-Interscience, New York, 1977) (Translated from the French by S. Hemley, N. Ostrowsky, and D. Ostrowsky).

[41]I. N. Levine, *Quantum Chemistry* (Prentice Hall, Upper Saddle River, NJ, 2000).

[42]R. W. Robinett, *Quantum Mechanics: Classical Results, Modern Systems, and Visualized Examples* (Oxford U. P., New York, 1997).

[43]A very similar question was asked in the 1996 Physics GRE (GR9677, Question No. 17) (Ref. 33). The correct response rate listed for that question is 40% from the student population (presumably mostly physics seniors intending to go to graduate school) taking the test.

[44]A. P. French and E. F. Taylor, ''Qualitative plots of bound state wave functions,'' Am. J. Phys. **39**, 961–962 (1971); *An Introduction to Quantum Mechanics*, M.I.T. Introductory Physics Series (Norton, New York, 1978).

[45]Question No. 97 of the 1996 Physics GRE (GR9677) has a listed correct response rate of 11% (Ref. 33).

[46]C. M. McCallum, ''Book and Media reviews,'' J. Chem. Phys. **74**, 343 (1997).

[47]R. L. Liboff, *Introductory Quantum Mechanics*, 2nd ed. (Addison-Wesley, Reading, PA, 1991).

[48]R. W. Robinett, ''Quantum and classical probability distributions for position and momentum,'' Am. J. Phys. **63**, 823–832 (1995); see also Ref. 42, pp. 114–115.

[49]Many of the background details necessary for the module under development are discussed in depth in M. A. Doncheski and R. W. Robinett, ''Comparing the classical and quantum probability densities for an asymmetric infinite well,'' Eur. J. Phys. **21**, 217–228 (2000).

[50]M. Andrews, ''Wave packets bouncing off walls,'' Am. J. Phys. **66**, 252–254 (1998).

[51]Unlike studies of student understanding of concepts related to such topics as introductory mechanics or E&M, where sample sizes of order $N\sim 100$ are often readily obtainable from large service courses, the enrollments in the kinds of courses we consider (ModPh-, UgQM1-, and GrQM1-type classes) are often 10–30 at many US institutions. It may well be that data from a number of different sites will have to be collected, compared, and integrated over in order to obtain similar sample sizes.

[52]The written responses have already been used extensively to assess detailed student understanding, especially during the test of the original V0.3 version, helping to suggest improvements which were implemented in V0.4. More detailed discussions of the written responses and many other issues will be included in an upcoming paper [E. Cataloglu and R. W. Robinett (in preparation)].