The Pennsylvania State University

The Graduate School

Department of Curriculum and Instruction

DEVELOPMENT AND VALIDATION OF AN ACHIEVEMENT TEST IN INTRODUCTORY QUANTUM MECHANICS:

THE QUANTUM MECHANICS VISUALIZATION INSTRUMENT (QMVI)

A Thesis in

Curriculum and Instruction

by

Erdat Cataloglu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2002

We approve the thesis of Erdat Cataloglu.

Date of Signature

Vincent N. Lunetta
Professor of Education
Thesis Co-Adviser
Chair of Committee

Richard W. Robinett
Professor of Physics
Thesis Co-Adviser

Peter A. Rubba
Professor of Education

Hoi Suen
Professor of Educational Psychology

Patrick W. Shannon
Professor of Language and Literacy Education
Coordinator for Graduate Studies in
Curriculum and Instruction

ABSTRACT

The purpose of this study was to construct a valid and reliable multiple-choice achievement test to assess students’ understanding of core concepts of introductory quantum mechanics. Development of the Quantum Mechanics Visualization Instrument (QMVI) occurred across four successive semesters in 1999-2001. During this time 213 undergraduate and graduate students attending the Pennsylvania State University (PSU) at University Park and Arizona State University (ASU) participated in this development and validation study. Participating students were enrolled in four distinct groups of courses: Modern Physics, Undergraduate Quantum Mechanics, Graduate Quantum Mechanics, and Chemistry Quantum Mechanics.

Expert panels of professors of physics experienced in teaching quantum mechanics courses and graduate students in physics and science education established the core content and assisted in the validating of successive versions of the 24-question QMVI. Instrument development was guided by procedures outlined in the *Standards for Educational and Psychological Testing* (AERA-APA-NCME, 1999).

Data gathered in this study provided information used in the development of successive versions of the QMVI. Data gathered in the final phase of administration of the QMVI also provided evidence that the intended score interpretation of the QMVI achievement test is valid and reliable. A moderate positive correlation coefficient of 0.49 was observed between the students’ QMVI scores and their confidence levels. Analyses of variance indicated that students’ scores in Graduate Quantum Mechanics and Undergraduate Quantum Mechanics courses were significantly higher than the mean scores of students in Modern Physics and Chemistry Quantum Mechanics courses ($p < 0.05$). That finding is consistent with the additional understanding and experience that should be anticipated in graduate students and junior-senior level students over sophomore physics majors and majors in another field. The moderate positive correlation coefficient of 0.42 observed between students’ QMVI scores and their final course grades was also consistent with expectations in a valid instrument. In addition, the Cronbach-alpha reliability coefficient of the QMVI was found to be 0.82, a relatively high reliability coefficient for an achievement test.

This study also provided data from which preliminary findings were drawn on students’ understanding of introductory quantum mechanics concepts. The data included limited information about students’ visual understanding of quantum mechanics concepts. Data suggested that the construct of quantum mechanics understanding is most likely multidimensional and the *Main Topic* defined as “Quantum Mechanics Postulates” may be an especially important factor for students in acquiring a successful understanding of quantum mechanics. Students’ difficulties with concepts including *momentum-space* and *time-dependence* were observed. Limited data on visualization indicated that students were able to relate their mathematical and verbal knowledge with visual representations of quantum mechanics concepts. However, data also suggested that students did not perform as well, in general, when they had to connect conceptual ideas with quantitative interpretation.

The results of this study suggest recommendations for further development of the QMVI, for development of understanding of introductory quantum mechanics, and for possible applications of the QMVI. The recommendations have potentially important implications for the teaching of introductory quantum mechanics and for the development of text and technology resources.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am indebted to many people at the Pennsylvania State University for their help in the completion of this work. The first expression of my deep appreciation goes to my thesis co-advisor and committee chair Dr. Vincent N. Lunetta, Professor of Science Education. Without Professor Lunetta's deep insights regarding science education that is truly embedded in theory and practice and his invaluable help and support for me in developing a multidimensional perspective in science education, it would have been impossible for me to be able to begin to understand the deep meaning of science education. He was a critical listener, a mentor, and an emotional supporter throughout my work. Without his help in editing, I could not have written this dissertation in my second language.

My deep appreciation also goes to my thesis co-advisor Dr. Richard W. Robinett, Professor of Physics. Due to his continuously high expectations and insightful inquiries into my academic work, I was able to apply psychometrics theory to practice in physics eucation. His gentle, positive and endless encouragement made this study and writing this dissertation most rewarding.

Dr. Hoi Suen, Professor of Psychometrics, provided me a powerful and clear theoretical perspective that guided my work, particularly my research methodology. That perspective grounded in psychometric theory and practical experience guided me throughout this entire study. Without the insightful course and the guidance of Dr. Peter Rubba, Professor of Science Education, it would not have been possible for me understand the critical aspects of my research. He taught me about educational research methodology issues that needed to be examined to ground this study in scientific methodology. Learning about qualitative research methodology from him, enabled me to comprehend and explore, especially in the early stages of study, the entire research agenda.

Special thanks are extended to Bugrahan Yalvac, Joseph Taylor, and Fatih Tasar for their assistance in the early stages of this study. Also, I would like to extend my appreciation to the student who volunteered their valuable time and made it possible to validate the QMVI.

My final words of appreciation go to a special person in my life; to my spouse Ummahan. Her everlasting love, passion for my academic growth, understanding, sacrifices, and support allowed me to keep going through some difficult times and complete this work.

CHAPTER 1

INTRODUCATION TO THE STUDY

1.1 The Purpose of the Study

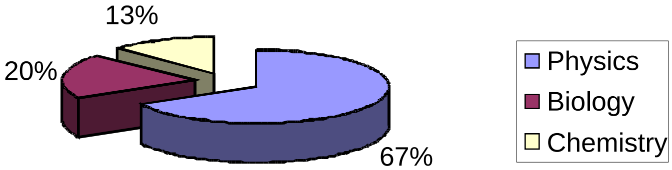
The principal purpose of this research was to develop a multiple-choice test to assess university students' understanding of introductory quantum mechanics. A rudimentary part of the research was to identify the core concepts that should be central in introductory university quantum mechanics courses. Since visualization of these concepts is now considered an important part of understanding in quantum mechanics, the test is designed to examine students' visualization of concepts as well as their understanding of more conventional mathematical and verbal representations. The Quantum Mechanics Visualization Instrument (QMVI) should be valid and reliable in assessing students' understanding of selected topics in quantum mechanics. The test includes assessment of understanding of selected visual concepts in quantum mechanics. A secondary goal is to examine students' understanding of more conventional quantum mechanics concepts.

1.1.1 Studies of Students' Understanding of Scientific Concepts

Since the early 1980s researchers have given special attention to students' understanding of science concepts studied in introductory science education. Early research focused especially on the understanding of concepts in physics, while more recent studies have examined multiple interrelated concepts both within and across science disciplines.

Pfundt and Duit (1991) listed more than 1000 studies that specifically addressed the issues of students' understanding of scientific concepts. Their reference list included studies that were published in refereed journals presented at conferences as well as unpublished studies. As shown in Figure 1.1, about 67% of these studies were within the domain of *physics* (including a few *earth and space* science studies). About 20% of the studies examined students' conceptual understanding were conducted in the domain of biology and about 13% in the domain of chemistry.

Figure 1.1 Distribution of studies of students' scientific concepts

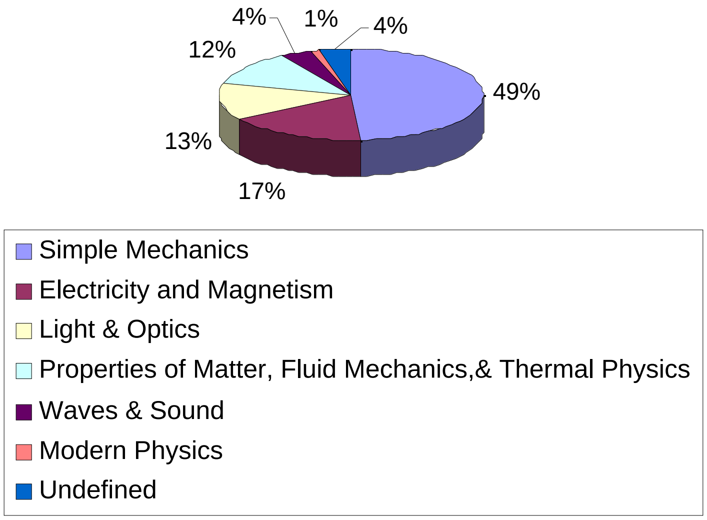


1.1.2 Studies of Students' Understanding of Physics Concepts

Wandersee et al. (1994) reported that of some 700 studies conducted primarily in the United States of America on high school or college students within the school subject of physics, about 300 (~43%) had been conducted on concepts in mechanics, including force and motion, gravity, velocity, and acceleration, about 160 (~23%) studies were conducted on students' understanding of scientific concepts in electricity, about 10% of these studies were each devoted to concepts of heat, optics, and the kinetic-molecular nature of matter and energy, and about 5% on physics concepts of the earth and space sciences. Only 10 (~1.4%) studies concerning students' understanding of concepts presented in modern physics, that is physics based on relativity and quantum theory, were conducted.

In a more recent publication McDermott and Redish (1999) reported a distribution of published studies that was similar to that of Wandersee et al. (1994). McDermott and Redish (1999) published a resource letter with annotated references to empirical studies, principally about students' understanding of science concepts. The letter also classified studies as “problem solving”, “the effectiveness of laboratory instruction and lecture demonstrations”, “the ability to apply mathematics in physics”, “students’ attitudes and beliefs”, and discussion of research into reasoning in physics education. As shown in Figure 1.2, their literature review of research on students’ understanding of scientific concepts in physics showed that about 49% of these studies had been conducted in mechanics followed by studies on “electricity and magnetism” with 17%. Studies on students understanding of “light and optics” and “properties of matter, fluid mechanics, and thermal physics” received almost the same proportion of attention by researchers, 13% and 12% respectively. Only about 4% of the research on issues of students’ understanding of scientific concepts in physics was devoted to “waves and sound”. Concepts in modern physics at 1% received the most limited attention in the literature.

Figure 1.2 Distribution on studies conducted on students’ understanding of scientific concepts in physics with respect to subject area.



As shown in Figure 1.2, investigations on students’ understanding of scientific concepts were mainly conducted in the domain of simple mechanics. Perhaps because introductory mechanics was traditionally the first domain of physics that was taught to students. Moreover, Wandersee et al. (1994) claimed that in fact studies on students’ understanding in introductory mechanics most probably were the most influential studies in many aspects in the area of students' understanding of scientific concepts research. Therefore, it seems to be appropriate to briefly summarize some important problem areas identified by these research studies.

A large number of studies (e.g., McCloskey et al, 1980; McCloskey, 1983 Watts, 1980; Clement, 1983; Clement, 1982; Minstrell, 1990; Minstrell, 1982; McDermott, 1998; Aguirre, 1989;) reported that a majority of students (junior-high school to college level) and graduate students even in some cases (e.g., Trowbridge, 1980; Trowbridge, 1981) had ideas about introductory mechanics that were different than the scientific concepts of experts. The term "alternative concepts", among many other descriptive terms, was mostly used to label students' incorrect concepts. Some common alternative concepts in kinematics, dynamics, rotational motion and gravity have been reported regularly in successive studies. For example, studies had revealed that students from a wide range of demographics have great difficulties in understanding very basic concepts such as the concept of force that was so central to dynamics. Common students' alternative concepts about the concept of force can be summarized as follows:

1. When a force acts on an object, then the object moves in the direction of the applied force.
2. In order to maintain constant velocity, a constant force must applied.
3. The velocity of the object is directly proportional to the applied force.
4. If the applied force is to be removed then objects slow down and eventually came to a stop.
5. If there no force is applied to an object, then the object is at rest (e.g., Clement, 1983).

US college students were the subject of McCloskey's (1983) study about curvilinear motion. The students were asked to predict by drawing the path of the ball after it was no longer constrained by a circular tube. Over half the students drew paths that showed that the ball would continue to move in a curved path. The students reasoned that the object gained such a "force" or "momentum" during its previous circular motion in the tube that it would continue to do what it was "trained" to do. However, some of the students believed also that this acquired "force" would eventually dissipate and afterwards the ball would move in a straight line. Caramazza (1980) also studied 50 US first year college students on curvilinear motion. 36% of the students in their sample had the same alternative concepts i.e., they also believed that after an object becomes free of circular constraint by a tube, it would continue on a curved path.

Some studies showed that students had alternative concepts of free falling objects and the relationship between their speed, weight and position. More specifically, students believed that heavier objects fall faster than lighter objects even when frictional forces were negligible (White, 1983; Watts, 1982).

These studies provide much evidence that large numbers of students who have completed introductory high school and college physics courses had alternative concepts in simple mechanics. In other words, these studies suggested that students left simple mechanics classes with a diverse set of alternative concepts. Similar conclusions were also reached in studying students' concepts of electricity, magnetism, light, sound, heat, and energy (see for example Driver et al., 1994). As a result of these studies, the field of science education has gained valuable information on students' understandings. This information enables science educators to develop and examine curriculum materials and instructional strategies to help students acquire more scientific concepts. Similar limitations are probably present in students' understandings of modern physics and quantum mechanics concepts, but to date, there have been insufficient studies of students' understanding in quantum mechanics to justify that inference with empirically (see Figure1.2.).

1.1.3 Visualization and Quantum Mechanics

In a publication issued by the Association of Computing Machinery, McCormick et al. (1987) wrote that visualization was a form of communication that transcended application and technological boundaries. The report suggested that visualization was a tool that could promote discovery, scientific understanding, and learning. Learning science was viewed as an active process in which students themselves construct meaningful understanding (Piaget,1972; Ausubel, 1968, Vygotsky, 1997). In this model of learning, knowledge is selected, discriminated, associated, and elaborated within a *previously existing* cognitive structure. Research pointed to the advantages of visualization in developing scientific understanding by providing additional opportunities for students to establish connections with preexisting knowledge structures (Brody, 1984, Dwyer,1972, Holliday, 1975, Rigney and Lutz, 1976). Gotwals (1995) pointed out that scientists made extensive use of visualization tools to translate data into visual images with the hope that consistent or inconsistent patterns will emerge and a new level of scientific understanding reached utilizing those images and visual models.

Many of the concepts and ideas that students encounter in modern physics and quantum mechanics courses do not lend themselves to students' natural intuition as do concepts in introductory mechanics (Driver, 1994). For example, students can experience the acceleration or deceleration of a car. They can conduct the Millikan oil drop experiment, observe effects, and compute the charge of an electron. . While students cannot "see" the charge on an electron, they can observe the effects of different amounts of charge in the relative velocities of the charged oil drops. However, the effects of concepts in modern physics and quantum mechanics can not be directly observed in the student's world of experience or even in the introductory laboratory. Thus, an important challenge for today is that many concepts in modern physics and quantum mechanics. if not most, are abstract and can only be described with complex mathematical expressions. Thus, engaging students with visual representations of quantum mechanics concepts, as is increasingly possible with new computing technologies, offers important contemporary opportunities for teaching and associated scholarship that may enhance conceptual understanding.

1.2 Assessment Instruments used to Investigate Students' Understanding of Scientific Concepts.

A detailed review of assessment instruments and strategies used to investigate students' understanding of scientific concepts will be presented in Chapter 2. However, it is appropriate to introduce and provide some preliminary information on how researchers generally have probed students' understanding of scientific concepts in this section.

Among the numerous instruments to probe students' understanding, *interviews* have frequently been used to investigate their understanding of scientific topics. (see, for example, the Handbook of research on science teaching and learning, 1994). Science education researchers have used a variety of interview techniques, for example the *clinical interview* technique involves a formal setting in which the student is confronted with a question, experiment, or task. The researcher has a limited interaction with the student during the interview period. A more open version of this interview technique is known as the method of *clinical exploration*. The researcher is free to have a dialogue with the student (Ginsberg & Opper, 1988).

Paper and pencil tests have been used regularly to assess students' understanding. Although interviews allow the researcher to probe a wide range of concepts, paper and pencil tests allow researchers to probe the understanding of many more students at a time. For example, Hestenes et al. (1992) developed a 29 item multiple-choice test, the Force Concept Inventory (FCI). The test was designed to determine whether students were able to distinguish between Newtonian and non-Newtonian concepts in introductory mechanics. The authors claimed that the FCI would classify students' alternative conceptions with respect to six Newtonian concept domains (kinematics, first law, second law, third law, superposition principle, and kinds of force). By 1998, the FCI had been administered to more than 6500 high-school and college students Hake, 1999. It should be noted, however, that some substantial criticisms have been raised about the reliability and validity of the FCI (Heller & Huffman, 1995; Hestenes & Halloun, 1995; Huffman & Heller, 1995). The criticisms are due, in part, to the failure of the developers to construct the FCI in accordance with well-established test development standards such as the *Standards for Education and Psychology Testing* (AERA-APA-NCME, 1999). A more detailed discussion of this topic follows in chapter 2.

1.3 The Significance of the Study

The development of a carefully designed, multiple-choice instrument to assess students' understanding of selected topics in quantum mechanics incorporating visual representations is important for several reasons. First, extending physics education research on conceptual understanding to a different domain of physics in a more advanced course is important. By assessing basic and more elementary concepts in physics, researchers have found that high school and first year college physics students in the USA and in many other countries have difficulties in understanding basic concepts presented in their formal classes (Wandersee et al., 1994). While conceptual understanding of basic physics concepts has been recognized as important to reach, research has shown clearly that many students may be able to pass their elementary physics courses successfully while retaining alternative ideas that are not the scientific concepts presented in those elementary courses Driver, 1994. However, little research on students' understanding of quantum mechanics and related test construction has been conducted. Therefore, there is substantial need for research on students' understanding in quantum mechanics and for the development of a valid assessment instrument that can support that research and the development of more effective teaching.

The Force Concept Inventory (Hestenes et al., 1992) and the Mechanics Base Line test (Hestenes et al., 1992) are examples of similar efforts in introductory mechanics. Since these two tests have been utilized widely, much baseline data have been obtained (Hake, 1998). In the process, researchers have investigated the effects of numerous variables (e.g., gender) that had been hypothesized to be relevant in learning and teaching introductory mechanics.

Historically, visual representations of quantum mechanics concepts have not been emphasized. However, with the development of faster (and less expensive) computers, the world wide web, and powerful software (e.g., Matlab™, Maple™, Mathematica™) students can more regularly be engaged in interacting with visual representations of quantum mechanics concepts and students' understanding of such representations should also be assessed. Physics instructors will increasingly develop and use these high technology resources to teach the visual aspect of quantum mechanics with the hope of extending their students' experiences in learning modern physics and quantum mechanics and their holistic understanding of quantum mechanics concepts. Thus, developing a valid and reliable multiple-choice type of instrument to assess students' understanding of concepts in modern physics and quantum mechanics with an emphasis on visual representation is a very important and timely task.

Students enrolled in introductory quantum mechanics courses are usually physics majors who have advanced background knowledge in physics and mathematics. Therefore physics educators and others tend to assume that students who have passed these courses have mastered the fundamental quantum mechanics concepts presented in these introductory quantum mechanics courses. Based upon the research conducted on students' concepts in other fields of physics, there is reason to be very skeptical of such an assumption, but up to this time, there has been little or no research on this topic. Thus, the development of a valid instrument to assess students' understanding of the core concepts in introductory quantum mechanics with an emphasis upon visual representation should be an important step in filling a serious gap in the literature.

1.4 Definition of Terms

To understand and interpret the design and analysis of this study, it is important to define the use of the following terms with precision.

Achievement tests: "(1) examinations in individual courses of instruction in schools of all kind at all levels, (2) measures of achievement (course examinations) used routinely by all instructors in particular units, and (3) commercially distributed tests of achievement used thought the country" (Nunnally, 1978)

Alternative concept: an understanding of a scientific concept either partially or completely different from the currently accepted scientific explanation.

Concept: a class of things, objects, events, ideas, or relations to each member of which we give the same label" (Roid and Haladyna, 1982, p. 147).

Construct Validity: "focuses primarily on the test score as a measure of the psychological characteristics of interest" (Standard for Education and Psychological Testing, 1999, p. 9). In this study the construct is defined as students' understanding of core introductory quantum mechanics concepts.

Content Validity: "the degree to which the sample of items, task, or questions on a test are representative of some defined universe or domain of content" (Standard for Education and Psychological Testing, 1999, p. 10)

General purpose test: "a test that is intended to be employed very widely with diverse samples of subjects in numerous studies relating to applied problems or in basic research" (Nunnally, 1978, p.256)

Quantum Mechanics: "a collection of postulates based on a huge number of experimental observations and the tool derived from those postulates." (Morrison, 1990, p.2)

Reliability: "the degree to which test scores are free from errors of measurement" (Standard for Education and Psychological Testing, 1999, p. 19).

Test Item: "an instruction or question that requires a student response and a rule for scoring" (Haladyna, 1997, p. 10)

Test Item Difficulty: the proportion of examinees in a group who answered the item correctly.

Validity: the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores. Traditional validity evidence categories are content, criterion, and construct validity. An ideal validation includes documentation of all three of the traditional categories. (Standard for Education and Psychological Testing, 1999, p. 9)

Visualization: The use of computer graphics to represent numerical data in two-dimensional and three-dimensional visual images.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This Chapter consists of two major sections. The first major section (2.1 - 2.5), reports on the historical development of quantum mechanics, teaching and learning of quantum mechanics, and studies of students' understanding in science education and particularly in physics education and its potential implications for introductory quantum mechanics courses. More specifically, the section in begins with a brief report of the historical development of the field of quantum mechanics and reports on the major experimental discoveries that led to the formulation of quantum mechanics theory. The second section in domain 1 explores conventional methods of teaching introductory quantum mechanics and the third section in domain 1 discusses various critiques that have been raised about conventional teaching approaches in quantum mechanics. Finally, the fourth section in domain 1 reports on recent attempts to restructure conventional quantum mechanics courses to promote greater conceptual understanding.

The second major section (2.6 – 2.9), discusses procedures for the construction of achievement tests. More specifically, section 2.6 in domain 2 examines test development and validation of test instruments. Three types of validity, namely content, criterion-related, and construct validity are reviewed. Section 2.7 in domain 2 deals with test reliability and describes the various types of test reliability. The last section (2.8), reports on objective tests utilized in science and physics education and discusses possible problems such objective test might have because these knowledge are highly relevant to the research undertaken in this research.

2.2 The Historical Development of Quantum Mechanics

At the beginning of the twentieth century, physicists were discovering new phenomena that were inconsistent with and not satisfactorily explained by the theories provided by classical physics. X-rays, photoelectricity, radioactivity, black-body radiation, and relativity are examples of discoveries in the late nineteenth and early twentieth centuries that were not explained by classical physics. Theories in classical physics led physicist to predict that there were free moving electrons in the space between atoms. Accordingly, electromagnetic waves "hitting" a metal surface were expected to excite the free electrons, and eject electrons near the surface of the metal when given enough energy. Theories in classical physics predicted that electrons emitted in this way, however, would be expected to have high kinetic energy with intense incident light and low kinetic energy with less intense incident light. In other words, the theories in classical physics suggested a directly linear relationship between the emitted electron's kinetic energy and the intensity of the incident light. However, experimental evidence indicated that the electrons' kinetic energies were independent of the intensity of the light. On the other hand, the electrons' energies were depended on the frequency of the light. Although the experimental results suggested that the energy absorbed by the electrons came directly from the incident light, no plausible explanation was available until Einstein in 1905 introduced the idea of energy quanta, meaning that the interchange of energy between matter and radiation took place in energy quanta. This idea not only implied that light energy was emitted in packets or quanta, but also meant that light waves "were" energy packets that are absorbed as units. The essence of this idea is that light was particulate in nature. However, interference, polarization, and the electromagnetic theory suggested a wave nature of light. Such new ideas, introduced under the umbrella of Quantum Mechanics, required major conceptual adjustments resulting in the development of ideas now associated with modern physics. Accordingly, atomic and subatomic particles were found to have both - wave like and particle like- properties.

2.3 Conventional Approaches to Teaching Quantum Mechanics

Conventional teaching of quantum mechanics generally involves one of two distinct approaches to introducing the students to the basic concepts of quantum mechanics. The first approach to introducing quantum mechanics can be described as a quantitative approach through which the students are introduced to the mathematical algorithms and processes used to solve quantum mechanics problems and in the process they are to become acquainted with the mathematical tools needed to understand quantum mechanics concepts. "The level of rigor [of mathematics] is what I think is needed to make a practicing quantum mechanic [sic] out of the student" (Shankar 1994, p. ix). Shankar strongly expressed his beliefs that students have to be introduced first to the relevant mathematical skills so that later the students "can give quantum theory their fullest attention without having to battle mathematical theorems at the same time" (p. x). A similar concern was shared by McMurray (1993). She wrote that "students should come to grips with the basic mathematical ideas that are necessary for proper understanding of quantum theory" (p. vi).

The idea of teaching the students the relevant mathematics first is not unique to quantum mechanics courses. Similar pedagogical approaches have been used in other domains of physics. For example, college freshmen students enrolled in a typical calculus based introductory mechanics course first learn about measurement and vectors. Both of these topics are mathematical in nature. Cohen et al (1978) studied the relationship between students' mathematics achievement and their physics achievement. They found out that students mathematics scores correlated highly with their physics scores. Other research conducted in this area, however, showed that mathematical skills are only one of several variables prerequisite to the understanding of physics concepts presented in a typical introductory mechanics course, and high scores on mathematics tests are not sufficient indicators of conceptual understanding in physics (Champagne, 1980; Hasmetoglu, 1994; Hudson and McIntire, 1977). Studies of students' understanding in introductory mechanics have also shown that students' understandings were highly fragmented and compartmentalized. That is, scientific concepts which have definite relationship in science (e.g., force and acceleration) are loosely, if at all, linked in students’ cognitive structure. Therefore, the acquired knowledge in mathematics might not transfer as easily as some have assumed with the quantitative teaching approach to the topics taught later in the physics course. The results of cognitive research on transfer of learning have indicated similar findings, thus supporting the findings in science education on transfer of learning. More specifically, cognitive research has shown that when a student was taught only in one subfield of a discipline, transfer of knowledge to related subfields was difficult and cannot be assumed to take place involuntarily. (Bransford et al, 1999 p. 50).

The second approach to introducing students to the basic concepts of quantum mechanics can be expressed as an historical-conceptual approach. In this approach, initial importance is given to the history of the development of the historical experiments, concepts, and theories that have led to the theory of quantum mechanics. The historical-conceptual approach to teaching introductory quantum mechanics, incorporates the "historical-development" of quantum mechanics and introduces students to challenges similar to those faced by physicists in the early 20th century. According to Liboff (1980) a review of the historical development of quantum mechanics and elements of classical mechanics are "important to a firm understanding of quantum mechanics" (p. vii). Kuwalski (2001) takes a similar position by emphasizing that the development of scientific knowledge is a logical process. He reasons that omitting or bypassing certain steps of the historical development would hinder students' conceptual growth and understanding in quantum mechanics.

Studies of students' conceptual understanding (e.g. Driver, 1985) in science education have reported that some students' misconceptions were similar in nature to those driving the development of various science fields by former members of the scientific community. For example, studies on students' conceptual understandings in electricity and magnetism showed that an important number of students viewed electricity as a "fluid", and a similar concept was held by the scientists in the late 18th century. Similar findings were obtained in the area of introductory mechanics as well. Research showed that many students explained motion in ways that were similar to

Aristotle’s explanations. For example, students believed that heavier objects would fall faster than lighter objects Sequeira and Leite (1991). Consequently, the similarity between students' conceptual structures and past scientific explanations led Wandersee (1985) to suggest that knowledge of the history of science could help science educators and science teachers anticipate the topics which the students would be likely to develop serious misconceptions.

2.4 Questioning the Practice of Teaching Quantum Mechanics

Recently, several physicists have expressed dissatisfaction with traditional teaching practices in quantum mechanics. For example, they have questioned the necessity of intense, high level mathematics in developing a conceptual understanding of topics, ideas, and principles that are basic in introductory quantum mechanics.

Taylor (1998) the 1998 AAPT Oersted Medal recipient pointed out that there is a need for the physics education community to reconsider the pedagogy used in teaching introductory quantum mechanics. He suggested that the core concepts of introductory quantum mechanics should be introduced at an earlier stage than is the current practice, indeed, as early as a college student has completed an introductory physics course. According to Taylor, an introductory course in quantum mechanics could be taught "without tensors, without differential equations, without formalism - and definitely without the Schrödinger equation!" (p. 369). "Anyone with a mastery of basic calculus and an introductory physics acquaintance with momentum and energy can now explore the boundaries of Nature" (p.369) by explaining that there are examples, although few in number, that can be used to introduce the core concepts of introductory quantum mechanics without introducing complex mathematical equations. Besides advocating for new curriculum development projects in introductory quantum mechanics, Edwin Taylor questioned the composition of the student audience in quantum mechanics courses as well. He argued that introductory quantum mechanics courses should not be perceived only as courses for students majoring in physics. He argued that the physics education community in the United States needs to broaden its student audience for introductory quantum mechanics courses and include students from other sciences, mathematics, and engineering.

Taylor is not alone in articulating these concerns and suggestions. Hobson (2000) and Howes (2000) have expressed similar positions. Howes, for example, also questioned the current student audience of physics courses. She implied that the physics community had taken an elitist position by restricted the student population who receives the "exciting stuff" at the cutting edge of modern physics. She suggested an elitist position perhaps also influenced the general pedagogy of physics education as well. As a result of this elitist view, she argued that non-physicists often understand physics to consist only of basic mechanics and electricity. She suggested that faculty need to modify these perceptions and find new ways to incorporate modern physics into introductory courses especially for students who are majoring in other sciences.

Hobson (2000) pointed out that relativity and quantum mechanics, ironically labeled as “modern physics” actually were developed at the beginning of the 20th century, now a hundred years ago. Yet, most science students and certainly non-science students never get the opportunity to learn topics treated in advanced modern physics courses. Only a relatively few number of students receive the opportunities to study important and interesting topics in modern physics such as "time, space, matter, cosmology, radiation, energy continuity, observation, causality locality and physical reality itself." (p.388). The real disappointment according to Hobson, is that students are taught to solve problems only in the Newtonian paradigm. "If the course is a good one they [students] learn the principles of the seventeenth-, eighteenth- and nineteenth-century physics. But classical physics and the classical world view that students tend to unconsciously absorb along with the physics are seriously and fundamentally flawed and far from our contemporary description of nature" (p. 388) argues Hobson (2000). He went on to argue that quantum mechanics should be introduced at earlier stages of college education and to a wider range of students than is the current practice.

2.5 New Approaches in Teaching Introductory Quantum Mechanics

In the previous section, experts in physics education have expressed legitimate concerns about current practices and perceptions in the teaching of introductory, non-relativistic quantum mechanics. In summary, these experts have pointed out that it would be possible to introduce the students to important topics and concepts treaded in introductory quantum mechanics without requiring prior advanced mathematical knowledge. Moreover, the target audience should be broadened to introduce students from other sciences to introductory quantum mechanics courses. Finally, they have pointed out that the nature of the teaching of conventional introductory quantum mechanics courses can introduce misconceptions.

New ideas exist to restructure traditional introductory quantum mechanics courses and establish new ways to introduce a broader range of students to topics in introductory quantum mechanics in a more meaningful way. In a continuing effort to reconstruct the conventional teaching practice of introductory quantum mechanics, Roussel (1999) searched for new ways to improve conceptual understanding in quantum mechanics. He intended to put the actual emphasis on concepts in introductory quantum mechanics rather than on the application of complex mathematics used by experts and frequently emphasized in typical introductory quantum mechanics courses. By introducing the computer as a problem solving and *visualization* tool, Roussel was able to provide more authentic problems that approximated realistic quantum mechanical systems. His redesigned course deviated substantially from a conventional course. He demonstrated that it was possible to reduce the technical mathematical skills necessary for the students to learn introductory quantum mechanics, by introducing powerful computer mathematics software that are readily available in most universities. However, fewer topics were included in the redesigned curriculum. For example, the topic on the *perturbation theory* was dropped from the curriculum. Less time was devoted to lecturing sessions as well, whereas more opportunities were provided for collaborative learning with computer software. During the collaboration sessions, the students consulted with each other or with the instructor who assumed a tutor role. Moreover, using computers as a learning tool enabled the students to visualize useful quantum mechanics concepts that Roussel claimed enhanced the learning of such concepts. (Roussel, 1999).

Cuppari et al. (1997) designed a course in which the core topics of quantum mechanics were introduced by considering models from classical mechanics such as harmonic oscillator (e.g. the periodic motion of a mass and spring) and then introduced Plank's constant as the variable of the model. This approach involved four steps. First, the students used classical examples of periodic motion. Second, the Planck constant h was introduced as the quantum of action. Third the limits between classical and quantum mechanics were discussed. Fourth, the energy quantization in microscopy systems was discussed qualitatively. Although a multiple-choice test was used to measure students' understanding after the treatment, no statistical results were reported. However, the authors reported that the results were encouraging and showed that it was possible to introduce core concepts that helped students gain understanding in quantum mechanics without utilizing complex mathematical equations.

Wise and Kelley (1977) made use of graphical representations to avoid the mathematical difficulties students usually experienced. Their goal was to develop visualized concepts of quantum mechanics and to present wave-particle duality as a unified picture rather then as a paradox. This allowed also the use of a more statistical interpretation of quantum mechanics rather then treating a single particle as if it had a spread in momentum and then constructing a wave packet to represent it. There were three stages in their course: a) construction of the uniform distribution of particles from the momentum distribution of the Fraunhofer pattern, b) extension of the method to distributions with non-zero average momentum, and c) incorporation of time development in the spatial distributions. The authors concluded that graphical representations were effective in conveying core ideas of quantum mechanics, although these were normally considered too sophisticated mathematically.

Hadzidaki, Kalkanis, & Stavrou (2000) explained that the birth of quantum mechanics introduced a new worldview or paradigm in physics. Thus, the authors argued that conceptual understanding of topics in quantum mechanics requires that students develop a new way of thinking about the physical world. The authors criticized contemporary physics education for failing to address this problem. They advocated reform in teaching quantum mechanics that would involve more qualitative teaching approach based on epistemological and pedagogical foundations rather than on mathematical problem-solving methods. In this way the basic ideas of quantum mechanics can be introduced much earlier than they are normally addressed in contemporary courses. They argued that a qualitative approach could a) limit the development of serious misconceptions of modern physics topics, b) introduce the core concepts of quantum mechanics to a wider range of students than would otherwise have the opportunity to learn these topics, c) help the students to form a more comprehensive understanding that is compatible with contemporary scientific understanding of the macroscopic and microscopic world.

2.5.1 Visualization

The previous section briefly reviewed dissatisfaction that has been expressed about teaching of introductory quantum mechanics and new approaches in teaching important topics in non-relativistic quantum mechanics courses. The emphasis was on utilizing new interactive technologies because computer based technologies hold great promise both for increasing access to knowledge, to dynamic visualization of concepts, and for promoting learning. For example, technologies can help students to visualize difficult to understand concepts such as the wave behavior of a particle and the dynamic representation of complex physical phenomena.

Research has pointed out the importance of visualization in learning and teaching and developing scientific understanding. These suggestions emanate in part from principles of the relatively new sciences of learning, such as in cognitive psychology and neuroscience. By utilizing non-invasive imaging technology (e.g. positron emission tomology) neuroscientists have explored and were able to clarify some of the learning mechanisms that took place in human and animal learning. For example, neuroscientists have discovered that visual experiences have a direct effect on the physiology of the central nervous system. The Visualization in Scientific Computing report (1997) stated that an estimate of 50% of the neurons in the brain are associated with vision. Furthermore, numerous studies in mathematics education and science education have shown the importance of visualization in teaching and learning of mathematics and science concepts. For example Aiken (1971), and Battista (1991) found that there was a positive relationship between visualization and achievement in mathematics. Researchers also documented a positive correlation between mathematical aptitude and visualization (Battista, 1999).

2.5.2 Visualization in Mathematics Education and Science Education

Interactive computer graphics and graphing calculators are among the new technologies that can greatly expand the scope and power of visualization in mathematics education and science education. Demana, Schoen, and Waits (1993) have reported on several recent research studies that examined effects on learning visualization and graphing technologies at the secondary and college level. A few of these studies discuss recent findings in educational research in the domain of visualization that have the potential to enhance students' conceptual understanding in quantum mechanics courses.

For example, Rich (1990) found that students who used graphing calculators for the entire year of precalculus were far better able to deal with issues of scale on a graph when compared with students in a more traditional instructional setting. Rich also found that students who were taught precalculus using graphing calculators better understood the connection between an algebraic representation and its graph. Moreover, the finding revealed that the students viewed graphs more globally and that they understood the importance of a function's domain, its asymptotic behavior, and its end behavior. Rich also found that precalculus students in traditional instruction made almost no use of graphs except in the units dealing explicitly with graphs in the courses. Browning (1988) found that high-school precalculus students who used graphing calculators for one year exhibited a significantly increased ability to deal with graphing in more advanced levels such as Van Hiele levels of analysis and ordering. Farell (1989) also observed that precalculus students who used graphing calculators demonstrated greater ability with higher-order thinking skills than traditional students. Dunham (1990) found that in college algebra classes in which students were required to use calculators, gender-related differences in performance on graphing items were reduced, while pretest performance on graphing items indicated that females had scored consistently lower than their male counterparts.

2.6 Research on Students' Understanding in Science and Physics

The purpose of this section is to review the relatively large body of literature on research on students' understanding in science and in physics. As it was pointed out earlier in Chapter 1, there has been a much international interest on students' understanding of ideas, concepts, and theories taught in science courses. Many studies have revealed that students at different ages with different backgrounds come to science courses with well established ideas that have a powerful influence on what they learn. The students’ ideas, in many cases, are in marked contrast with the currently accepted scientific concepts that students are expected to learn in these science courses. Abimbola (1988), Gilbert and Swift (1985) have built a strong case for using the term alternative conception to refer to these type of students' understanding. The term refers to experience based explanations constructed by a student to explain a range of natural phenomena plausibly. The term also implies that alternative conceptions are contextually valid and rational from the stand point of the student

As a result of students' alternative conception research or "Alternative Conception Movement" research (Millar, 1989) the following seven claims about *alternative conceptions* are widely accepted

1. Students come to science instruction with well-established ideas concerning natural phenomena;
2. Students alternative conceptions cut across age, ability, gender and nationality;
3. Students' alternative conceptions are resistant to change by conventional teaching strategies;

4. Students' alternative conceptions often parallel explanations of natural phenomena offered by previous generations of scientists;
5. Students' alternative conceptions have their roots in personal experiences, perception, culture, language, school and schooling;
6. Teachers have alternative conceptions that are similar to those of their students;
7. Students' alternative conceptions conflict with the knowledge presented in formal instruction.

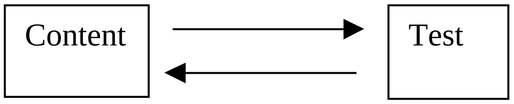
2.7 Objective Test Development & Validation of an Instrument

The procedure of determining the validity of an objective test is a very important task because "validity has always been regarded as the most fundamental and important in psychometrics" (Thorndike, 1971). Ideally, objective test construction needs to accumulate and establish evidence of *content*, *criterion-related*, and *construct* validity. Every effort is needed to be made to determine validity. However in his essay titled the "Historical and epistemological bases of validity" Cronbach (1988) pointed out that full investigation of validation of an instrument might not be possible. The validation process is best seen as unending. Both philosophical and practical arguments justifying the unending nature of the validation process exists. In test construction, however, practical concerns need to be taken into consideration as well. For example, time and resources often limit the collection of evidence for purposes of validation. At a minimum, however, investigators need to establish at least one type of validity for each instrument constructed. Instruments serve three major functions: 1) establishment of a statistical relationship with a particular variable, 2) assessment of a specified universe of content, and 3) measurement of psychological traits. Corresponding to these three functions there are three types of validity: 1) predictive validity, 2) content validity, and 3) construct validity (Nunnally, 1978). For an achievement test, content validity needs to be established. For an aptitude test, criterion-related validity needs to be established and for a personality or cognitive style-type test construct validity needs to be established. Section 3.1.2 reports on the validation type and process as it pertain to this research.

2.7.1 Content Validity

The determination of content validity most often involves judgmental procedures. Content experts review items on an achievement test and render judgments about whether the items presented on the achievement test are actually measuring the conceptual and procedural knowledge in that particular subject. The process of content validation is an interactive one. It starts by defining the content topics establishing the basis for the development and selection of the test items for the instrument. Then, content experts of the test review each test question and judge its suitability in terms of coverage of specific content. The experts review each test question and render a judgment as to whether it measures the specific conceptual and procedural knowledge to be assessed by the test. Figure 2.1 shows the arrow pointing from the content to the test indicating that the content served as the basis for the development of the test. The arrow pointing towards the test plan shows the content validation procedure of the test.

Figure 2.1 Judgmental procedure of content validation



The judging of test items with respect to the content can be conducted in a number of ways. An ideal way would be to constitute a group of experts and ask them to decide which particular part of the test content an item measures. If a number of such experts, working independently, indicates that a particular test item is measuring a certain part of the test content and that question indeed had been constructed and marked to measure that part of the content, then this can be taken as evidence of the content validity of that item. This process is repeated for every item on the test. If a major percentage of the experts agree on what part of the content each item measures and this is in agreement with the test developers intentions, then one can conclude with a fair degree of confidence that the content validity of the test has been established. If, however, there are disagreements among the experts who are reviewing the items then those items need to be reexamined and some form of correction must be take place. This might involve modification of the item or even possibly discarding it.

The procedure above is perhaps the strict way of establishing the content validity of an achievement test. Modifications of this procedure can be developed which can also establish the content validity of a test.

2.7.2 Criterion-related Validity

The procedure for establishing the criterion-related validity of a test is to administer the test to a group of students and compare the data with an external independent variable that is collected for each student (the criterion). The criterion is being considered to be related to the construct being measured. For example, an earlier version of a valid and reliable test could be the criterion when constructing a new test about the same content. Criterion-related validity can be categorized as *predictive validity* and *concurrent validity*. Predictive validity is the extent to which the test scores can be used to predict the score from a criterion measurement procedure that will take at place at some future point in time (Suen, 1990, p. 141). Concurrent validity "is the relationship between the scores of the test and the criterion measure taken at the same time" (Suen, 1990, p. 141). The correlation coefficient (r) is then computed between scores on the test and scores on the criterion measure for the group. The correlation coefficient is referred to as a *validity coefficient* because it describes the relationship between a test and a criterion measure.

2.7.3 Construct Validity

The determination of construct validity of a measure frequently requires a number of studies. The two principal types of studies that usually conducted are: *studies of group differences* and *studies of correlates* of the test under study. Studies of group differences involve administering the instrument to groups who are expected, on the basis of theory developed from previous research, to score at different levels on the instrument and then to compare the results obtained with the expectations. If the results are consistent with expectations, evidence is accumulated that demonstrates the construct validity of the instrument. Studies of correlates involve statistical analyses to investigate the internal structure of a test. The aim is to provide empirical evidence between the theoretical construct and the internal structure of the test.

2.8 Reliability

Reliability concerns the extent to which a series of measurements are consistent. Bohrnstedt (1970) wrote:

> “What is meant by reliability? Perhaps the best synonym is consistency. If no true change occurs in a given attitude an individual holds, does the attitude scale consistency yield the same ordering for him relative to the others? If not, the scale is unreliable” (p. 83).

There are four empirical methods to estimate the reliability of an instrument: (1) the test retest, (2) the equivalent form, (3) the split-half, and (4) coefficient alpha (Suen, 1990).

2.8.1 Test-Retest

This method of estimating reliability corresponds most closely to the conceptual notion of reliability. The same instrument is administered to the same group of persons at two different times, and the correlation between the two sets of scores is computed. This coefficient, usually called the Pearson r, is the reliability estimate.

Shaw and Wright (1967) reported that "the test-retest has the advantage of holding constant the items used, thus eliminating unreliability due to difference between items, which occurs in the equivalent-forms method. It also has the advantage of acquiring only a single scale; no additional forms are needed" (p. 16).

According to Bohrnstedt (1970), the following are disadvantages of test-retest reliability:

1. Different results may occur depending upon the length of time between measurement and remeasurement. The longer the time interval, the lower the reliability estimate. For one thing, when the time interval is short, persons may remember how they answered statements on the first administration and thus may appear to be more consistent than they actually are.

2. A second problem with test-retest reliability estimates is that individuals' true scores have a greater probability of actually changing, the longer the time between test and retest. It is clear that if individuals have truly changed, a low test-retest correlation does not necessarily mean that the reliability of an instrument is low.

3. Another problem a researcher must face when using any test-retest procedure has been called the reactivity problem. Reactivity refers to the fact that a respondent's sensitivity or responsiveness to the variable under study may be enhanced by the measurement of that variable.

2.8.2 Equivalent-forms

This method of estimating reliability requires two forms that are considered equivalent. The two forms are then administered to a group of subjects, and the two sets of scores are correlated to obtain an estimate of reliability. "The major advantage of this method is that the effects of time interval and of responding to one scale upon the response to the other are minimized" (Shaw and Wright, 1967, p. 17).

A disadvantage of this method is that at least two forms are necessary and that the correlation coefficient reflects not only consistency of measurement, but also the degree to which the two forms actually do measure the same attitude. Therefore, the acceptability of this estimate of reliability depends upon the degree of equivalence of the two forms (Shaw and Wright, 1967). According to Bohrnstedt "it is practically impossible to find two completely parallel sets of items" (1970, p. 87).

2.8.3 Split-half

The split-half method estimates reliability by treating each of two parts of the instrument as a separate scale. The scales may be selected by treating odd-numbered items as one scale and even-numbered items as another, or by randomly selecting items to constitute the subscales. Regardless of the method of choosing the subscales, the reliability estimate is the correlation between the scores of the separate scales. The Spearman-Brown prophecy formula is applied to the obtained correlation to estimate the reliability of the total scale.

The split-half approach was criticized, first by Cronbach (1951). He wrote: “Instead of giving a single coefficient for the test, the procedure gives different coefficients depending on which items are grouped when the test is split in two parts. If one split may give a higher coefficient than another, one can have little faith in whatever result is obtained from single split . . . . Such a coefficient is a property of pair of tests, not a single test” (p. 298) . Again, because of the arbitrary way in which the halves are chosen, this procedure is generally not recommended for determining reliability (Bohrnstedt, 1970).

2.8.4 Coefficient Alpha

Nunnally (1978) defined coefficient alpha as "the expected correlation of one test with an alternative form containing the same number of items (p. 214). He wrote that coefficient alpha provides a good estimate of reliability in most situations, since the major source of measurement error is because of the sampling of content.

2.9 Objective Tests in Science Education and Physics Education

Objective tests are utilized in science education at all levels and domains. Multiple-choice, the most common form of objective tests, are used regularly for a variety of reasons. They are often used to measure students' content knowledge, to rank, to place, to diagnose, to dismiss, and to certify students. They are also used to diagnose the efficiency of instruction the influence of the curriculum and the teacher.

Multiple-choice tests consist of a stem, usually the question, that is followed by distracters, which are the multiple-choice alternative responses. Usually, only one of the distracters is the correct response.

In order to examine students' understanding using multiple-choice test, Tamir (1971) focused not on the correct responses students gave on multiple-choice items but focused on their wrong answers. Tamir argued as follows: “When the test constructor writes the alternatives, he follows his own associations, his own idiosyncrasies, and his own thought patterns. But isn't it possible that students' idiosyncrasies and associations are quite different? Isn't it true that major objective of testing is to detect misunderstandings and misconceptions of the student? If this is true, why not first ask the students for alternative answers and construct the item accordingly?” (Tamir, 1971 p.306)

Thus, the construction of distracters for multiple-choice test items should be based on typical students' ideas. Students’ responses to interviews, essay questions, and to other open-ended questions could serve as sources for the data. The aim of having a multiple- choice test for Tamir was to diagnose student understanding rather than to classify and/or assess achievement. This type of a multiple-choice test in the scientific literature is commonly referred to as a diagnostic instrument.

Treagust (1986) proposed a strategy to develop a multiple-choice diagnostic test. Treagust suggested the same data sources as Tamir did that could serve for the test construction. (i.e., open-ended questions, literature, and individual interviews with students). The data collected in this way, can provide valuable information about students' misconceptions about a particular scientific area well defined in terms of propositional knowledge statements that the test is supposed to measure. Moreover, these data can be used to construct the distracters. Instead of just utilizing a multiple-choice test that consisted of a stem (question) and a few alternative responses, Treagust proposed to add an additional tier to the test. The first tier consists of multiple-choice responses related to the propositional statements. The second tier consists of a multiple-choice set of reasons for the answer given in the first tier. Furthermore, space was provided to enable students to express reasons different from the reasons provided on the test.

Haslam and Treagust (1987) used such a test to study 13-15 years old students understanding of photosynthesis. An exemplary item from a two-tier diagnostic tool test they developed follows:

Figure 2.2 Exemplary two-tier item

Which gas is taken by green plants in large amounts when there is no light energy at all?

1. Carbon dioxide gas
2. Oxygen gas

The reason for my answer is because:

A) This gas is used in photosynthesis which occurs in green plants all the time
B) This gas is used in photosynthesis which occurs in green plants when there is no light energy at all
C) This gas is used in respiration which only occurs in green plants there is no light energy to photosynthesize
D) This gas is used in respiration which takes place in green plants
E) Other explain: ____________________________________

To the best of my knowledge, no two-tier type diagnostic test has yet been constructed and used to test students' physics conceptions.

Halloun and Hestenes (1985a, 1985b) were among the first to construct a physics diagnostic test. It elicited students' basic knowledge of mechanics and was constructed using similar steps suggested by Tamir (1971). The test distracters were initially selected to assess the students qualitative conceptions of motion and its causes and to identify common misconceptions (referred to *common sense knowledge* by the authors) as reported in earlier research. That is, the item distracters were chosen from students' responses that were identified as misconceptions in earlier studies.

Licht and Thijs (1990) also used the same method to construct two diagnostic instruments in physics, one in introductory mechanics, the other one in electricity. The introductory mechanics test consisted of 13 items and the electricity test had 20 items.

The Force Concept Inventory or FCI Hestenes (1992) is one of the well-known and regularly used diagnostic multiple-choice test in physics education. It is perhaps the only diagnostic test that has been translated into several different languages.

The diagnostic tests mentioned to this point have the following three points in common: a) the wording is simple; b) the majority of the items include illustrations; and c) they do not include items that include the rote recall and use of mathematical formulas.

2.9.1 The Force Concept Inventory

As mentioned in the previous section, the FCI is one of the frequently used diagnostic tests in physics education. Hence, controversial positions exist about the validity of this diagnostic test.

The authors report that the FCI has been administered by several hundred physics teachers to a population of more than ten thousand high-school and college students in the United States of America. The FCI is composed of 29 multiple-choice questions each with five distracters. The goal of the FCI is to assess students' understanding of Newton's force concept, a central concept in introductory mechanics. Questions on the FCI are qualitative in nature and most of them are accompanied by a figure. The questions on the FCI usually do not require quantitative manipulation, however they require a forced choice between Newtonian and non-Newtonian concepts. In order to maximize the reliability of the FCI the authors tried to construct the test by minimizing the false positives (i.e., students select the correct answer for the wrong reason) and maximize the false negatives (i.e., an incorrect answer was chosen by a student who actually knew the correct answer).

In order to provide evidence for content validity the authors constructed the items on the FCI based on common students' misconceptions that had been extensively reported in previous studies. For example, the scenario about the drifting rocket in space, initially with the engines turned off and then the engines turned on and off the four associated questions were first used by Clement (1981) to investigate student's understanding of the force concept. A similar version of the question of a person pushing a heavy box on the floor and the question about the trajectory of a cannonball after leaving the cannon were used by Osborne and Gilbert (1980) to investigate student's understanding of motion. Question number 12 on the FCI, which probes students' understanding of Newton's third law was previously used by Minstrell (1982) to show that the majority of the students' believed that inanimate objects could not exert force. Another question that asks the student to identify the force(s) acting on a falling brick from a not so high building on the FCI, was employed by Vionnet (1979) as an open-ended question to investigate students' understanding of the force concept. Physics professors who were asked to examine the FCI reported no serious concerns, therefore, the authors reported that the content validity was "beyond reasonable doubt" Hestenes, 1995, p. 506.

The FCI was administered in 1992 to more than 1500 high-school students and more than 500 university students. Data analysis showed that the FCI score was highly correlated with Newtonian understanding. The data analysis also provided evidence that a FCI score of 85% could be interpreted as the Newtonian Mastery threshold. The authors indicate that a score of 85% and above means that the student is a confirmed Newtonian thinker. This means that the student has developed a universal force concept and learned to identify active and passive agents of force. The student has developed coherent dynamic concepts, including vector concepts of velocity, acceleration and force; the student has developed a coherent understanding of Newton's Third Law. The authors indicated that a FCI score between 60% and 85% could be interpreted to mean that the student has begun to understand Newtonian mechanics. The authors concluded that the scores obtained from the FCI could be used in the following three ways:

1. to interpret the effectiveness of instruction
2. to diagnose students' understanding
3. to make placement decision such as group formation

Serious concerns were raised in the science education literature about the three ways the FCI scores were suggested to be used by physics teachers and science education researchers. The following section discusses these concerns.

2.9.2 Concerns about the FCI

Huffman and Heller (1995) and Heller and Heller and Huffman, 1995 acknowledge the content validity of the FCI and comment that Hestenes et al. "did an outstanding job in selecting common-sense distracters" (p. 507) and thus provided sufficient evidence for content validity. However, they questioned the construct validity of the test. After administering the FCI to students who were enrolled in an introductory college course, Heller and Huffman ran a post- factor analysis on the collected data. The result of the factor analysis showed that the items on the FCI were loosely related to each other and that the items on the FCI did not load or cluster on the 6 domains initially suggested by Hestenes et al. Consequently, Heller and Huffman concluded that the FCI might actually not measure students' understanding of a force concept. This has important consequences. For example, students who scored between 60% and 85% on the FCI do not have a more (or less for that matter) coherent understanding of the force concept than students who scored below 60%. As a result, physics teachers and instructors should be cautious about concluding that the student FCI score truly measures student understanding of Newtonian force concept. Moreover, Heller and Huffman refuted the idea that the FCI scores provided thresholds thus, stages towards Newtonian understanding and that a high gain in FCI score does not warrant that the students are developing desired Newtonian understanding, the essential goal of an introductory mechanics course.

In spite of the evidence provided by Heller and Huffman (1995) that raises serious questions on the FCI’s construct validity, Hestenes and Halloun firmly defend their initial claim and that the test scores could be used as pointed out earlier in section 2.8.1. However, sections 2.6 clearly reports that a test can only measure what it purports to measure and the FCI purports to measure the Newtonian notion of force. Therefore, any interpretation of the FCI test scores about student’s understanding of classical mechanics is beyond the aim of the test and is not a valid interpretation. Contrary to the authors' suggestions, the FCI cannot be an effective measure to determine the effectiveness of an introductory mechanics course in terms of students' understanding of introductory mechanics concepts..

Despite the controversial opinions about the extend of interpretation of the FCI scores, the Force Concept Inventory seems to be still the choice of instructors and teachers to examine students' understanding, to diagnose course effectiveness and group students in introductory mechanics. For example, Hake (1998) used the FCI to determine the effectiveness of traditional lecturing and courses that were more student centered. The sample consisted of high-school, college and university students (n = 6542) who were enrolled in 62 different introductory mechanics courses. Fourteen (n = 2084) of those 62 courses were identified as traditional courses whereas 48 (n=4458) were identified as more student oriented courses. Data analysis showed that the average gain for the traditional courses was 0.23 (std. dev.= +/- 0.04) and that average gain for the conceptual oriented courses was 0.48 (std. dev.= +/- 0.14), almost two standard deviations above that of the traditional courses. The authors claimed that the results clearly suggested that student centered courses did promote conceptual understanding beyond that obtained in traditional classes.

2.10 Summary

Findings of empirical research in science education and physics education showed that conventional teaching strategies unfortunately do not highly promote conceptual understanding especially at high-school and first year college level. Consequently, students might leave basics physics courses with alternative concepts that are markedly in contrast with the accepted scientific concepts (section 2.6). These findings suggest that students enrolled in higher level courses such as introductory quantum mechanics courses, might be faced with similar conceptual problems. Experts in quantum mechanics have already started questioning the effect of conventional teaching methodologies currently used by many instructors and with good reasons have voiced their concerns and dissatisfaction towards this method of teaching (sections 2.3 -

2.4). New teaching approaches are suggested and even tried out. However, as it was reported in section 2.5 most of these approaches failed to report empirically the effectiveness of their new teaching methods. Constructing a valid and reliable instrument assessing students' achievement in introductory quantum mechanics, for example, might help to assess the effectiveness of instruction.

Since the purpose of this study was to develop a multiple-choice test in introductory quantum mechanics sections 2.7 – 2.8 reviewed important concepts pertaining to test construction. Chapter 3 will document on how these important concepts were implemented in this study.

CHAPTER 3

3.1 Methodology and Procedure

3.1.1 Introduction

The principal purpose of this research is to develop a multiple-choice test to assess college students' understanding of introductory quantum mechanics. A rudimentary part of the research was be to identify the core concepts that should be central in introductory university quantum mechanics courses. Since visualization of these concepts has now come to be considered an important part of understanding, the test will examine student' visualization of concepts as well as their understanding of more conventional mathematical and verbal representations. The Quantum Mechanics Visualization Instrument (QMVI) should discriminate levels of achievement of test takers. A secondary goal has been to identify the level of college students' understanding of selected topics in quantum mechanics including their understanding of visual representations after completing an introductory course in quantum mechanics.

3.1.2 Conventional Test Construction Procedure and the QMVI

Although the construction of a valid and reliable multiple-choice test is a complex process, well-established methods exist to guide the development (see sections 2.7 -2.8). The American Educational Research Association (AERA) has published a widely used standard for developing valid and reliable tests: the Standards for Educational Psychological Testing (1999).

In order to provide evidence about instrument validity and reliability, generally, instrument development starts by identifying the psychometric purpose of the instrument. That is, is the instrument going to be a general-purpose test or a special purpose test (see Chapter 1 for definition of the terms). The quantum mechanics visualization instrument (QMVI) is a general-purpose instrument. Because the OMVI was designed to be used with a widely diverse sample of students. More specifically, the QMVI aims to probe sophomore through first year graduate students' achievement of introductory quantum mechanics concepts. The QMVI can either be administered as an in class test or handed out as a homework assignment.

Open ended or essay and multiple choices tests are two types of tests widely used in assessing scholastic achievement. According to Ooesterhof (1996) multiple-choice type tests are one of the most frequent used test type. They have five basic advantages. Multiple-choice tests:

1. are versatile, i.e., they can measure multiple levels of specific cognitive skill
2. allow well distributed content sampling
3. can be quickly and objectively scored (hence the name objective test)
4. can provide valuable diagnostic information
5. are better liked by students when compare to other measures (Mehrens, & Lehmann, 1991; Meier, 1994; Cunningham, 1998).

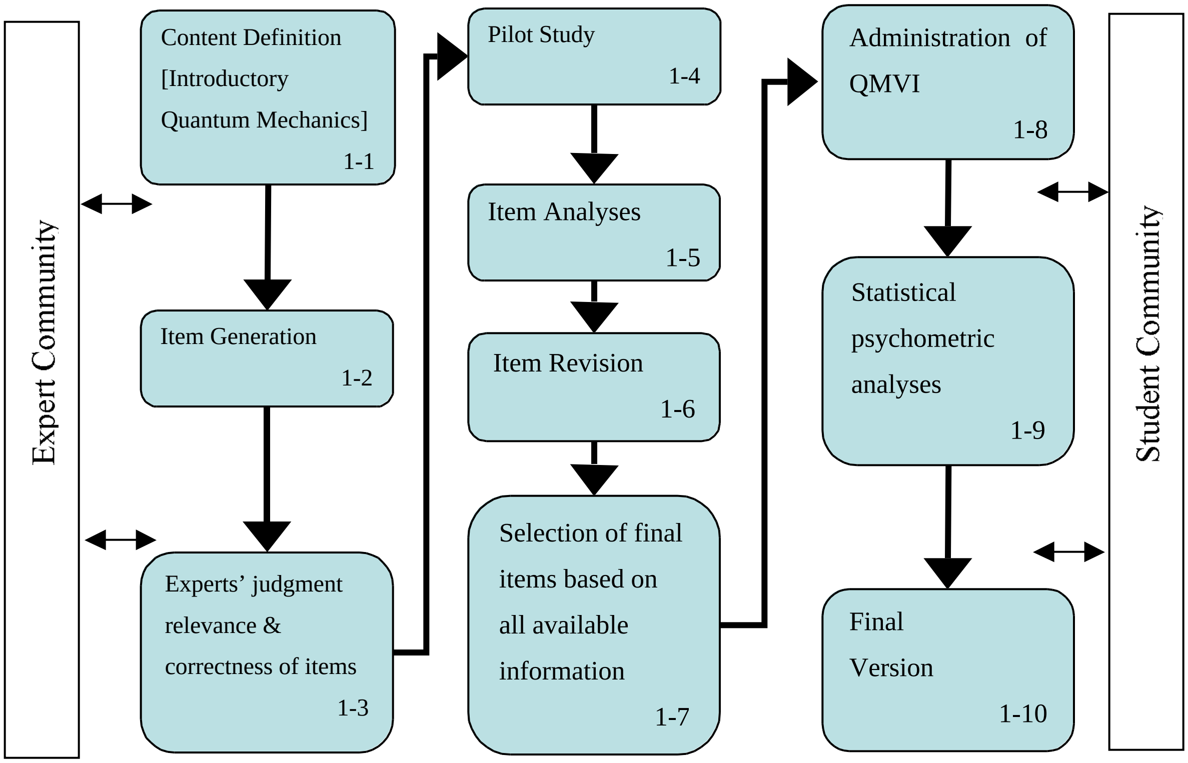
Despite the popularity of multiple-choice tests, limitations do exist. For example, multiple-choice tests tend to focus student attention on information in isolation by testing one element at a time. The larger context and structure of relationships between and among the elements can get lost. Multiple-choice test cannot assess the kinds of complex, multifaceted problem solving skills that are relevant to real world (Stiggins, 1994).

Recent research in science education, therefore, suggested extending multiple-choice tests into several tiers. This increases the probability of gaining more information about students' actual understanding when responding to questions. The additional data collected, can be used either to create a separate diagnostic assessment tool or used to improve the instrument already at hand (see section 2.9 for a more detailed discussion) (Treagust, 1989; 1988, 1986; Haslam and Treagust, 1987).

In response to these suggestions from the literature the QMVI was constructed as a three tier multiple-choice test. The first tier included the base question that is the stem and distracters. (i.e., ordinary multiple-choice test), the second tier was an open-ended question, asking the students for a brief, one or two sentence long explanation about their reasoning when selecting a particular distracters. The second tier, in the long run, will serve the purpose of creating a database on actual student answers which in turn could be used to develop a diagnostic tool as suggested above in the literature. The third tier was a 4 level Likert type scale that probed further for the students' confidence level for each item. Although not a very crucial tier and an uncommon practice in test construction, the third tier added another data dimension which was utilized in the QMVI construction. The third was a self assessment by the student about how well they perceived they did on a particular question. Therefore, a *moderate correlation* coefficient should exist between the students score on tier one and on tier three.

After determining the purpose and the type of the QMVI instrument the procedure continued through content definition, item generation, creating the draft version of the instrument, first draft administration of the modified multiple-choice test, item revision and refinement based on field test data, statistical psychometric analysis, and the general study of the instrument's validity and reliability (Nunnally, 1978; Rubba and Andersen, 1978; Suen, 1990). The development of the QMVI followed a construction process consistent to these standards. Suen's IPC examination development process (1998) was selected as the model that would guide QMVI development. Figure 3.1 represents a flowchart and the major sequences followed in developing the QMVI.

Figure 3.1 QMVI Development Process



As mentioned earlier, test construction is a continuous process that occurs in phases. This study is an initial attempt to establish evidence towards valid and reliable instrument. The process shown in figure 3.1 should be viewed as an initial phase of developing the QMVI and box 1-10 indicates only the end of this initial phase.

3.1.2a Validity and the QMVI

Section 2.7 introduced and reviewed the concept of validity. Because the determination and documentation of validity is a crucial part in constructing any instrument, the following section extends in section 2.7, that is the concept of test validity and more importantly how it applied to the construction of the QMVI.

Traditionally, validity has been defined as “the extent to which a test measures what it purports to measure” (Suen, 1990, p 134). However, recent interpretations of validity have a more comprehensive perspective to it. “Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationale support the adequacy and appropriateness of inferences and actions based on test scores or other modes of measurement” (Messick,1989 as cited in Suen, 1990, p.134). The Standards for Educational and Psychological Testing (AERA-APA-NCME, 1985) although not as comprehensive as Messick, provides a similar definition of validity that has set the current standard in this field. The standard defines validity as “the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores” (p.9). By definition, a test or a test score does not inherit validity as the traditional definition might somewhat suggest. Rather validity is “the reasonableness of using the test score for a particular purpose or for a particular inference [to be made]” (Suen, 1990, p.134). Since validity can be defined as the degree of appropriateness of a particular use of that test scores, it follows that test validation is a process of collecting or accumulating evidence which supports the intended use of the test scores. Consequently, validity should be conceptualized as matter of degree rather than as an all- or-none entity and viewed as a continuous process (Nunnally, 1978, Suen, 1990). This process involves an extended process of instrument planning, developing, field testing, checking and rechecking to obtain the highest degree of validity possible.

To obtain evidence of validity in constructing an instrument, several validation procedures exist. According to the Standards for Educational and Psychological Testing (1999) there are three procedures for use in validation studies that provide for collecting evidence for specific inferences to be made from test scores. They are construct-related evidence, content- related evidence, and criterion-related (predictive and concurrent) evidence of validity.

> “It should be emphasized that these three types of validity are not distinct categories and that they are not three types of validity. Rather, they are results of three different approaches to the establishment of validity … Dependent on the intended use of the test score, not all three types of studies are needed. Certain types of evidence obtained from certain types of studies may be more appropriate than others for certain score interpretation and usage.” (Suen, 1990, p.135).

The Quantum Mechanics Visualization Instrument (QMVI) aimed to assess college students’ understanding of topics in introductory quantum mechanics. In other words, the test scores reflected students’ achievement. Therefore, the QMVI was classified as an achievement test. Moreover, the QMVI as mentioned earlier was a general-purpose instrument because it does not target highly specialized groups or subgroups of the population of quantum mechanics students. An instrument meeting these two criteria, that is, assessing achievement and being a general-purpose instrument, needs to establish evidence of content validity (Nunnally, 1978). Hence evidence of content validity must be documented in the construction process of the QMVI.

Content validity is the degree to which the items on a test adequately represent the content of the discipline or field of study (Suen, 1990). In the construction of the QMVI, the items needed to be representative with respect to the content college students encounter in conventional modern physics and introductory level quantum mechanics courses.

Section 3.2.2 discusses in detail the content definition procedure and the how the core content of quantum mechanics was established.

3.1.2b Test Reliability and the QMVI

Reliability is a complex concept concerning the consistency of measurement over time. Suen (1990) defined reliability as “the strength of the relationship between the observed score and the true score” (p. 28). The observed score is a composite score of the true score and random score. A test’s reliability coefficient can be estimated by four empirical methods: 1) Test-Retest, 2) equivalent forms, 3) split-half, and 4) coefficient alpha (see section 2.8 for more details).

3.2 Development of the QMVI – A Procedural Overview

The QMVI was designed as a general-purpose achievement test including visual representations on selected topics in quantum mechanics. The following section reports on the test construction procedure, which were necessary to provide evidence of validity. "The validity of an achievement test can be established only by subjective judgments regarding the extent to which the test measures what it intended to measure-in other words the extent to which it reflects the content" (Tikelman, 1971, p. 50)

3.2.1 Instrument Specification

Quantum mechanics understanding, as defined for the purpose of development of this instrument, consisted of basic knowledge and technical terminology in quantum mechanics, the understanding of quantum mechanics theory, and the knowledge to interpret visual representations of quantum mechanical phenomena. Students are usually introduced to actual laboratory investigations in quantum mechanics after they have first an introductory quantum mechanics course. Therefore, no attempt was made to include knowledge associated with conducting experiments in quantum mechanics. Assessment of philosophical interpretation of quantum mechanics theory was also not included. The QMVI was aimed at college level students. More specifically, the QMVI targeted sophomore to first year graduate students who had completed at least one modern physics or quantum mechanics course. The test was designed to be administered in 90 minutes or less either in class or as a homework assignment. The resulting scores are to be used for evaluating the achievement of individual students.

3.2.2 Content Definition

As stated in the previous section the selection of topics and the definition of the content of introductory quantum mechanics was a very important process in developing and validating the QMVI. Students who pursue their degree in physics or closely related field encounter at least one course in modern physics or quantum mechanics. Nevertheless, in Fall 1999 through Spring 2000 there was no accepted standard curriculum upon which the QMVI could be based. Therefore, it was necessary to define the core quantum mechanics content upon which an appropriate instrument could be constructed. The core content was to be perceived as central by experts in teaching college level introductory modern physics and introductory quantum mechanics courses in order to assess those topics, concepts, and basic theories that high achieving students should have mastered after completing an introductory quantum mechanics course.

The identification of core content with respect to introductory level quantum mechanics was important because it was to serve two purposes, i.e., prescription and guidance in construction of the instrument. First, the content outline defined the curriculum area hence, the subject matter that was included in the QMVI. Second, the content outline provided guidance in suggesting possible test items for the QMVI. The content outline promoted balance on the items that should be included in an achievement test that reflects well the topics taught in conventional introductory level quantum mechanics courses.

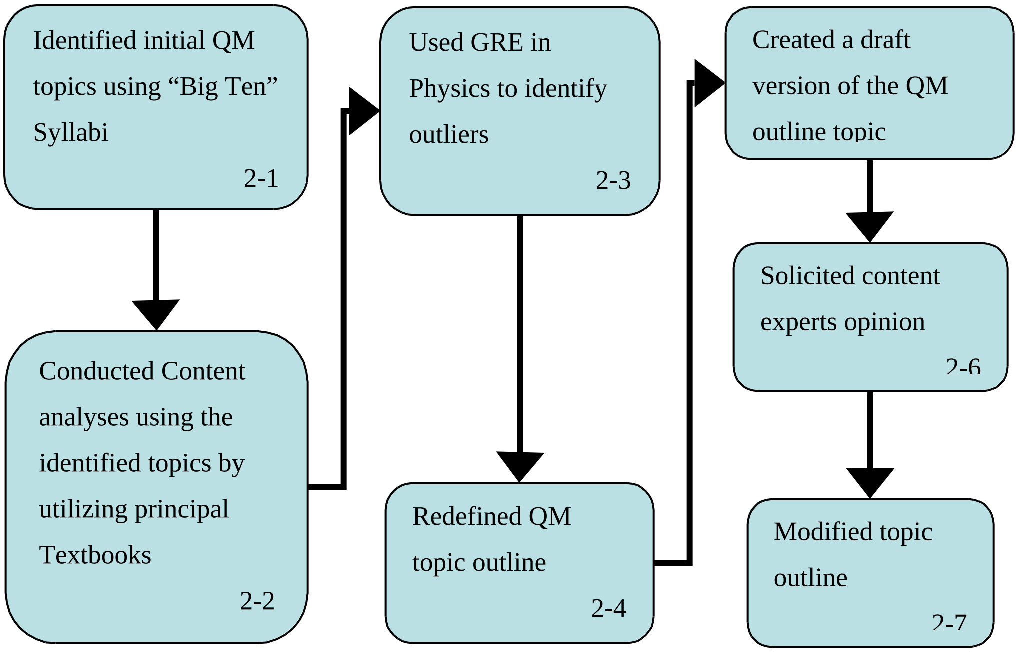
In order to establish the content of introductory quantum mechanics, a content analysis technique has been utilized in this study. Content analysis can be defined as a "research techniques for the objective, systematic, and quantitative description of the manifested content of communication” (Gall, 1996, p.357).

A common procedure used to establish a tentative content outline with relative weights assigned to topics is to analyze about 10-12 widely used textbooks and associated student assistance guides (Peters et al., 2000) because they convey much about the curriculum. After determining a draft outline of topics, a recommended procedure is to determine the median number of pages devoted to each topic in order to provide an index of importance as judged from textbook content. Analysis of the syllabi from representative universities was another source of data representative of the content to be assed by the QMVI (Tinkelman, 1971). In addition, the advice of experts in teaching introductory quantum mechanics at the college level was sought. To these ends, this study examined the following documents to establish the content outline and the relative weight of each topic:

a.) Syllabi of introductory level quantum mechanics courses in the "Big Ten" universities.
b.) Introductory level quantum mechanics textbooks.
c.) The Graduate Record Examinations (GRE) test in physics.
d.) Content experts at the Pennsylvania State University in collaboration with consultants at other universities.

Figure 3.2 below is an elaboration of box 1-1.

Figure 3.2. The identification of core content in introductory quantum mechanics.



Before elaborating more on the identification process of core content topics in introductory quantum mechanics, figure 3.2 introduces the main procedure followed to establish the content outline for introductory quantum mechanics. Briefly, as can be seen from the figure above, the “Big Ten” syllabi were used to identify initial conceptual domains and topics. Second, using these identified domains and topics descriptive statistics was conducted on textbooks to establish a more detailed draft content outline with respective weights given to the topics addressed. Third, the Graduate Record Examinations test in physics (GRE) was used as an additional data source. Fourth, content experts were asked to validate the draft content outline in introductory quantum mechanics. Finally, the draft content outline was modified according to the feedback obtained from the content experts. A more elaborate discussion follows on the sections below.

3.2.2a Analysis of “Big Ten” Syllabi

In order to establish the content outline of introductory level quantum mechanics, an *initial* attempt had to be made to identify possible topics addressed frequently in conventional introductory mechanics courses at the college level. For several reasons, samples of the eleven “Big Ten” universities (i.e., University of Illinois, Indiana State University, University of Iowa, University of Michigan, Michigan State University, University of Minnesota, Northwestern University, Ohio State University, Pennsylvania State University, Purdue, University of Wisconsin) syllabi of introductory level quantum mechanics courses were chosen as the starting point. First, quantum mechanics courses offered in the “Big Ten” universities logically seemed to be a representative population when considering the sample students in this study initiated at the Pennsylvania State University. Secondly, the syllabi provided condensed information that outlined the topics the instructors intended to address in their introductory quantum mechanics courses. Additionally, the syllabi identified information about the main textbooks and supplementary books the instructors perceived to be important in their courses.

The syllabi were analyzed to generate coding categories which were later used to analyze the textbooks (discussed in the next section). The coding categories were determined by identifying the topics and counting the frequency their appearance in the syllabi. As a result of analysis of the syllabi, 21 topics were identified. They are listed in alphabetical order in Table 3.1.

Table 3.1 Topics and their respective frequency counts in the "Big Ten" syllabi.

Topics	Frequency
1. Angular Momentum	4
2. Atomic Physics	1
3. Bohr Atom	3
4. Delta function	2
5. Elementary particles	2
6. Harmonic oscillator	7
7. Hydrogen Atom	7
8. Molecular Physics	1
9. Nuclear Physics	2
10. Operator methods	3
11. Perturbation theory	4
12. Postulates of Quantum Mechanics	2
13. Potential and barrier problems	6
14. Quantum Statistics	1
15. Scattering in one dimension	6
16. Schrödinger Equation	10
17. Schrödinger equation in three dimension	2
18. Solid state	2
19. Tunneling	6
20. Uncertainty Principle	2
21. Wavefunction	8

As a result of the analysis, 21 key topics were identified for further investigation. Table 3.1 indicates that harmonic oscillator, hydrogen atom, potential and barrier problems, scattering in one dimension, Schrödinger equation, tunneling, and wavefunction seemed to be the principal topics constituting an introductory quantum mechanics course across the eleven universities because of their high frequency counts. This result is not too surprising, but it should be noted that only two professors exclusively mentioned the postulates of quantum mechanics in their syllabi. Nevertheless, the purpose of this analysis was to obtain keywords towards a more elaborate textbook analysis. Also, syllabi are documents that serve the purpose of a brief tentative overview about courses.

The syllabi analysis identified also the most frequently used textbooks in introductory quantum mechanics courses. Supplementary books which were not related directly to quantum mechanics were not included in the analysis. Table 3.2 depicts the frequency distribution of the central textbooks; they are listed in an alphabetical order.

Table 3.2 Frequency count of textbooks used in the "Big Ten" syllabi.

Authors	Title	Frequency
1. Bransden & Joachain (1989)	Introduction to quantum mechanics	1
2. Griffiths (1995)	Introduction to quantum mechanics	5
3. Liboff, (1980)	Introductory quantum mechanics	1
4. Sakurai & Tuan (1994)	Modern quantum mechanics	1
5. Merzbacher (1970)	Quantum mechanics	1
6. Peebles (1992)	Quantum mechanics	1
7. Gasiorowicz (1974)	Quantum physics	3
8. Eisberg & Resnick (1985)	Quantum physics of atoms, molecules, solids, nuclei, and particles	2

The most frequently required principal textbook was Griffiths' Introduction to quantum mechanics (1995). Followed by Quantum physics by Gasiorowicz (1974) and Quantum physics of atoms, molecules, solids, nuclei, and particles by Eisberg & Resnick (1985).

3.2.2b Analysis of Introductory Quantum Mechanics Textbooks

After obtaining initial topics by conducting the analysis of the syllabi, the analysis was extended to introductory quantum mechanics textbooks. The approach to establish the draft content for introductory level quantum mechanics using the textbooks was a semiquantative one. First, the Pennsylvania State university’s computerized library database was searched. Like many advanced search engines the Pennsylvania State university’s computerized library database had the feature for advanced search options. Using these options the search was limited to the textbooks. The keyword “quantum mechanics” was used initially resulting in the identification of 240 books. This made it necessary to limit the search even further. Therefore, the following keywords and their combination were used additionally to obtain a more manageable booklist - Quantum Mechanics, Modern Physics, Introduction, Introductory, Quantum Theory, and Quantum Physics. The second search narrowed down the potential booklist for the content analysis to 25 including the three most popular introductory quantum mechanics textbook shown in Table 3.2. All of the textbooks were located at the library. In order to identify the textbooks that were at the appropriate level, each textbook’s preface was subjectively analyzed. That is, the textbook authors, usually did mention the target students they had intended to reach. As a result of this subjective analysis, 14 books were chosen for a more detailed content analysis. The topics listed in Table 3.1 were then used in the more elaborate content analysis on these textbooks. The content of each textbook was analyzed based on the identified topics (listed in Table 3.1). The relative weight of each topic was computed by determining the percent of pages devoted to that particular topic in each book with the overall number of pages in that particular book. As a result of the syllabi analysis, the following 5 main and 30 subcategories were established (see Table 3.3).

Table 3.3 Respective weights of each topic in introductory quantum mechanics textbooks.

Main Topic	Subtopic	Textbook	Weights
A. Historical development & terminology	1. Blackbody radiation 2. Photoelectric effect 3. Bohr Atom 4. Wave Particles duality 5. deBroglie hypothesis 6. Uncertainty Principle 7. Probability & semi-classical behavior, waves	Brandt, S. & Dahmen, H.D	0%
		French, A. P. & Taylor E. F	19%
		Gasiorowicz, S	27%
		Landshoff, P, Metherell, A., & Rees, G	10%
		Liboff, R	11%
		McMurry, S. M.	8%
		Morrison, M. A.	8%
		Peleg, Y., Pnini, R., & Zaarur, E	0%
		Robinett, R,W.	5%
		Shankar, R.	14%
		Thaller, B	0%
		Winter, R, G.	11%
		Griffiths, D.J	9%
		Eisberg,R.M, & Resnick E.M	8%
B. Quantum Mechanics Postulates	8. Postulates of quantum mechanics I-V 9. Observable and operators 10. Eigenfunctions and eigenvalues 11. Dirac Delta function 12. State function and expectation values 13. Hilbert space and its properties	Brandt, S. & Dahmen, H.D	0%
		French, A. P. & Taylor E. F	8%
		Gasiorowicz, S	25%
		Landshoff, P, Metherell, A., & Rees, G	18%
		Liboff, R	0%
		McMurry, S. M.	20%
		Morrison, M. A.	26%
		Peleg, Y., Pnini, R., & Zaarur, E	13%
		Robinett, R,W.	18%
		Shankar, R.	19%
		Thaller, B	16%
		Winter, R, G.	38%
		Griffiths, D.J	22%
		Eisberg,R.M, & Resnick E.M	18%

Table 3.3 Continued

Main Topic	Subtopic	Textbook	Weights
C. Schrödinger equation	14. Wave function of a single particle 15. the Schrödinger equation (time dependent & time independent) 16. scalar products of wave function 17. normalization and probability density	Brandt, S. & Dahmen, H.D	15%
		French, A. P. & Taylor E. F	18%
		Gasiorowicz, S	11%
		Landshoff, P, Metherell, A., & Rees, G	16%
		Liboff, R	20%
		McMurry, S. M.	21%
		Morrison, M. A.	16%
		Peleg, Y., Pnini, R., & Zaarur, E	10%
		Robinett, R,W.	14%
		Shankar, R.	10%
		Thaller, B	20%
		Winter, R, G.	12%
		Griffiths, D.J	24%
		Eisberg,R.M, & Resnick E.M	18%
D. Application of Schrödinger equation in one dimension	18. Particle in a box (infinite hard wall) 19. Particle in a time-independent potential 20. Continuity conditions 21. Unbound and bound well 22. Harmonic oscillator 23. Wave packet and scattering (time-independent) 24. Probability density and probability current 25. Scattering by a one dimensional well 26. Tunneling	Brandt, S. & Dahmen, H.D	27%
		French, A. P. & Taylor E. F	22%
		Gasiorowicz, S	22%
		Landshoff, P, Metherell, A., & Rees, G	24%
		Liboff, R	22%
		McMurry, S. M.	22%
		Morrison, M. A.	24%
		Peleg, Y., Pnini, R., & Zaarur, E	23%
		Robinett, R,W.	28%
		Shankar, R.	21%
		Thaller, B	40%
		Winter, R, G.	37%
		Griffiths, D.J	20%
		Eisberg,R.M, & Resnick E.M	18%

Table 3.3 Continued

Main Topic	Subtopic	Textbook	Weights
E. Advanced Applications	27. Further applications of Schrödinger equation in, two and three dimension 28. Wentzel-Kremes-Brillouin Method 29. Time-independent perturbation theory 30. Angular momentum	Brandt, S. & Dahmen, H.D	32%
		French, A. P. & Taylor E. F	12%
		Gasiorowicz, S	15%
		Landshoff, P, Metherell, A., & Rees, G	28%
		Liboff, R	17%
		McMurry, S. M.	8%
		Morrison, M. A.	19%
		Peleg, Y., Pnini, R., & Zaarur, E	42%
		Robinett, R,W.	30%
		Shankar, R.	36%
		Thaller, B	0%
		Winter, R, G.	0%
		Griffiths, D.J	21%
		Eisberg,R.M, & Resnick E.M	20%

Table 3.4, below, is a summary of the overall normalized weight for the main five topics identified and listed in Table 3.3. Note that C. Schrödinger equation and D. Application of Schrödinger equation in one dimension made up 49% of an introductory quantum mechanics course.

Table 3.4 Results of content analysis conducted on 14 textbooks in quantum mechanics.

Main Topics	Number of Books	Median	Normalized Weight
A. Historical development & terminology –nomenclature	14	9%	10%
B. Quantum mechanics Postulates	14	18%	20%
C. Schrödinger equation	14	20%	22%
D. Application of Schrödinger equation in one dimension	14	24%	27%
E. Advanced Applications	14	18%	20%

3.2.2c Analysis of Graduate Record Examinations in Physics

The reason the Graduate Record Examinations (GRE)–Physics was included in the content analysis was to make sure that no important topics were left out. Although, the GRE-physics test was designed as a general physics test, it claimed to assess important topics in physics, including quantum mechanics. Therefore, a test of conceptual understanding in introductory quantum mechanics courses should include the quantum mechanics content that the GRE-physics test currently assess. In other words, if the analysis of GRE – physics content revealed topics not included in Table 3.3 it might indicate that a fundamental topic might be missing from the attempt to provide a complete introductory quantum mechanics content in this study.

The Graduate Record Examinations (GRE) –physics was developed and published by Educational Testing Service (ETS). The GRE – physics test is a standardized subject test that aims to asses potential graduate students understanding of fundamental principles in physics. The GRE-physics was designed to test the basic knowledge and understanding of the most important topics in physics. Its scores are being used to predict success as well as graduate admission decisions by the universities and colleges.

The content of the GRE-physics test as reported by ETS (see Figure 3.1.) is based on the first three years of undergraduate physics at the college level. The approximate percentage of the content topics is as follows: “classical mechanics (20%), fundamentals of electromagnetism (18%), atomic physics (10%), physical optics and wave phenomena (9%), quantum mechanics (12%), thermodynamics and statistical mechanics (10%), special relativity (6%), and laboratory methods (6%). The remaining 9% of the test covers advanced topics such as nuclear and particle physics, condensed matter physics, and astrophysics” (GRE –Physics, 1996, p.8)

The GRE-physics test did not provide a detailed content outline for introductory quantum mechanics other than that it is 12% of the overall test. However, the ETS provides three actual full length GRE- physics tests. (ETS, 1996) Each question of these tests were analyzed and appropriately classified using the topics listed in Table 3.3. The result of the GRE-physics content is depicted in Table 3.5.

Table 3.4 Content Distribution of GRE-Physics of Quantum Mechanics.

Topic	Weight
A. Historical development & terminology	18%
B. Quantum mechanics Postulates	18%
C. Schrödinger equation	22%
D. Application of Schrödinger equation in one dimension	29%
E. Advanced Applications	13%

The analysis of the GRE-physics tests revealed that there were no missing topics in the quantum mechanics content analysis conducted earlier in this study. When GRE-physics topics were compared with the topics listed in table 3.3 the result revealed that the GRE-physics emphasized the domain of “Historical development & terminology” more than did the proposed introductory quantum mechanics topics in this study whereas less emphasize was given in the main topic of “Advanced Applications” (see Table 3.6.).

Table 3.5 Comparison of Weight of the GRE-physics Test Content and the Proposed Introductory Quantum Mechanics Content.

Topic	I	II
A. Historical development & terminology	18%	10%
B. Quantum mechanics Postulates	18%	20%
C. Schrödinger equation	22%	22%
D. Application of Schrödinger equation in one dimension	29%	27%
E. Advanced Applications	13%	20%

I: GRE-physics quantum mechanics content weight
II: Proposed introductory quantum mechanics content weight

3.2.2d Content Experts

The final step in establishing an adequate core content for introductory quantum mechanics was to obtain the expert opinion on the proposed core content. Therefore, three content experts in the department of physics at the Pennsylvania State University were asked to review the proposed content and the draft version of the QMVI (see section 3.2.3 for the item validation procedure). All three professors had experience in teaching modern physics and quantum mechanics courses. A formal letter of request, the proposed content and the draft version of the QMVI was submitted to the panel of experts. The experts were requested to return their feedback within ten days. Comments and suggestions of two content experts were received directly recorded on the documents they had received. The third expert provided oral feedback.

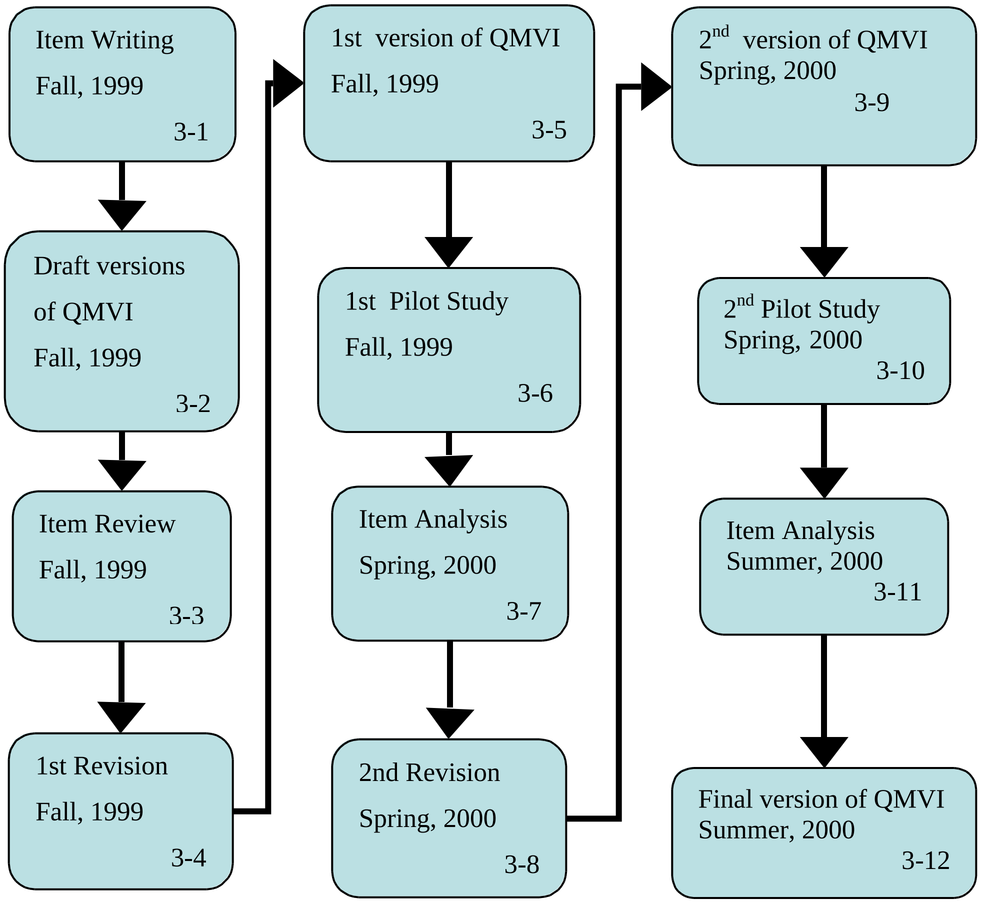
No substantial suggestions were made regarding the 5 main topics and their relative weights. However, the two subtopic “Dirac Delta function” and “Hilbert space and its properties” in Main Topic B “Quantum Mechanics Postulates” were removed because theses two topics were perceived to me more advanced.

3.2.3 Item Generation, Item Validation, and Test Revision

The previous sections provided a detailed explanation about the determination of a core content in introductory quantum mechanics upon which the QMVI items have been based on. Sections 3.2.3a through 3.2.3f describe the item generation and item validation procedure which includes development and revision procedures of the QMVI.

The flowchart shows the major steps of item generation and test revision in the development of the QMVI

Figure 3.3 Major steps in item generation, validation, and test revision.



3.2.3a Item Writing

All items constituting the draft versions of the QMVI were written by Dr. Robinett, who is a content expert in the field of quantum mechanics with extensive teaching experience. He also is an author of a textbook in quantum mechanics (Robinett, 1997).

3.2.3b Item Review

Before any pilot study was conducted on the draft version of the instrument, Tinkelman (1971) recommended that the test items should be reviewed from three different perspectives: (a) the technical, with particular attention to principles, including those relating to item form; (b) the subject matter, with attention to appropriateness of content and to accuracy of the scoring key; and (c) the editorial, with attention to appropriate overall format and editorial consistency from one item to another. (Thorndike, 1971, p.76). The reason for this kind of item review of versions of the QMVI is to develop a quality instrument.

Figure 3.4 Item review & validation procedure on the QMVI

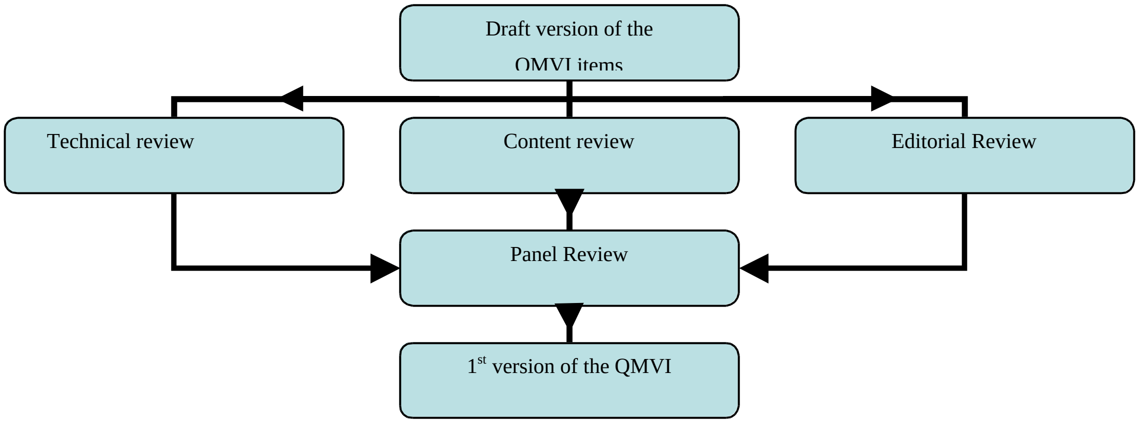


Figure 3.4 is an elaboration of boxes 3-2, 3-3, and 3-4.on page 52. Figure 3.4 shows all three-review steps suggested above were conducted in the late fall and early spring semester of the academic year 1999 - 2000 at the Pennsylvania State University. The technical review was conducted by four graduate students in Science Education, the content review was conducted by two professors in the department of physics, and the editorial review was conducted by an English language expert.

A) Technical Review

The technical review was conducted in by four (including the researcher) Ph.D. candidate graduate students in Science Education at the Pennsylvania State University with expertise in physics education. A letter which explained the procedure (Appendix B) and the first draft version of the QMVI was submitted. Accordingly, the graduate students independently reviewed each item of the draft version of the QMVI for clarity and comprehension based on similar criteria provided by Edward (1957). Comments and suggestions were recorded directly on the first draft version of the QMVI. Later the second draft version QMVI was then resubmitted to the graduate students with the same criteria. All of them approved the changes and had no further suggestions.

B) Content Review

The content review was conducted several times by three professors between Fall 1999 and Spring 2000 in the department of physics at the Pennsylvania State University. The professors were provided with the QMVI and content outline. Their task was to review each item for content accuracy. They were also asked to check for item accuracy.

C) Editorial Review

An English Language expert reviewed the items of the QMVI for grammatical, spelling, and syntax errors. As a result of this review process, minor suggestions were made to enhance the clarity of some statements. For example, the sentence in question number two “Order the time spent in each small position bin (dx), t(I), t(II), t(III), and t(IV), from largest to smallest as the particle moves back and forth in the well” was changed to “order the time spent in each small position bin (dx), t(I), t(II), t(III), and t(IV), from longest to shortest as the particle moves back and forth in the well”.

D) Panel Review

A panel of one content expert and two Ph.D. candidate graduate students in science education with expertise in physics education met regularly during the late fall and early spring semester of the academic year 1999 – 2000. As a result of this ongoing content validation process in the series of reviews mentioned above, changes were made to the items on the draft version of the QMVI, in order to reflect the content adequately, obtain a scientifically correct instrument, to improve visual and verbal clarity, and to obtain a coherent look (see appendix A).

3.2.3c First Pilot Study of the QMVI

The 1st version of the QMVI (Figure 3.3, box 3-5, p. 52) was administered at the end of the fall semester 1999 at the Pennsylvania State University to a total number of 67 students. The students were offered extra credit by their individual course professors for taking part of the pilot study.

More specifically, 45 sophomore students enrolled in a modern physics course, 17 senior students enrolled in a introduction to quantum mechanics course, and 5 first year graduate students enrolled a quantum mechanics course took the 1st version of the QMVI. These groups of students also represented the target population for the final version of the QMVI instrument.

3.2.3d Item Analysis on the First Version of the QMVI

The purpose of conducting the item analysis was to improve the reliability of the QMVI by identifying the items that did not assess students’ understanding properly and had problems. The item analysis and the revision procedure on the first version of the QMVI was conducted during earlier period of Spring 2000.

As mentioned briefly previously, the item analysis provides information that enhanced the reliability of the total score. In order to conduct an item analysis, item indices need to be computed. These indices are known as item difficulty index and item discrimination index.

The *item difficulty index* is the proportion of correct answers given by the students to an item and can have a value between 0 and 1. The higher the item difficulty index, the easier the item. The item difficulty index is also know as the p-value.

The *item discrimination* index is the difference between the proportion of correct responses of the high-ability group and the low-ability group. Usually 27% of the high scoring examines are defined as the high-ability group. Whereas, 27% of the low scoring group is defined as the low-ability group. The item discrimination index can have a value between -1 and 1. The reliability of the test score is maximized when the discrimination value ideally is 1 for all items, meaning that all high scoring examines scored correctly whereas the low scoring group missed all correct answers.

If the item discrimination index value is close to zero or negative, it indicates that a higher proportion of the low-ability group students chose that particular distracter. If the item discrimination index value is positive and close to one, it indicates that a higher proportion of the high-ability group students chose that particular distracter. Therefore, it follows that for the correct answer the item discrimination index value should have a positive value close to one. For all distracters the discrimination index value should be negative. However, if the item discrimination index value is close to one or negative for the correct answer and a positive value for any of the distracters, it can be concluded that the item did not assess students understanding properly.

Table 3.6 Results of the Item Analysis on the 1st Version of the QMVI

Number of Items	25
Number of Students	67
Mean	9.9
Variance	18.2
Standard Deviation	4.3
Skew	0.85
Minimum	3
Maximum	24
Median	9.0
Alpha	0.75
SEM	2.14
Mean Item difficulty (p-value)	0.4

Table 3.7 shows the overall results of the item analysis conducted on the 1st version of the QMVI after the first pilot study (Figure 3.3, box 3-7, p.53) . The results were very satisfactory. The reliability coefficient Cronbach alpha was 0.75, a satisfactory index value for an achievement test. The highest possible score of the QMVI was 25. Therefore, the desired mean score for the QMVI was 15 when including the factor of random guessing. However, the mean item difficulty was 0.4 indicating that the QMVI was a slightly difficult test. This was also reflected in the test means score of 9.9, which was about 5 points lower, than the desired mean.

As a result of a more detailed item analysis, questions number 7, 9, 14, 16, 20, and 21 were dropped from the test based on the item analysis theory explained above. The discussion that follows explains why these questions were replaced.

Table 3.7 Result of Item Analysis for Question Number 7.

p-value	Choices	High-Ability Group	Low-Ability Group	Item Discrimination Index	Correct Answer
0.358	A	0.556	0.167	0.389	
	B	0.333	0.556	-0.223	*
	C	0.056	0.056	0	
	D	0.056	0.111	-0.055	
	E	0	0.111	-0.111	
	Other	0	0	0	

As can be seen from Table 3.8, the item discrimination index value for the correct response was –0.223 indicating that a higher portion of the low-ability group students scored higher whereas a lower portion of the high-ability group students missed the correct answer. Moreover, the high-ability group students tended to choose distracter A, clearly indicating that the item did not work properly. Therefore, the panel decided to replace the question.

Table 3.8 Result of Item Analysis for Question Number 9.

p-value	Choices	High-Ability Group	Low-Ability Group	Item Discrimination Power	Correct Answer
0.478	A	0	0.222	-0.222	
	B	0	0	0	
	C	0.111	0.333	-0.222	
	D	0.833	0.278	0.555	*
	E	0.056	0.111	-0.055	
	Other	0	0.056	-0.056	

Question number 9 had a relatively high item discrimination index value of 0.555 and a moderate p-value making it an appealing question at the first sight. However, analysis on the second tier of the QMVI revealed that a majority of sophomore (Modern Physics) students had complaints and problems understanding the questions due to technical wording. Therefore, the panel decided to change the question.

Table 3.9 Result of Item Analysis for Question Number 14

p-value	Choices	High-Ability Group	Low-Ability Group	Item Discrimination Index	Correct Answer
0.209	A	0.667	0.056	0.611	
	B	0.111	0	0.111	*
	C	0.056	0.167	-0.111	
	D	0.111	0.444	-0.333	
	E	0.056	0.222	-0.166	
	Other	0	0.111	-0.111	

As can be seen from Table 3.10, the item discrimination index value for the correct response was 0.111 indicating a low item discrimination power. Moreover, the high ability students chose distracter A over the correct answer. This statistic indicated a problem with the item. Therefore, the panel decided to change the question.

Table 3.10 Result of Item Analysis for Question Number 16

p-value	Choices	High-Ability Group	Low-Ability Group	Item Discrimination Index	Correct Answer
0.030	A	0.056	0.056	0	
	B	0	0.056	-0.056	
	C	0.056	0	0.056	*
	D	0.889	0.389	0.5	
	E	0	0.389	-0.389	
	Other	0	0.111	-0.111	

As can be seen from Table 3.11, the item discrimination index value for the correct response was only 0.056 indicating a very low item discrimination power. Moreover, the high-ability group students chose distracter D over the correct answer. The p-value of 0.30 showed that question number 16 was difficult. Therefore, the panel decided to change the question.

Table 3.11 Result of Item Analysis for Question Number 20

p-value	Choices	High-Ability Group	Low-Ability Group	Item Discrimination Power	Correct Answer
0.209	A	0	0.056	-0.056	
	B	0.444	0.167	0.277	*
	C	0.389	0.333	0.056	
	D	0.111	0.167	-0.056	
	E	0.056	0.167	-0.111	
	Other	0	0.111	-0.111	

The item discrimination index value was slightly below a moderate value for question number 20. The p-value also showed that this question was a rather difficult one. In order to increase the mean p-value of the QMVI the panel decided to replace this question.

Table 3.12 Result of Item Analysis for Question Number 21

p-value	Choices	High Group	Low Group	Item Discrimination Power	Correct Answer
0.119	A	0.056	0.222	-0.166	
	B	0.222	0.232	-0.010	
	C	0.056	0.111	-0.055	
	D	0.556	0.157	0.399	
	E	0.111	0.222	-0.111	*
	Other	0	0.056	-0.056	

Table 3.12 shows a p-value of 0.119 meaning that question number 21 was a difficult one. The item discrimination index is -0.111 indicating that a higher portion of the low-ability group students chose the correct answer over the high-ability group students. Therefore, question number 21 was replaced from the test.

In summary, questions number 7, 9, 14, 16, 20, and 21 were replaced with new questions for the second pilot study. In addition, as a result of the item analysis, modifications were made on question number 2, 11, 12, 17 and 22. For example the correct answer for question number 21 was “None of the above”, it was changed with the correct solution.

3.2.3e Second Pilot Study of the QMVI

The second version of the QMVI (Figure 3.3, box 3-9, p. 56) was administered at the end of the spring semester 2000 at the Pennsylvania State University to 44 students. The students were offered extra credit by their individual course professors for taking part in the pilot study. More specifically, the sample consisted of 28 sophomore students enrolled in a modern physics course and 16 senior students enrolled in an introductory quantum mechanics course.

3.2.3f Item Analysis on the Second Version of the QMVI

The results of the item analysis on the second version of the QMVI (3-11) are shown in Table 3.13 below. The overall results were very satisfactory. A alpha value of

0.5 is suggested for a typical achievement test. For a multiple-choice test with 5 alternatives the suggested p-value is 0.6 due to the 20% random guessing change. Table 3.13 shows that the third version of the QMVI has a mean p-value of 0.45 approaching the suggested p-value. When compared with the second version we observe an improvement from 0.4 to 0.45 on the p-value (see Table 3.7). This was also reflected in the mean score which increased from 9.9 (see Table 3.7) to 11.3. An improvement was also observed in the test reliability coefficient alpha which increased from 0.75 to 0.83. However, the range of score (i.e. the difference between the maximum and minimum score) decreased from 21 to 19.

Table 3.13 Results of Item Analysis on the 3rd Version of QMVI

N of Items	25
N of Examinees	44
Mean	11.3
Variance	25.9
Std. Dev.	5.10
Skew	0.31
Minimum	3
Maximum	22
Median	10
Alpha	0.83
SEM	2.092
Mean Item difficulty (p-value)	0.45

In conclusion, an improvement in favor of the second version of the QMVI was observed when compared the statistics of the first versions. As a result, the panel decided that no further modifications were to be made on the first phase of the construction of the QMVI. The QMVI was administered in Spring 2000, Fall 2000, and Spring 2001.

3.2.4 Content Analyses on QMVI Version 2

After the final administration of the QMVI version 2 in spring 2001, an expert panel was formed to content validate the QMVI version 2. The panel was asked to classify each question on the QMVI as “Important”, “Not Important but Relevant”, or “Not Relevant at All” the criteria was the core content of introductory quantum mechanics defined earlier in this chapter. As a result of this validation procedure all but question number 1 on the QMVI were classified with 80% or higher consent as “Important”. Therefore question number 1 had to be excluded resulting in a 24 item-QMVI.

3.3 General Description of the Student Samples and Data Collection Procedure

3.3.1 The Samples

The participants were drawn from a pool of volunteer undergraduate and graduate students enrolled in either a modern physics or non-relativistic quantum mechanics course. Initially, students attending the Pennsylvania State University (PSU) at University Park and then Arizona State University (ASU) took part in this study. The students who took part in this study were enrolled in four distinct courses. A description of these four courses follows:

Modern Physics, henceforth referred as *ModPh*, is a post-classical physics course, typically taken by physics, chemistry, astronomy, and engineering students. The course is a 3 credit, one semester course and is usually required of all physics majors. It is typically taken in the fourth-semester at the sophomore level. The course covers much of the modem physics curriculum including topics such as special relativity, the concepts and mathematical formalism of quantum mechanics, both in one- and three-dimensional model systems, and the applications of quantum theory to topics ranging from atomic/molecular, nuclear, particle, and condensed matter physics to astrophysics.

Undergraduate introductory quantum mechanics course henceforth referred as *UgQM*, is a one semester 4-credit non-relativistic quantum mechanics course where students are typically introduced to basic postulates of quantum mechanics; the Schrödinger wave equation; stationary states; variational method; scattering in one dimension; orbital angular momentum; hydrogen atom; numerical methods. The ModPh and a mathematics course in ordinary and partial differential equations are prerequisites for this UgQM junior-senior level course in quantum mechanics.

Graduate quantum mechanics course in physics henceforth referred as *GrQM*, is a more advanced 3-credit course designed mainly for first year graduate students in physics. Postulates of quantum mechanics, Hilbert space methods, one dimensional potentials, spin systems, harmonic oscillator, angular momentum, and the hydrogen atom are usually the main topics covered in this one semester course. UgQM (or its equivalent) is a prerequisite for this course.

Chemistry quantum mechanics course henceforth referred as *ChemQM*, is a one semester 3-credit non-relativistic quantum mechanics course where students are typically introduced to the principles of quantum mechanics course its applications to the study of chemical bonding and the structure of simple molecules. Topics include an introduction to the principles of quantum mechanics; approximation methods, including the variation method and perturbation theory, and applications of these principles and techniques to chemical bonding. A prerequisite course on chemical principles, including properties of matter and fundamentals of chemical thermodynamics is required.

3.3.2 Data Collection Procedure

The study was conducted in several phases. In phase 1, the core content of non-relativistic introductory quantum mechanics was determined. The core content of non-relativistic introductory quantum mechanics can be conceptualized as the content students would encounter during a fairly, standard non-relativistic introductory quantum mechanics course. Therefore, the predetermination of the core content was an essential process in order to be able to properly assess the core concepts with the QMVI. Common textbooks and syllabi used in Big Ten universities were used as the data source to establish a draft core content. The proposed draft core content with slight modifications was than reviewed by an expert panel of three professors who had extensive experience in teaching introductory quantum mechanics courses.

In phase 2, a draft version of the items which would constitute the QMVI were created based up on the core content created in Phase 1. The QMVI item writing and review procedure QMVI was completed mid-semester, 1999. The first version of the QMVI was administered at the end of the fall semester, 1999. After the data were analyzed, the item difficulty index and the item discrimination index were computed for each item of the QMVI version 1. As a result of this item analysis, 6 questions were dropped from the first version of QMVI (question number 7, 9, 14, 16, 20, and 21). New questions were generated to replace these six questions in the second version of the QMVI. (For a more detailed account on the procedure and subsequent changes on the QMVI please refer to section 3.2.3d).

Figure 3.1 Composite of the Sample

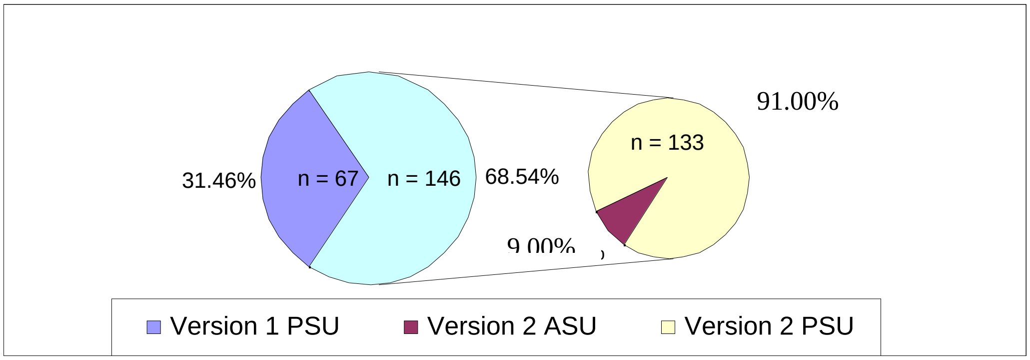


Figure 4.1 presents information on the numbers of students that took part in OMVI version one and two respectively. A total of 213 students took part in this test development and validation study. 67 (31.46%) of these students took version 1of the QMVI whereas 146 (68.54%) took version 2. Out of the 146 students who took version 2 of the QMVI, 133 (91.00%) students were from the Pennsylvania State University and 13 (9.00%) were from the Arizona State University.

In phase 3, the second version of the QMVI was administered at the end of each semester for 3 successive semesters to undergraduate and graduate volunteer students enrolled in either a modern physics courses or introductory quantum mechanics courses at the Pennsylvania State University (PSU) first and then at the Arizona State University (ASU).

Table 3.14 General Overview of the Samples

			Number of Students								
		Groups[1] _	ModPh		UgQM		GrQM		ChemQM		
Semester	University	Version									Total
			Number of students	Section[†] Number	Number of students	Section Number	Number of students	Section Number	Number of students	Section Number	
Fall, 2000	PSU	1	45	1	17	2	5	3			67
Spring, 2000	PSU	2	28	4	16	5					44
Fall, 2000	PSU	2	9	6	19	7	13	8	14	9	55
Spring,2001	PSU	2	34	10	0						34
Spring, 2001	ASU	2			9	11	4	12			13
Total version 1 & 2			116		61		22		14		213
Total version 2			71		44		17		14		146

Table 4.1 presents information on the 12 groups of students that from which data were collected in developing and validating the QMVI. Cells in Table 4.1 report the number of students in each group and the respective section number is shown also. For example, in fall 1999 the first version of the QMVI was administered to three groups of students at Penn State (a total of 67 students). More specifically, 45 of these students were enrolled in the Modern physics course (ModPh), 17 of these students were enrolled in the undergraduate introductory quantum mechanics course (UgQM), and 5 students were enrolled in the graduate quantum mechanics course (GrQM). After the item analysis on the first version of the QMVI changes were made in the items in order to improve the quality of the achievement test. As a result of these changes, version 2 of the QMVI was developed. Version 2 was administered in Spring 2000, Fall 2000, and Spring 2001 to 8 distinct groups with 143 students in all version 2 groups. 130 of the Version 2 students were enrolled at the Pennsylvania State University (i.e., sections 6 to 10) and 13 students were enrolled at the Arizona State University (i.e., sections 11 and 12).

---

[1] Used as independent variables in this study

Figure 3.2 QMVI Version 2, Distribution of Number of Students With Respect to the 4 Distinct Courses

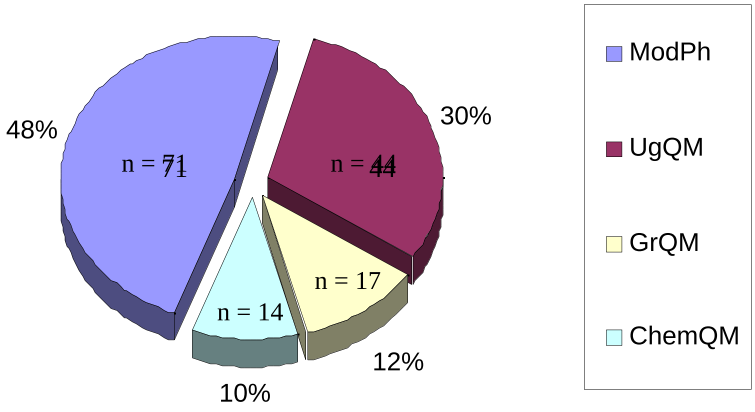


Figure 4.2 shows that among all the students who took the second version of the QMVI, 48% (n = 71) were sophomore students enrolled in the modern physics courses, 30% (n = 44) were junior-senior students enrolled in the undergraduate introductory quantum mechanics courses, 13% (n =17) graduate students were enrolled in a graduate level introductory quantum mechanics physics course, and 10% (n = 14) students were enrolled in an introductory graduate quantum mechanics chemistry course.

3.4 Summary

This chapter, outlined methodology for construction of a valid and reliable achievement test in introductory quantum mechanics that was employed in this study. The model presented two main procedures, content definition and item generation and test revision consisting of 6 steps that the investigator followed.

To define the content of introductory quantum mechanics: (1) “Big Ten” syllabi, principal textbooks and the Graduate Record Examinations test in physics were used as a resources to establish the draft content outline, (2) a panel of experts in the field of college physics judged the draft content outline.

To construct the test: (1) items were created based on the content defined in the previous procedure, (2) a panel of experts in college physics and science education judgments were used to create the first version of the instrument, (3) the item analysis results of the first pilot study were used to compute coefficient alpha, item discrimination index value and item difficulty index value, (4) a panel of experts in college physics and science education judgments were used revise the first version of the instrument, which resulted in the second version of the instrument, (5) the item analysis results of the second pilot study were used to compute coefficient alpha, item discrimination index value and item difficulty index value, (7) panel of experts in physics judged content validity on QMVI version 2 and as a result of this analyses question number 1 was excluded from the final version of the QMVI.

In Chapter IV, part 1, results is presented from the administration of the Quantum Mechanics Visualization Instrument (QMVI). In addition, the concurrent validation procedure of the QMVI will be discussed. Part 2 of chapter 4, examines student understanding as measured by the QMVI following in different kinds of introductory quantum mechanics courses.

CHAPTER 4

Presentation and Analysis of Findings

4.1 Introduction

The principle purpose of the study was to establish a valid and reliable achievement instrument in introductory quantum mechanics. This chapter presents: a) data analyses conducted to determine the validity and reliability of the Quantum Mechanics Visualization Instrument (QMVI), and b) data analyses from which preliminary findings are drawn on students' understanding of introductory quantum mechanics concepts.

Section 4.2 presents results of the data analyses conducted in the validation study; these results are used in Chapter 5 to develop a scientifically sound validation argument to support the uses of the QMVI test scores as an achievement test. The following results of the data analyses are presented: a) the Pearson product-moment correlation between students' QMVI scores and their perceptions of item difficulty; b) the one-way analysis of variance of the four course mean scores; and c) the Pearson product-moment correlation between students' QMVI scores and their final course grades.

Section 4.3 of this chapter summarizes the item analysis and reports the Cronbach alpha reliability coefficient of the QMVI. Section 4.4 presents the data analyses on the five Main Topics from which preliminary findings are drawn on students' understanding of introductory quantum mechanics.

4.2 Validity of the QMVI

Test validity is not an all or none concept, rather it is the degree to which the interpretation of a test score is believed to accurately represent what the test score is purported to measure. Unfortunately, due to the nature of validity there is no one absolute statistical method to determine test validity. Consequently, test validation is an ongoing process of collecting qualitative and quantitative evidence from many different resources.

The first part of the validation study was qualitative in nature. Core content had to be established to provide a “blueprint” of concepts and procedures upon which the QMVI instrument could be constructed.

The following section utilizes a quantitative approach to provide further evidence for the validity of the QMVI. Since the QMVI is an instrument that measures students’ achievement in non-relativistic introductory quantum mechanics it is expected that certain relationships of correlations do exist with external variables as well as with internal variables (independent variables).

More specifically, the following correlation should exist:

1. A moderate positive correlation between students’ confidence level rating and QMVI mean score
2. A moderate positive correlation between students’ course grade and QMVI mean score
3. UgQM and GrQM student groups should out perform the ModPh student group.

4.2.1 Correlation Between Students’ QMVI score and Their Confidence Level

To acquire evidence associated with the validity of the QMVI, students were asked to rate their confidence on a four level Likert scale for each question they answered on the test. The assumption underlying the statistical analysis is that students who were able to understand the question, could then judge their ability to answer that particular question correctly or incorrectly. Of course, a student who did not understand the question well or does not know the topic might also have some problems determining how well he or she performed on that particular question. If a positive correlation between the students' QMVI scores and their confidence levels exists, one can deduce that the students believed they understood what they were reading. More specifically, a student who scored relatively high on the QMVI also should have fairly high confidence in the correctness of his or her responses. While a student who scored low would be expected to have a relatively low level. The following section examines the correlation between students’ QMVI scores and their confidence level: a) first at the four distinct course level (ModPh, UgQM, GrQM, and ChemQM), b) and then at the section level (4 – 12).

4.2.1a Correlation Between QMVI Score and Confidence Level for the Modern Physics Course

Figure 4.1 is a scatter plot of modern physics students’ QMVI scores versus their respective confidence levels. The scatter plot depicts a positive trend. In general, students with higher QMVI scores had higher levels of confidence.

Figure 4.1 ModPh Students' Confidence Levels versus Their QMVI scores.

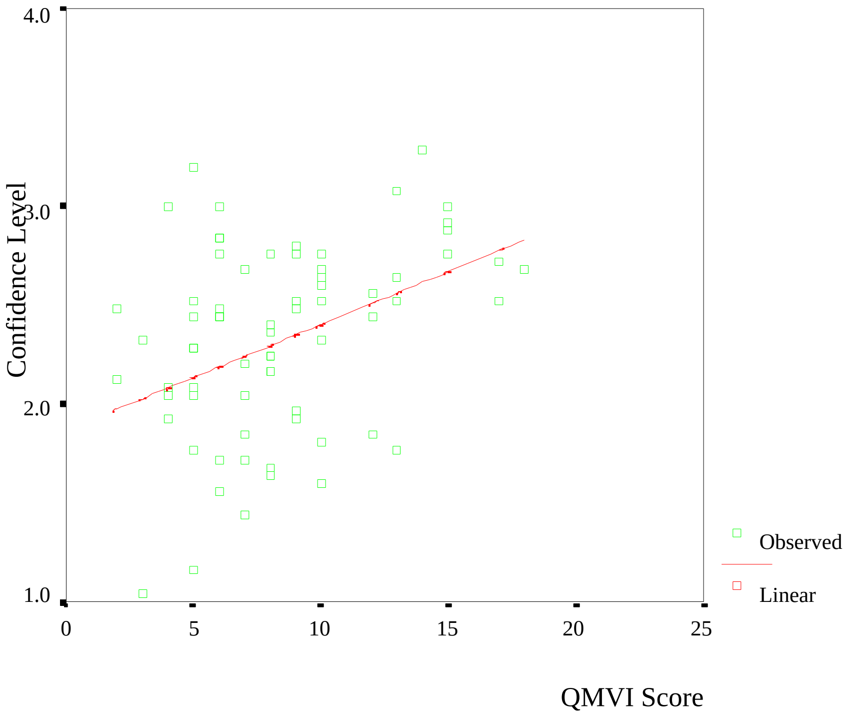


Table 4.1 QMVI Score versus Confidence: *Modern Physics*

		QMVI Score
Confidence	Pearson Correlation	.393
	Sig. (2-tailed)	.001
	N	71

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.1 presents the results of the Pearson product moment correlation between the QMVI score of the Modern Physics group and their confidence level. A statistically significant positive correlation of 0.393 was found at $p < 0.01$.

4.2.1b Correlation Between QMVI Score and Confidence Level for the Undergraduate Quantum Mechanics Course - Junior-Senior Group

Figure 4.2 is a scatter plot between undergraduate introductory quantum mechanics group students' QMVI scores versus their respective confidence level. The scatter plot depicts a positive trend, that is, in general students with higher QMVI scores had a higher confidence level.

Figure 4.2 UgQM Students' Confidence Level versus Their QMVI Scores

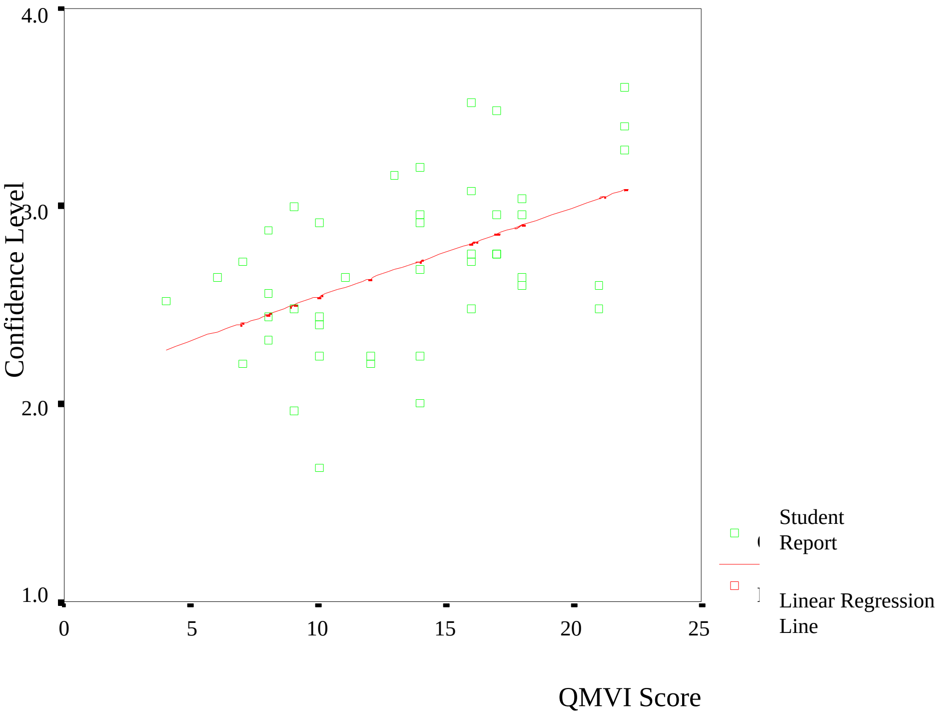


Table 4.2 QMVI Score versus Confidence: Undergraduate Introductory Quantum Mechanics

		QMVI Score
Confidence	Pearson Correlation	.501
	Sig. (2-tailed)	.001
	N	44

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.2 presents the results of the Pearson product moment correlation between the undergraduate introductory quantum mechanics group and their confidence level. A statistically significant positive correlation of 0.501 was found at an alpha level of 0.01.

4.2.1c Correlation Between QMVI Score and Confidence Level for the Graduate Introductory Quantum Mechanics Course

Figure 4.3 is a scatter plot between graduate introductory quantum mechanics group students' QMVI scores versus their respective confidence level. The scatter plot depicts a positive trend, that is, in general students with higher QMVI scores had a higher confidence level.

Figure 4.3 GrQM Students' Confidence Level versus Their QMVI Scores

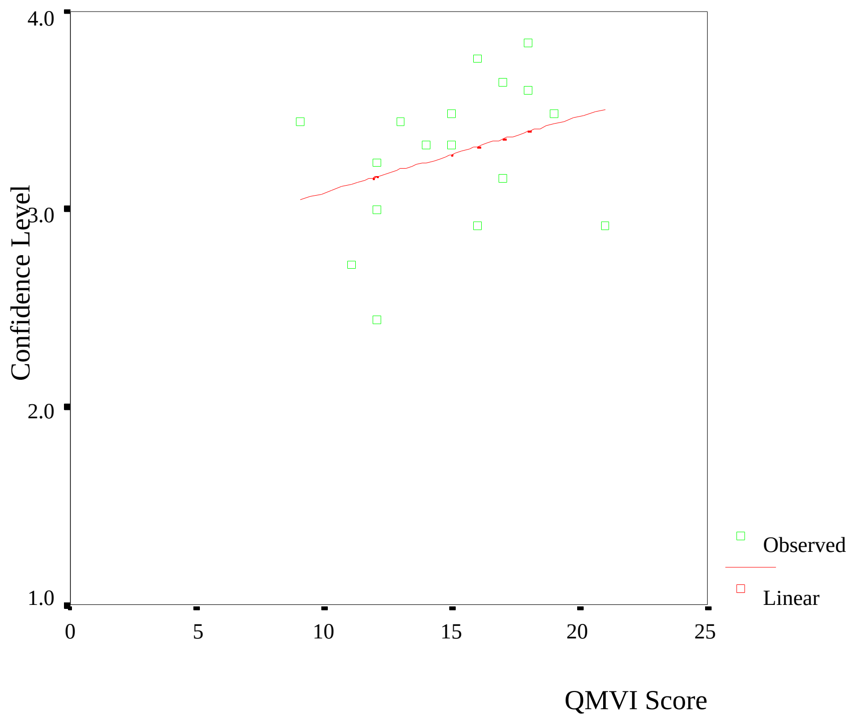


Table 4.3 QMVI Score versus Confidence: *Graduate Introductory Quantum Mechanics*

		QMVI_S
Confidence	Pearson Correlation	.325
	Sig. (2-tailed)	.204
	N	17

Table 4.3 presents the results of the Pearson product moment correlation between the undergraduate introductory quantum mechanics group and their confidence level. A positive correlation of 0.325 was found which was not statistically significant at $p < 0.01$.

4.2.1d Correlation Between QMVI Score and Confidence Level for the Graduate Chemistry Quantum Mechanics Course

Figure 4.4 is a scatter plot between graduate introductory quantum mechanics chemistry group students' QMVI scores versus their respective confidence level. The scatter plot depicts a positive trend, that is, in general students with higher QMVI scores had a higher confidence level.

Figure 4.4 ChemQM Students' Confidence Level versus Their QMVI Scores

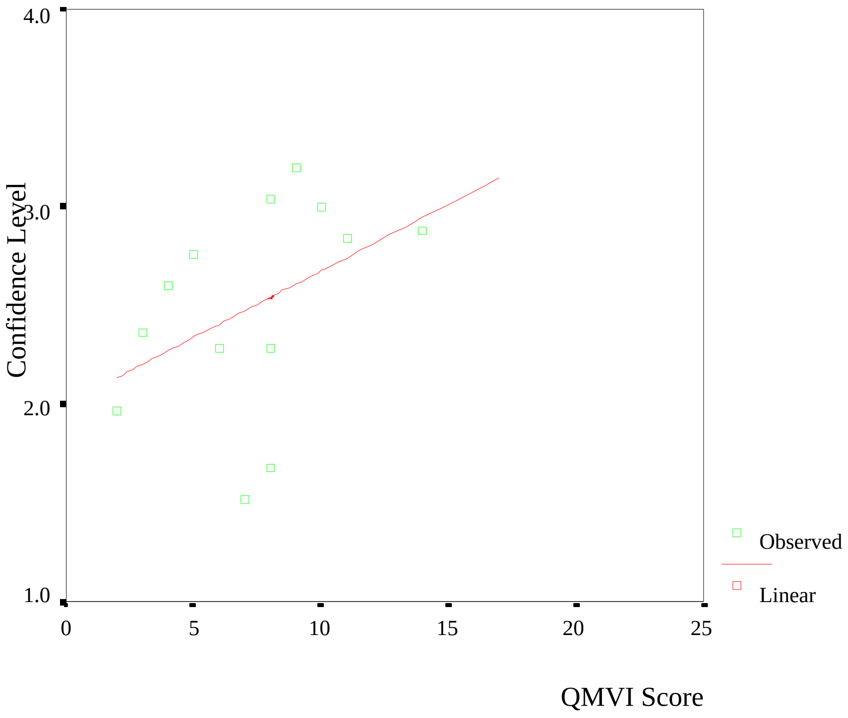


Table 4.4 QMVI Score versus Confidence: *Chemistry Quantum Mechanics*

		QMVI Score
Confidence	Pearson Correlation	.425
	Sig. (2-tailed)	.148
	N	13

Table 4.4 presents the results of the Pearson product moment correlation between the undergraduate introductory quantum mechanics chemistry group and their confidence level. A positive correlation of 0.425 was found which was not statistically significant at an alpha level of 0.01.

4.2.1e Correlation Between QMVI Score and Confidence Level for the All Courses

Figure 4.5 is a scatter plot between all students' QMVI scores versus their respective confidence level. The scatter plot depicts a positive trend, that is, in general students with higher QMVI scores had a higher confidence level.

Figure 4.5 All Students' Confidence Level versus Their QMVI Scores

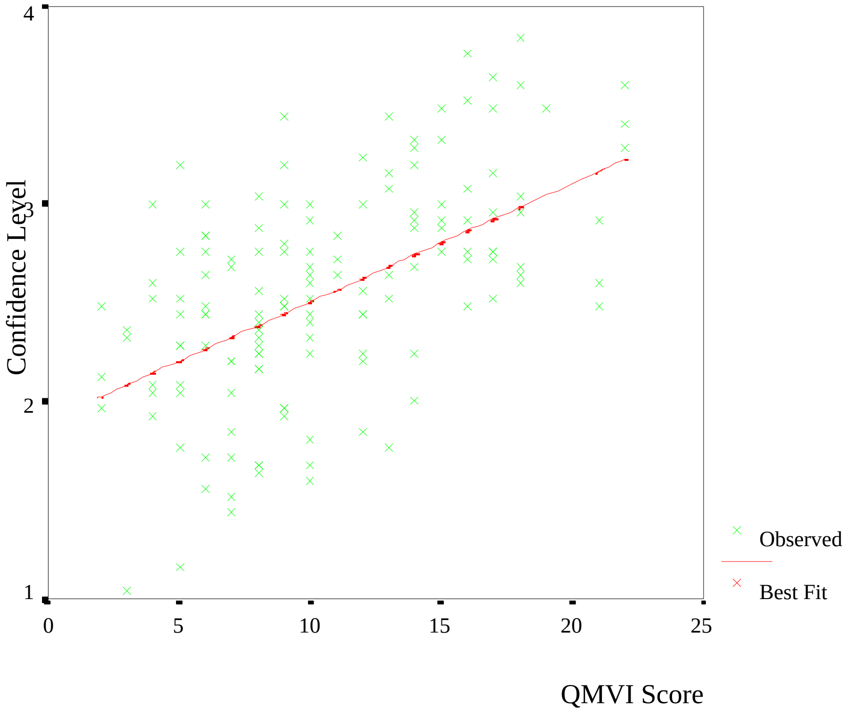


Table 4.5 QMVI Score versus Confidence: ***All Courses***

		QMVI Score
Confidence	Pearson Correlation	.490
	Sig. (2-tailed)	.000
	N	146

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.5 presents the results of the Pearson product moment correlation between the all students' QMVI score and their confidence level. A statistically significant positive correlation of 0.490 was found at an alpha level of 0.01.

4.2.2 Comparisons Between Students' QMVI Mean Score by Student Groups and Sections

Figure 4.6 is a boxplot of the QMVI scores by sections. Figure 4.6 shows the mean score distribution as well as the variance of QMVI scores across for section. The Modern Physics groups in this study scored consistently lower than the Graduate Quantum Mechanics and Undergraduate Quantum Mechanics groups.

Figure 4.6 QMVI Mean Scores of Course

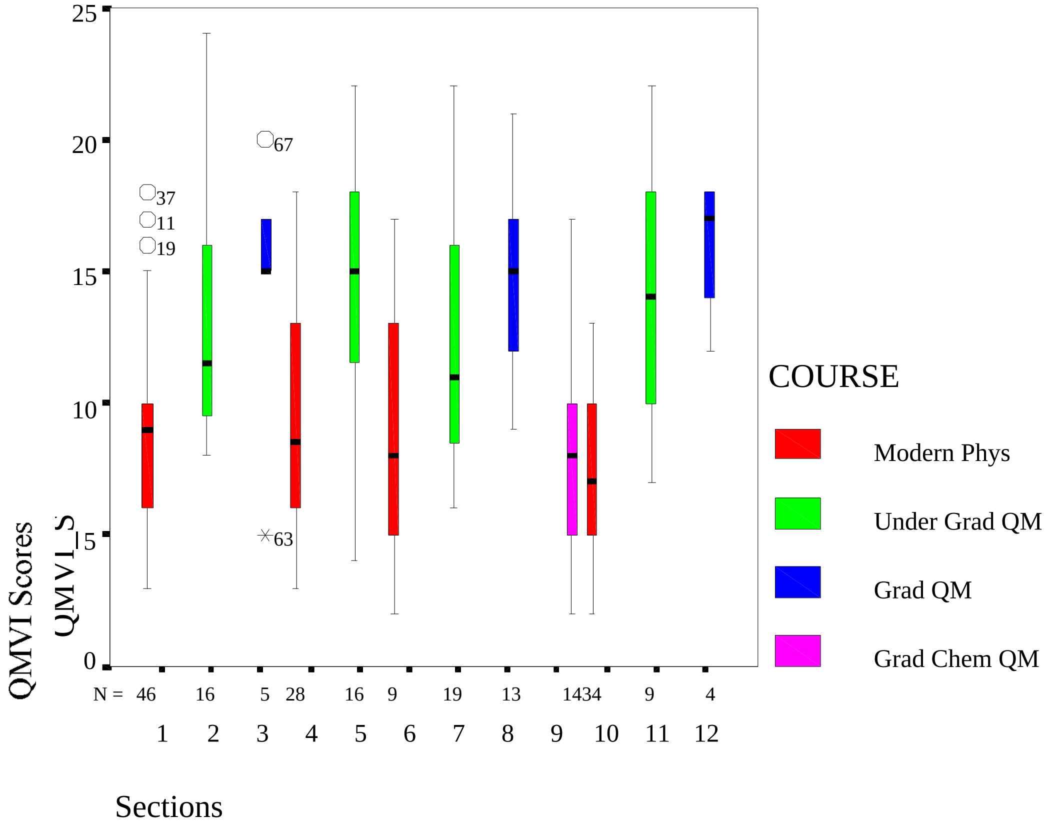


Table 4.6 ANOVA for QMVI Mean Scores of the Four Groups.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1160.360	3	386.787	23.17	.000
Within Groups	2370.160	142	16.691		
Total	3530.521	145			

A one-way analysis of variance (ANOVA) was used to determine whether or not there were statistical differences across the scores of students in different courses at alpha = 0.05 level. Table 4.6 summarizes the analysis of variance for the QMVI mean scores. The F value of 23.17 is significant at $p < 0.05$. This means that there is at least one QMVI course mean score that is statistically significant different among the four distinct QMVI course means.

To determine which of the possible paired comparisons were statistically significant the Tukey HSD was used. This statistic compares the obtained mean difference value with a critical value. Table 4.7 presents the result of the Tukey HSD follow up test for the four different QMVI course mean scores.

Table 4.7 Tukey HSD Test of Significances for the Four Distinct QMVI Course Mean Scores

		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
(I) Mean Score	(J) Mean Socre				Lower Bound	Upper Bound
ModPh	UgQM	-5.1674	.7839	.000	-7.1812	-3.1537
	GrQM	-6.6901	1.1031	.000	-9.5242	-3.8561
	ChemQM	.3099	1.1947	.994	-2.7594	3.3791
UgQM	ModPh	5.1674	.7839	.000	3.1537	7.1812
	GrQM	-1.5227	1.1667	.560	-4.5200	1.4746
	ChemQM	5.4773	1.2536	.000	2.2567	8.6979
GrQM	ModPh	6.6901	1.1031	.000	3.8561	9.5242
	UgQM	1.5227	1.1667	.560	-1.4746	4.5200
	ChemQM	7.0000	1.4745	.000	3.2120	10.7880
ChemQM	ModPh	-.3099	1.1947	.994	-3.3791	2.7594
	UgQM	-5.4773	1.2536	.000	-8.6979	-2.2567
	3.00	-7.0000	1.4745	.000	-10.7880	-3.2120

* The mean difference is significant at the .05 level.

Table 4.7 presents the results of the Tukey HSD follow-up test for the QMVI course mean scores. Students from the graduate quantum mechanics course (GrQM) and students from the undergraduate introductory quantum mechanics courses QMVI mean scores' were significantly higher ($p < 0.05$) than both the modern physics and graduate chemistry quantum mechanics courses' QMVI mean scores. However, there was no statistical difference between the means scores of the graduate quantum mechanics course and undergraduate introductory quantum mechanics course were not statistically significant at the 0.05 level. There were also no statistically significant differences between the mean scores of Modern Physics students and Chemistry Quantum Mechanics students.

4.2.3 Correlation Between Students' QMVI Mean Score and Course Grade

The Office of Regulatory Compliance at the Pennsylvania State University required not only the explicit consent of students who took part in this OMVI test development and validation study but also required the researchers to obtain a separate consent from the same students to use external data such as their course grades. Therefore, students who participated in this study were given the right to consent only for their QMVI scores or for both their QMVI scores and their course grades.

Table 4.8 Student Consent for QMVI Score Only or Both QMVI Score and Course Grade

			Consent	
Section	Course	Institution	QMVI Only	QMVI & Course Grade
4	ModPh	PSU	28	25
5	UgQM	PSU	9	9
6	ModPh	PSU	16	16
7	UgQM	PSU	19	19
8	GrQM	PSU	13	13
9	ChemQM	PSU	14	14
10	ModPh	PSU	34	30
11	UgQM	ASU	9	-
12	GrQM	ASU	4	-
Total			146	126

Table 4.8 shows that not all students who consented to take part in the QMVI test gave consent for their course grades to be released. 126 students out of 146 (86.3%) students consented to the use of both their QMVI scores and course grades and 20 students (13.7%) gave consent only for their QMVI scores. Therefore, the following analysis includes only those 126 students who gave permission for the use of their course grades. The QMVI scores were used by some instructors to determine the student's final course grade. Therefore, the correlation coefficient computed in this section will reflect a systematic error.

Table 4.9 Correlation by Section: QMVI Scores versus Course Grades

Sections	Courses		# of students	Mean	SD	R	Sig
4	ModPh	QMVI score	25	9.36	4.14	0.40*	0.05
		Grade	25	3.42	0.77		
5	UgQM	QMVI score	9	9.00	5.10	0.43	0.24
		Grade	9	3.11	0.74		
6	ModPh	QMVI score	16	14.63	5.14	0.67*	0.00
		Grade	16	3.07	0.95		
7	UgQM	QMVI score	19	12.22	4.26	0.77*	0.00
		Grade	19	3.30	0.66		
8	GrQM	QMVI score	13	14.69	3.35	0.80*	0.00
		Grade	13	3.59	0.53		
9	ChemQM	QMVI score	14	8.00	4.13	0.49*	0.08
		Grade	14	3.50	0.62		
10	ModPh	QMVI score	30	7.26	2.83	0.24	0.21
		Grade	30	3.21	0.56		

Table 4.9 presents the results of the Pearson product moment correlation for each section between students' QMVI scores and their final course grades. A moderate positive correlation coefficient ranging between 0.24 (section 10) to 0.80 (section 8) is observed. For sections 4, 6, 7, 8, and 9 statistically significant correlation was found to exist between the students' QMVI score and their final course grades ($p < 0.1$). For section 5 and 10 the correlation coefficient was not statistically significant ($p > 0.1$). However, it should be noted that section 5 had only 9 students.

Table 4.10 Correlation by Group: QMVI Scores versus Course Grades

Courses		# of students	Mean	SD	R	Sig.
ModPH	QMVI score	64	8.31	3.79	0.37*	0.03
	Grade	64	3.28	0.67		
UgQM	QMVI score	35	13.48	4.78	0.64*	0.00
	Grade	35	3.19	0.80		
GrQM	QMVI score	13	14.69	3.35	0.80*	0.00
	Grade	13	3.59	0.53		
ChemQM	QMVI score	14	8.00	4.13	0.49*	0.08
	Grade	14	3.50	0.62		

Table 4.10 presents the results of the Pearson product moment correlation for groups between students' QMVI scores and their final course grades. A moderate to high positive correlation coefficient ranging between 0.49 (ChemQM) to 0.80 (GrQM) is observed. All correlation coefficients by group between the students' QMVI score and their final course grades were found to be statistically significant ($p < 0.1$).

Table 4.11 Correlation Between QMVI Scores versus Course Grades

All students	# of students	Mean	SD	R	Sig.
QMVI score	126	10.62	4.93	0.42*	0.00
Grade	126	3.31	0.70		

Table 4.11 presents the results of the Pearson product moment correlation between students' QMVI scores and their final course grades. A moderate positive correlation coefficient of 0.42 was observed. This correlation coefficient between the students' QMVI score and their final course grades was found to be statistically significant correlation ($p < 0.1$).

4.3 Results of Item Analysis

The 24 items on the QMVI were subjected to the ITEMAN and SPSS computer programs. ITEMAN data files (ASCII format) of test item responses were produced by manual data entry. ITEMAN can be used to score and analyze multiple-choice and true/false test data, as well as survey (e.g., Likert-type rating scale) data. ITEMAN computes summary statistical information for each scale, including the mean, variance, standard deviation, skew, and kurtosis of total scores, the minimum and maximum score, and median score.

SPSS or Statistical Package for Social Sciences is an advanced statistical program that performs a variety of statistical analysis, including the computation of test reliability coefficient. This program is available at the Pennsylvania State University Computation Center.

Table 4.12 shows the results of the item analysis conducted on the second version of the QMVI to a sample of 146 students. An overall improvement can be observed from version one to version two and the results are very promising. For example, the test reliability coefficient Alpha has improved from 0.75 to 0.82. The data skewness has decreased from 0.85 to 0.41 and the test mean has shown an absolute improvement of 5.21%. The range of the scores, that is the difference between the maximum and minimum score, remained the same. Moreover the mean item difficult remained almost the same with a value of 0.4. The ideal item difficulty index for an achievement test would be 0.5, the value of 0.4 indicates a slightly difficult test.

Table 4.12 Overall Item Descriptive Statistics for the QMVI

	Version 2	Version 1
Number of Items	24	25
Number of Examinees	146	67
Mean	10.00	9.9
Variance	21.89	18.2
Std. Dev.	4.68	4.3
Skew	0.41	0.85
Minimum	2	3
Maximum	21	24
Median	9.00	9.00
Alpha	0.82	0.75
Mean Item Total	0.42	0.4

Table 4.13 Descriptive statistics for each Item on the QMVI

Item Number	Item Mean Difficulty	Standard Deviation
1	.51	.50
2	.36	.48
3	.47	.50
4	.91	.29
5	.56	.50
6	.29	.46
7	.27	.45
8	.51	.50
9	.22	.42
10	.15	.36
11	.49	.50
12	.64	.48
13	.17	.38
14	.35	.48
15	.32	.47
16	.27	.45
17	.74	.44
18	.40	.49
19	.18	.39
20	.54	.50
21	.39	.49
22	.38	.49
23	.26	.44
24	.12	.32

Table 4.13 shows descriptive statistics for each item on the QMVI. The item mean difficulty value ranged between 0.12 to 0.91 and their standard deviations ranged between 0.32 and 0.29. 7 questions (1, 4, 5, 8, 12, 17, & 20; 29%) of the QMVI had item mean difficulty indices equal or higher than 0.5. 8 questions (2, 3, 11, 14, 15, 18, 21, & 22; 33%) of the QMVI had item mean difficulty indices ranging between 0.3 and 0.5. 5 questions (6, 7, 9, 16, & 23; 21%) of the QMVI had item mean difficulty indices ranging between 0.2 and 0.3. 4 questions (10, 13, 14, & 24; 17%) of the QMVI had item mean difficulty indices ranging between 0.1 and 0.2.

Figure 4.7 QMVI Item Mean Score Distribution of the 4 Distinct Courses.

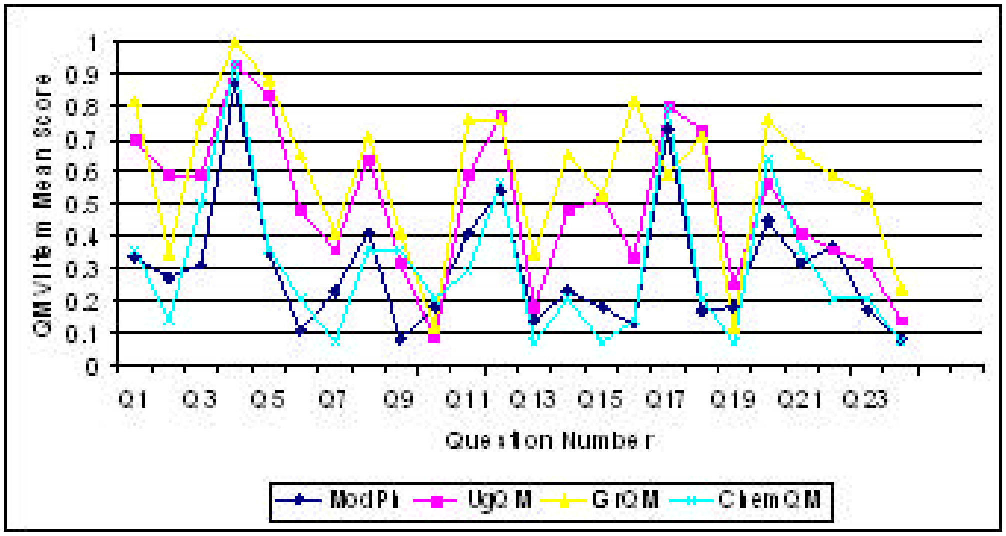


Figure 4.7 displays the QMVI item mean score distribution with respect to the four distinct courses. The graph displays clearly a parallel trend in the performance of students in the four courses. In general, students enrolled in the graduate quantum mechanics course outperformed the other three student groups at the end of their formal course. Moreover the ModPh group had the lowest item mean scores expect for question number 10 where they slightly did better than all the other courses.

The one-way analysis of variance (ANOVA) was used in order to determine whether or not there were statistical differences on each of the 24 QMVI item mean scores among the four distinct courses. Table 4.14 is the result of the analysis of the variance summary table for the 24 QMVI item mean scores.

Table 4.14 ANOVA for All Items With Respect to Each Distinct Course

		Sum of Squares	df	Mean Square	F	Sig.
Q1	Between Courses	5.762	3	1.921	8.875	.000
	Within Courses	30.731	142	.216		
	Total	36.493	145			
Q2	Between Courses	3.612	3	1.204	5.671	.001
	Within Courses	30.148	142	.212		
	Total	33.760	145			
Q3	Between Courses	3.950	3	1.317	5.775	.001
	Within Courses	32.378	142	.228		
	Total	36.329	145			
Q4	Between Courses	.259	3	8.643E-02	1.060	.368
	Within Courses	11.583	142	8.157E-02		
	Total	11.842	145			
Q5	Between Courses	8.883	3	2.961	15.536	.000
	Within Courses	27.063	142	.191		
	Total	35.945	145			
Q6	Between Courses	6.020	3	2.007	11.719	.000
	Within Courses	24.315	142	.171		
	Total	30.336	145			
Q7	Between Courses	1.419	3	.473	2.431	.068
	Within Courses	27.622	142	.195		
	Total	29.041	145			
Q8	Between Courses	2.413	3	.804	3.351	.021
	Within Courses	34.080	142	.240		
	Total	36.493	145			
Q9	Between Courses	2.616	3	.872	5.535	.001
	Within Courses	22.370	142	.158		
	Total	24.986	145			
Q10	Between Courses	.307	3	.102	.791	.501
	Within Courses	18.378	142	.129		
	Total	18.685	145			
Q11	Between Courses	2.786	3	.929	3.912	.010
	Within Courses	33.707	142	.237		
	Total	36.493	145			
Q12	Between Courses	1.884	3	.628	2.797	.042
	Within Courses	31.877	142	.224		
	Total	33.760	145			
Q13	Between Courses	.771	3	.257	1.830	.144
	Within Courses	19.948	142	.140		
	Total	20.719	145			
Q14	Between Courses	3.574	3	1.191	5.713	.001
	Within Courses	29.611	142	.209		
	Total	33.185	145			
Q15	Between Courses	4.746	3	1.582	8.394	.000
	Within Courses	26.761	142	.188		
	Total	31.507	145			

Table 4.14 Continued

Q16	Between Courses	7.111	3	2.370	15.347	.000
	Within Courses	21.930	142	.154		
	Total	29.041	145			
Q17	Between Courses	.560	3	.187	.963	.412
	Within Courses	27.549	142	.194		
	Total	28.110	145			
Q18	Between Courses	10.572	3	3.524	20.353	.000
	Within Courses	24.586	142	.173		
	Total	35.158	145			
Q19	Between Courses	.444	3	.148	.974	.407
	Within Courses	21.563	142	.152		
	Total	22.007	145			
Q20	Between Courses	1.607	3	.536	2.196	.091
	Within Courses	34.646	142	.244		
	Total	36.253	145			
Q21	Between Courses	1.464	3	.488	2.082	.105
	Within Courses	33.282	142	.234		
	Total	34.747	145			
Q22	Between Courses	1.145	3	.382	1.636	.184
	Within Courses	33.135	142	.233		
	Total	34.281	145			
Q23	Between Courses	2.000	3	.667	3.625	.015
	Within Courses	26.110	142	.184		
	Total	28.110	145			
Q24	Between Courses	.358	3	.119	1.157	.328
	Within Courses	14.662	142	.103		
	Total	15.021	145			

The F values of 1.06, 2.43, 0.79, 1.83, 0.96, 0.97, 2.20, 2.10, 1.64, and 1.16 for question numbers 4, 7, 10, 13, 17, 19, 20, 21, 22, and 24 respectively indicated that there was no statistical significant difference between the QMVI item mean scores among the four distinct courses at the alpha 0.05 level. For all the other questions there was at least one statistical difference between the QMVI item mean scores among the four distinct student groups at the alpha 0.05 level.

4.4 Student Understanding: Overall, Course, and Section Results

Sections 4.2 and 4.3 reported the results of statistical analyses on the validity and reliability of the QMVI. This section examines the understanding of the student samples as measured by the QMVI.

These analyses provide information that can suggest improvements in the QMVI and support the design of more effective quantum mechanics textbooks and courses. More specifically, students' achievements in the 5 Main Topics are analyzed. In addition, performance on each question making up a Main Topic is examined separately and compared with performance on the other questions of the same Main Topic. Moreover, students' achievement among and within the courses is also examined. Note that all the data reported in this section is taken from version 2 of the QMVI. Note also that the content definition process is described in detail in Chapter 3.

Table 4.15 shows the numbers of the specific questions associated with each Main Topic and the percent weighting of each of the topics on the QMVI. The percent weights of each of the core content topics on the QMVI were intended to be the same as the topic weight assigned to each of those quantum mechanics topics by the expert panel. In the QMVI version 2, the *Historical Development & Terminology* and *Schrödinger Equation* topics are slightly overrepresented whereas the *Advanced Application* and *Quantum Mechanics* topics are slightly underrepresented.

Table 4.15 Table of Specification

Main Topic	QMVI Question Numbers	Number of Questions	QMVI Content Representation	Expert's Content Representation
A. Historical Development & Terminology	1, 2,3,6	4	13.3%	10%
B. Quantum Mechanics Postulates	9, 14, 15, 16, 19	5	16.7%	20%
C. Schrödinger Equation	4,5,6,7,8, 10,18,19	8	26.7%	22%
D. Schrödinger Application Equation - one dimension	10,11,12,13, 14,15,17,18	8	26.7%	27%
E. Advanced Applications	20,21,22,23,24	5	16.7%	20%

Figure 4.8 Student Understanding by Course

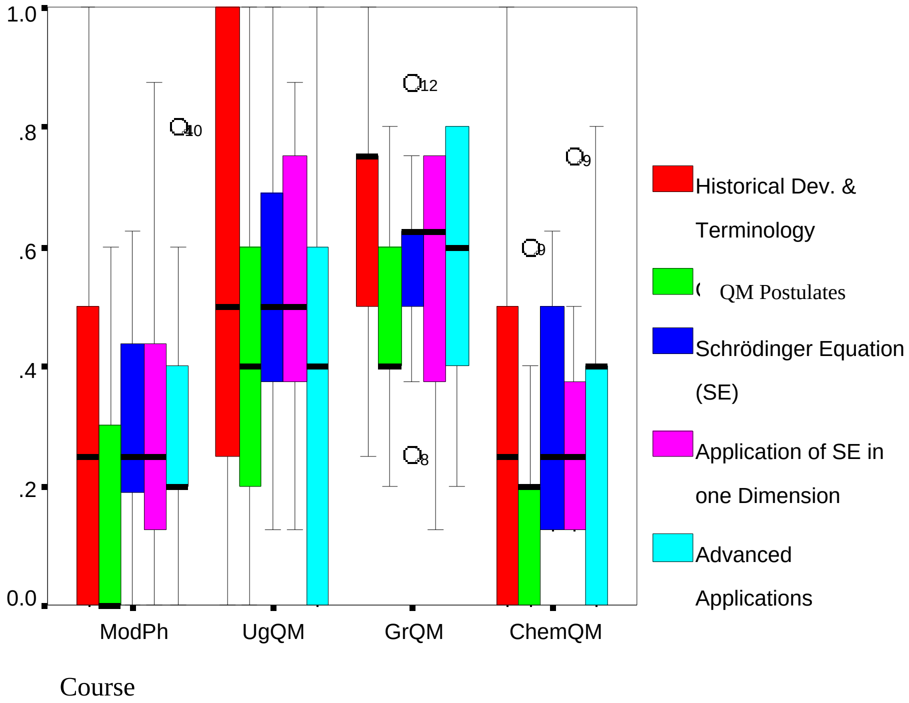


Figure 4.8 is a visual "boxplot" and Table 4.16 is a report of the performance data of students in each of the courses in the various Main Topics. A visual examination of Figure 4.8 and Table 4.16 reveal several interesting findings. First, for all *Main Topics* (A-E) students in the *Graduate Quantum Mechanics* course (GrQM) seem to have a smaller OMVI score variance than the other students' QMVI scores in all the other courses. Although the GrQM achievement may appear at first glance to be similar for Main Topic C and D with achievement in the *Undergraduate Quantum Mechanics* (UgQM) course, the data indicate a more uniform understanding suggested by less variance in the GrQM course. For example, a comparison of the standard deviation values between UgQM and GrQM (Table 4.16) shows Main Topics C and D lower standard deviation values in the GrQM course. Second, a comparison of students' performance across Main Topics A-E, shows that students' achievement scores were the lowest on Main Topic B. Among the five Main Topics, Main Topic B covers the

fundamental and perhaps the most abstract content including the postulates of quantum mechanics (see Chapter 3 section 3.2 for details). Third, performance in the graduate Chemistry Quantum Mechanics (ChemQM) course on the QMVI was consistently lower than in the UgQM and the GrQM courses. Students in the ChemQM course showed a similar level of understanding across the five Main Topics as did students in the *Modern Physics* (ModPh) course.

Table 4.16 Student Course Mean Scores by Main Topic

		Main Topic				
Courses		A	B	C	D	E
ModPh	N	71	71	71	71	71
	Mean	.2570	.1606	.3134	.3222	.2789
	Median	.2500	.0000	.2500	.2500	.2000
	Std. Deviation	.2927	.1901	.1593	.2141	.2097
	Range	1.00	.60	.63	.88	.80
UgQM	N	44	44	44	44	44
	Mean	.5909	.3818	.5398	.5199	.3591
	Median	.5000	.4000	.5000	.5000	.4000
	Std. Deviation	.3620	.2920	.1977	.2222	.2912
	Range	1.00	1.00	.88	.75	1.00
GrQM	N	17	17	17	17	17
	Mean	.6471	.5059	.5735	.5588	.5529
	Median	.7500	.4000	.6250	.6250	.6000
	Std. Deviation	.2176	.1886	.1535	.1932	.2065
	Range	.75	.60	.63	.63	.60
ChemQM	N	14	14	14	14	14
	Mean	.3036	.1714	.3036	.3036	.3000
	Median	.2500	.2000	.2500	.2500	.4000
	Std. Deviation	.3422	.1729	.1816	.1748	.2320
	Range	1.00	.60	.50	.63	.80

4.4.1 Main Topics: Analyses by Section, Course, and University

The following sections present a synthesis of students' performance on each Main Topic analyzed by student *section*, *course*, and *university*. More specifically the data matrix of the four courses, made up of 8 sections from two different universities are analyzed initially with respect to the five Main Topics and then further analyzed with respect to the questions which made up these Main Topics. The initial procedure involved generating detailed descriptive statistics for each section of the four courses for each main Topic. Then, the data were analyzed by comparing the mean scores of the four courses. Subsequently, students’ mean scores were compared across the sections with within each course. Finally, performance in the questions within each Main Topic was compared. It is not appropriate to include a complex and long data matrix of the various comparison analyses in the text of this dissertation. Instead, this section reports significant deviations which result from the comparison analyses between the course means, section means, and question means, and provides possible explanations to suggest why these deviations might have occurred.

4.4.1a Main Topic A: Historical Development & Terminology

Questions 1, 2, 3, & 6 shown in Table 4.15 (see page 91) focus on aspects of “Historical Development & Terminology” (i.e., Blackbody radiation, Photoelectric effect, Bohr Atom, Wave Particles duality, deBroglie hypothesis, Uncertainty Principle, & semi-classical behavior, waves). The first three questions assess students’ understanding of semi-classical behavior of particles whereas the last question assesses students’ understanding the Wave-Particle duality and the probabilistic issues associated with this concept. Figure 4.9 and Table 4.17 show the achievement results in Main Topic A for each individual section.

Figure 4.9 Main Topic A: Student Achievement by Section

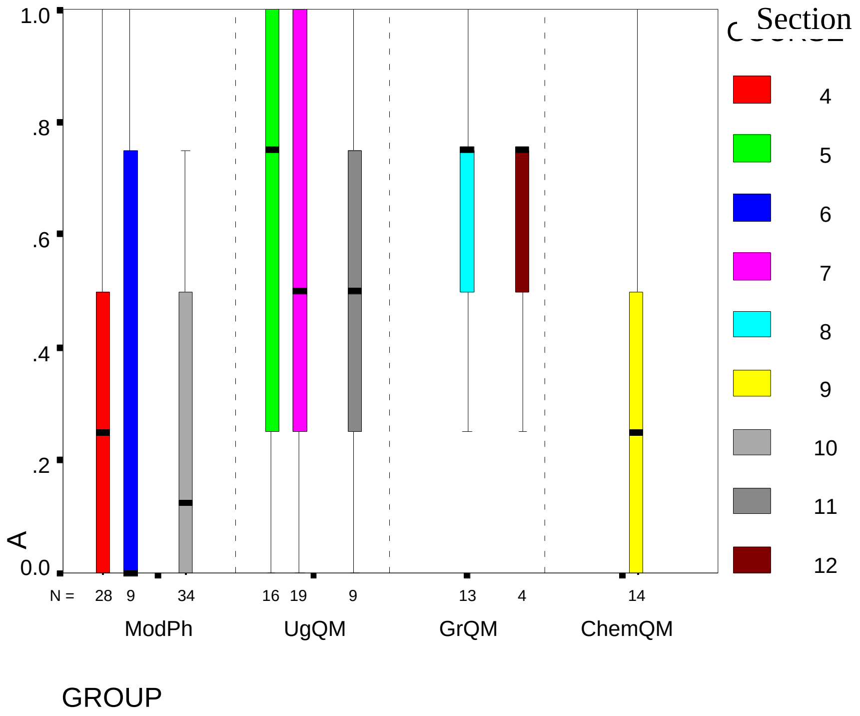


Table 4.17 Main Topic A: Students' Mean Scores by Sections

Section	Course	University	N	Mean	Median[2]	Std. Deviation
4	ModPh	PSU	28	.2946	.2500	.3047
5	UgQM	PSU	16	.6406	.7500	.3532
6	ModPh	PSU	9	.3056	.0000	.4104
7	UgQM	PSU	19	.5921	.5000	.3652
8	GrQM	PSU	13	.6538	.7500	.2174
9	ChemQM	PSU	14	.3036	.2500	.3422
10	ModPh	PSU	34	.2132	.1250	.2472
11	UgQM	ASU	9	.5000	.5000	.3953
12	GrQM	ASU	4	.6250	.7500	.2500
Total			146	.4075	.2500	.3536

---

[2] is represented by the black horizontal line in the boxplot

The results of the more detailed analysis shown in Table 4.17 are consistent with the results of the more global analysis discussed in the previous section. Both sections of the Graduate Quantum Mechanics course (sections 8, & 12) consistently scored higher than did students in the Modern Physics course (sections 4, 6, & 10), the Undergraduate Quantum Mechanics course (sections 5, 7, & 11) and the Chemistry Quantum Mechanics (section 9) course indicating that the students' mean score for the Main Topic A "Historical Development & Terminology" was higher for students with higher levels of education. Indeed, an ANOVA confirms that the GrQM course's mean score on Historical Development & Terminology was statistically significant higher than the ModPh and ChemQM at the $p < 0.05$ level (see Appendix E). Moreover, the mean scores and especially the variance of the scores for the two different institutions (PSU & ASU) were similar as well, reinforcing the idea that students develop a better understanding over the years, somewhat independent of universities and/or instructor (We must be very cautious in interpreting this information due to the small sample size and the number of successive analyses).

Students' achievement on each question was compared across each of the four questions constituting Main Topic A. The results show that the GrQM students performed better on all questions. However, in all courses distinctly lower scores were observed for question number 6 (see Appendix F). A possible reason for these relatively lower scores is that this is a cross topic question. It assesses students' understanding of the probability concept (Main Topic A) and also requires a quantitative understanding of computing the expectation value of a wave function (Main Topic C).

In summary, data analyzed by section, course and university showed that higher levels of formal education were reflected in higher mean scores in Main Topic A especially in areas such as semi-classical behavior, wave, and probability. However, mean scores were consistently lower across all courses when similar questions had quantitative components.

4.4.1b Main Topic B: Quantum Mechanics Postulates

Questions 9, 14, 15, 16, and 19 assess students' understanding of concepts represented in Main Topic "Quantum Mechanics Postulates". As in Main Topic A, Figure 4.10 and Table 4.19 also show that GrQM students' achievement was higher than in all the other courses with a decreased variance in the scores. Even when the data were analyzed by section the same statement holds true. However, an analysis of variance (ANOVA) yielded no statistically significant differences in the mean scores between the GrQM sections (8 & 12) and UgQM sections (5, 7, &11) at $p > 0.05$ level (see Appendix E).

Figure 4.10 Main Topic B: Student Achievement by Section

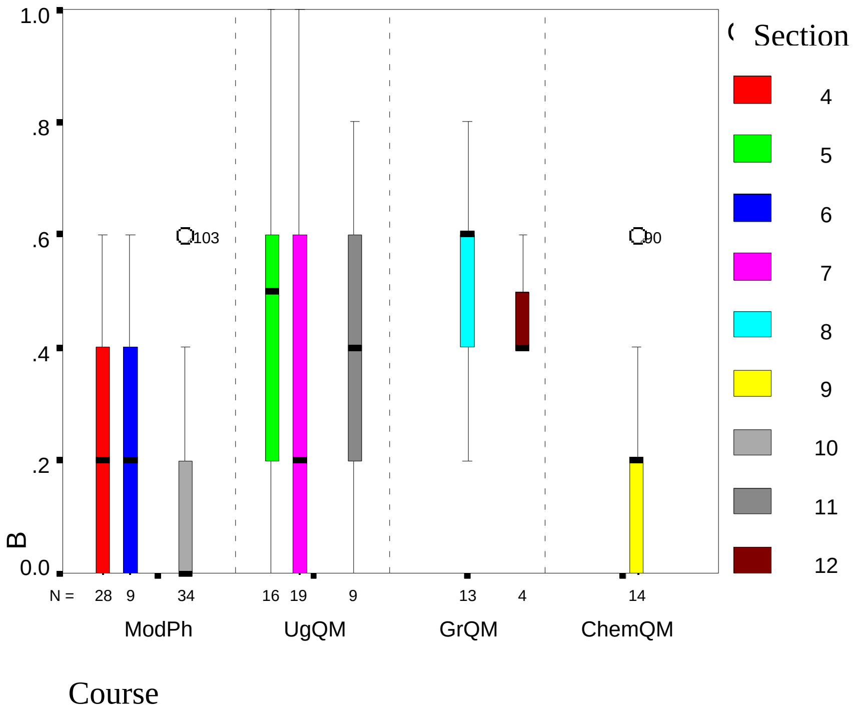


Table 4.18 Main Topic B: Students' Mean Scores by Sections

Sections	Courses	Universities	N	Mean	Median	Std. Deviation
4	ModPh	PSU	28	.2357	.20	.1967
5	UgQM	PSU	16	.4625	.50	.2802
6	ModPh	PSU	9	.2222	.20	.2108
7	UgQM	PSU	19	.3263	.20	.3070
8	GrQM	PSU	13	.5231	.60	.2088
9	ChemQM	PSU	14	.1714	.20	.1729
10	ModPh	PSU	34	0.080	.00	.1487
11	UgQM	ASU	9	.3556	.40	.2789
12	GrQM	ASU	4	.4500	.40	0.100
Total			146	.2685	.20	.2578

Although, GrQM students' mean scores were again higher than were the mean scores of students in all three other courses and although there was a lower overall mean score for main Topic B, questions 16 and 19 need a more detailed report (see appendix 5-3 for details). For question number 16, an unusual gain in mean score for the GrQM course was observed. The overall mean score for this particular question is 0.27, whereas the GrQM group's mean score is substantially higher at 0.82. The question asks students to solve a one-dimensional potential energy function which should also satisfy the related “half-well” solution. One reason for the high score might be that the graduate students were able to recognize the odd-parity solution for the Schrödinger Equation for a one-dimensional symmetric potential well.

For question 19 the mean scores, although slightly higher for students in the UgQM and GrQM courses, the mean scores were essentially equal for all four courses and were at the random guessing level; the overall mean score was 0.18 (see appendix 5-3). The question asks students to deal with the stationary state solutions in momentum-space probability distributions. Although an essential topic, the concept of momentum-space seems to be a difficult construct for the students to understand.

In summary, data analyzed by section, course and university showed that higher levels of formal education were reflected in higher mean scores in Main Topic B as well. However, the overall mean score for Main Topic B ranks the lowest among the other four Main Topics, suggesting adjustment in expectation or in teaching momentum-space is a difficult construct for students to master; even for the Graduate Quantum Mechanics (GrQM) group or it is a concept that has not been equally carefully well emphasized in standard quantum mechanics courses.

4.4.1c Main Topic C: Schrödinger Equation

Questions 4, 5, 6, 7, 8, 10, 18, and 19 are designed to assess students' understanding of concepts associated with in Main Topic C (the wave function of a single particle, the Schrödinger equation (time dependent & time-independent), scalar products of the wave function, normalization and probability density). Figure 4.11 and Table 4.19 reveal performance data consistent with the data presented in the Main topics A and B, that is, GrQM students' mean scores were higher than of the mean scores of students in the other courses and a smaller variance among the GrQM scores was observed as well. When students' achievement was statistically compared within each of the courses (i.e., across section of the same course) students in the GrQM and in the UgQM sections seemed to perform at similar levels. Indeed, the results of an ANOVA confirmed that there was no statistically significant difference in the mean scores between the GrQM sections (8 & 12) and UgQM sections (5, 7, &11; $p > 0.05$). Moreover, there was no statistically significant difference in the mean scores between the two courses as well. However, the results of the ANOVA showed that UgQM and GrQM mean score were statistically higher than those in the ModPh and ChemQM courses ($p < 0.05$, see Appendix E).

Figure 4.11 Main Topic C: Student Achievement by Section

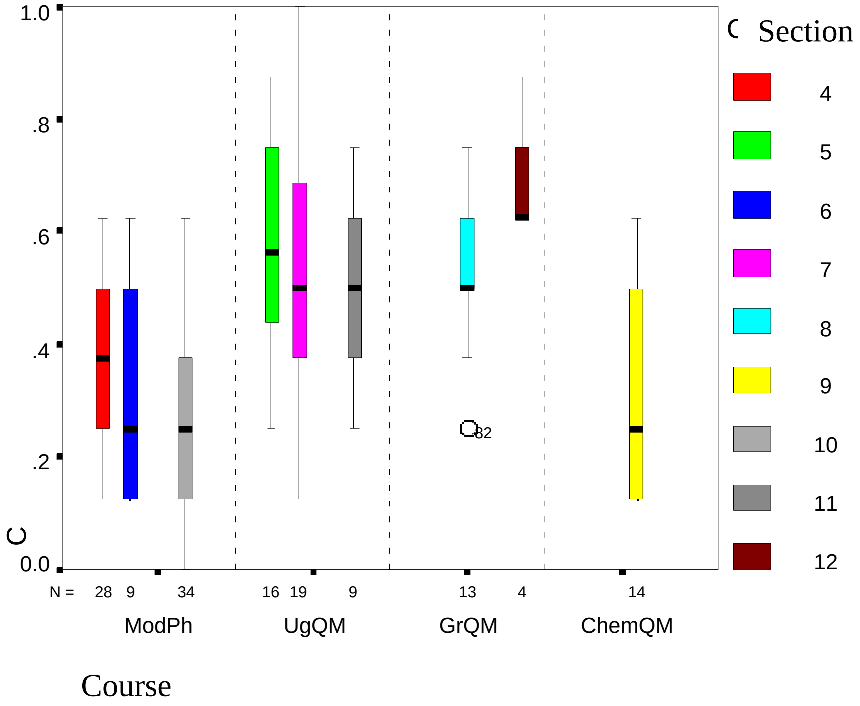


Table 4.19 Main Topic C: Students' Mean Scores by Sections

Sections	Courses	Universities	N	Mean	Median	Std. Deviation
4	ModPh	PSU	28	.3393	.3750	.1438
5	UgQM	PSU	16	.5703	.5625	.1766
6	ModPh	PSU	9	.3194	.2500	.1887
7	UgQM	PSU	19	.5461	.5000	.2252
8	GrQM	PSU	13	.5385	.5000	.1478
9	ChemQM	PSU	14	.3036	.2500	.1816
10	ModPh	PSU	34	.2904	.2500	.1649
11	UgQM	ASU	9	.4722	.5000	.1743
12	GrQM	ASU	4	.6875	.6250	.1250
Total			146	.4110	.3750	.2081

The analysis of students' achievement on each individual question revealed that the questions could be group into two clusters. For questions 4, 5, 8, & 18 students' performance ranged from high to moderate (0.91, 0.56, 0.51, & 0.40) whereas for questions 6, 7, 10, & 19 students' performance was observed to be low. For the first group of questions a more visual understanding is required with little or no mathematical manipulation. For the second group of questions not only is visual understanding is required and students also need to employ limited mathematical manipulation to solve the problems. The performance results suggest that many students do develop a visual understanding of Wave Function of a single particle, scalar products of wave function, and normalization and probability density of Wave Function. However, students have difficulties when mathematical manipulation is required to successfully solve similar kinds of questions.

4.4.1d Main Topic D: Application of Schrödinger Equation in one Dimension

Questions 10,11,12,13, 14,15,17 and 18 are all designed to assess students' understanding of concepts represented in Main Topic D (Particle in a box, Particle in a time-independent potential, Continuity conditions, Unbound and bound well, Harmonic oscillator, Wave packet and scattering (time-independent), Probability density and probability current, Scattering by a one-dimensional well, and Tunneling). The data reported in figure 4.12 and Table 4.20 again shows that GrQM students' understanding was higher than in all the other courses; however a smaller variance in the scores was not as obvious for this Main Topic. When students' achievement were statistically compared within courses (i.e., across sections) significant differences was not observed between the GrQM and the UgQM. Indeed, the results of an ANOVA confirmed that there was also no statistically significant difference in mean score between the GrQM sections (8 & 12) and thr UgQM sections (5, 7, &11; $p < 0.05$). Moreover, there was no statistically significant difference across courses mean scores as well. However, the results of the ANOVA showed that the UgQM and the GrQM mean scores were significantly higher than the mean scores of the ModPh and the ChemQM courses at ($p < 0.05$, see Appendix E).

Figure 4.12 Main Topic D: Student Achievement by Section

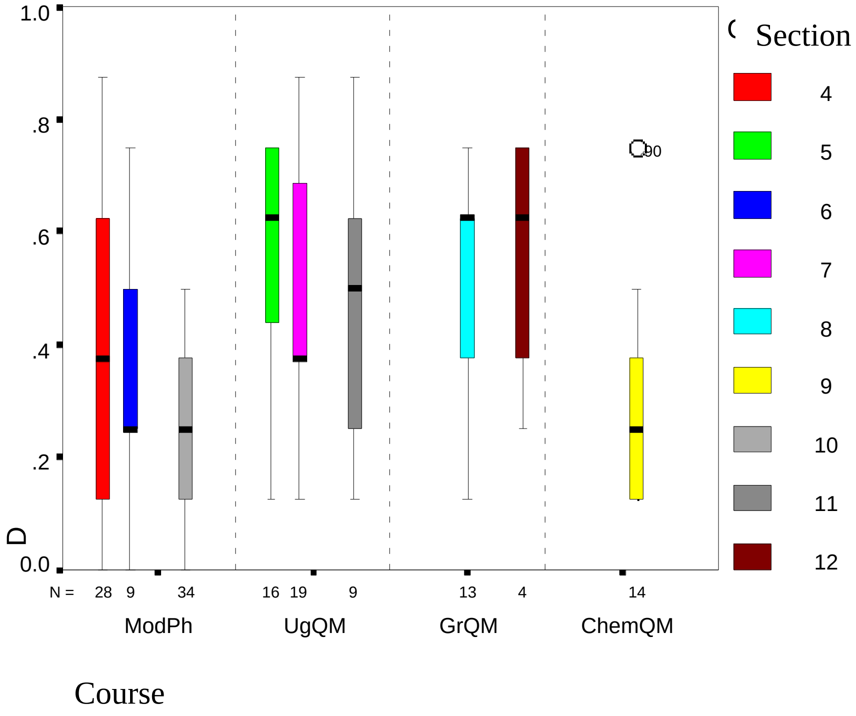


Table 4.20 Main Topic D: Students' Mean Scores by Sections

Sections	Courses	Universities	N	Mean	Median	Median	Std. Deviation
4	ModPh	PSU	28	.3929	.3750	.3750	.2694
5	UgQM	PSU	16	.5625	.6250	.6250	.1936
6	ModPh	PSU	9	.3611	.2500	.2500	.2292
7	UgQM	PSU	19	.4868	.3750	.3750	.2278
8	GrQM	PSU	13	.5577	.6250	.6250	.1883
9	ChemQM	PSU	14	.3036	.2500	.2500	.1748
10	ModPh	PSU	34	.2537	.2500	.2500	.1249
11	UgQM	ASU	9	.5139	.5000	.5000	.2684
12	GrQM	ASU	4	.5625	.6250	.6250	.2394
Total			146	.4075	.3750	.3750	.2340

The analysis of students' achievement for the Main Topic D "Application of Schrödinger equation in one dimension" showed an increase in course mean score with an increase in students' level of education. However, for question 10, 13, and 17 a different pattern emerged. For question 10 and 13 students' performance in all courses were at the random guessing level. Interestingly, both topics are usually discussed and used as examples widely in Modern Physics and Quantum mechanics courses. Question 10 is a question about a particle in a box with impenetrable walls. A widely used problem in modern physics and quantum mechanics it is often used to introduce an application of the fundamental concept, namely quantization. The problem, which was presented slightly out of its standard context, required students to solve the time-independent Schrödinger Equation for an arbitrary energy eigenvalue E with the boundary condition provided visually. The unique representation may have confused many students on this question. Question 14, asks the students to solve a fairly standard problem of incident particles (with energy E) on to a step wall (with potential V) where V < E. However, the objective response alternatives for the item were provided as visual representations. Students appeared to have problems identifying the correct shape of the reflected and transmitted waves. For question 17, the GrQM students were out performed by the other 3 groups. However, the relatively high mean scores ($\mu_{\text{ModPh}} = 0.73$, $\mu_{\text{UgQM}} = 0.80$, $\mu_{\text{GrQM}} = 0.59$, $\mu_{\text{ChemQM}} = 79$) and the small difference between the maximum and minimum mean score on this question indicated that few students participating in this study had difficulties with a question involving the Pauli principle and spin theory.

In summary, data analyzed by section, course and university showed that course mean score increased with the level of education for Main Topic D as well. The overall mean score for Main Topic D ranks second among the other Main Topics. The results suggested that students achieved a satisfactory understanding ($\mu_{\text{UgQM}} = 0.52$ $\mu_{\text{GrQM}} = 0.56$). However, students did exhibit problems when ‘known” questions were asked in different visual contexts.

4.4.1e Main Topic E. Advanced Applications

Questions, 20,21,22,23, and 24 are designed to assess students’ understanding of concepts represented in Main Topic E i.e., further applications of Schrödinger equation in two and three dimensions, Wentzel-Kremes-Brillouin Method, time-independent perturbation theory, Angular momentum. Figure 4.13 and Table 4.21 shows the GrQM course mean score was higher than students’ means scores in all the other courses consistent with the pattern observed in the other main topics. However, the results of an ANOVA confirmed that there were no statistical significant mean score differences among the four groups (see Appendix E for the detailed results).

Figure 4.13 Main Topic E: Student Achievement by Section

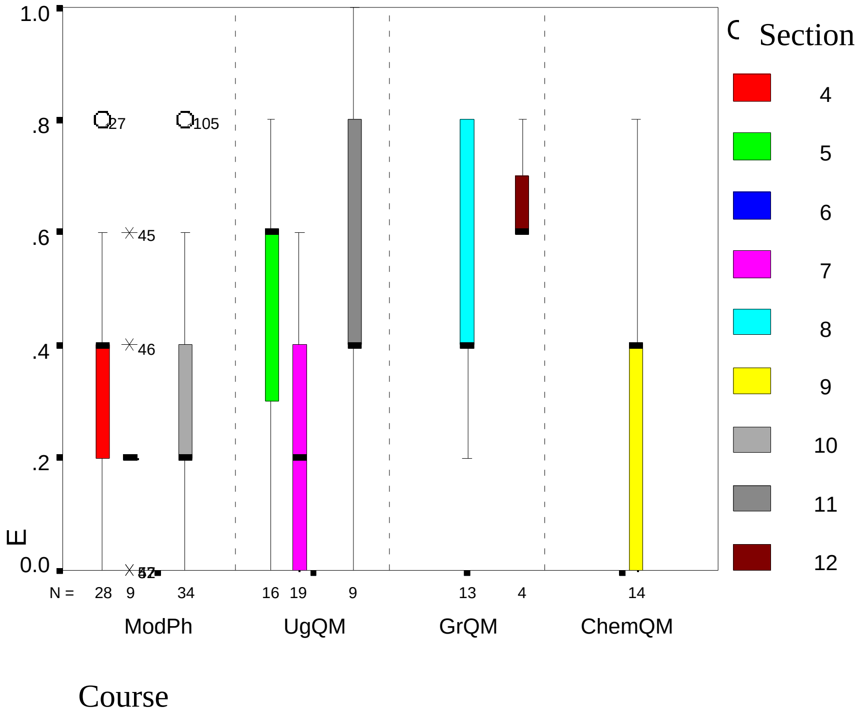


Table 4.21 Main Topic E: Students' Mean Scores by Sections

Sections	Courses	Universities	N	Mean	Median	Std. Deviation
4	ModPh	PSU	28	.3214	.4000	.2200
5	UgQM	PSU	16	.4625	.6000	.2802
6	ModPh	PSU	9	.2222	.2000	.1856
7	UgQM	PSU	19	.2000	.2000	.2108
8	GrQM	PSU	13	.5231	.4000	.2242
9	ChemQM	PSU	14	.3000	.4000	.2320
10	ModPh	PSU	34	.2588	.2000	.2061
11	UgQM	ASU	9	.5111	.4000	.3180
12	GrQM	ASU	4	.6500	.6000	.1000
Total			146	.3370	.4000	.2516

An analysis of Table 4.21 reveals the greatest fluctuation across section mean scores in Main Topic E. The range in section mean scores for UgQM was 0.31. For Main Topics A-D, UgQM and GrQM students' mean scores were statistically significantly higher than were the mean scores for ModPh and ChemQM courses. However, this situation was not true for the scores on Main Topic E. The results of the ANOVA showed that there were no statistically significant differences between the means scores of the courses. Although UgQM students' course means scores showed an increase over the mean scores of the ModPh course, the dispersion in UgQM scores suggests that many students have problems in understanding this Main Topic.

Students performed at the random guessing level for Question 24 most probably because the question was too difficult in the sense that it involved a time dependent situation. One reason for this poor performance might be that *time-dependence* is not normally addressed in depth in introductory quantum mechanics courses although it is part of the conventional curriculum. Time-dependence was included in the definition of the core content of introductory quantum mechanics developed by the expert panel in this study because it was perceived to be an important introductory topic.

4.5 Summary

In addition to the evidence documentation of content validity outlined in Chapter 3.2, chapter 4 sections 4.1 to 4.3 presented additional evidence of validity from the results of statistical analyses conducted between the QMVI mean scores and three external variables: *students' confidence level*, *level of education*, and *students' final course grade*. The results of the analyses of the internal consistence of the QMVI also provided important additional evidence of validity.

Section 4.4 examined the students' understanding in introductory quantum mechanics with regard to the five Main Topics which were earlier defined in chapter 3. Students' understanding was further analyzed on questions bases and their educational level and experience.

Chapter 5 will culminate the results of the data analyses presented in this chapter. The first part of chapter 5 will provide syntheses of the QMVI validity based on the results reported in sections 4.2 to 4.3 of the chapter. The second part of chapter 5 will discuss further students' understanding in introductory mechanics. Finally, chapter 5 will conclude with recommendations and implications based on the findings of this validation study in introductory quantum mechanics.

CHAPTER 5

5.1 Overview of the test design and development protocol

The principal purpose of this study was to develop a valid and reliable achievement instrument, the Quantum Mechanics Visualization Instrument (QMVI), to assess students' understanding in introductory quantum mechanics. The study is significant for several reasons. First, extending physics education research to examine students' understanding in courses beyond introductory mechanics is important. Research on students' understanding in quantum mechanics and the development of a valid and reliable assessment instrument can support such an effort. Second, this study provides the physics education community with a valid and reliable multiple-choice type instrument to assess students' understanding of core topics in introductory quantum mechanics which ultimately will contribute to scholarly understanding of learning in quantum mechanics and to the quality of instruction in physics education. Third, the Quantum Mechanics Visualization Instrument (QMVI) includes visual representations of quantum mechanics concepts. That can facilitate greater understanding of the role of concept visualization concepts in the learning and teaching of quantum mechanics and it can also promote increased use of appropriate visualization technologies in the teaching and learning of quantum mechanics.

The Standards for Educational Psychological Testing published by the AERA-APA-NCME (1999) guided the development and validation of the QMVI. Accordingly, the following principal steps were pursued, sometimes in successive iterations:

1. Establish the core content of introductory quantum mechanics
2. Develop draft questions based on the core content
3. Solicit content experts' opinions on the questions and modify them accordingly

4. Construct and pilot successive versions of the QMVI
5. Analyze the data and revise the QMVI
6. Identify characteristics of student samples

5.2 Validity and Reliability

The QMVI is an achievement test; therefore, a student's test score should reflect that student's understanding of the core concepts of quantum mechanics. During and after instrument development, data were accumulated and used in examining the validity of the intended score interpretations of the QMVI. Four principal sources of evidence were utilized: test content, student confidence, student year/level, and student course grade. This section reviews and synthesizes various components of evidence in a coherent examination of QMVI validity.

5.2.1 Evidence Based on Test Content

The QMVI is an achievement test and as described in Chapter 2, section 2.7.1 the definition of the content domain assessed by this instrument is very important for validation purposes. The content validation procedure and how evidence of content validity was obtained is discussed in detail in Chapter 3, section 3.1.2a and 3.2.2. As suggested by the AERA Standards, a diverse panel of experts was utilized to ensure the test content (chapter 3, section 3.2.3) and content domain (chapter 3, section 3.2.4) validity of the QMVI. Test content refers to the format of questions, the grammar and descriptions for the procedures of administering and scoring the QMVI and the content domain is the careful description of the content with a classification of areas of the content assessed by specific items. Employing this procedure ensured consistency between the core content and the QMVI. The standardized procedures used in developing the QMVI suggest that the QMVI has content validity.

5.2.2 Evidence Based on Student Variables

As outlined earlier, analyses of the relationship of test scores to variables external to the test are important sources of validity evidence. This section reports three such external variables: Students'

1. Confidence Level
2. College Year/Level
3. Course Grade

In chapter 4 section 4.3 a positive correlation was hypothesized between students' confidence level and their QMVI score. The Pearson correlation coefficient between students' confidence level and their QMVI score for all the four groups combined was found to be 0.49 indicating a moderate positive correlation. The correlation coefficients with each group were found to be moderately positive correlated too (rModPh = 0.39, rUgQM = 0.50, rGrQM = 0.33, rChem = 0.43). The correlation coefficients by groups and as a whole imply that students were generally able to predict their performance on a particular question. Moreover, the correlation coefficient increased with students' year of schooling, indicating that students at higher levels were able to predict more accurately their performance. The moderate positive correlation between students' confidence level and their QMVI score provides validity evidence because more confidant students were more successful in understand the questions and responding correctly.

Students in this validation study formed four distinct groups of examinees. These groups were identified in this study using the approximate name of the course in which the students were enrolled: Modern Physics (ModPh), Undergraduate Quantum Mechanics (UgQM), Graduate Quantum Mechanics (GrQM) and Chemistry Quantum Mechanics (ChemQM). Examining test scores relationships between the test scores of students in these different courses can provide additional evidence for test validity. The Analysis of Variance results presented in chapter 4 section 4.3.2 show that UgQM and GrQM students' QMVI scores were significantly higher than were the QMVI scores of students in the ModPH and ChemQM courses ($p < .05$). UgQM and GrQM students had received more formal education in quantum mechanics and related courses. Therefore, students in these two groups should be expected to out perform the ModPh group on the QMVI if the QMVI assesses understanding of introductory quantum mechanics.

The third external variable used in this validation study was students' course grade. The underlying assumption is that both students' QMVI score and course grades are a measure of a similar construct of understanding of quantum mechanics. Therefore, a moderate positive relationship between the QMVI test scores and students' course should exist. This kind of evidence is referred as convergent evidence (Standards, 1999). The Pearson correlation coefficient between students' QMVI score and course grade was 0.42. Thus, this moderate positive coefficient is another source of evidence of the test score interpretations of the QMVI as valid achievement test.

5.2.3 Reliability

The QMVI has a relatively high Cronbach alpha coefficient of 0.82. The QMVI mean test score is a 10 out of a possible 24. This is a secondary index that supports the instruments and its intended use as an achievement measure, since an ideal value for the index for an achievement instrument such as the QMVI would be 12. The QMVI mean item difficulty index was found to be 0.4 indicating a slightly difficult instrument. For the 24 questions on the QMVI, the item mean difficulty indices ranged between 0.12 to 0.91, yielding in an absolute range of almost 0.8 out of a possible 1.0. A high absolute range indicates that the items of the QMVI on average differentiated students' understanding. In summary, the QMVI items reliably discriminates between the low and high achieving students.

5.2.4 Reliability and Validity Synthesis

The present study was designed to construct and validate an achievement test in introductory quantum mechanics. The previous section discussed the existing evidence which supports the intended interpretation of the QMVI test scores as an achievement instrument. Another prerequisite in constructing a highly valid test is that the test scores are not influenced systematically by external variables such gender. In other words, the QMVI should not measure students' understanding concepts other than quantum mechanics concepts. One method to screen for such variables is to use the procedure mentioned in chapter 3 where a panel of experts screens the questions. Additionally, the AERA-APA-NCME Standards (1999) emphasize the importance incorporating student data during test development and validation process. The moderate positive correlation between the students' confidence levels and their QMVI scores is another source of validity evidence. This particular source of evidence can be examined in the future to provide more detailed information using techniques such as interviewing students' perceptions and their understanding of each question.

The results discussed in section 5.2.3 showed that the QMVI differentiates students with more formal education in quantum mechanics and physics. These students performed better than did those with less experience. Of course, students with more education in quantum mechanics would be expected to have achieved a higher level of understanding of topics presented in quantum mechanics. Hence, the fact that the QMVI can differentiate students' understanding is important additional evidence of validity.

The moderate positive correlation between individual students' course grades and their QMVI scores is another source of validity evidence. This moderate positive correlation suggests that the QMVI assesses a construct of quantum mechanics understanding. Both, course grades and the QMVI scores are summative in nature. Thus, to a certain degree the QMVI score and course grade reflects each student's understanding of the topics covered in the quantum mechanics courses. A student who received a high course grade in quantum mechanics would normally be expected to achieve a higher QMVI score than a student who had received a lower final course grade (course grades are a composite of many data points and factors; test scores are only one of those factors).

In addition, the QMVI was constructed with an emphasis on visual representations of the core concepts in quantum mechanics. Chapter 2 sections 2.4 and 2.5 pointed out that historically and traditionally quantum mechanics courses have been taught mostly in a quantitative manner with limited visual representation of the concepts. Since some of the questions in the QMVI are more visual in nature than are those on traditional quantum mechanics tests, and since the positive correlation between students' QMVI score and course grade was established, one possible results of this positive correlation suggest that students tend to develop a construct of visual understanding of quantum mechanics concepts despite the lack of explicit attention to visualization in standard quantum mechanics courses. Furthermore, the results in chapter 4.5.1, particularly section 4.5.1c discusses evidence that students construct a visual understanding of wave functions of a single particle. Also discussed in chapter 2 sections 2.5.1 and 2.5.2 was the importance of visualization in learning and teaching for scientific understanding of quantum mechanics concepts. The finding that students were capable of constructing visual understanding of quantum mechanics concepts is important. Developing a scientific understanding of the construct of quantum mechanics with visual dimensions has potentially important implications in teaching and learning quantum mechanics which will be discussed further in the implications section of this chapter.

The reliability of the QMVI is similar to or exceeds the reliability value of comparable tests in physics education. For example, the reliability value of the well known and frequently used Force Concept Inventory (FCI, Hestenses et al., 1992) is 0.80 and the reliability value of the Conceptual Survey of Electricity and Magnetism (CSEM, Maloney et al., 2001) is 0.75. The QMVI Cronbach alpha reliability coefficient of 0.82, slightly higher than in the other two tests, indicates a high internal structure of the test. Moreover, this relatively high reliability coefficient (Maloney et al., 2001) indicates that the QMVI minimizes false positives, that is, students selecting the correct answer for the wrong reason and maximizes the false negatives, that is, students selecting incorrect answers but actually knowing the correct answer. It should also be pointed out that the reliability coefficient is directly proportional to the number of questions in a test. The number of question is less for the QMVI ($n_{QMVI} = 24$) when compared with the other two well known tests ($n_{FCI} = 29$, $n_{CSEM} = 32$). Thus, there is substantial evidence that the QMVI is a valid and reliable achievement instrument in quantum mechanics.

5.3 Students’ visual understanding

As described in chapter 2, section 2.3, there currently are two principal approaches used in teaching introductory quantum mechanics courses. These are a mathematical approach and a historical-conceptual approach. Some strengths and shortcomings of these approaches were discussed in section 2.4, and the problems and need identified by some quantum mechanics expert instructors were reported as well. Section 2.5 also reported some innovative approaches in teaching quantum mechanics. Some instructors, for example, are making increased use of technology that can help student interact with models and visual representations of relationships.

Visualization is currently hypothesized by cognitive scientists to play an important role in the meaningful learning of abstract concepts in physics and related fields (Gotwals, 1995). The terms *visualization* and *visual understanding* have appeared with increasing regularity in the physics education literature, and they have been used in different ways. These ways have included, but have not been limited to, the understanding of spatial concepts and spatial constructs, graphic and symbolic representations, and pattern formation. These topics have received special attention in cognitive scholarship in physics education and in mathematics education in recent years. There has been much effort, research, and debate on the nature of visualization in the teaching and learning of science and mathematics in the recent past. One common perception is that visualization is an important and deep cognitive function that is vital in students’ understanding of concepts associated with visual models and representations.

For these reasons some QMVI items incorporated visual representations; OMVI items are not focused only on the more traditional verbal and mathematical understandings and skills included in conventional quantum mechanics course texts. Question number 12 for example, presents what has commonly been referred to as a “particle in a box” question. The question asks for the lowest energy state (ground state) of the particle. The item in the QMVI, however, departs in part from the classical mathematics tradition by providing the student with a visual representation of the probability density, |_ (x)|^2, versus x of the particle and the energy the particle has at that particular state as well. To be able to solve the problem, a student must understand the visual representation of the probability density function to deduce the state of the particle. As reported in Table 4.11, the majority of the students (64% for question #12) were able to answer the question correctly. The results reported in chapter 4, section 4.5, indicate that students were able to use their conceptual understanding of quantum mechanics while using and manipulating the visual information provided in some of the items. Although, this study was not especially designed to examine students’ visual understanding in quantum mechanics, the findings of the study do provide some limited information on students’ visualization which will be described in the next section.

Research aiming to determine the nature of visual understanding of quantum mechanics concepts and the role these concepts can play in learning and in the understanding of quantum mechanics is especially warranted at this time. This validation study and the QMVI provides a bases for further research in the visual understanding of quantum mechanics. Future versions of the QMVI can be constructed to provide more detailed information on students’ visual understanding and the interactions of these understandings with specific mathematical and verbal quantum mechanics concepts. Findings from such versions of the QMVI can be used in developing and examining strategies and resources for teaching quantum mechanics concepts with increased attention to visualization and to using computer technologies that can engage students in more frequent and substantive interaction with visual representations.

5.4 The nature of students' understanding of quantum mechanics concepts

As reported in chapter 4, section 4.2.2 students in four courses were involved in this validation study. These four courses were Modern Physics (ModPh), Undergraduate Physics (UgQM), Graduate Quantum Mechanics (GrQM), and Chemistry Quantum Mechanics (ChemQM). The analyses of students' understanding (Chapter 4, section 4.5) revealed that students' understanding was lowest for students in the ModPh course and highest for students in GrQM course for all 5 Main Topics.

Among the 5 Main Topics, Main Topic B incorporated the most abstract concepts and also the most sophisticated mathematical ideas of the five Main Topics. For the four courses presented in this study, the QMVI mean scores were persistently the lowest for Main Topics B "Quantum Mechanics Postulates." This result suggests that the students had experienced the greatest difficulties in understanding Main Topic B. As described in chapter 3, Main Topic B includes the postulates of quantum mechanics I-V, that underlie observable and operators, eigenfunctions and eigenvalues, and state function and expectation values. When compared across the four student groups, a distinct increase in students' QMVI mean score was observed with level of education. The ModPh students performed at the random guessing level, the UgQM students had significantly higher scores, and the GrQM students had a satisfactory level of understanding. One plausible explanation for the conceptual difficulties ModPh students had, might be that Main Topic B required higher conceptual understanding than did the other four Main Topics. Given the fact that GrQM students had more courses related to quantum mechanics, they most likely did develop a deeper understanding of the quantum mechanics related concepts included in Main Topic B probably resulting from a larger network of concepts of quantum mechanics and related topics such as mathematics. The current data collected in this study, however, does not enable the researcher to have definitive information on how the GrQM students developed a deeper conceptual understanding and what external mechanisms played important roles during this process. Nevertheless, Main Topic B appears to play an important, perhaps an overarching role in the conceptual understanding of quantum mechanics concepts. Hence, more research is needed to clarify the relationship Main Topic B plays in the role of conceptual understanding of quantum mechanics concepts. Information gathered from such research could probably inform more effective instruction.

GrQM students appeared to understand Main Topic D, application of the Schrödinger equation in one dimension, certainly at a more satisfactory level than Main Topic B. For Main Topic A, historical development and terminology and Main Topic C, the Schrödinger equation, the GrQM group seemed to acquire a satisfactory conceptual understanding as well.

As defined in chapter 3, section 3.2, the QMVI in its current stage of development is a test that measures only one principal construct that is broadly defined as the understanding of quantum mechanics. Numerous studies in physics education and especially in science education have shown that complex cognitive structures underlie students' understanding of relatively simple concepts. For example, it has been suggested that an expert understanding of the main construct of introductory mechanics is composed of six sub-constructs (Hestenes et al, 1992). Based on previous research and the complexity of the constituent concepts, quantum mechanics is almost certainly as complex as introductory mechanics, and it is probable that the construct of quantum mechanics contains sub-constructs as well. Moreover, the results showed that Main Topic B seems to be distinctly different from the other four main topics. Main Topic B might play an important role in successful understanding of quantum mechanics concepts; QMVI data in this study suggest that Main Topic B may be an overarching sub-construct. Further research is needed to understand the detailed construct of quantum mechanics understanding. However, it seems that the construct of quantum mechanics understanding is at least two dimensional, i.e., Main Topic B and the other 4 Main Topics.

Exploratory factor analysis or principle component analysis could be employed to study the construct and sub-constructs of quantum mechanics understanding more carefully. Factor analytic statistical techniques would enable further exploration of the nature of the construct of quantum mechanics understanding by identifying inter-related items on the QMVI.

5.5 Limitations of the study

There were several limitations to this study which restrict the generalization of its results. The results of the present study have been generated by the validation procedure outlined in Chapter 3 during four successive semesters (Fall 1999 to Spring 2001) at the Pennsylvania State University and Arizona State University. The sample consisted of sophomore to graduate level students enrolled in modern physics, undergraduate quantum mechanics, graduate quantum mechanics, and graduate chemistry quantum mechanics courses. The sample size was 146 students in all groups combined. Ideally, more students should be involved in a validation study (Nunnally, 1978). Due to the nature of the concepts and the unique population, it would have been extremely difficult to administer several iterations to hundreds of students as recommended by some experts. Due to the small sample size and due to the limited number of universities involved in the study, however, caution is warranted in interpreting results, especially in Chapter 4 section 4.5 a-e. While the students who have participated in this study were most likely representative of groups in similar universities, caution is warranted in applying these findings to other populations since, for example, the students were predominantly male.

As noted in chapter 2, section 2.7 and chapter 3 section 3.1, the construction and validation procedure of an achievement instrument such as the QMVI is very dependent upon proper content definition by an expert panel. While there was consultation and input from multiple universities, the expert panel in this study (chapter 3 sections 3.2.2d and 3.2.3) was composed only of experts from the Physics Department and Science Education program at the Pennsylvania State University.

Furthermore, students majoring in chemistry, certain engineering fields, and materials sciences regularly study modern physics and topics in quantum mechanics. However, as noted in chapter 3, section 3.2 the core content of the QMVI was developed based principally on contemporary physics curricula in quantum mechanics. Hence, the core content of introductory quantum mechanics as defined and reflected in the QMVI in this study is limited to the physics curriculum. Information should be gathered about the content of the topics taught in introductory quantum mechanics courses in related fields to inform the definition of quantum mechanics and to inform the further development of the QMVI.

Data gathered in this validation study were gathered principally from volunteers whose final course grades were not based upon their performance on the QMVI. Data also should be gathered in more naturalistic environments such as in classrooms in which the QMVI score is a specified proportion of the class grade.

5.6 Summary of Outcomes

This validation study provides several important findings:

1. There is considerable evidence that the QMVI has content validity. Findings that support this conclusion are: a) The expert panel in the field of physics approved the core content of introductory quantum mechanics defined in this study, b) 80% or more of the panel of judges in physics and science education approved the content appropriateness of the QMVI items, c) the panel of judges in physics and science education approved the QMVI items using protocols similar to Edwards' 14 criteria, and d) the QMVI (version 2) discriminated between high and low achieving students and correlated positively with students' final course grades when used with 126 students.

2. The 24-item QMVI has acceptable reliability. Findings that support this conclusion are: a) a Cronbach alpha coefficient of 0.82; b) an item mean difficulty of 0.4 with a QMVI score range of 19 out of a possible 24.

3. The QMVI is a one construct scale by definition (see Chapter 3, section 3.2.1); it measures the construct of the understanding of introductory quantum mechanics. However, the results reported in the students' understanding section suggest sub-constructs of students' understandings of introductory quantum mechanics. Findings that support this conclusion are: a) students in graduate quantum mechanics

and undergraduate quantum mechanics outperformed student in modern physics in all five Main Topics, b) Main Topic B, “quantum mechanics postulates” appears to be an overarching sub-construct in introductory quantum mechanics; and c) Factor analysis is recommended to determine relationship among the QMVI sub-constructs associated with students’ understanding of quantum mechanics.

4. The validation study provided data from which preliminary findings were drawn on students’ verbal, mathematical, and visual understanding of introductory quantum mechanics concepts. Limited data on students’ understanding suggest: a) students achieve understanding in areas such as semi-classical particle behavior, waves and probability, but display more limited understanding of probability and quantitative concepts; b) the topic of momentum-space is a difficult construct for students; c) students are able to construct and manipulate visual understanding of selected quantum mechanics concepts; and d) students seem to have difficulties when items involve time-dependent situations.

5.7 Recommendations for further research

Further research is needed to support the development of improved versions of the QMVI. In addition, scholarship associated with use of the QMVI can provide needed information about the relationships between instruction, text and technology resources, and students’ understanding of introductory quantum mechanics. Several specific recommendations emanate from this validation research study.

Section 5.7.1 reports recommendations pertaining to the improvement of the QMVI. These recommendations are based upon the methodology of *test validation* and the results presented in chapters 3 and 4. In section 5.7.2 recommendations pertaining to the nature of student understanding in quantum mechanics are discussed. Section 5.7.3 reports recommendations pertaining to the future use of the QMVI in furthering understanding in physics education.

5.7.1 Recommendations Pertaining to Improvement of the QMVI

As discussed in the limitations section, the research methodology followed in this validation study of the QMVI and the data collection procedures of necessity limit the generalizability and the nature of the findings. Expanding representation on the panels of quantum mechanics experts and examining data for larger numbers of more diverse students will play an important role in extending the validity and the generalizability of the QMVI. Hence, the following recommendations pertaining to the improvement of the QMVI are suggested.

Expert panel

- Extend beyond faculty and graduate students at the Pennsylvania State University.
- Include applied and theoretical physicists;
- Include experts from related fields such as chemistry, materials sciences, electrical engineers, the history and philosophy of physics, science education, etc;

Participating students

- Alternative forms of information and feedback should be gathered from participating students to ascertain their rationales for specific responses on the QMVI and to identify alternative concepts they may have constructed.

5.7.2 Recommendations Pertaining to the Construct of Quantum Mechanics Understanding

The results of this study provided interesting preliminary data on students' understanding in introductory quantum mechanics. Students' understanding of introductory quantum mechanics concepts are important and should be investigated further. Hence, the following recommendations associated with examining students' understanding in introductory quantum mechanics also emanate from this study.

- Students’ understandings of quantum mechanics concepts appear to be multidimensional in nature. The construct of the QMVI should be further examined by employing exploratory factor analysis.
- Main Topic B “Quantum Mechanics Postulates” appears to include concepts with which all students, especially ModPh students have substantial difficulties. The relationship of the quantum mechanics postulates to the understanding of specific quantum mechanics topics should be examined.
- The nature of the relationship between visual understanding of specific quantum mechanics concepts to the verbal and mathematical understanding of these concepts should be investigated in greater detail.
- Students’ difficulties with specific problematic concepts such as momentum-space and time-dependence should be identified and examined.

Identification of principal sub-constructs important in understanding quantum mechanics concepts could help to improve the teaching and learning of quantum mechanics concepts. Such studies should inform the development of more effective textbooks and technology related teaching materials.

5.7.3 Recommendations to Examine Effects of Teaching and Instructional Resources

The first obvious areas of potentially helpful studies are “causal – comparative” studies. The aim of such studies is “to determine existing differences in the behavior or status of groups”.

- The QMVI should be administered in different universities in different settings and in different nations to further assess students’ understanding of introductory quantum mechanics. The results obtained from early studies of this kind can be compared with the baseline data provided in this study.

- The QMVI should be administered to more diverse groups (e.g., gender, ethnicity, institutional, etc.) to examine the effects of culture and context.
- The QMVI should be administered to groups of student groups in quantum mechanics related fields beyond physics such as in electrical engineering, materials science, chemistry, etc., and the results obtained from such studies could be used to inform the development of courses and teaching resources
- The QMVI should be used to study the relationship between the students' understanding of introductory quantum mechanics and other ancillary concepts and skills in mathematics, understanding of classical mechanics, and understanding of statistical physics, etc.

A second area of study is research that will examine the effects of instruction and learning resources by manipulating treatment. For example:

- The QMVI could be used to study the effects of specific teaching strategies on students' understanding of introductory quantum mechanics.
- The QMVI could be used to study the effects of various textbooks, alternative media (e.g., hypertext), and computer software used in introductory quantum mechanics courses.

The results of such studies have potential to improve knowledge and understanding of the teaching and learning of quantum mechanics and to inform the development of improved teaching materials that can enhance students' understanding of quantum mechanics concepts.

5.8 Final Remarks

Much empirical research has been conducted in the teaching and learning of physics. Over the years important results have been reported which continue to influence knowledge and assumptions about teaching and learning physics and that inform the development of teaching practices, textbooks, and instructional resources. Most of the research to date has focused on entry level physics with special attention to introductory mechanics and with lesser attention to geometric optics, electricity etc. This study resulted in the development of a valid instrument; the Quantum Mechanics Visualization Instrument (QMVI) was constructed to examine students' understanding in quantum mechanics. The development and initial use of the QMVI provided preliminary data that can serve as a foundation for needed further research and development on the teaching and learning of quantum mechanics.

REFERENCES

Abimbola, I. O. (1988). The problem of terminology in the study of student conceptions in science. *Science Education,* 72, 175-184.

Aguirre, J. M., & Rankin, G. (1989). College students' conceptions about vector kinematics. *Physics Education*, 24, 290-294.

Aiken, L. R. (1971). Intellective Variables and Mathematics Achievement: Directions for Research. *Journal of School Psychology,* 9, 2.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington, DC: Author.

Battista, M. T. (1991). Research into Practice: Using Spatial Imagery in Geometric Reasoning. *Arithmetic Teacher,* 39, 18-21.

Battista, M. T. (1999). Geometry Results from the Third International Mathematics and Science Study. *Teaching Children Mathematics,* 5, 367-73.

Battista, Michael T ., & Clements, Douglas H (1991). Research into Practice: Using Spatial Imagery in Geometric Reasoning. *Arithmetic Teacher,* 39, 18-21.

Bohrnstedt, G. W. (1970) Reliability and validity assessment in attitude measurement. In Summers, G.F., (Ed.) *Attitude Measurement*, Chicago: Rand McNally,

Brandt, S., & Dahmen, H.D. (1994). *Quantum mechanics on the personal computer*. (3rd ed.). New York: Springer-Verlag.

Bransford, J. D., A. L. Brown, et al., Eds. (1999). How people learn: Brain, mind, experience, and school. Washington, D.C., National Academy Press.

Brody, P.J. (1984). Research of pictures in instructional texts. Educational Communication and Technology Journal, 29, 93-100.

Browning, C. (1988). Characterizing levels of understanding of functions and their graphs. In Romberg, T.A., Fennema, E., & Carpenter, T.P. (1993). *Integrating research on he graphical representation of functions*. New Jersey: Lawrence Erlbaum Associates.

Caramazza, A., McCloskey, M., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science,* 210, 1139-1141.

Champagne, A. B., Klopfer, L. E., & Anderson, J. (1980). Factors influencing the learning of classical mechanics. *American Journal of Physics*, 48, 1074-1079.

Clement, J. (1981). Analogy generation in scientific problem solving. Proceedings of the Third Annual Meeting of the Cognitive Science Society.

Clement, J. (1983). Students' alternative conceptions in mechanics: A coherent system of preconceptions ? Proceedings of the International Seminar Misconceptions in Science and Mathematics. H. Helm, Novak, J. Ithaca, N. Y., Cornell University: 310-315.

Cohen, H.D., Hullman, D.F., & Agne, R.M. (1978). Cognitive level and college physics achievement. *American Journal of Physics*, 46, 1026-1029

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

Cronbach, Lee J (1988) Internal Consistency of Tests: Analyses of Old and New. Psychometrika; v53 n1 p63-70 Mar 1988.

Cunningham, G. K., (1998). *Assessment in the classroom*. Washington, DC: The Falmer Press.

Cuppari, A., G. Rinaudo, et al. (1997). Gradual Introduction of Some Aspects of Quantum Mechanics in a High School Curriculum. *Physics Education*, 32, 302-308.

Demana, Franklin; Waits, Bert K (1993). The Particle--Motion Problem. Mathematics Teacher; v86 n4 p288-92 Apr 1993

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, 23, 5-12.

Driver, R., Guesne, E. , Tiberghien, A. (1985). *Children's ideas in science*. Milton Keynes: Open University Press.

Dwyer, F.M (1972). The effect of overt responses in improving visually programmed science instruction. *Journal of Research in Science Teaching*, 9, 47-55.

Edwards, A. L. (1957). *Techniques of attitude scale construction*. New York: Appleton –Century –Crafts.

Eisberg, R.M, & Resnick E.M. (1985). *Quantum physics of atoms, molecules, solids, nuclei, and particles*. (2nd ed.). New York: Wiley

Farrell, A. (1989). Teaching and learning behaviors in technology-oriented precalculus classrooms. In Romberg, T.A., Fennema, E., & Carpenter, T.P. (1993). *Intergrading research on the graphical representation of functions*. New Jersey: Lawrence Erlbaum Associates.

French, A. P. & Taylor E. F. (1978) *An introduction to quantum physics*. New York: Norton.

Gall, M. D., Borg, W.R. & Gall, J.P. (1996). *Educational Research: An introduction*. New York, Longman.

Gasiorowicz, S (1974) *Quantum physics*. New York: Wiley.

Gilbert, J. K., & Swift, D. J. (1985). Towards a Lakatosian analysis of the Piagetian and alternative conceptions research programs. *Science Education*, 69, 681-696.

Gotwals, R.R. (1995). Scientific visualization in chemistry, better living through chemistry, better chemistry through pictures: Scientific visualization for secondary chemistry students. In Thomas, D.A. (pp. 153 -179). Charlottesville: Association for the Advancement of Computing in Education.

Griffiths, D.J. (1995). *Introduction to quantum mechanics*. New Jersey: Prentice Hall

Hadzidaki, P., Kalkanis, G., & Stavrou, D. (2000). Quantum mechanics: a systematic component of the modern physics paradigm. *Physics Education*, 35, 386-392.

Hake, R. R. (1998). Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66, 64-74.

Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Needham Heights: Allyn & Bacon.

Halloun, I.A., & Hestenes, D. (1985b). Common sense concepts about motion. American Journal of Physics, 53, 1043_1055.

Halloun, I.A., & Hestenes, D. (1985b). The initial knowledge state of college physics students American Journal of Physics, 53, 1056_1055.

Haslam, F., & Treagust, D.F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *Journal of Biological Education* 21(3): 203-211.

Hasmetoglu, H. (1994). *Effect of students' cognitive operational level, mathematical skills and preconceptions on achievement in introductory mechanics*. Ankara: Middle East Technical University.

Heller, P. & Huffman, D. (1995). Interpreting the force concept inventory. A reply to Hestenes and Halloun. *The Physics Teacher,* 33, 503, 507-511.

Hestenes D., Wells M., & Swackhamer G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-166.

Hestenes, D., & Halloun (1995). Interpreting the force concept inventory. A response to march 1995 critique by Huffman and Heller. *The Physics Teacher,* 33, 502,504-506.

Hobson, A. (2000). Teaching Modern Physics in introductory courses. *The Physics Teacher,* 38, 388.

Holliday, W.G. (1975). The effects of verbal and adjunct pictorial-verbal information in science instruction. *Journal of Research in Science Teaching*, 12, 77-83.

Howes, R. (2000). Modern Physics. *The Physics Teacher,* 38, 73.

Hudson, H.T., & McIntire W.R. (1977). Correlation between mathematical skills and success in physics. *American Journal of Physics*, 45, 470-471.

Huffman, D., & P. Heller (1995). What does the force concept inventory actually measure? *The Physics Teacher,* 33, 138-143.

Landshoff, P, Metherell, A., & Rees, G. (1997) *Essential quantum physics*. (2nd ed.). Cambridge: Cambridge University Press.

Liboff, R. (1980). *Introductory quantum mechanics*. San Francisco: Holden-Day.

Licht, P., & Thijs, G.D. (1990) Method to trace coherence and persistence of preconceptions. *International Journal of Science Education*, 12, 403-416.

Maloney, D.P., O'Kuma, T.L., Hielggelke, C.J., & Van Heuvelen, A. (2001). Surveying students' conceptual knowledge of electricity and magnetism. American Journal of research, 69, s12-s23.

McCloskey, M. (1983). Intuitive physics. *Scientific American,* 248, 114-122.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science,* 210, 1139-1141.

McCormick B.H., DeFanti T.A., & Brown M.D. (1987) Special issue on visualization in scientific computing. Computer Graphics, ACM SIGGRAPH, 21, 6.

McDermott, L. (1998). Comments on learning and understanding key concepts in electricity. Connecting research in physics education. A. Tiberghien, Jossem, E., Barojas, J. Ohio, ICPE Books: 1-3.

McDermott, L.C & Redish, E.F. (1999). Resource Letter: Per-1:Physics education research. *American Journal of Physics,* 67, 755-767.

McMurry, S. M. (1994) *Quantum mechanics*. Wokingham: Addison-Wesley Pub. Co.

Mehrens W. A., & Lehmann I. J., (1991). *Measurement and the evaluation in education and psychology* (4th ed.). Chicago: Ted Buchholz.

Meier S. T., (1994). *The chronic crisis in psychological measurement and assessment: A historical survey*. San Diego: Academic Press.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). Washington, DC: American Council on Education and National Council on Measurement in Education.

Millar, R. (1989). Constructive criticisms. *International Journal of Science Education,* 11, 587-596.

Minstrell, J. (1982). Explaining the at rest condition of an object. *The Physics Teacher,* 20, 10-14.

Minstrell, J. (1990). A teaching system for diagnosing student' conceptions and prescribing relevant instruction. Paper prepared for a session of the American Educational Research Association titled Classroom perspectives on conceptual change teaching, Boston.

Morrison, M. A. (1990). *Understanding quantum physics*: *A user's manual*. New Jersey, Prentice Hall.

Muino, P. L. (2000). Introducing the uncertainty principle using diffraction of light waves. *Journal of Chemical Education,* 77, 1025-1027.

Nunnally, J.C. (1978). *Psychometric theory*. New York, McGraw-Hill.

Oosterhof A., (1996). *Developing and using classroom assessments*. New Jersey: Prentice Hall.

Osborne, R. and J. K. Gilbert (1980). A method for investigating concept understanding in science. *European Journal of Science Education,* 2, 311-321.

Peleg, Y., Pnini, R., & Zaarur, E. (1998): *Schaum's outline of theory and problems of quantum mechanics*. New York: McGraw-Hill.

Pfundt, H., & Duit, R. (1991). Bibliography: Students' alternative frameworks and science education (3rd ed.). Kiel: Institute for Science Education

Rich, B.S. (1990). The effect of the use of graphing calculators on the learning of functions concepts in precalculus mathematics. In Romberg, T.A., Fennema, E., & Carpenter, T.P. (1993). *Intergrading research on the graphical representation of functions*. New Jersey: Lawrence Erlbaum Associates.

Rigney, J.W., & Lutz, K.A. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitude. *Journal of Educational Psychology,* 68, 305-311.

Robinett, R,W. (1997) *Quantum mechanics :Classical results, modern systems, and visualized examples*. New York: Oxford University Press.

Roid, G. H. & T. M. Haladyna (1982). *A technology for test-item writing*. New York, Academic Press Inc.

Roussel, M. R. (1999). Redesigning the quantum mechanics curriculum to incorporate problem solving using a computer algebra system. *Journal of Chemical Education*, 76, 1373-1377.

Rubba, P.A., & Andersen, H.O. (1978). Development of an instrument to assess secondary school students` understanding of the nature of scientific knowledge. *Science Education*, 62, 449-458.

Sequeira, M. F. & Leite, L. (1991). Alternative conceptions and history of science in physics teacher education. *Science Education*, 75, 45-56.

Shankar, R. (1994). *Principles of quantum mechanics*. (2nd ed.). New York: Plenum Press.

Shaw, M.E., & Wright, J.M. (1967). *Scales for the measurement of attitudes*. New York: McGraw-Hill.

Stiggins, R.J. (1994) *Student-centered classroom assessment*. New York: Macmillan.

Suen, H. K. (1990). *Principles of test theories*. New Jersey: Lawrence Erlbaum Associates.

Tamir, P. (1971). An alternative approach to the construction of multiple choice test items. *Journal of Biological Education*, 5, 305-307.

Taylor, F. E. (1998). The boundaries of Nature: Special and general relativity and quantum mechanics, a second course in physics: Edwin F. Taylor's acceptance speech for the 1998 Oersted Medal presented by the American Association of Physics Teachers, 6 January 1998. *American Journal of Physics Teachers,* 66, 369-376.

Thaller, B (2000): *Visual quantum mechanics: selected topics with computer-generated animations of quantum-mechanical phenomena*: New York: Springer/TELOS.

Thorndike, R.L., Ed. (1971). *Educational Measurement*. Washington, D.C.: American Council on Education.

Tinkelman, R. L. (1971). Planning the objective test. Educational Measurement. In. Thorndike.R.L. *Educational Measurement*. (pp. 55-60). Washington, D.C., American Council on Education.

Treagust, D. F. (1986). Evaluating students' misconceptions by means of diagnostic multiple choice items. *Research in Science education,* 16, 199-207.

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education,* 10, 159-169.

Trowbridge, D. E., McDermott, L. C. (1980). Investigation of student understanding of the concept of velocity in one dimension. *American Journal of Physics*, 48, 1020-1028.

Trowbridge, D. E., McDermott, L. C. (1981). Investigation of student understanding of the concept of acceleration in one dimension. *American Journal of Physics*, 49, 242-253.

Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. *European Journal of Science Education,* 1, 205-221.

Vygotsky, L. S. (1997). *Educational psychology*. Boca Raton, Florida: St. Lucie Press.

Wandersee, J. H. (1985). Can the history of science help science educators anticipate students' misconceptions? *Journal of Research in Science Teaching*, 23, 581-597.

Wandersee, J., Mintzes, J., & Novak, J. (1994). Research on alternative conceptions in science. In D. Gabel (Ed.), Handbook of research on science teaching and learning. New York: Macmillan

Watts, M. (1980). Suggestions for generating and designing cards for the interview about instances approach. Working paper presented by the Institute for Educational Technology, University of Surrey.

Watts, M. (1982). Gravity - Don't take it for granted! *Physics Education*, 17, 116-121.

White, B. (1983). Sources of difficulty in understanding Newtonian dynamics. *Cognitive Science,* 7, 41-65.

Winter, R, G.(1979). *Quantum physics*. Belmont, CA: Wadsworth Pub. Co

Wise, M.N. & Kelley T.G. (1977). Fundamental quantum mechanics: a graphic presentation. *American Journal of Physics*, 45, 384-394.

Piaget, J. (1972). *The principles of genetic epistemology*. New York, Basic Books.

Ausubel, D. P. (1969). *Readings in school learning*. New York: Holt, Rinehart and Winston.

APPENDIX A: LETTERS TO EXPERT PANEL

Letter to content expert reviewer, suggested content outline in introductory quantum mechanics, letter to graduate student reviewers and evaluation form, the first draft version of QMVI

Dear Dr. …..,

We are a group at Penn State University developing science concept tests (and instructional strategies). At present we are particularly interested in the nature of students' understanding of central quantum mechanics concepts following formal instruction in relevant physics courses.

In order to reach some of our goals, it is crucial for us to establish a set of core concepts in introductory quantum mechanics upon which most experts in the field can agree. By core concepts we mean central concepts students are expected to understand after completing introductory courses in modern physics and quantum mechanics (Instruction in these courses would normally involve students in using these concepts). In order to develop an initial list of core quantum mechanics concepts from which to begin, the following quantum mechanics and modern physics resources were carefully reviewed:

- Syllabi of introductory QM courses offered at *Big Ten* universities;
- Principal textbooks used in the introductory QM courses
- GRE and other similar problem / solution handbooks on QM topics.

We are now conducting a Delphi Study in which we are soliciting information from experts like yourself in the teaching community. From that input, we expect to reach an understanding among experts of the core quantum mechanics concepts normally included in modern physics and introductory quantum mechanics courses. We hope you will be able to take a few minutes to provide us with information about your perspectives on central quantum mechanics concepts by completing the brief survey that is enclosed. We are asking you and a small number of colleagues to examine a list of suggested core concepts and to suggest the addition or deletion of concepts on the list. You are also asked to adjust the percentage (of emphasis and course time) that are included if you think adjustments are warranted. Your input will contribute significantly toward establishing a field of valid core concepts that represent perceptions of the physics education community about the conceptual content of introductory courses in quantum mechanics.

A revised version of the concepts list that represents the compiled input of those who participate in this survey will be sent to you for your approval at the conclusion of this round of data collection from colleagues. We thank you in advance for participating in this important Delphi study. Please try to complete and return this survey to:

Dr. Richard W. Robinett, Assistant Department Head, Physics Department of Physics The Pennsylvania University University Park, PA 16802	Or	Mr. Erdat Cataloglu Research Assistant 111B Kern Bldg The Pennsylvania University University Park, PA 16802

no later than April 21, 2000.

Sincerely Yours,

Dr. Richard W. Robinett

Erdat Cataloglu

Dr. Richard W. Robinett, Assistant Department Head, Physics Department of Physics The Pennsylvania University University Park, PA 16802 (814) 863-0965	Dr. Vincent N. Lunetta Professor of Science Education 166 Chambers Penn State University University Park, PA 16802 (814) 865-2237	Mr. Erdat Cataloglu Research Assistant 111B Kern Bldg Penn State University University Park, PA 16802 (814) 865-4211

Dear Expert,

We are a group at the Pennsylvania State University studying the nature of students' understanding of central quantum mechanics concepts following formal instruction in relevant physics courses. As a part of the process we are developing a Quantum Mechanics achievement test. In order to reach some of our goals, it is important for us to establish *content validity* of the Quantum Mechanics achievement test under construction.

We are now conducting a content analysis in which we are soliciting information from experts like yourself in the physics community.

You have been provided with a package that contains,

1. The Quantum Mechanics Visualization Instrument (QMVI)
2. The Expert Judge Form
3. The Content Outline

We are asking you and a small number of colleagues to examine *each* question on the QMVI and judge for *each question* if it is “**Important**”, “**Not Important but Relevant**”, or “**Not Relevant at All**” *based* on the *Content Outline* provided with the packedge you received. Please mark directly your responses on the Expert Judge Form. After you have finished your judging, please write your full name, sign the Expert Judge Form, and return it to 303J Osmond Lab (Dr. Richard Robinett’s Office). If you have any questions about any of these materials or procedure, please contact us using the information provided below.

Sincerely yours

Dr. Richard W. Robinett

Erdat Cataloglu

Dr. Richard W. Robinett,
303J Osmond Lab
Assistant Department Head, Physics
Department of Physics
The Pennsylvania University
University Park, PA 16802
Email: rick@phys.psu.edu
Ph: 863-0965

Or

Mr. Erdat Cataloglu
Research Assistant
113 Kern Bldg
The Pennsylvania University
University Park, PA 16802
Email: exc18@psu.edu
Ph: 865-1500
Ph: 867-6218

Main Topic	Subtopic	Weights
A. Historical development & terminology	8. Blackbody radiation 9. Photoelectric effect 10. Bohr Atom 11. Wave Particles duality 12. deBroglie hypothesis 13. Uncertainty Principle 14. Probability & semi-classical behavior, waves	10%
B. Quantum mechanics	14. Postulates of quantum mechanics I-V 15. Observable and operators 16. Eigenfunctions and eigenvalues 17. Dirac Delta function 18. State function and expectation values 19. Hilbert space and its properties	20%
C. Schrödinger equation	18. Wave function of a single particle 19. the Schrödinger equation (time dependent & time independent) 20. scalar products of wave function 21. normalization and probability density	22%
D. Application of Schrödinger equation in one dimension	27. Particle in a box (infinite hard wall) 28. Particle in a time-independent potential 29. Continuity conditions 30. Unbound and bound well 31. Harmonic oscillator 32. Wave packet and scattering (time-independent) 33. Probability density and probability current 34. Scattering by a one dimensional well 35. Tunneling	27%
E. Advanced Applications	31. Further applications of Schrödinger equation in, two and three dimension 32. Wentzel-Kremes-Brillouin Method 33. Time-independent perturbation theory 34. Angular momentum	20%

Dear ……

Thank you for agreeing to serve as a reviewer of the Quantum Mechanics Visualization Instrument (QMVI). This instrument aims to assess students' understanding of general quantum mechanics concepts.

Although well established criteria were used in constructing the draft version of the QMVI we need your expertise to even further improve the instrument. Therefore we ask you to review the draft version of the QMVI for clarity and comprehension by college physics students. Please, record your comments directly on the draft instrument.

The revised items will be resubmitted to you and subsequently revised until clarity is achieved. I am asking you to return the packet to my mailbox by Thursday, November 16, 1999.

The following questions will serve you in your task in reviewing the clarity and comprehension of each item present in the draft version.

1. Is the level of reading appropriate to the students ability?
2. Does the question clearly present the problem? That is does the student exactly understand what is being asked.
3. Are all options parallel in type of content?
4. Do the options avoid repetitive words?
5. Is extraneous content excluded from the stem?
6. Are adjectives or adverbs emphasized when they reverse or significantly alter the meaning of the stem or option?
7. Is the grammar in each option consistent with the stem?
8. Does the item exclude options equivalent to "all of the above" and "non-of the above"?
9. Does the item provide clues to the correct answer. (e.g., response length, grammar, repetition of key words, common associations, etc.?
10. Does the item have "sometimes", "never", "always" "all" and :"none"?
11. Does the item focus on a central idea?
12. Does the item hint a correct answer to another item?
13. Is the item an opinion based item?
14. Does the item use correct grammar, punctuation, capitalization, and spelling?
15. Does the item involve heavy reading that might interfere with the actual problem being asked?

16. Does the item include controversial material?

If you have any questions regarding this process please feel free to call me at 814-867-6218 or email me at exc18@psu.edu. Once again thank you for your valuable time and expertise.

Best Regards

Erdat Cataloglu

APPENDIX B: INFORMED CONSENT FROM

Informed Consent Form

Informed consent form for behavioral research study

Title of Project: Fostering students' conceptual understanding in quantum mechanics

Person in Charge:	Dr. Richard W. Robinett, Assistant Department Head, Physics Department of Physics The Pennsylvania University University Park, PA 16802 (814) 863-0965	Mr. Erdat Cataloglu Science Education The Pennsylvania University University, Park, PA 16802 (814) 865-4211

1. This section provides an explanation of the study in which you will be participating:

A. The study in which you will be participating is part of research intended to foster college students' conceptual understanding in quantum mechanics. By conducting this research we hope to gain amend understanding on students' level of quantum mechanics concepts after completing a quantum mechanics course.

B. If you agree to take part in this research, you will be asked to authorize the researchers to access your grades, homework assignments or related course artifacts, take written notes on classroom observations, and perform interviews which will be audio taped and then destroyed following the completion of the research.

2. This section describes your right as a research participant:

A. You may ask any questions about the research procedures, and these questions will be answered. Further questions should be directed to Dr. Richard W. Robinett (rick@phys.psu.edu) or Mr. Erdat Cataloglu (exc18@psu.edu).

B. Your participation in this research is confidential. Only the people in charge will have access to your identity and to information that can be associated with your identity. In the event of publication of this research, no personally identifying information will be disclosed. To make sure your participation is confidential, only a code numbers will be used. Only the researchers can match names with code.

C. Your participation is voluntary. You are free to stop participating in the research at any time, or decline to answer any specific question without penalty.

D. This study involves minimal risk, that is, no risk to your physical or mental health beyond those encountered in the normal course of everyday life.

3. This section indicates that you are giving your informed consent to participate in the research.
Participant:

- I agree to participate in a scientific investigation fostering students' conceptual understanding in quantum mechanics, as an authorized part of the education and research program of the Pennsylvania State University.

- I understand the information given to me, and I have received answers to any questions I may have about the research procedure. I understand and agree to the conditions of this study as described.

- To the best of my knowledge and belief, I have no physical or mental illness or difficulties that would increase the risk to me of participation in this study.

- I understand that I will receive no compensation for participating.

- I understand that my participation in this research is voluntary, and that I may withdraw from this study at any time by notifying the person in charge.

- I am 18 years of age or older, and/or a full time student of the Pennsylvania State University.

- I understand that I will receive a signed copy of this consent form

Signature:______________________________ Date:__________________

Researcher:
I certify that the informed consent procedure has been followed, and that I have answered any questions from the participant above as fully as possible.

Signature:______________________________ Date:__________________

Signature:______________________________ Date:__________________

I, hereby authorize Dr. Richard Robinett and Erdat Cataloglu to gain access to my course artifacts and grades.

Name: _______________________ Social Security Number: _______________________

Signature:______________________ Date: ______________________________________

I, hereby authorize the registrar to release my academic records to Dr. Richard Robinett and Mr. Erdat Cataloglu.

Name _______________________ Social Security Number: _______________________

Signature:______________________ Date:____

APPENDIX C: ITEM ANALYSES RESULTS ON QMVI VERSION 1

Results of Item Analysis of 1st version of the QMVI

Q#	p-value	Choices	Proportion of Responses	High Group	Low Group	Item Discrimination Power	Correct Answer
1	0.567	A	0.060	0	0.167	-0.167	
		B	0.567	0.722	0.167	0.555	*
		C	0.343	0.222	0.667	-0.445	
		D	0.015	0.056	0	0.056	
		E	0.015	0	0	0	
		Other	0.000	0	0	0	
2	0.582	A	0.000	0	0	0	
		B	0.060	0.111	0	0.111	
		C	0.284	0.278	0.333	-0.055	
		D	0.582	0.611	0.444	0.167	*
		E	0.075	0	0.222	-0.222	
		Other	0.000	0	0	0	
3	0.657	A	0.657	0.833	0.278	0.555	*
		B	0.045	0	0.056	-0.056	
		C	0.254	0.111	0.5	-0.389	
		D	0.015	0.056	0.056	0	
		E	0.015	0	0.056	-0.056	
		Other	0.015	0	0.056	-0.056	
4	0.925	A	0.015	0	0	0	
		B	0.030	0.056	0.056	0	
		C	0.015	0	0.056	-0.056	
		D	0.015	0	0.056	-0.056	
		E	0.925	0.944	0.833	0.111	*
		Other	0.0	0	0	0	
5	0.403	A	0.194	0.111	0.111	0	
		B	0.075	0	0.056	-0.056	
		C	0.239	0.111	0.5	-0.389	
		D	0.403	0.722	0.222	0.5	*
		E	0.045	0.056	0.111	-0.055	
		Other	0.045	0	0	0	
6	0.328	A	0.328	0.5	0.167	0.333	*
		B	0.134	0.222	0	0.222	
		C	0.119	0.111	0.222	-0.111	
		D	0.299	0.167	0.556	-0.389	
		E	0.015	0	0	0	
		Other	0.104	0	0.056	-0.056	
7	0.358	A	0.358	0.556	0.167	0.389	
		B	0.507	0.333	0.556	-0.223	*
		C	0.030	0.056	0.056	0	

		D	0.060	0.056	0.111	-0.055	
		E	0.030	0	0.111	-0.111	
		Other	0.015	0	0	0	

Q#	p-value	Choices	Proportion of Responses	High Group	Low Group	Item Discrimination Power	Correct Answer
8	0.433	A	0.015	0	0	0	
		B	0.030	0.056	0.056	0	
		C	0.149	0.056	0.278	-0.222	
		D	0.433	0.667	0.278	0.389	*
		E	0.358	0.222	0.389	-0.167	
		Other	0.015	0	0	0	
9	0.478	A	0.209	0	0.222	-0.222	
		B	0.060	0	0	0	
		C	0.194	0.111	0.333	-0.222	
		D	0.478	0.833	0.278	0.555	*
		E	0.045	0.056	0.111	-0.055	
		Other	0.015	0	0.056	-0.056	
10	0.284	A	0.284	0.556	0.167	0.389	*
		B	0.149	0.056	0.167	-0.111	
		C	0.328	0.278	0.222	0.056	
		D	0.224	0.111	0.389	-0.278	
		E	0.015	0	0.056	-0.056	
		Other	0.000	0	0	0	
11	0.164	A	0.075	0	0.222	-0.222	
		B	0.657	0.833	0.444	0.389	
		C	0.030	0	0.111	-0.111	
		D	0.060	0	0.111	-0.111	
		E	0.164	0.167	0.056	0.111	*
		Other	0.015	0	0.056	-0.056	
12	0.627	A	0.179	0.056	0.333	-0.277	
		B	0.104	0	0.167	-0.167	
		C	0.627	0.944	0.333	0.611	*
		D	0.030	0	0.056	-0.056	
		E	0.045	0	0.056	-0.056	
		Other	0.015	0	0.056	-0.056	
13	0.403	A	0.030	0	0.056	-0.056	
		B	0.403	0.556	0.333	0.223	*
		C	0.075	0	0.222	-0.222	

		D	0.119	0.056	0.056	0	
		E	0.373	0.389	0.333	0.056	
		Other	0.000	0	0	0	
14	0.269	A	0.269	0.667	0.056	0.611	
		B	0.090	0.111	0	0.111	*
		C	0.075	0.056	0.167	-0.111	
		D	0.373	0.111	0.444	-0.333	
		E	0.164	0.056	0.222	-0.166	
		Other	0.030	0	0.111	-0.111	

Q#	p-value	Choices	Proportion of Responses	High Group	Low Group	Item Discrimination Power	Correct Answer
15	0.104	A	0.119	0.056	0.167	-0.111	
		B	0.418	0.111	0.389	-0.278	
		C	0.284	0.5	0.278	0.222	
		D	0.104	0.222	0.056	0.166	*
		E	0.060	0.111	0.056	0.055	
		Other	0.015	0	0.056	-0.056	
16	0.597	A	0.030	0.056	0	0.056	
		B	0.015	0	0.056	-0.056	
		C	0.030	0.056	0	0.056	*
		D	0.597	0.889	0.389	0.5	
		E	0.299	0	0.389	-0.389	
		Other	0.030	0	0.111	-0.111	
17	0.463	A	0.463	0.833	0.222	0.611	*
		B	0.045	0	0.056	-0.056	
		C	0.224	0.056	0.333	-0.277	
		D	0.224	0.111	0.278	-0.167	
		E	0.015	0	0	0	
		Other	0.030	0	0.111	-0.111	
18	0.149	A	0.134	0.111	0.167	-0.056	
		B	0.403	0.333	0.278	0.055	
		C	0.149	0.111	0.222	-0.111	
		D	0.119	0.111	0.111	0	
		E	0.149	0.333	0.111	0.222	*
		Other	0.045	0	0.111	-0.111	
19	0.269	A	0.269	0.556	0.056	0.5	*
		B	0.060	0.111	0	0.111	
		C	0.269	0.111	0.278	-0.167	
		D	0.358	0.222	0.556	-0.334	
		E	0.015	0	0	0	
		Other	0.030	0	0.111	-0.111	

20	0.209	A	0.060	0	0.056	-0.056	
		B	0.209	0.444	0.167	0.277	*
		C	0.448	0.389	0.333	0.056	
		D	0.119	0.111	0.167	-0.056	
		E	0.134	0.056	0.167	-0.111	
		Other	0.030	0	0.111	-0.111	
21	0.119	A	0.060	0.056	0.222	-0.166	
		B	0.164	0.222	0.232	-0.010	
		C	0.030	0.056	0.111	-0.055	
		D	0.597	0.556	0.157	0.389	
		E	0.119	0.111	0.222	-0.111	*
		Other	0.030	0	0.056	-0.056	

Q#	p-value	Choices	Proportion of Responses	High Group	Low Group	Item Discrimination Power	Correct Answer
22	0.299	A	0.104	0.056	0	0.056	
		B	0.209	0.222	0.333	-0.111	
		C	0.164	0.222	0.167	0.055	
		D	0.299	0.389	0.167	0.222	*
		E	0.209	0.111	0.278	-0.167	
		Other	0.015	0	0.056	-0.056	
23	0.448	A	0.119	0.056	0.167	-0.111	
		B	0.209	0.056	0.333	-0.277	
		C	0.075	0.056	0.056	0	
		D	0.448	0.833	0.222	0.611	*
		E	0.104	0	0.111	-0.111	
		Other	0.045	0	0.111	-0.111	
24	0.239	A	0.119	0.167	0.056	0.111	
		B	0.075	0	0	0	
		C	0.239	0.5	0.111	0.389	*
		D	0.299	0.056	0.389	-0.333	
		E	0.209	0.278	0.278	0	
		Other	0.060	0	0.167	-0.167	
25	0.224	A	0.030	0	0	0	
		B	0.090	0.167	0	0.167	
		C	0.224	0.389	0.222	0.167	*
		D	0.328	0.333	0.222	0.111	
		E	0.254	0.111	0.389	-0.278	

		Other	0.075	0	0.167	-0.167	

APPENDIX D: QMVI - FINAL VERSION

The Quantum Mechanics Visualization Instrument (QMVI)

Designed by Dr. Richard W. Robinett

Contributors to the development of this instrument included:
Mr. Erdat Cataloglu
Dr. Vincent N. Lunetta
Dr. Peter A. Rubba
Dr. Hoi Suen
Members of the Expert Panels in Physics and Science Education
Participating Faculty and Students

Quantum Mechanics Visualization Instrument (QMVI)

The survey you are about to complete is designed to probe students' understanding of concepts in quantum theory, especially their ability to visualize quantum mechanical ideas and phenomena, and how this ability changes over the undergraduate career. This test is being given to students ranging from undergraduates in their sophomore year (modern physics type classes) to first year graduate students. Because of this, the range of difficulty you encounter in the instrument may seem to vary dramatically from question to question. Look at each item carefully and use whatever knowledge you can bring to bear to answer it as best as you can.

The test consists of 25 multiple choice questions. For each item, please read the question, using the accompanying graphic, if appropriate, and then

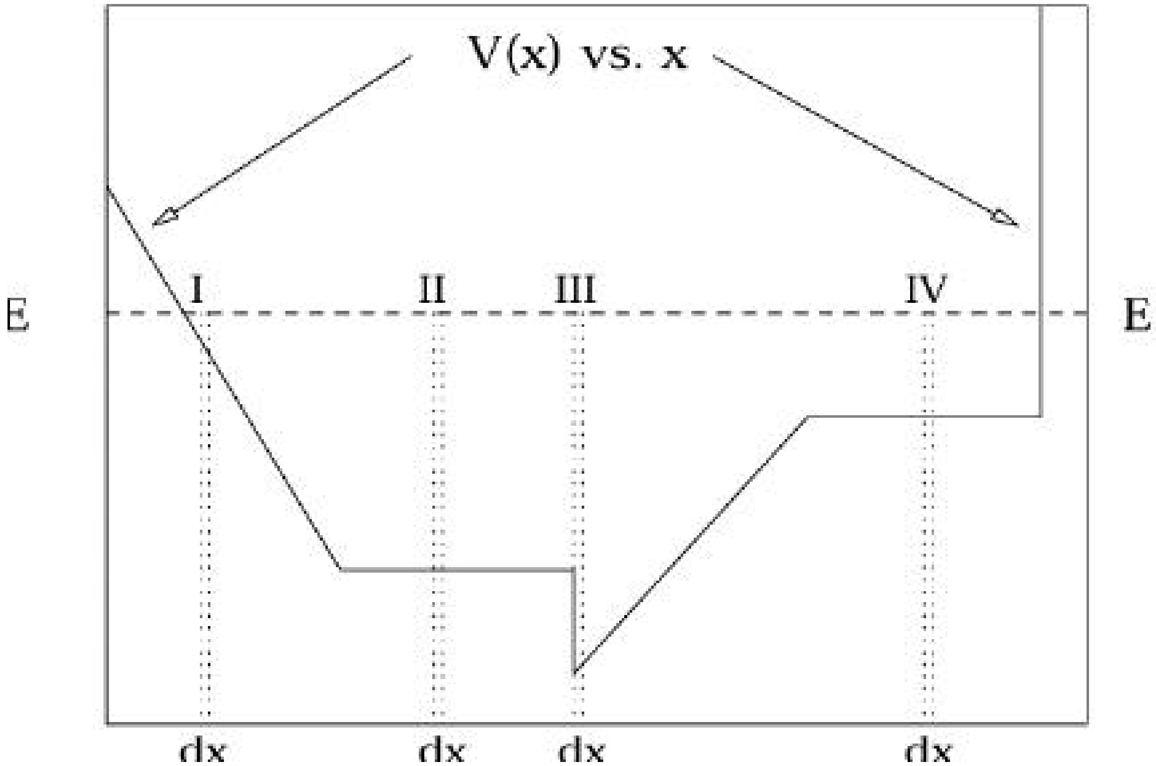
- Circle the best answer (a), (b), (c), (d), or (e)
- Briefly explain (in one or two sentences) your reasoning in picking that answer in the space provided
- Circle the statement at the bottom of the page which best describes how you feel about your answer, judging how confidant you are in your response.

Good luck and have fun!

Name: ____

Soc-Sec-No:____

Course: ____



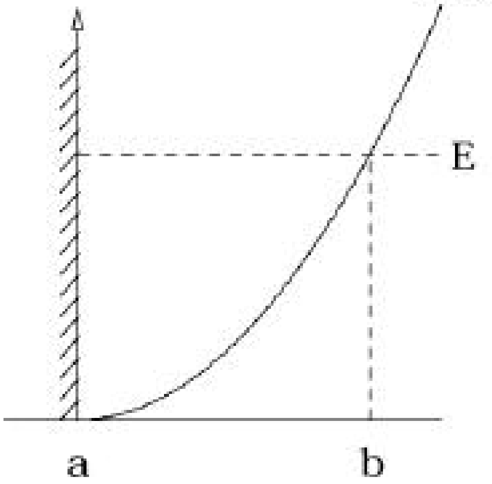
1. The plot above shows a function of **potential energy** versus position (*V(x)* versus *x*) for a one-dimensional system. A dashed horizontal line indicates the value of energy *E* for a particle moving in this potential well, corresponding to a bound state system. Small regions, each of equal width *dx*, are indicated at several locations in the well. Order the **time spent** in each small position bin *(dx)*, *t(I* , *t(II)*, *t(III)*, and *t(IV)*, from **longest** to **shortest** as the particle moves back and forth in the well.

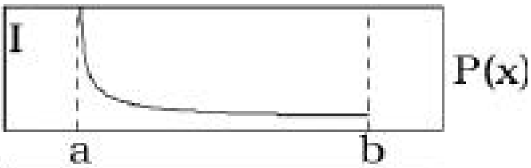
- *(a)* *t(III)> t(II)> t(IV)> t(I)*
- *(b)* *t(I)> t(IV) =t(II)>t(III)*
- *(c)* *t(III) > t(II) = t(IV)> t(I)*
- *(d)* *t(I)> t(IV)> t(II)> t(III)*
- *(e)* *t(I) = t(II) = t(III) = t(IV)* since *E* is constant

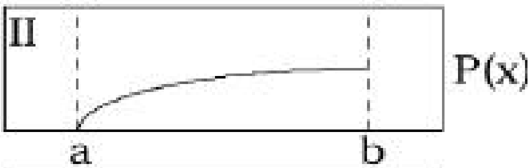
↑ Explain your answer in 1-2 sentences in the space above. ↑

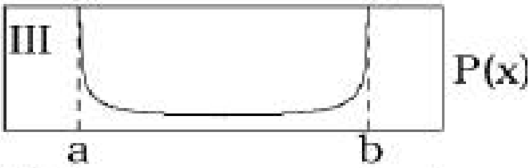
↓ Circle the statement below which describes how you feel about your answer. ↓

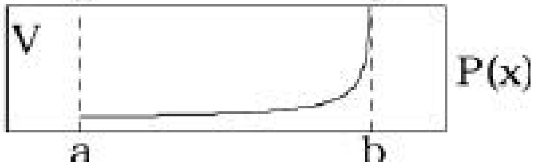
Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**











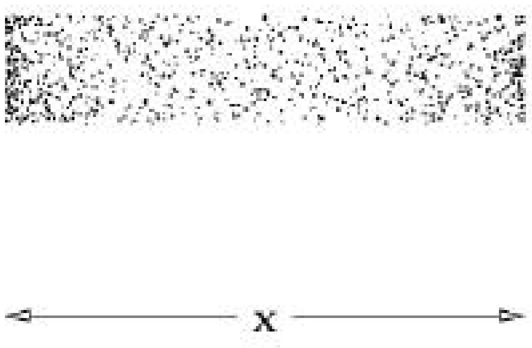
2. A particle of energy *E* moves between the classical turning points, *a* and *b*, in a one-dimensional potential as shown in the figure on the left above. The turning point at *a* is defined by an impenetrable wall, while the one at *b* is given by the intersection with the *V(x)* curve shown. Which of the **classical probability distributions**, *P(x)*, shown on the right corresponds to this system. (Recall that the classical probability distribution is defined such that the probability of finding the particle in the small interval ($x,x+dx$) is given by $d\mathrm{Prob}(x,x+dx) = P(x)dx$.)

(a) I
(b) II
(c) III
(d) IV
(e) V

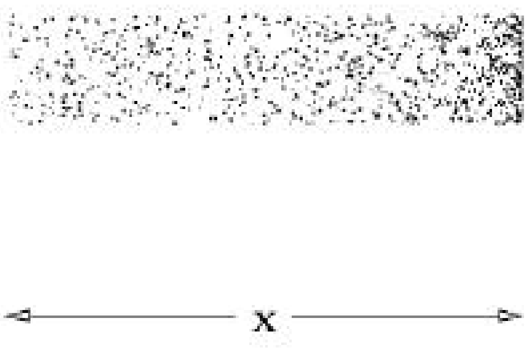
↑ Explain your answer in 1-2 sentences in the space above. ↑
↓ Circle the statement below which describes how you feel about your answer. ↓

Very certain	**somewhat certain**	**somewhat uncertain**	**very uncertain**

I



II



III

IV

3. The position of four objects with differing types of motion in one-dimension are captured at a large number of random times in the computer-generated 'snapshots' shown above. (The vertical spread in the dots is simply added to make the density of dots more clearly visible.) Two possible one-dimensional motions are:

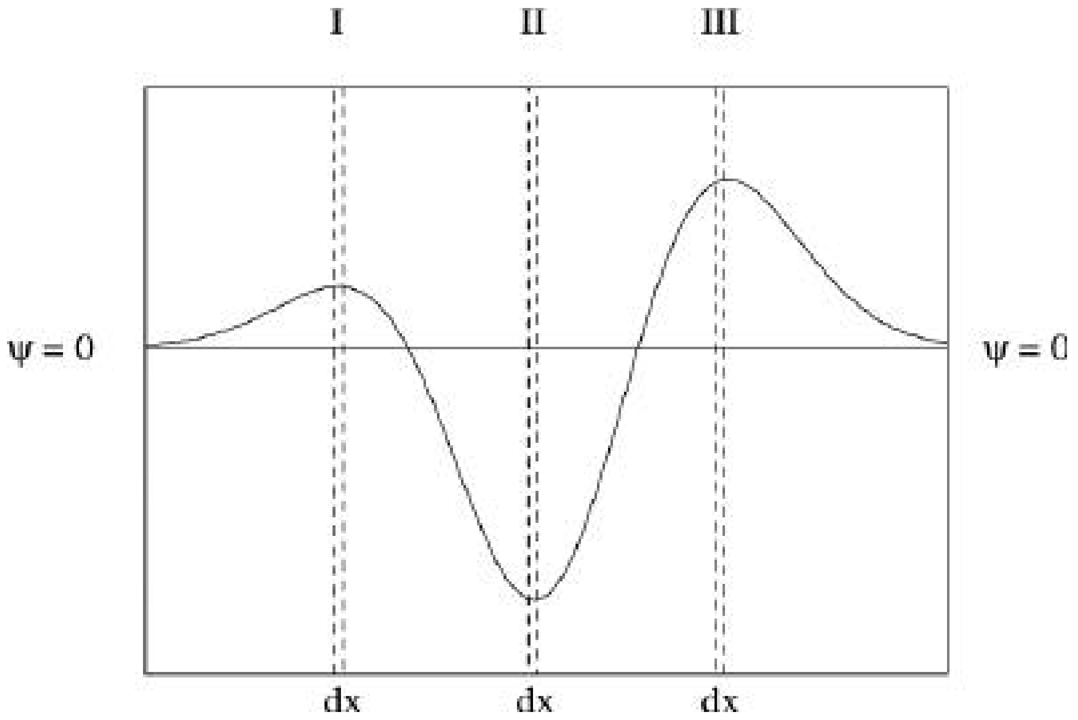
A. A mass oscillating at the end of a spring
B. A mass undergoing uniform acceleration to the right, starting from rest

Which motion (A, B) goes with which 'snapshot' (I, II, III, IV)?

(a) A: IV B: II
(b) A: IV B: III
(c) A: I B: II
(d) A: I B: III
(e) None of the above

↑ 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**

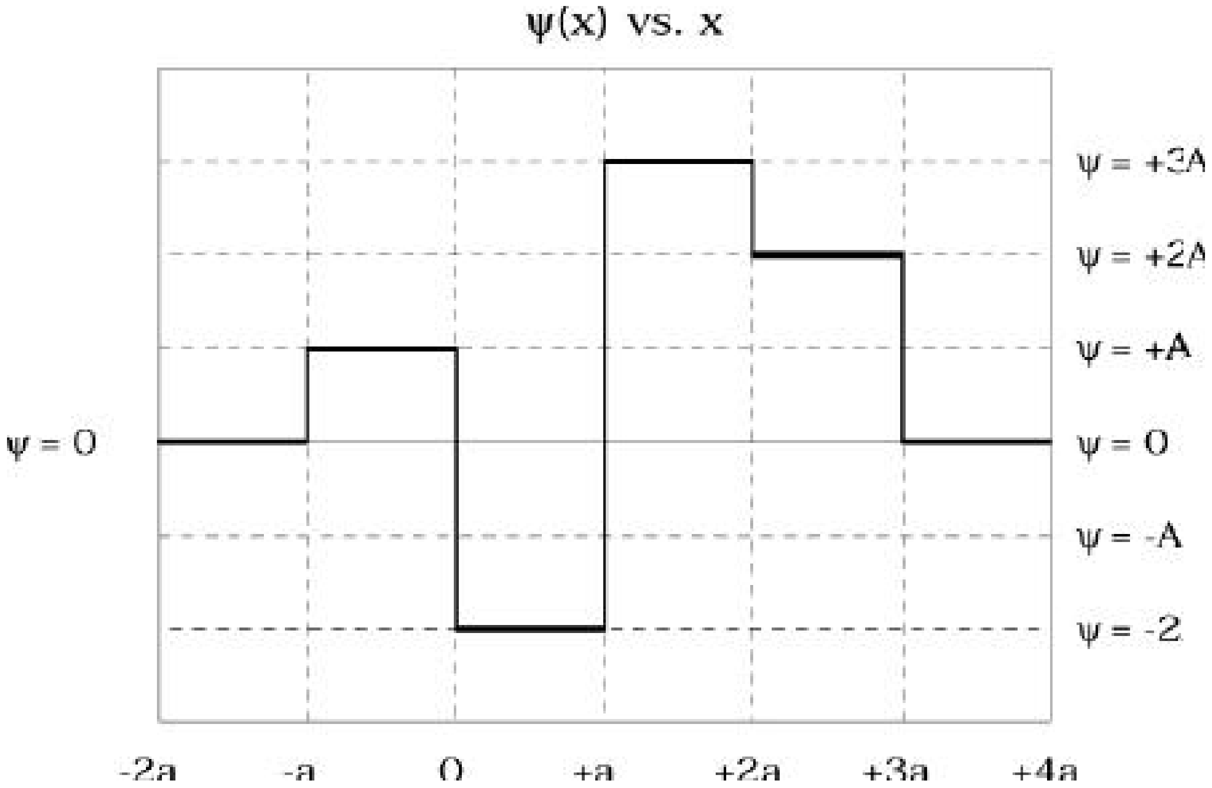


4. The plot above shows a one-dimensional wave function, $\psi(x)$ versus x. The labels, *I*, *II*, and *III*, indicate regions about which measurements of the position of the particle can be made. Order the probabilities of finding the particle described by $\psi(x)$ in a small region of width dx near each labeled point, from **biggest** to **smallest**.

(a) P(III) > P(I) > P(II)
(b) P(II) > P(I) > P(III)
(c) P(III) > P(II) > P(I)
(d) P(I) > P(II) > P(III)
(e) P(II) > P(III) > P(I)

↑ 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



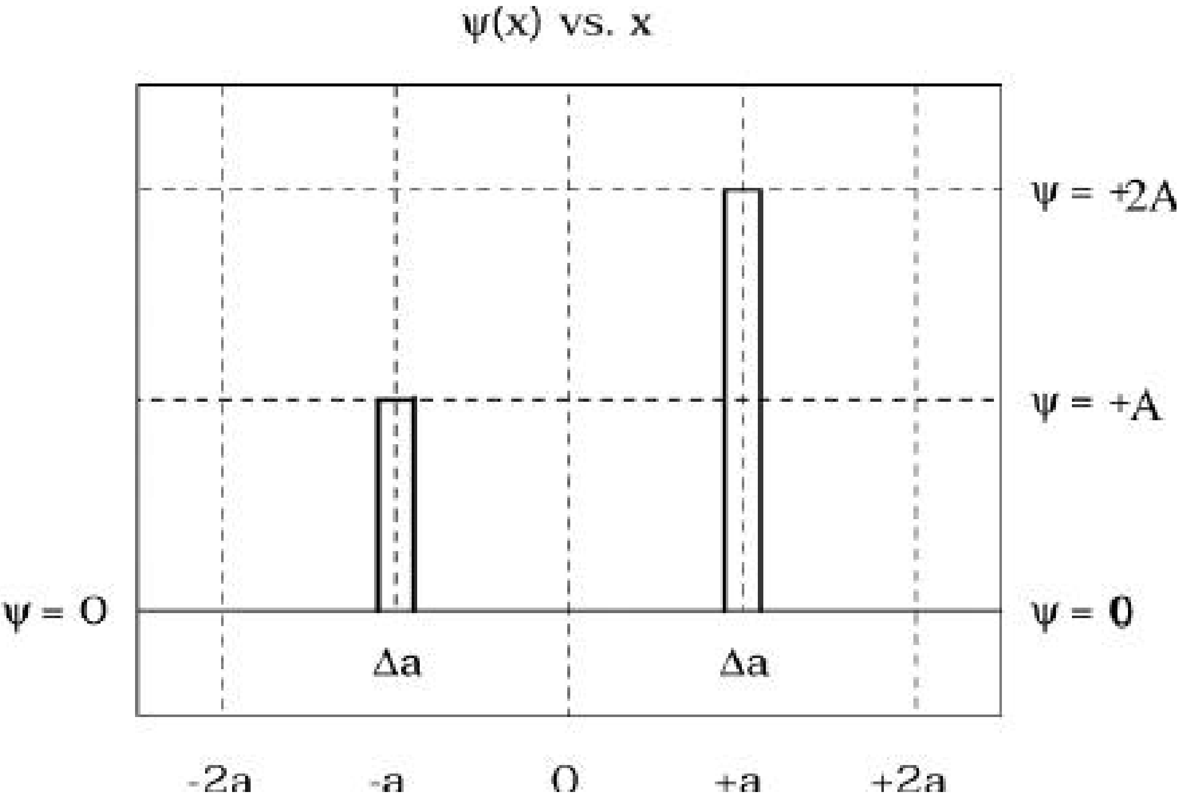
5. The picture above shows a plot of a (rather artificial) one-dimensional wave function, ψ*(x)* versus *x*, over the range *(-2a,+4a)* . The wave function vanishes for all other values of *x*. What is the probability that a measurement of the position of the particle would find it in the range *(+2a,+3a)*, i.e., from *x=+2a to x=+3a?*

(a) *2/9*
(b) *1/6*
(c) *1/4*
(d) *1/2*
(e) *1/3*

↑ Explain your answer in 1-2 sentences in the space above. ↑

↓ Circle the statement below which describes how you feel about your answer. ↓

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**



6. The figure above shows a plot of a (rather artificial) wave function, ψ*(x)* versus *x*, given by the two 'spikes' shown, each of width Δa . The wave function vanishes for all other values of *x* not shown. Which of the expressions below is closest to the value of the **spread** or **uncertainty** in the position variable, namely Δx . Recall that

$$\Delta x \equiv \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

where $\langle ... \rangle$ denotes **average** or **expectation** value.

(a) Δx = 3a/5
(b) Δx = 2a
(c) Δx = 2Δ a
(d) Δx = 4a/5
(e) Δx = Δ a

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

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

7. The position-space wave function of a particle at some instant is given by the functional form

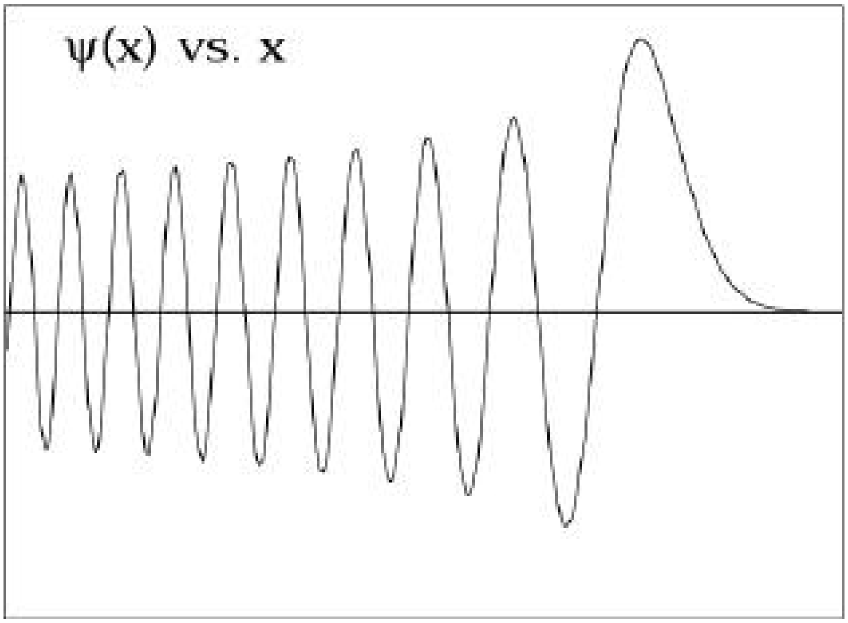
$$\psi(x) = \frac{1}{\left[L^2 + (x - x_0)^2/\alpha^2\right]}$$

where L,x_0, and _ are all constants. Which expression below would be closest to the value of the **spread** or **uncertainty** in **momentum**, i.e.,Δp , at that time. (You are not expected to evaluate any integrals for this problem.)

(a) Δp ~ ħ /(_L)
(b) Δp ~ ħ /_
(c) Δp ~ ħ /(_^{2} x_0)
(d) Δp ~ ħ /(_^{2} L^2)
(e) Δp ~ ħ /(_ x_0)

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

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

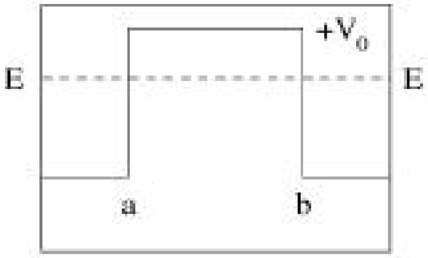


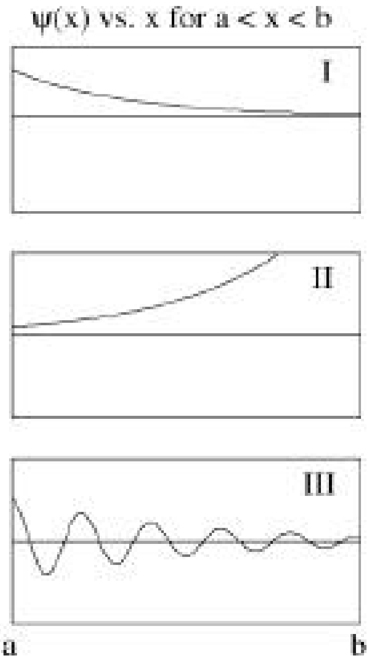
8. A plot of a position-space energy eigenstate, $\psi(x)$ versus x, is shown above. To which classical system would this solution of the time-independent Schrödinger equation most likely correspond?

(a) A free particle, moving at constant velocity to the right.
(b) A free particle, moving at constant velocity to the left.
(c) A particle undergoing uniform acceleration to the right, starting from rest.
(d) A particle undergoing uniform acceleration to the left, starting from rest.
(e) An otherwise free particle, moving to the right, but colliding with an infinite wall.

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

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**





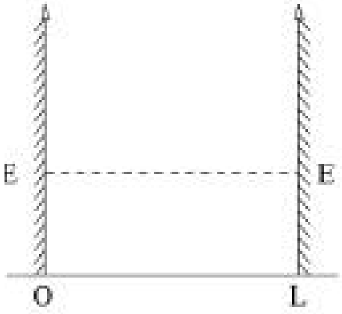
9.The figure on the left above indicates a particle of energy E near the boundary of a 'step up, then down' potential located between x=a and $x=b$ as shown, with $0 < E < +V_0$. If you solve the time-independent Schrödinger equation in the region (a,b) using this potential and this energy, which of the possible wave functions on the right could you find?

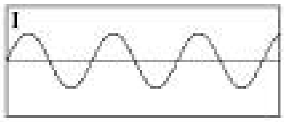
(a) *I* and *II* only.
(b) *III* only.
(c) *I* only.
(d) *I* and *III* only.
(e) *I*, *II* and *III* are all possible.

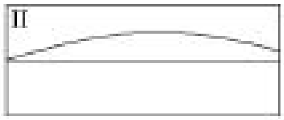
↑ Explain your answer in 1-2 sentences in the space above. ↑

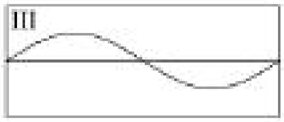
↓ Circle the statement below which describes how you feel about your answer. ↓

Very certain	**somewhat certain**	**somewhat uncertain**	**very uncertain**









10. You are given a computer program which is designed to solve the time-independent Schrödinger equation (TISE)

$$-\frac{\hbar d^2\psi(x)}{2m\ dx} + V(x)\varphi(x) = E\psi(x)$$

for a particle inside the infinite well shown above (defined by impenetrable walls at $x = 0$, L and with $V(x) = 0$ in between) for any value of $E > 0$. The program finds solutions to this particular TISE which also satisfy the boundary condition $\psi(0) = 0$ and outputs wavefunctions in a graphical format. A convenient reference value of energy is defined to be $E_{REF} = \hbar^2\pi^2/2mL^2$. Consider the following statements about the waveforms I -- IV shown above:

I: You could get Fig. I as output for some value of E >> E_{REF}
II: You could get Fig. II as output for some value of E < E_{REF}
III: You could get Fig. III as output for E = $2E_{REF}$
IV: You could get Fig. IV as output for some value of E.

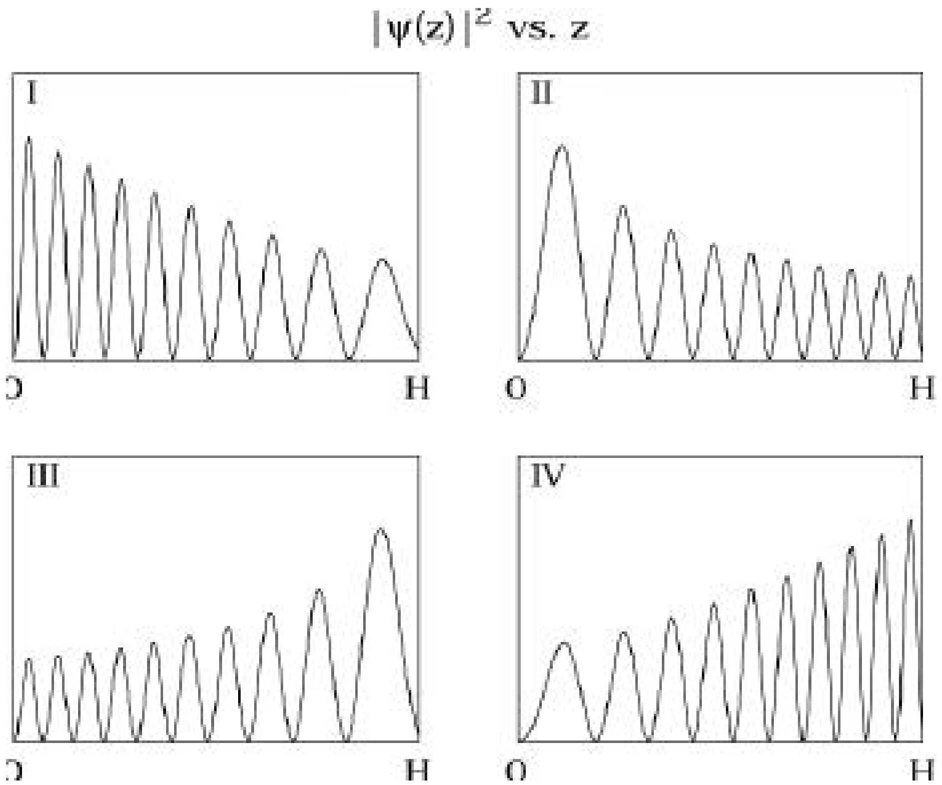
Which of these statements is (are) true?

(a) III and IV only
(b) III only
(c) I and II only
(d) I and III only
(e) None of them are true

↑ Explain your answer in 1-2 sentences in the space above. ↑

↓ Circle the statement below which describes how you feel about your answer. ↓

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

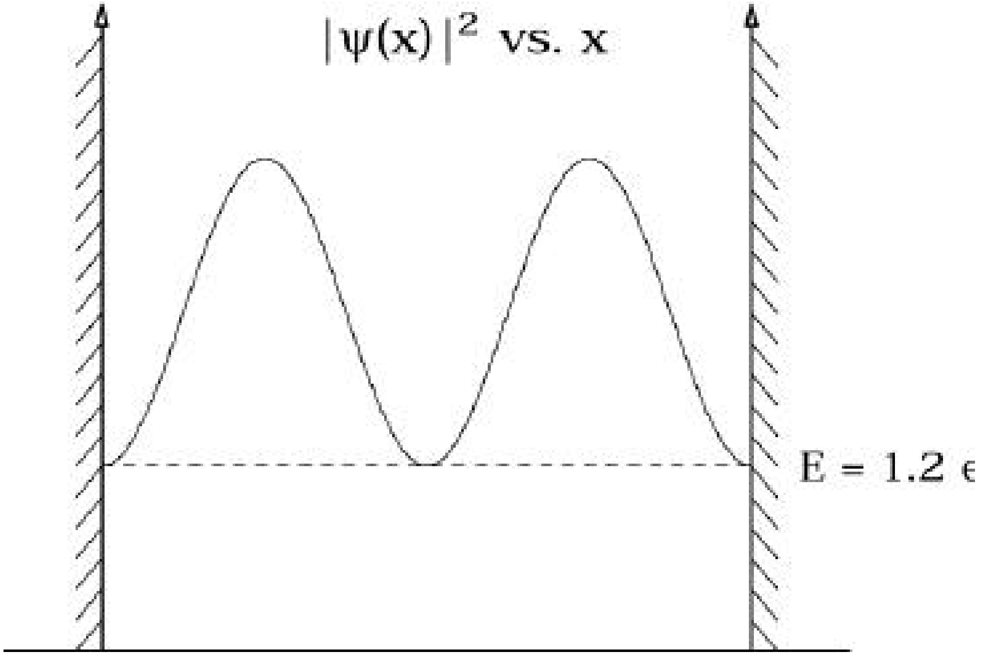


11.A particle is dropped from a height *H* under the influence of gravity and bounces, without loss of energy, from a flat surface (at $z=0$) back up to the same height. Which of the plots above would be the best representation of the quantum mechanical position-space probability density, $|\psi(z)^2|$versus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



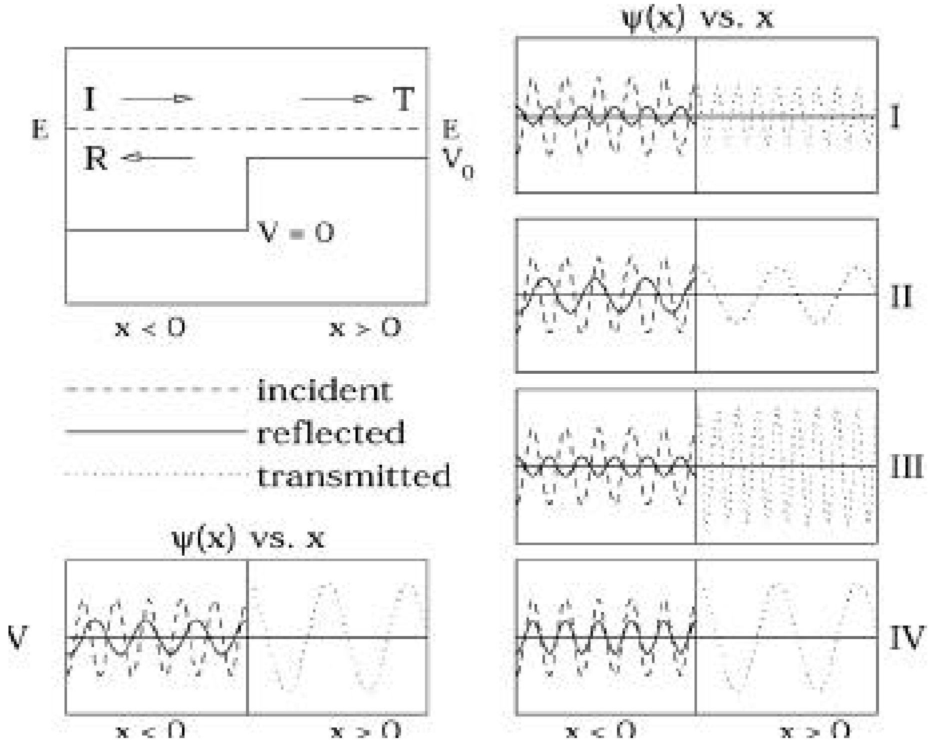
12 . A particle in an infinite well (where $V(x) = 0$ between the otherwise impenetrable walls) is in the energy eigenstate corresponding to the probability density $|\psi(x)|^2$ versus x shown above. The energy of this state is E = 1.2eV. What is the energy of the lowest possible allowed energy state in this well?

(a) 0.0eV
(b) 0.3eV
(c) 0.4eV
(d) 0.6eV
(e) 1.2eV

↑ Explain your answer in 1-2 sentences in the space above. ↑

↓ Circle the statement below which describes how you feel about your answer. ↓

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**



Figures I, II, III, IV, and V show possible solutions to the scattering problem illustrated in the upper left corner in terms of incident (dashed), reflected (solid) and transmitted (dotted) plane wave solutions.

13. A beam of particles of energy E is incident, from the left, on a step potential of height $V_0 < E$ as shown in the figure on the next page. This problem is often analyzed using plane-wave solutions of the time-independent Schrödinger equation with incident (I), reflected (R) and transmitted (T) components represented by the solution

$$\psi(x) = \begin{cases} Ie^{ikx} + \mathrm{R}e^{-ikx} & for\ x \le 0 \\ Te^{iqx} & for\ x \ge 0 \end{cases}$$

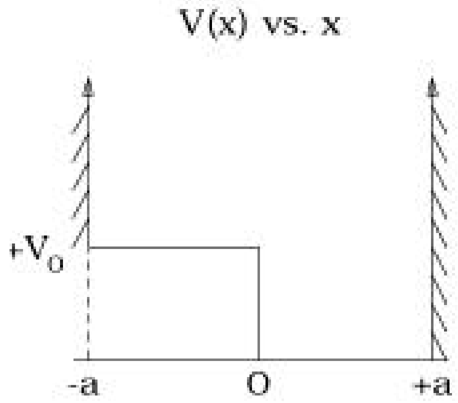
where k = $\sqrt{2mE/\hbar^2}$ and q = $\sqrt{2m(E-V_0)/\hbar^2}$.

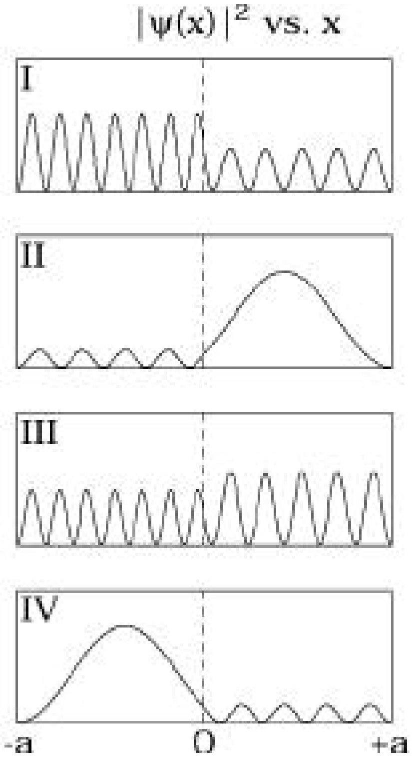
Which of the waveforms shown on the left figure (namely I, II, III, IV, or V) best represents the true solution in this approach?

(a) I
(b) II
(c) III
(d) IV
(e) V ↑ **Explain your answer in 1-2 sentences in the space above.** ↑

↓ **Circle the statement below which describes how you feel about your answer.** ↓

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**



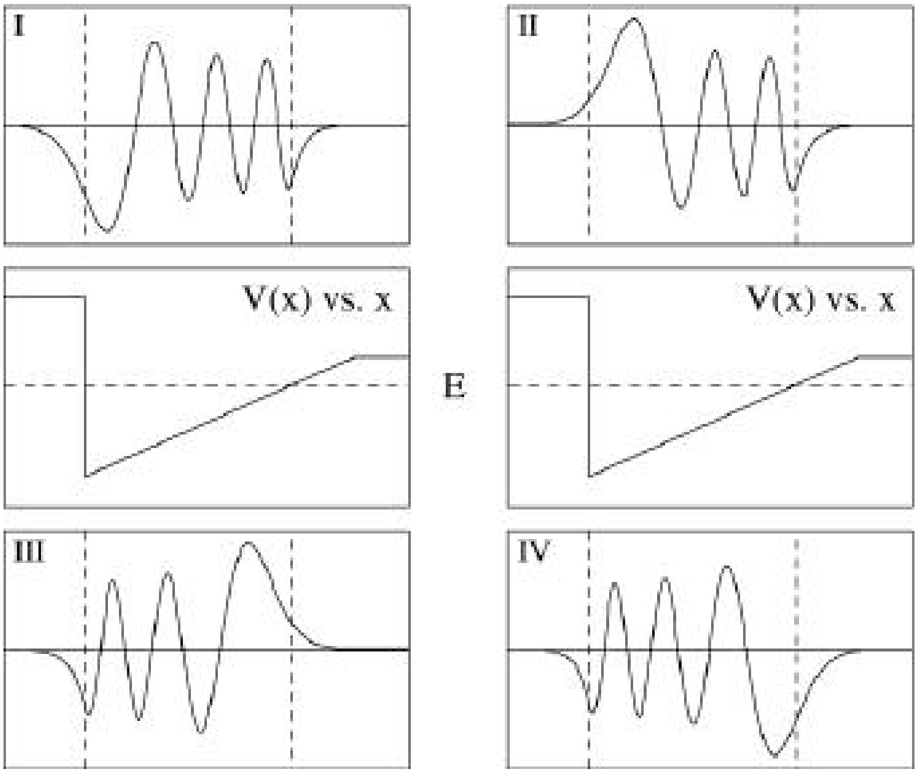


14.The plot on the left shows a potential energy function, $V(x)$ versus x, corresponding to an `asymmetric' infinite well. The infinite well is of width 2a, with impenetrable walls at $x = \pm a$, and where $V(x) = +V_0$ for $-a<x<0$ and $V(x) = 0$ for $0<x<+a$. Of the figures on the right, which is/are most likely to be physically acceptable energy eigenstate solutions for the time-independent Schrödinger equation for this well?

(a) *I and IV* only
(b) *I* only
(c) *II* and *III* only
(d) *II* only
(e) IV only

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

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

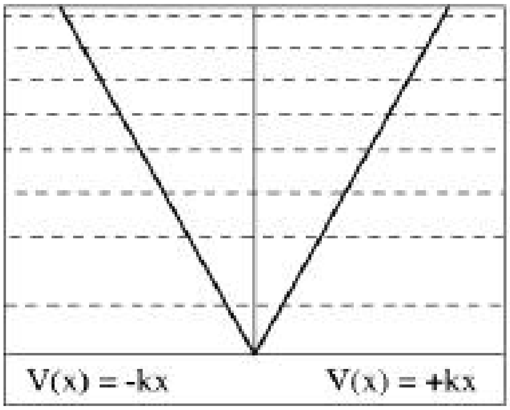


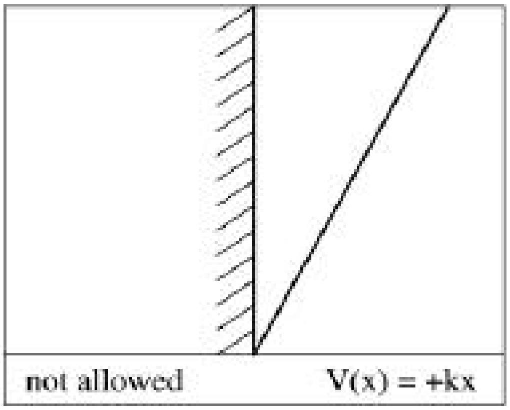
15.The two identical pictures in the center of the figure above show a one-dimensional potential, *V(x)* versus *x*, as well as a horizontal dashed line indicating a quantized energy level corresponding to the **sixth excited state**. Which of the wave function plots (shown on the top or bottom rows) is most likely to correspond to this state? Note that the vertical dashed lines in the $\psi(x)$ versus *x* pictures correspond to the locations of the classical turning points in the *V(x)* plot. (Recall that the lowest energy is conventionally called the **ground state**, the next higher level corresponds to the **first excited state** and so forth.)

(a) *I*
(b) *II*
(c) *III*
(d) *IV*
(e) There is something inconsistent about each one so that none of them is an acceptable wave function for this state.

↑ 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**





16.The picture on the left above shows a one-dimensional potential energy function given by V(x) = k|x| (solid lines) as well as the eight lowest allowed energy eigenstates (dashed horizontal lines.) These energies are given, in terms of a quantity __ by the values

E =(1.0, 2.3, 3.2, 4.1, 4.8, 5.5, 6.2, 6.8, …)__

The plot on the right shows the 'half-well' version of this potential, namely that given by

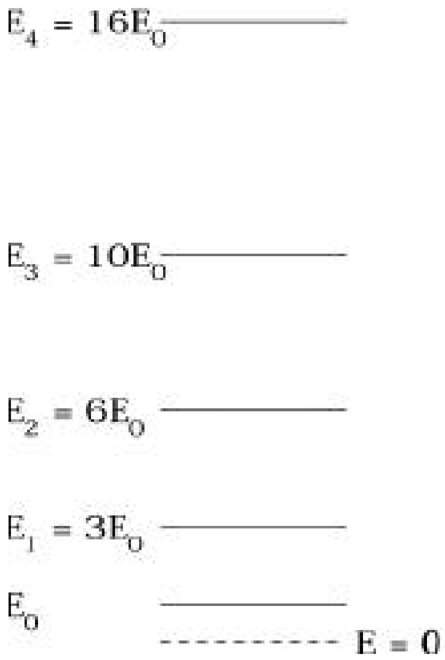
$$V(x) = \begin{cases} \infty & for\ x < 0 \\ +kx & for\ x > 0 \end{cases}$$

so that the potential is the same for *x>0* , but the particle is not allowed in the region *x<0* . The most likely pattern of energy eigenvalues in the 'half-well' is given by

(a) E = (2.3, 4.1, 5.5, 6.8,...)__
(b) E = (1.0, 3.2, 4.8,6.2,...)__
(c) E = (1.0, 2.3, 3.2, 4.1,...)__(i.e., the same energy spectrum)
(d) E = (1.0, 2.3, 3.2, 4.1,...)(2__) (i.e., the energies are all doubled)
(e) E = (1.0, 2.3, 3.2, 4.1,...)(__/2)(i.e., the energies are all halved)

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

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

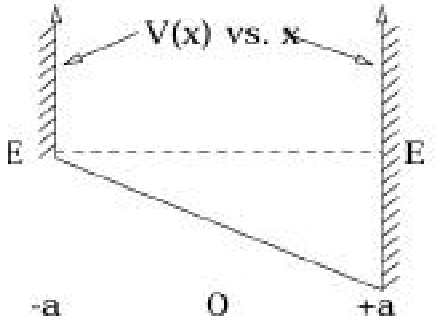


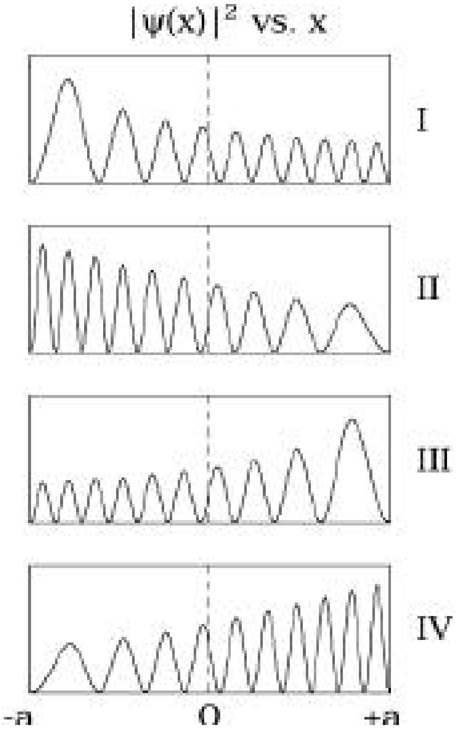
17. The spectrum of a one-dimensional bound state system is shown in the figure above, starting with the ground state (E_0) up through the 4th excited state (E_4). Five indistinguishable electrons are known to be in this well; their mutual interactions can be neglected. What is the lowest possible total energy that such a state can have?

(a) 0
(b) 5 E_0
(c) $14E_0$
(d) $36E_0$
(e) $80E_0$

↑ 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





18. On the left above is a picture of an infinite well potential, with impenetrable walls at ± *a*, but with a slanted bottom so that one end is deeper than the other. A particular quantized energy level is indicated by the horizontal dashed line. Which of the wave functions on the right (plots of $|\psi(x)|^2$ versus x are shown) is most likely to correspond to this energy level? (For the wave function pictures on the right, the vertical dashed line indicates the center of the well at $x = 0$.)

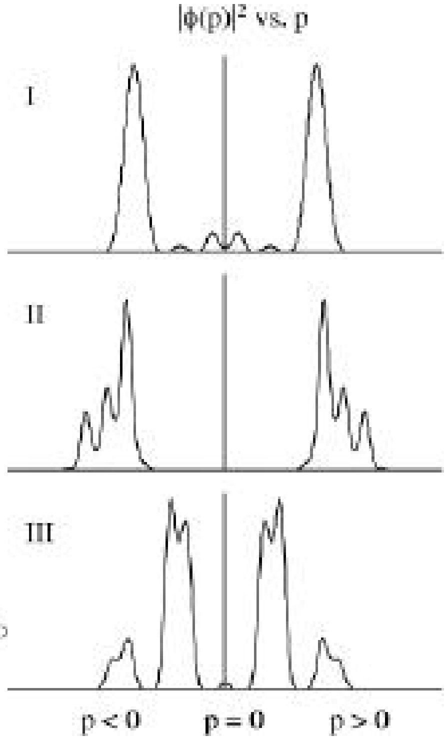
(a) *I*
(b) *II*
(c) *III*
(d) *IV*
(e) They are all possible solutions, depending on the exact numerical parameters of the well shown.

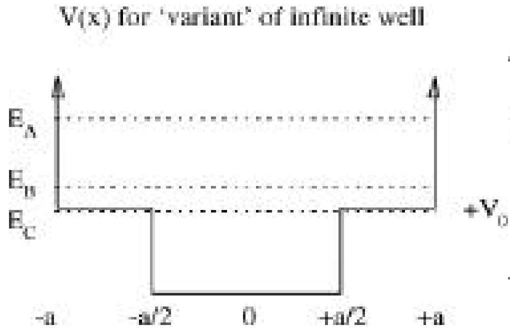
↑ Explain your answer in 1-2 sentences in the space above. ↑

↓ Circle the statement below which describes how you feel about your answer. ↓

Very	**somewhat**	**somewhat**	**very**

certain	**certain**	**uncertain**	**uncertain**



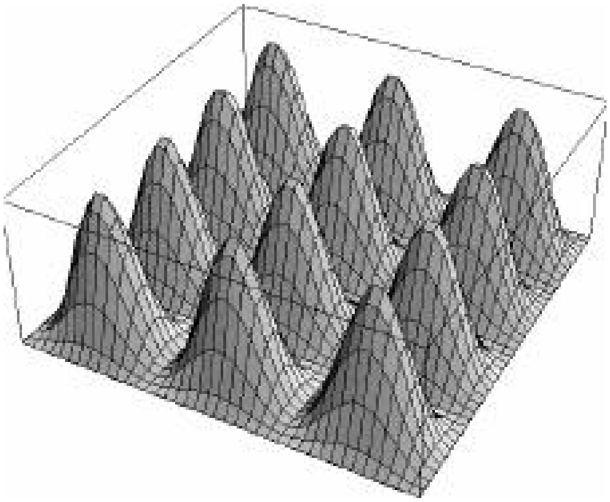


19. On the left above, there is a picture of three different quantized energy levels (E_A, E_B, E_C) in a variant of an infinite square well: it's an infinite well of width $2a$ with impenetrable walls at $x = \pm a$, but where $V(x) = 0$ for $-a/2 < x < +a/2$ and $V(x) = +V_0$ in the rest of the well. On the right are three different **momentum-space probability distributions** , plots of $|\phi(p)|^2$ versus p, corresponding to the three levels given by E_A , E_B , and E_C . Which momentum distribution (*I*, *II*, *III*) goes with which energy level? (Recall that the probability of finding the particle with momentum in the range $(p,p+dp)$ is given by $d\text{Prob}(p,p+dp) = |\phi(p)|^2 dp$.)

(a) E_A: I, E_B: III, E_C: II
(b) E_A: II, E_B: I, E_C: III
(c) E_A: III, E_B: II, E_C: I
(d) E_A: I, E_B: II, E_C: III
(e) E_A: II, E_B: III, E_C: I

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

Very	somewhat	somewhat	very
certain	**certain**	**uncertain**	**uncertain**



20.The plot above shows the position-space probability density of $|\psi\,(x,y)|^2$ versus (x,y), for an energy eigenstate of a particle of mass m in a **two-dimensional** infinite square well of length **L** on each side. The potential energy for such a well is given by

$$V(x,y)=\begin{cases} 0 & \text{for } 0<x,y<L \\ \infty & \text{for } x>L, x<0, y>L, \text{ or } y<0 \end{cases}$$

What is the energy eigenvalue corresponding to this state?

(a) E = $7\hbar^2\pi^2/2mL^2$

(b) E = $6\hbar^2\pi^2/2mL^2$

(c) E = $13\hbar^2\pi^2/2mL^2$

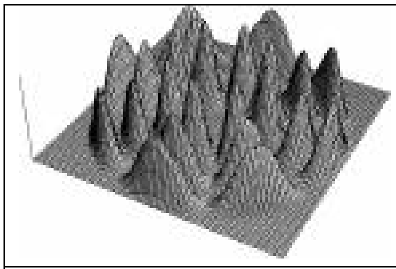
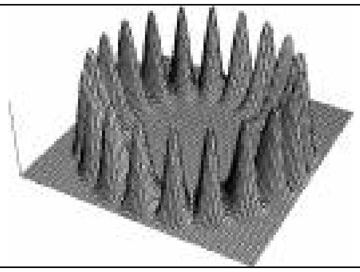
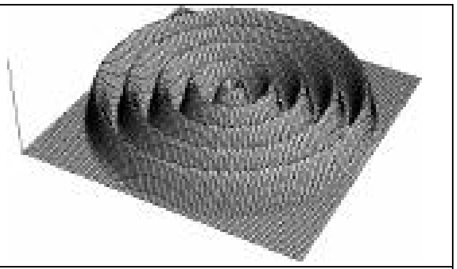
(d) E = $12\hbar^2\pi^2/2mL^2$

(e) E = $25\hbar^2\pi^2/2mL^2$

↑ Explain your answer in 1-2 sentences in the space above. ↑

↓ Circle the statement below which describes how you feel about your answer. ↓

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

		
I	II	III

21.The plots above illustrate the position-space probability density for three different states for a particle in a **two-dimensional infinite circular well** . Such a potential is defined by

$$V(x,y) = V(r) = \begin{cases} 0 & for\ r < a \\ \infty & for\ r > a \end{cases}$$

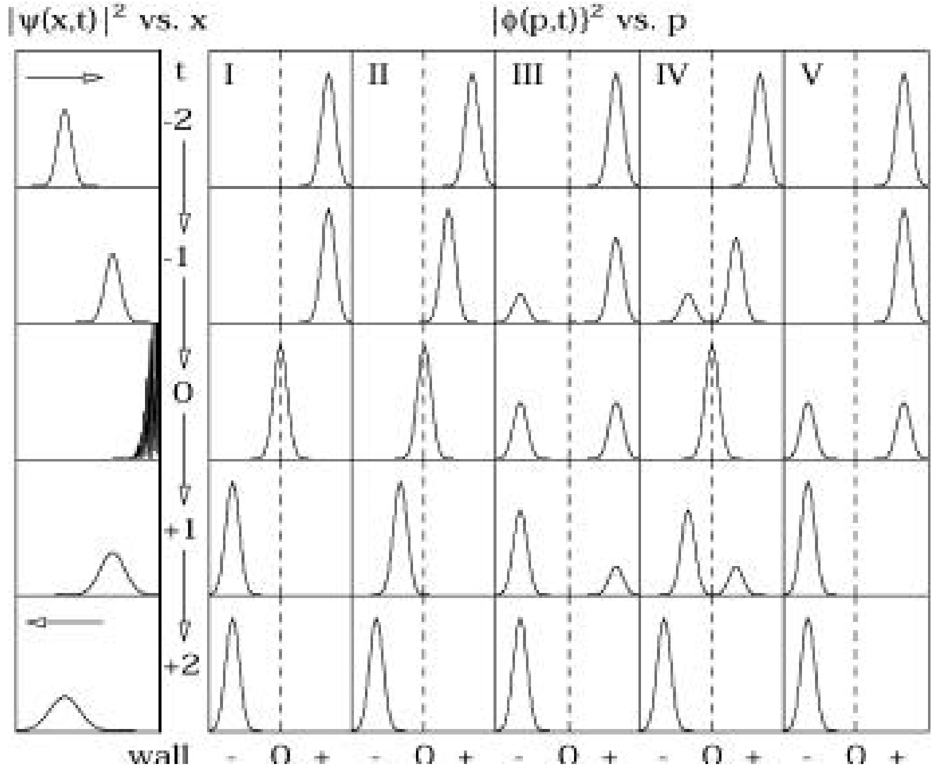
The three different energy eigenstates above have the same **total energy**. Order the three states, *I*, *II*, and *III*, in terms of their **rotational** or **angular kinetic energy**, *RKE*, from **highest** to **lowest**.

(a) RKE(II) > RKE(III) > RKE(I)
(b) RKE(III) > RKE(I) > RKE(II)
(c) RKE(I) = RKE(II) = RKE(III)
(d) RKE(II) > RKE(I) > RKE(III)
(e) RKE(III) > RKE(II) > RKE(I)

↑ Explain your answer in 1-2 sentences in the space above. ↑

↓ Circle the statement below which describes how you feel about your answer. ↓

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

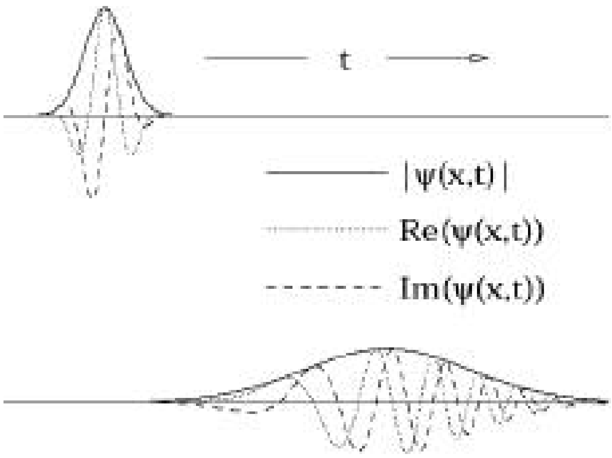


22. A position-space wave packet is constructed which represents an otherwise free particle which moves to the right, hits an impenetrable wall, 'bounces', and eventually moves freely to the left (i.e., $V(x) = 0$ to the left of the infinite wall.) Five 'snapshots' of the behavior of $|\psi (x,t)|^2$ versus x are shown in the left most column for times before (t < 0), during (t ≈ 0), and after (t > 0) the collision. Which of the corresponding `time-lapse' plots of the **momentum-space probability distributions**, $|\phi (p,t)|^2$ versus p, shown in the other vertical columns (labeled I, II, III, IV, and V) best represents the corresponding collision as seen in momentum-space? (Positive (+) and negative (-) values of momentum are indicated, with $p = 0$ denoted by the vertical dashed lines.)

(a) I
(b) II
(c) III
(d) IV
(e) V

↑ 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**



23. 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 ψ*(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 $|\psi(x,t)|$ 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	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**

24. The pictures on the next page illustrate the time evolution of a wave packet in a one-dimensional infinite square well of width *L*, formed from energy eigenstates $u_n(x)$ with quantized energies E_n, by the linear combination

$$\psi(x,t) = \sum_{n=1}^{\infty} a_n u_n(x) e^{-iE_n t/\hbar}$$

with appropriate coefficients, a_n. (Plots of $|\psi(x,t)|^2$ versus *x* for various times are shown.) The wave packet is designed to travel initially to the right, returning to something like its initial state after a time t_{CL} (its classical period) and two bounces with the walls, as shown by the figure on the left. The behavior of the wave packet after *4* and *8* classical periods is also shown, on the right. For even later times, the wave packet will:

(a) Permanently reform into two, distinct peaks, reflecting the presence of left-moving and right-moving momentum components.

(b) Reform exactly into the initial state after a time which depends on the initial spread (Δx) and/or average energy ($\overline{E}$) of the wave packet, repeating the cycle thereafter.

(c) Reform exactly into the initial state after a time which is determined by the width of the well *(L)* and the mass of the particle *(m)*, and Planck's constant ($\hbar$) repeating the cycle thereafter.

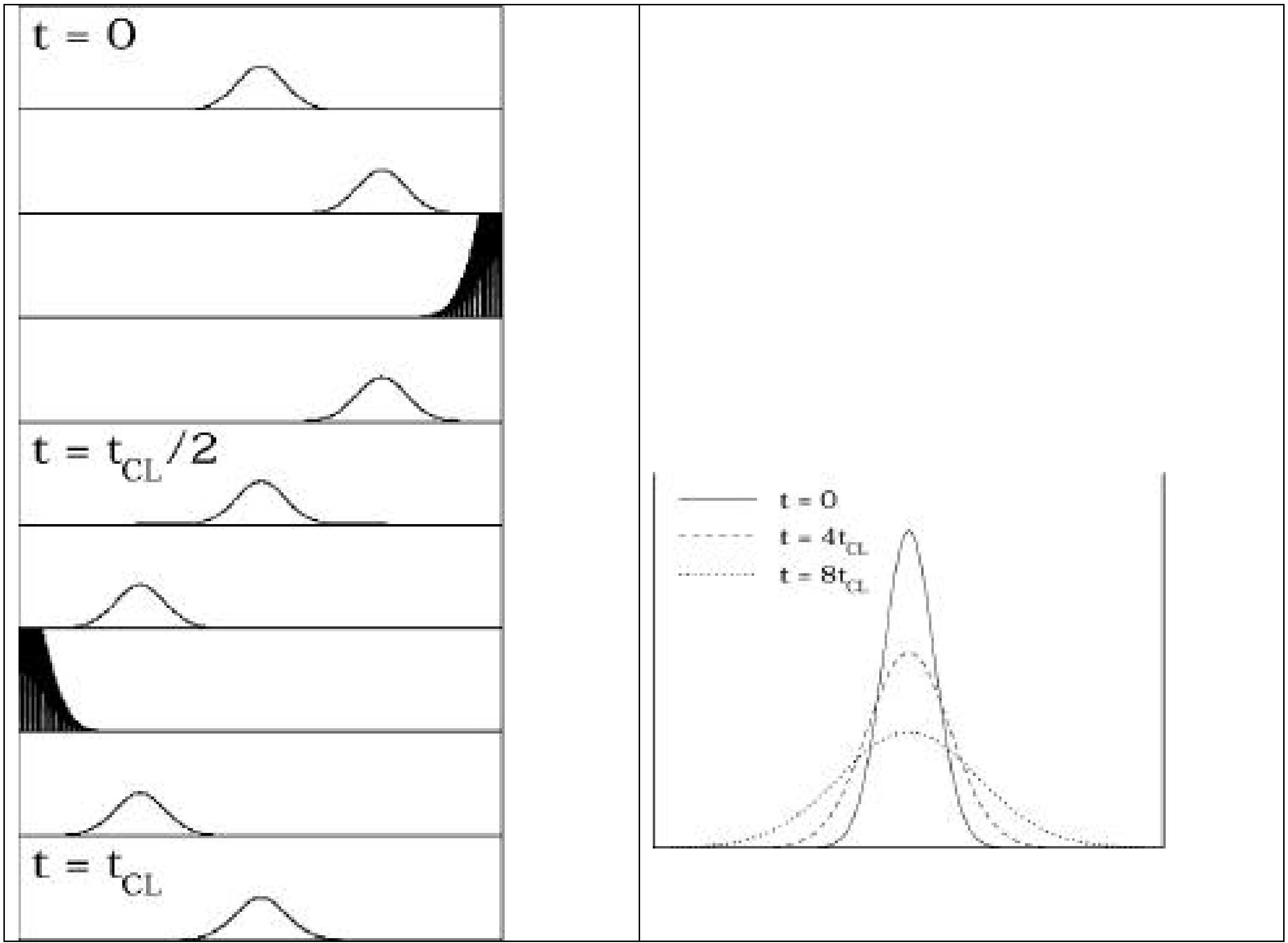
(d) Continue to bounce, spread and flatten, approaching a **non-zero**, constant value of $|\psi(x)|^2$ over the entire well for all later times.

(e) Continue to bounce, spread and flatten, approaching a **vanishing** value of $|\psi(x)|^2$ over the entire well for all later times.

↑ Explain your answer in 1-2 sentences in the space above. ↑

↓ Circle the statement below which describes how you feel about your answer. ↓

Very	**somewhat**	**somewhat**	**very**
certain	**certain**	**uncertain**	**uncertain**



Problem 24. Plots of $|\psi(x,t)|^2$ versus *x* for a wave packet in an infinite well. Plots over various times during the first classical period (left) and after 4 and 8 periods (right) are shown.

APPENDIX E: ANOVA ON COURSES' MEAN SCORE BY MAIN TOPICS

The one-way analysis of variance (ANOVA) was used to determine whether or not there were statistical differences among the four courses' QMVI mean scores at a 0.05 alpha level. Table 5-2.1 is the analysis of variance for the QMVI mean scores. The F values of 14.34 for Main Topic A, 16.42 for Main Topic B, 22.39 for Main Topic C, 12.07 for Main Topic D, and 6.23 for Main Topic E are significant at the level of 0.05. This means that there is at least one QMVI *course* mean score that is statistically significant different among the four distinct groups.

Table E.1 ANOVA by Course

Main Topic		Sum of Squares	df	Mean Square	F	Sig.
A	Between Groups	4.214	3	1.405	14.34	.000
	Within Groups	13.913	142	9.798E-02		
	Total	18.127	145			
B	Between Groups	2.482	3	.827	16.42	.000
	Within Groups	7.153	142	5.037E-02		
	Total	9.635	145			
C	Between Groups	2.017	3	.672	22.39	.000
	Within Groups	4.263	142	3.002E-02		
	Total	6.280	145			
D	Between Groups	1.613	3	.538	12.07	.000
	Within Groups	6.326	142	4.455E-02		
	Total	7.939	145			
E	Between Groups	1.073	3	.358	6.23	.001
	Within Groups	8.107	142	5.709E-02		
	Total	9.180	145			

Table E.2 Follow-up Tukey HSD test

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
A	ModPh	UgQM	-.3339	6.006E-02	.000	-.4882	-.1796
		GrQM	-.3900	8.452E-02	.000	-.6071	-.1729
		ChemQM	-4.65E-02	9.153E-02	.957	-.2817	.1886
	UgQM	ModPh	.3339	6.006E-02	.000	.1796	.4882
		GrQM	-5.61E-02	8.939E-02	.923	-.2858	.1735
		ChemQM	.2873	9.605E-02	.015	4.059E-02	.5341
	GrQM	ModPh	.3900	8.452E-02	.000	.1729	.6071
		UgQM	5.615E-02	8.939E-02	.923	-.1735	.2858
		ChemQM	.3435	.1130	.013	5.327E-02	.6337
	ChemQM	ModPh	4.653E-02	9.153E-02	.957	-.1886	.2817
		UgQM	-.2873	9.605E-02	.015	-.5341	-4.05E-02
		GrQM	-.3435	.1130	.013	-.6337	-5.3272E-02

Table E.2 Continued

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
B	ModPh	UgQM	-.2213	4.306E-02	.000	-.3319	-.1106
		GrQM	-.3453	6.060E-02	.000	-.5010	-.1896
		ChemQM	-1.0865E-02	6.563E-02	.998	-.1795	.1577
	UgQM	ModPh	.2213	4.306E-02	.000	.1106	.3319
		GrQM	-.1241	6.409E-02	.213	-.2887	4.059E-02
		ChemQM	.2104	6.887E-02	.012	3.346E-02	.3873
	GrQM	ModPh	.3453	6.060E-02	.000	.1896	.5010
		UgQM	.1241	6.409E-02	.213	-4.0594E-02	.2887
		ChemQM	.3345	8.100E-02	.000	.1264	.5425
	ChemQM	ModPh	1.087E-02	6.563E-02	.998	-.1577	.1795
		UgQM	-.2104	6.887E-02	.012	-.3873	-3.3463E-02
		GrQM	-.3345	8.100E-02	.000	-.5425	-.1264
C	ModPh	UgQM	-.2264	3.324E-02	.000	-.3118	-.1410
		GrQM	-.2601	4.678E-02	.000	-.3803	-.1400
		ChemQM	9.809E-03	5.067E-02	.997	-.1204	.1400
	UgQM	ModPh	.2264	3.324E-02	.000	.1410	.3118
		GrQM	-3.3757E-02	4.948E-02	.904	-.1609	9.336E-02
		ChemQM	.2362	5.317E-02	.000	9.961E-02	.3728
	GrQM	ModPh	.2601	4.678E-02	.000	.1400	.3803
		UgQM	3.376E-02	4.948E-02	.904	-9.3360E-02	.1609
		ChemQM	.2700	6.253E-02	.000	.1093	.4306
	ChemQM	ModPh	-9.8089E-03	5.067E-02	.997	-.1400	.1204
		UgQM	-.2362	5.317E-02	.000	-.3728	-9.9614E-02
		GrQM	-.2700	6.253E-02	.000	-.4306	-.1093
D	ModPh	UgQM	-.1977	4.050E-02	.000	-.3017	-9.3666E-02
		GrQM	-.2366	5.699E-02	.000	-.3831	-9.0226E-02
		ChemQM	1.861E-02	6.172E-02	.990	-.1400	.1772
	UgQM	ModPh	.1977	4.050E-02	.000	9.367E-02	.3017
		GrQM	-3.8937E-02	6.028E-02	.917	-.1938	.1159
		ChemQM	.2163	6.477E-02	.005	4.993E-02	.3827
	GrQM	ModPh	.2366	5.699E-02	.000	9.023E-02	.3831
		UgQM	3.894E-02	6.028E-02	.917	-.1159	.1938
		ChemQM	.2553	7.618E-02	.004	5.955E-02	.4510
	ChemQM	ModPh	-1.8612E-02	6.172E-02	.990	-.1772	.1400
		UgQM	-.2163	6.477E-02	.005	-.3827	-4.9928E-02
		GrQM	-.2553	7.618E-02	.004	-.4510	-5.9553E-02
E	ModPh	UgQM	-8.0218E-02	4.584E-02	.298	-.1980	3.756E-02
		GrQM	-.2741	6.452E-02	.000	-.4398	-.1083
		ChemQM	-2.1127E-02	6.987E-02	.990	-.2006	.1584
	UgQM	ModPh	8.022E-02	4.584E-02	.298	-3.7556E-02	.1980
		GrQM	-.1939	6.823E-02	.023	-.3691	-1.8555E-02
		ChemQM	5.909E-02	7.332E-02	.852	-.1293	.2474
	GrQM	ModPh	.2741	6.452E-02	.000	.1083	.4398
		UgQM	.1939	6.823E-02	.023	1.855E-02	.3691
		ChemQM	.2529	8.623E-02	.018	3.140E-02	.4745
	ChemQM	ModPh	2.113E-02	6.987E-02	.990	-.1584	.2006
		UgQM	-5.9091E-02	7.332E-02	.852	-.2474	.1293
		GrQM	-.2529	8.623E-02	.018	-.4745	-3.1403E-02

* The mean difference is significant at the .05 level.

APPENDIX F: ANOVA ON SECTION MEAN SCORES BY MAIN TOPICS

The one-way analysis of variance (ANOVA) was used to determine whether or not there were statistical differences among the 8 section mean scores when analyzed with respect to the 5 Main Topics at a 0.05 alpha level. Table 5-3.1 is the analysis of variance for the QMVI mean scores. The F values of 5.58 for Main Topic A, 7.99 for Main Topic B, 9.10 for Main Topic C, 5.68 for Main Topic D, and 5.00 for Main Topic E are significant at the level of 0.05. This means that there is at least one QMVI *course* mean score that is statistically significant different among the four distinct groups.

Table F.1 Result of ANOVA by Section

Main Topic		Sum of Squares	df	Mean Square	F	Sig.
A	Between Groups	4.457	8	.557	5.583	.000
	Within Groups	13.670	137	9.978E-02		
	Total	18.127	145			
B	Between Groups	3.068	8	.383	7.999	.000
	Within Groups	6.567	137	4.794E-02		
	Total	9.635	145			
C	Between Groups	2.179	8	.272	9.096	.000
	Within Groups	4.101	137	2.994E-02		
	Total	6.280	145			
D	Between Groups	1.976	8	.247	5.676	.000
	Within Groups	5.963	137	4.352E-02		
	Total	7.939	145			
E	Between Groups	2.076	8	.259	5.003	.000
	Within Groups	7.105	137	5.186E-02		
	Total	9.180	145			

Table F2 Follow-up Tukey HSD test

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) Course				Lower Bound	Upper Bound
A	4	5	-.3460	9.8E-02	.014	-.6530	-3.89E-02
		6	-1.0913E-02	.1210	1.000	-.3863	.3645
		7	-.2975	9.4E-02	.041	-.5887	-6.24E-03
		8	-.3592	.1060	.020	-.6880	-3.04E-02
		9	-8.9286E-03	.1034	1.000	-.3296	.3118
		10	8.141E-02	8.0E-02	.985	-.1686	.3314
		11	-.2054	.1210	.749	-.5808	.1701
		12	-.3304	.1688	.574	-.8541	.1934
	5	4	.3460	9.8E-02	.014	3.893E-02	.6530
		6	.3351	.1316	.210	-7.3172E-02	.7433
		7	4.852E-02	.1072	1.000	-.2839	.3810
		8	-1.3221E-02	.1179	1.000	-.3791	.3526
		9	.3371	.1156	.085	-2.1509E-02	.6956
		10	.4274	9.5E-02	.000	.1304	.7244
		11	.1406	.1316	.979	-.2676	.5489
		12	1.563E-02	.1766	1.000	-.5321	.5633
	6	4	1.091E-02	.1210	1.000	-.3645	.3863
		5	-.3351	.1316	.210	-.7433	7.317E-02
		7	-.2865	.1278	.378	-.6830	.1099
		8	-.3483	.1370	.212	-.7732	7.657E-02
		9	1.984E-03	.1350	1.000	-.4166	.4206
		10	9.232E-02	.1184	.997	-.2750	.4596
		11	-.1944	.1489	.930	-.6563	.2674
		12	-.3194	.1898	.757	-.9082	.2693

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
A	7	4	.2975	9.4E-02	.041	6.242E-03	.5887
		5	-4.8520E-02	.1072	1.000	-.3810	.2839
		6	.2865	.1278	.378	-.1099	.6830
		8	-6.1741E-02	.1137	1.000	-.4144	.2909
		9	.2885	.1113	.189	-5.6566E-02	.6336
		10	.3789	9.1E-02	.001	9.823E-02	.6595
		11	9.211E-02	.1278	.999	-.3044	.4886
		12	-3.2895E-02	.1738	1.000	-.5719	.5061
	8	4	.3592	.1060	.020	3.037E-02	.6880
		5	1.322E-02	.1179	1.000	-.3526	.3791
		6	.3483	.1370	.212	-7.6570E-02	.7732
		7	6.174E-02	.1137	1.000	-.2909	.4144
		9	.3503	.1217	.094	-2.7101E-02	.7277
		10	.4406	.1030	.001	.1211	.7601
		11	.1538	.1370	.971	-.2710	.5787
		12	2.885E-02	.1806	1.000	-.5314	.5891
	9	4	8.929E-03	.1034	1.000	-.3118	.3296
		5	-.3371	.1156	.085	-.6956	2.151E-02
		6	-1.9841E-03	.1350	1.000	-.4206	.4166
		7	-.2885	.1113	.189	-.6336	5.657E-02
		8	-.3503	.1217	.094	-.7277	2.710E-02
		10	9.034E-02	.1003	.993	-.2208	.4015
		11	-.1964	.1350	.876	-.6150	.2222
		12	-.3214	.1791	.686	-.8769	.2341
	10	4	-8.1408E-02	8.1E-02	.985	-.3314	.1686
		5	-.4274	9.6E-02	.000	-.7244	-.1304
		6	-9.2320E-02	.1184	.997	-.4596	.2750
		7	-.3789	9.1E-02	.001	-.6595	-9.8E-02
		8	-.4406	.1030	.001	-.7601	-.1211
		9	-9.0336E-02	.1003	.993	-.4015	.2208
		11	-.2868	.1184	.272	-.6540	8.052E-02
		12	-.4118	.1670	.249	-.9297	.1061
	11	4	.2054	.1210	.749	-.1701	.5808
		5	-.1406	.1316	.979	-.5489	.2676
		6	.1944	.1489	.930	-.2674	.6563
		7	-9.2105E-02	.1278	.999	-.4886	.3044
		8	-.1538	.1370	.971	-.5787	.2710
		9	.1964	.1350	.876	-.2222	.6150
		10	.2868	.1184	.272	-8.0519E-02	.6540
		12	-.1250	.1898	.999	-.7138	.4638
	12	4	.3304	.1688	.574	-.1934	.8541
		5	-1.5625E-02	.1766	1.000	-.5633	.5321
		6	.3194	.1898	.757	-.2693	.9082
		7	3.289E-02	.1738	1.000	-.5061	.5719
		8	-2.8846E-02	.1806	1.000	-.5891	.5314
		9	.3214	.1791	.686	-.2341	.8769
		10	.4118	.1670	.249	-.1061	.9297
		11	.1250	.1898	.999	-.4638	.7138

Table F.2 Continued

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
B	4	5	-.2268	6.8E-02	.027	-.4396	-1.4E-02
		6	1.349E-02	8.4E-02	1.000	-.2467	.2737
		7	-9.0602E-02	6.5E-02	.901	-.2925	.1113
		8	-.2874	7.3E-02	.003	-.5153	-5.9E-02
		9	6.429E-02	7.2E-02	.993	-.1580	.2866
		10	.1534	5.6E-02	.132	-1.9947E-02	.3267
		11	-.1198	8.4E-02	.887	-.3801	.1404
		12	-.2143	.1170	.661	-.5773	.1487
	5	4	.2268	6.8E-02	.027	1.396E-02	.4396
		6	.2403	9.1E-02	.173	-4.2686E-02	.5232
		7	.1362	7.4E-02	.660	-9.4246E-02	.3666
		8	-6.0577E-02	8.1E-02	.998	-.3142	.1930
		9	.2911	8.1E-02	.009	4.254E-02	.5396
		10	.3801	6.6E-02	.000	.1743	.5860
		11	.1069	9.1E-02	.962	-.1760	.3899
		12	1.250E-02	.1224	1.000	-.3671	.3921
	6	4	-1.3492E-02	8.4E-02	1.000	-.2737	.2467
		5	-.2403	9.1E-02	.173	-.5232	4.269E-02
		7	-.1041	8.8E-02	.962	-.3789	.1707
		8	-.3009	9.5E-02	.041	-.5953	-6.4E-03
		9	5.079E-02	9.4E-02	1.000	-.2394	.3409
		10	.1399	8.2E-02	.744	-.1147	.3944
		11	-.1333	.1032	.934	-.4535	.1868
		12	-.2278	.1316	.727	-.6359	.1803
	7	4	9.060E-02	6.5E-02	.901	-.1113	.2925
		5	-.1362	7.4E-02	.660	-.3666	9.425E-02
		6	.1041	8.8E-02	.962	-.1707	.3789
		8	-.1968	7.8E-02	.233	-.4412	4.768E-02
		9	.1549	7.7E-02	.537	-8.4312E-02	.3941
		10	.2440	6.3E-02	.003	4.944E-02	.4385
		11	-2.9240E-02	8.9E-02	1.000	-.3040	.2456
		12	-.1237	.1204	.983	-.4973	.2499
	8	4	.2874	7.3E-02	.003	5.944E-02	.5153
		5	6.058E-02	8.2E-02	.998	-.1930	.3142
		6	.3009	9.4E-02	.041	6.371E-03	.5953
		7	.1968	7.9E-02	.233	-4.7677E-02	.4412
		9	.3516	8.4E-02	.001	9.008E-02	.6132
		10	.4407	7.1E-02	.000	.2193	.6622
		11	.1675	9.5E-02	.706	-.1270	.4620
		12	7.308E-02	.1252	1.000	-.3152	.4614
	9	4	-6.4286E-02	7.2E-02	.993	-.2866	.1580
		5	-.2911	8.0E-02	.009	-.5396	-4.3E-02
		6	-5.0794E-02	9.4E-02	1.000	-.3409	.2394
		7	-.1549	7.7E-02	.537	-.3941	8.431E-02
		8	-.3516	8.4E-02	.001	-.6132	-9.0E-02
		10	8.908E-02	6.9E-02	.937	-.1266	.3047
		11	-.1841	9.354E-02	.566	-.4743	.1060
		12	-.2786	.1241	.377	-.6636	.1065

Table F.2 Continued

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
	10	4	-.1534	5.5E-02	.132	-.3267	1.995E-02
		5	-.3801	6.6E-02	.000	-.5860	-.1743
		6	-.1399	8.2E-02	.744	-.3944	.1147
		7	-.2440	6.2E-02	.003	-.4385	-4.9E-02
		8	-.4407	7.1E-02	.000	-.6622	-.2193
		9	-8.9076E-02	6.9E-02	.937	-.3047	.1266
		11	-.2732	8.2E-02	.025	-.5278	-1.8E-02
		12	-.3676	.1157	.040	-.7266	-8.7E-03
	11	4	.1198	8.4E-02	.887	-.1404	.3801
		5	-.1069	9.1E-02	.962	-.3899	.1760
		6	.1333	.1032	.934	-.1868	.4535
		7	2.924E-02	8.9E-02	1.000	-.2456	.3040
		8	-.1675	9.5E-02	.706	-.4620	.1270
		9	.1841	9.4E-02	.566	-.1060	.4743
		10	.2732	8.3E-02	.025	1.863E-02	.5278
		12	-9.4444E-02	.1316	.999	-.5025	.3137
	12	4	.2143	.1170	.661	-.1487	.5773
		5	-1.2500E-02	.1224	1.000	-.3921	.3671
		6	.2278	.1316	.727	-.1803	.6359
		7	.1237	.1204	.983	-.2499	.4973
		8	-7.3077E-02	.1252	1.000	-.4614	.3152
		9	.2786	.1241	.377	-.1065	.6636
		10	.3676	.1157	.040	8.671E-03	.7266
		11	9.444E-02	.1316	.999	-.3137	.5025
C	4	5	-.2310	5.4E-02	.001	-.3992	-6.3E-02
		6	1.984E-02	6.6E-02	1.000	-.1858	.2255
		7	-.2068	5.1E-02	.002	-.3663	-4.7E-02
		8	-.1992	5.8E-02	.018	-.3793	-1.9E-02
		9	3.571E-02	5.6E-02	.999	-.1400	.2114
		10	4.884E-02	4.4E-02	.974	-8.8E-02	.1858
		11	-.1329	6.6E-02	.540	-.3386	7.2E-02
		12	-.3482	9.2E-02	.005	-.6351	-6.1E-02
	5	4	.2310	5.4E-02	.001	6.284E-02	.3992
		6	.2509	7.2E-02	.015	2.725E-02	.4745
		7	2.426E-02	5.8E-02	1.000	-.1578	.2064
		8	3.185E-02	6.4E-02	1.000	-.1685	.2322
		9	.2667	6.3E-02	.001	7.034E-02	.4631
		10	.2799	5.2E-02	.000	.1172	.4426
		11	9.809E-02	7.2E-02	.912	-.1255	.3217
		12	-.1172	9.6E-02	.954	-.4172	.1828
	6	4	-1.9841E-02	6.6E-02	1.000	-.2255	.1858
		5	-.2509	7.2E-02	.015	-.4745	-2.7E-02
		7	-.2266	7.0E-02	.033	-.4438	-9.4E-03
		8	-.2190	7.5E-02	.084	-.4517	1.370E-02
		9	1.587E-02	7.4E-02	1.000	-.2134	.2452
		10	2.900E-02	6.5E-02	1.000	-.1722	.2302
		11	-.1528	8.2E-02	.632	-.4058	.1002
		12	-.3681	.1040	.012	-.6906	-4.6E-02

Table F.2 Continued

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
	7	4	.2068	5.1E-02	.002	4.725E-02	.3663
		5	-2.4260E-02	5.9E-02	1.000	-.2064	.1578
		6	.2266	7.001E-02	.033	9.443E-03	.4438
		8	7.591E-03	6.228E-02	1.000	-.1856	.2008
		9	.2425	6.1E-02	.002	5.345E-02	.4315
		10	.2556	5.0E-02	.000	.1019	.4093
		11	7.383E-02	7.0E-02	.980	-.1433	.2910
		12	-.1414	9.5E-02	.862	-.4367	.1538
	8	4	.1992	5.8E-02	.018	1.906E-02	.3793
		5	-3.1851E-02	6.5E-02	1.000	-.2322	.1685
		6	.2190	7.5E-02	.084	-1.3700E-02	.4517
		7	-7.5911E-03	6.2E-02	1.000	-.2008	.1856
		9	.2349	6.7E-02	.013	2.818E-02	.4416
		10	.2480	5.6E-02	.000	7.302E-02	.4230
		11	6.624E-02	7.5E-02	.994	-.1665	.2990
		12	-.1490	9.9E-02	.853	-.4559	.1578
	9	4	-3.5714E-02	5.7E-02	.999	-.2114	.1400
		5	-.2667	6.3E-02	.001	-.4631	-7.0E-02
		6	-1.5873E-02	7.4E-02	1.000	-.2452	.2134
		7	-.2425	6.1E-02	.002	-.4315	-5.3E-02
		8	-.2349	6.7E-02	.013	-.4416	-2.8E-02
		10	1.313E-02	5.5E-02	1.000	-.1573	.1836
		11	-.1687	7.4E-02	.353	-.3979	6.064E-02
		12	-.3839	9.8E-02	.003	-.6882	-8.0E-02
	10	4	-4.8845E-02	4.4E-02	.974	-.1858	8.8E-02
		5	-.2799	5.2E-02	.000	-.4426	-.1172
		6	-2.9003E-02	6.5E-02	1.000	-.2302	.1722
		7	-.2556	5.0E-02	.000	-.4093	-.1019
		8	-.2480	5.6E-02	.000	-.4230	-7.3E-02
		9	-1.3130E-02	5.5E-02	1.000	-.1836	.1573
		11	-.1818	6.4E-02	.114	-.3830	1.940E-02
		12	-.3971	9.1E-02	.000	-.6807	-.1134
	11	4	.1329	6.6E-02	.540	-7.2705E-02	.3386
		5	-9.8090E-02	7.2E-02	.912	-.3217	.1255
		6	.1528	8.1E-02	.632	-.1002	.4058
		7	-7.3830E-02	7.0E-02	.980	-.2910	.1433
		8	-6.6239E-02	7.5E-02	.994	-.2990	.1665
		9	.1687	7.3E-02	.353	-6.0641E-02	.3979
		10	.1818	6.4E-02	.114	-1.9398E-02	.3830
		12	-.2153	.1040	.494	-.5378	.1072
	12	4	.3482	9.2E-02	.005	6.135E-02	.6351
		5	.1172	9.6E-02	.954	-.1828	.4172
		6	.3681	.1040	.012	4.556E-02	.6906
		7	.1414	9.5E-02	.862	-.1538	.4367
		8	.1490	9.8E-02	.853	-.1578	.4559
		9	.3839	9.8E-02	.003	7.966E-02	.6882
		10	.3971	9.1E-02	.000	.1134	.6807
		11	.2153	.1040	.494	-.1072	.5378

Table F.2 Continued

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
D	4	5	-.1696	6.5E-02	.189	-.3724	3.315E-02
		6	3.175E-02	7.9E-02	1.000	-.2162	.2797
		7	-9.3985E-02	6.2E-02	.849	-.2863	9.835E-02
		8	-.1648	7.0E-02	.310	-.3820	5.234E-02
		9	8.929E-02	6.8E-02	.929	-.1225	.3011
		10	.1392	5.3E-02	.180	-2.5959E-02	.3043
		11	-.1210	7.9E-02	.849	-.3690	.1269
		12	-.1696	.1115	.846	-.5155	.1762
	5	4	.1696	6.5E-02	.189	-3.3153E-02	.3724
		6	.2014	8.6E-02	.332	-6.8237E-02	.4710
		7	7.566E-02	7.0E-02	.979	-.1439	.2952
		8	4.808E-03	7.7E-02	1.000	-.2368	.2464
		9	.2589	7.6E-02	.020	2.211E-02	.4957
		10	.3088	6.3E-02	.000	.1126	.5050
		11	4.861E-02	8.6E-02	1.000	-.2210	.3182
		12	.0000	.1166	1.000	-.3617	.3617
	6	4	-3.1746E-02	7.9E-02	1.000	-.2797	.2162
		5	-.2014	8.6E-02	.332	-.4710	6.824E-02
		7	-.1257	8.4E-02	.861	-.3876	.1361
		8	-.1966	9.0E-02	.423	-.4772	8.402E-02
		9	5.754E-02	8.9E-02	.999	-.2189	.3340
		10	.1074	7.8E-02	.908	-.1351	.3500
		11	-.1528	9.8E-02	.830	-.4578	.1523
		12	-.2014	.1254	.802	-.5902	.1875
	7	4	9.398E-02	6.2E-02	.849	-9.8353E-02	.2863
		5	-7.5658E-02	7.1E-02	.979	-.2952	.1439
		6	.1257	8.4E-02	.861	-.1361	.3876
		8	-7.0850E-02	7.5E-02	.990	-.3038	.1621
		9	.1833	7.3E-02	.235	-4.4653E-02	.4112
		10	.2332	5.9E-02	.003	4.781E-02	.4185
		11	-2.7047E-02	8.4E-02	1.000	-.2889	.2348
		12	-7.5658E-02	.1148	.999	-.4316	.2803
	8	4	.1648	7.0E-02	.310	-5.2342E-02	.3820
		5	-4.8077E-03	7.790E-02	1.000	-.2464	.2368
		6	.1966	9.0E-02	.423	-8.4021E-02	.4772
		7	7.085E-02	7.5E-02	.990	-.1621	.3038
		9	.2541	8.0E-02	.042	4.880E-03	.5034
		10	.3040	6.8E-02	.000	9.300E-02	.5150
		11	4.380E-02	9.0E-02	1.000	-.2368	.3244
		12	-4.8077E-03	.1193	1.000	-.3748	.3652
	9	4	-8.9286E-02	6.8E-02	.929	-.3011	.1225
		5	-.2589	7.6E-02	.020	-.4957	-2.2113E-02
		6	-5.7540E-02	8.9E-02	.999	-.3340	.2189
		7	-.1833	7.3E-02	.235	-.4112	4.465E-02
		8	-.2541	8.0E-02	.042	-.5034	-4.8802E-03
		10	4.989E-02	6.6E-02	.998	-.1556	.2554
		11	-.2103	8.9E-02	.307	-.4868	6.615E-02
		12	-.2589	.1183	.413	-.6258	.1079

Table E.2 Continued

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
	10	4	-.1392	5.3E-02	.180	-.3043	2.596E-02
		5	-.3088	6.3E-02	.000	-.5050	-.1126
		6	-.1074	7.8E-02	.908	-.3500	.1351
		7	-.2332	5.9E-02	.003	-.4185	-4.78E-02
		8	-.3040	6.8E-02	.000	-.5150	-9.30E-02
		9	-4.9895E-02	6.6E-02	.998	-.2554	.1556
		11	-.2602	7.8E-02	.025	-.5028	-1.7637E-02
		12	-.3088	.1103	.115	-.6509	3.323E-02
	11	4	.1210	7.9E-02	.849	-.1269	.3690
		5	-4.8611E-02	8.6E-02	1.000	-.3182	.2210
		6	.1528	9.8E-02	.830	-.1523	.4578
		7	2.705E-02	8.4E-02	1.000	-.2348	.2889
		8	-4.3803E-02	9.0E-02	1.000	-.3244	.2368
		9	.2103	8.9E-02	.307	-6.6155E-02	.4868
		10	.2602	7.8E-02	.025	1.764E-02	.5028
		12	-4.8611E-02	.1254	1.000	-.4375	.3402
	12	4	.1696	.1115	.846	-.1762	.5155
		5	.0000	.1166	1.000	-.3617	.3617
		6	.2014	.1254	.802	-.1875	.5902
		7	7.566E-02	.1148	.999	-.2803	.4316
		8	4.808E-03	.1193	1.000	-.3652	.3748
		9	.2589	.1183	.413	-.1079	.6258
		10	.3088	.1103	.115	-3.3231E-02	.6509
		11	4.861E-02	.1254	1.000	-.3402	.4375
E	4	5	-.1411	7.1E-02	.560	-.3624	8.029E-02
		6	9.921E-02	8.7E-02	.969	-.1714	.3699
		7	.1214	6.8E-02	.686	-8.8516E-02	.3314
		8	-.2016	7.6E-02	.171	-.4387	3.541E-02
		9	2.143E-02	7.4E-02	1.000	-.2098	.2526
		10	6.261E-02	5.8E-02	.978	-.1177	.2429
		11	-.1897	8.7E-02	.423	-.4603	8.097E-02
		12	-.3286	.1217	.148	-.7061	4.898E-02
	5	4	.1411	7.1E-02	.560	-8.0288E-02	.3624
		6	.2403	9.4E-02	.216	-5.4029E-02	.5346
		7	.2625	7.7E-02	.020	2.283E-02	.5022
		8	-6.0577E-02	8.5E-02	.999	-.3243	.2032
		9	.1625	8.3E-02	.579	-9.5992E-02	.4210
		10	.2037	6.9E-02	.077	-1.0463E-02	.4178
		11	-4.8611E-02	9.4E-02	1.000	-.3429	.2457
		12	-.1875	.1273	.868	-.5824	.2074
	6	4	-9.9206E-02	8.7E-02	.969	-.3699	.1714
		5	-.2403	9.4E-02	.216	-.5346	5.403E-02
		7	2.222E-02	9.2E-02	1.000	-.2636	.3080
		8	-.3009	9.8E-02	.059	-.6071	5.433E-03
		9	-7.7778E-02	9.7E-02	.997	-.3796	.2240
		10	-3.6601E-02	8.5E-02	1.000	-.3014	.2282
		11	-.2889	.1073	.151	-.6219	4.408E-02
		12	-.4278	.1368	.046	-.8522	-3.3228E-03

Table E.2 Continued

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
Main Topic	(I) Course	(J) course				Lower Bound	Upper Bound
	7	4	-.1214	6.8E-02	.686	-.3314	8.852E-02
		5	-.2625	7.7E-02	.020	-.5022	-2.2833E-02
		6	-2.2222E-02	9.2E-02	1.000	-.3080	.2636
		8	-.3231	8.2E-02	.003	-.5773	-6.8841E-02
		9	-1.0000E-01	8.0E-02	.946	-.3488	.1488
		10	-5.8824E-02	6.5E-02	.993	-.2611	.1435
		11	-.3111	9.2E-02	.021	-.5969	-2.5291E-02
		12	-.4500	.1253	.010	-.8386	-6.1431E-02
	8	4	.2016	7.6E-02	.171	-3.5408E-02	.4387
		5	6.058E-02	8.5E-02	.999	-.2032	.3243
		6	.3009	9.8E-02	.059	-5.4330E-03	.6071
		7	.3231	8.1E-02	.003	6.884E-02	.5773
		9	.2231	8.7E-02	.211	-4.8979E-02	.4951
		10	.2643	7.4E-02	.011	3.392E-02	.4946
		11	1.197E-02	9.8E-02	1.000	-.2943	.3183
		12	-.1269	.1302	.988	-.5308	.2769
	9	4	-2.1429E-02	7.4E-02	1.000	-.2526	.2098
		5	-.1625	8.3E-02	.579	-.4210	9.599E-02
		6	7.778E-02	9.7E-02	.997	-.2240	.3796
		7	1.000E-01	8.0E-02	.946	-.1488	.3488
		8	-.2231	8.7E-02	.211	-.4951	4.898E-02
		10	4.118E-02	7.2E-02	1.000	-.1831	.2655
		11	-.2111	9.7E-02	.426	-.5129	9.067E-02
		12	-.3500	.1291	.144	-.7505	5.045E-02
	10	4	-6.2605E-02	5.8E-02	.978	-.2429	.1177
		5	-.2037	6.9E-02	.077	-.4178	1.046E-02
		6	3.660E-02	8.5E-02	1.000	-.2282	.3014
		7	5.882E-02	6.5E-02	.993	-.1435	.2611
		8	-.2643	7.4E-02	.011	-.4946	-3.3924E-02
		9	-4.1176E-02	7.2E-02	1.000	-.2655	.1831
		11	-.2523	8.5E-02	.076	-.5171	1.249E-02
		12	-.3912	.1204	.032	-.7645	-1.7812E-02
	11	4	.1897	8.7E-02	.423	-8.0970E-02	.4603
		5	4.861E-02	9.4E-02	1.000	-.2457	.3429
		6	.2889	.1073	.151	-4.4081E-02	.6219
		7	.3111	9.2E-02	.021	2.529E-02	.5969
		8	-1.1966E-02	9.8E-02	1.000	-.3183	.2943
		9	.2111	9.7E-02	.426	-9.0668E-02	.5129
		10	.2523	8.5E-02	.076	-1.2492E-02	.5171
		12	-.1389	.1368	.985	-.5633	.2856
	12	4	.3286	.1217	.148	-4.8981E-02	.7061
		5	.1875	.1273	.868	-.2074	.5824
		6	.4278	.1368	.046	3.323E-03	.8522
		7	.4500	.1253	.010	6.143E-02	.8386
		8	.1269	.1302	.988	-.2769	.5308
		9	.3500	.1291	.144	-5.0455E-02	.7505
		10	.3912	.1204	.032	1.781E-02	.7645
		11	.1389	.1368	.985	-.2856	.5633

* The mean difference is significant at the .05 level.

Table F.3 Main Topic A: Descriptive Summary

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Lower Bound	Upper Bound
Q1	ModPh	71	.34	.48	5.65E-02	.23	.45
	UgQM	44	.70	.46	6.96E-02	.56	.84
	GrQM	17	.82	.39	9.53E-02	.62	1.03
	ChemQM	14	.36	.50	.13	7.00E-02	.64
	Total	146	.51	.50	4.15E-02	.42	.59
Q2	ModPh	71	.27	.45	5.29E-02	.16	.37
	UgQM	44	.59	.50	7.50E-02	.44	.74
	GrQM	17	.35	.49	.12	9.97E-02	.61
	ChemQM	14	.14	.36	9.71E-02	-6.68E-02	.35
	Total	146	.36	.48	3.99E-02	.28	.44
Q3	ModPh	71	.31	.47	5.53E-02	.20	.42
	UgQM	44	.59	.50	7.50E-02	.44	.74
	GrQM	17	.76	.44	.11	.54	.99
	ChemQM	14	.50	.52	.14	.20	.80
	Total	146	.47	.50	4.14E-02	.38	.55
Q6	ModPh	71	.11	.32	3.78E-02	3.73E-02	.19
	UgQM	44	.48	.51	7.62E-02	.32	.63
	GrQM	17	.65	.49	.12	.39	.90
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.29	.46	3.79E-02	.22	.37

Table F.4 Result of ANOVA by Question

		Sum of Squares	df	Mean Square	F	Sig.
Q1	Between Groups	5.762	3	1.921	8.875	.000
	Within Groups	30.731	142	.216		
	Total	36.493	145			
Q2	Between Groups	3.612	3	1.204	5.671	.001
	Within Groups	30.148	142	.212		
	Total	33.760	145			
Q3	Between Groups	3.950	3	1.317	5.775	.001
	Within Groups	32.378	142	.228		
	Total	36.329	145			
Q6	Between Groups	6.020	3	UgQM7	11.719	.000
	Within Groups	24.315	142	.171		
	Total	30.336	145			

Table F.5 Follow-up Tukey HSD test

Question 1	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			
GrQM	*			*
ChemQM			*	
Question 2	ModPh	UgQM	GrQM	ChemQM
ModPh		*		
UgQM	*			*
GrQM				
ChemQM		*		

Question 3	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			
GrQM	*			
ChemQM				

Question 6	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			
GrQM	*			*
ChemQM			*	

Table F.5 Main Topic B: Descriptive Summary

		N	Mean	Std. Dev	Std. Error	95% Confidence Interval for Mean	
						Lower Bound	Upper Bound
Q9	ModPh	71	8.45E-02	.28	3.32E-02	1.82E-02	.15
	UgQM	44	.32	.47	7.10E-02	.17	.46
	GrQM	17	.41	.51	.12	.15	.67
	ChemQM	14	.36	.50	.13	7.00E-02	.64
	Total	146	.22	.42	3.44E-02	.15	.29
Q14	ModPh	71	.23	.42	4.99E-02	.13	.32
	UgQM	44	.48	.51	7.62E-02	.32	.63
	GrQM	17	.65	.49	.12	.39	.90
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.35	.48	3.96E-02	.27	.43
Q15	ModPh	71	.18	.39	4.62E-02	9.09E-02	.28
	UgQM	44	.52	.51	7.62E-02	.37	.68
	GrQM	17	.53	.51	.12	.26	.79
	ChemQM	14	7.14E-02	.27	7.14E-02	-8.29E-02	.23
	Total	146	.32	.47	3.86E-02	.24	.39
Q16	ModPh	71	.13	.34	3.98E-02	4.75E-02	.21
	UgQM	44	.34	.48	7.23E-02	.20	.49
	GrQM	17	.82	.39	9.53E-02	.62	1.03
	ChemQM	14	.14	.36	9.71E-02	-6.68E-02	.35
	Total	146	.27	.45	3.70E-02	.20	.35
Q19	ModPh	71	.18	.39	4.62E-02	9.09E-02	.28
	UgQM	44	.25	.44	6.60E-02	.12	.38
	GrQM	17	.12	.33	8.05E-02	-5.31E-02	.29
	ChemQM	14	7.14E-02	.27	7.14E-02	-8.29E-02	.23
	Total	146	.18	.39	3.22E-02	.12	.25

Table F.6 Result of ANOVA by Question

		Sum of Squares	df	Mean Square	F	Sig.
Q9	Between Groups	2.616	3	.872	5.535	.001
	Within Groups	22.370	142	.158		
	Total	24.986	145			
Q14	Between Groups	3.574	3	1.191	5.713	.001
	Within Groups	29.611	142	.209		
	Total	33.185	145			
Q15	Between Groups	4.746	3	1.582	8.394	.000
	Within Groups	26.761	142	.188		
	Total	31.507	145			
Q16	Between Groups	7.111	3	2.370	15.347	.000
	Within Groups	21.930	142	.154		
	Total	29.041	145			
Q19	Between Groups	.444	3	.148	.974	.407
	Within Groups	21.563	142	.152		
	Total	27	145			

Table F.7 Follow-up Tukey HSD test

Question 9	ModPh	UgQM	GrQM	ChemQM
ModPh	*	*		
UgQM	*			
GrQM				
ChemQM				

Question 14	ModPh	UgQM	GrQM	ChemQM
ModPh	*	*		
UgQM	*			
GrQM	*			*
ChemQM			*	

Question 15	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			*
GrQM	*			*
ChemQM		*	*	

Question 16	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*		*	
GrQM	*	*		*
ChemQM			*	

Question 19	ModPh	UgQM	GrQM	ChemQM
ModPh				
UgQM				
GrQM				
ChemQM				

Table F.7 Main Topic C: Descriptive Summary

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Lower Bound	Upper Bound
Q4	ModPh	71	.87	.34	3.98E-02	.79	.95
	UgQM	44	.93	.25	3.84E-02	.85	1.01
	GrQM	17	ModPh	.00	.00	ModPh	ModPh
	ChemQM	14	.93	.27	7.14E-02	.77	1.08
	Total	146	.91	.29	2.37E-02	.86	.96
Q5	ModPh	71	.35	.48	5.71E-02	.24	.47
	UgQM	44	.84	.37	5.58E-02	.73	.95
	GrQM	17	.88	.33	8.05E-02	.71	1.05
	ChemQM	14	.36	.50	.13	7.00E-02	.64
	Total	146	.56	.50	4.12E-02	.48	.64
Q6	ModPh	71	.11	.32	3.78E-02	3.73E-02	.19
	UgQM	44	.48	.51	7.62E-02	.32	.63
	GrQM	17	.65	.49	.12	.39	.90
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.29	.46	3.79E-02	.22	.37
Q7	ModPh	71	.23	.42	4.99E-02	.13	.32
	UgQM	44	.36	.49	7.34E-02	.22	.51
	GrQM	17	.41	.51	.12	.15	.67
	ChemQM	14	7.14E-02	.27	7.14E-02	-8.29E-02	.23
	Total	146	.27	.45	3.70E-02	.20	.35
Q8	ModPh	71	.41	.50	5.88E-02	.29	.53
	UgQM	44	.64	.49	7.34E-02	.49	.78
	GrQM	17	.71	.47	.11	.46	.95
	ChemQM	14	.36	.50	.13	7.00E-02	.64
	Total	146	.51	.50	4.15E-02	.42	.59
Q10	ModPh	71	.18	.39	4.62E-02	9.09E-02	.28
	UgQM	44	9.09E-02	.29	4.38E-02	2.50E-03	.18
	GrQM	17	.12	.33	8.05E-02	-5.31E-02	.29
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.15	.36	2.97E-02	9.20E-02	.21
Q18	ModPh	71	.17	.38	4.48E-02	7.97E-02	.26
	UgQM	44	.73	.45	6.79E-02	.59	.86
	GrQM	17	.71	.47	.11	.46	.95
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.40	.49	4.08E-02	.32	.48
Q19	ModPh	71	.18	.39	4.62E-02	9.09E-02	.28
	UgQM	44	.25	.44	6.60E-02	.12	.38
	GrQM	17	.12	.33	8.05E-02	-5.31E-02	.29
	ChemQM	14	7.14E-02	.27	7.14E-02	-8.29E-02	.23
	Total	146	.18	.39	3.22E-02	.12	.25

Table F.8 Result of ANOVA by Question

		Sum of Squares	df	Mean Square	F	Sig.
Q4	Between Groups	.259	3	8.643E-02	1.060	.368
	Within Groups	11.583	142	8.157E-02		
	Total	11.842	145			
Q5	Between Groups	8.883	3	2.961	15.536	.000
	Within Groups	27.063	142	.191		
	Total	35.945	145			
Q6	Between Groups	6.020	3	UgQM7	11.719	.000
	Within Groups	24.315	142	.171		
	Total	30.336	145			
Q7	Between Groups	1.419	3	.473	2.431	.068
	Within Groups	27.622	142	.195		
	Total	29.041	145			
Q8	Between Groups	2.413	3	.804	3.351	.021
	Within Groups	34.080	142	.240		
	Total	36.493	145			
Q10	Between Groups	.307	3	.102	.791	.501
	Within Groups	18.378	142	.129		
	Total	18.685	145			
Q18	Between Groups	10.572	3	3.524	20.353	.000
	Within Groups	24.586	142	.173		
	Total	35.158	145			
Q19	Between Groups	.444	3	.148	.974	.407
	Within Groups	21.563	142	.152		
	Total	2UgQM7	145			

Table F.9 Follow-up Tukey HSD test

Question 4 no significant difference diff

Question 5	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			*
GrQM	*			*
ChemQM		*	*	

Question 6	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			
GrQM	*			*
ChemQM			*	

Question 7 non –sig follow-up
Question 8 non –sig follow-up
Question 10 non – sig follow-up

Question 18	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			*
GrQM	*			*
ChemQM		*	*	

Question 19 non sig follow-up

Table F.10 Main Topic D: Descriptive Summary

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Lower Bound	Upper Bound
Q10	ModPh	71	.18	.39	4.62E-02	9.09E-02	.28
	UgQM	44	9.09E-02	.29	4.38E-02	2.50E-03	.18
	GrQM	17	.12	.33	8.05E-02	-5.31E-02	.29
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.15	.36	2.97E-02	9.20E-02	.21
Q11	ModPh	71	.41	.50	5.88E-02	.29	.53
	UgQM	44	.59	.50	7.50E-02	.44	.74
	GrQM	17	.76	.44	.11	.54	.99
	ChemQM	14	.29	.47	.13	1.50E-02	.56
	Total	146	.49	.50	4.15E-02	.41	.58
Q12	ModPh	71	.54	.50	5.96E-02	.42	.65
	UgQM	44	.77	.42	6.39E-02	.64	.90
	GrQM	17	.76	.44	.11	.54	.99
	ChemQM	14	.57	.51	.14	.27	.87
	Total	146	.64	.48	3.99E-02	.56	.72
Q13	ModPh	71	.14	.35	4.16E-02	5.79E-02	.22
	UgQM	44	.18	.39	5.88E-02	6.32E-02	.30
	GrQM	17	.35	.49	.12	9.97E-02	.61
	ChemQM	14	7.14E-02	.27	7.14E-02	-8.29E-02	.23
	Total	146	.17	.38	3.13E-02	.11	.23
Q14	ModPh	71	.23	.42	4.99E-02	.13	.32
	UgQM	44	.48	.51	7.62E-02	.32	.63
	GrQM	17	.65	.49	.12	.39	.90
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.35	.48	3.96E-02	.27	.43
Q15	ModPh	71	.18	.39	4.62E-02	9.09E-02	.28
	UgQM	44	.52	.51	7.62E-02	.37	.68
	GrQM	17	.53	.51	.12	.26	.79
	ChemQM	14	7.14E-02	.27	7.14E-02	-8.29E-02	.23
	Total	146	.32	.47	3.86E-02	.24	.39
Q17	ModPh	71	.73	.45	5.29E-02	.63	.84
	UgQM	44	.80	.41	6.15E-02	.67	.92
	GrQM	17	.59	.51	.12	.33	.85
	ChemQM	14	.79	.43	.11	.54	1.03
	Total	146	.74	.44	3.64E-02	.67	.81
Q18	ModPh	71	.17	.38	4.48E-02	7.97E-02	.26
	UgQM	44	.73	.45	6.79E-02	.59	.86
	GrQM	17	.71	.47	.11	.46	.95
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.40	.49	4.08E-02	.32	.48

Table F.11 Result of ANOVA by Question

		Sum of Squares	df	Mean Square	F	Sig.
Q10	Between Groups	.307	3	.102	.791	.501
	Within Groups	18.378	142	.129		
	Total	18.685	145			
Q11	Between Groups	2.786	3	.929	3.912	.010
	Within Groups	33.707	142	.237		
	Total	36.493	145			
Q12	Between Groups	1.884	3	.628	2.797	.042
	Within Groups	31.877	142	.224		
	Total	33.760	145			
Q13	Between Groups	.771	3	.257	1.830	.144
	Within Groups	19.948	142	.140		
	Total	20.719	145			
Q14	Between Groups	3.574	3	1.191	5.713	.001
	Within Groups	29.611	142	.209		
	Total	33.185	145			
Q15	Between Groups	4.746	3	1.582	8.394	.000
	Within Groups	26.761	142	.188		
	Total	31.507	145			
Q17	Between Groups	.560	3	.187	.963	.412
	Within Groups	27.549	142	.194		
	Total	28.110	145			
Q18	Between Groups	10.572	3	3.524	20.353	.000
	Within Groups	24.586	142	.173		
	Total	35.158	145			

Table F.12 Follow-up Tukey HSD test

Question 10 no significant difference

Question 11	ModPh	UgQM	GrQM	ChemQM
ModPh			*	
UgQM				
GrQM	*			*
ChemQM			*	

Question 12	ModPh	UgQM	GrQM	ChemQM
ModPh		*		
UgQM	*			
GrQM				
ChemQM				

Question 13 no significant difference

Question 14	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			
GrQM	*			*
ChemQM			*	

Question 15	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			*
GrQM	*			*
ChemQM		*	*	

Question 17 no significant difference

Question 18	ModPh	UgQM	GrQM	ChemQM
ModPh		*	*	
UgQM	*			*
GrQM	*			*
ChemQM		*	*	

Table F.13 Main Topic E: Descriptive Summary

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Lower Bound	Upper Bound
Q20	ModPh	71	.45	.50	5.95E-02	.33	.57
	UgQM	44	.57	.50	7.55E-02	.42	.72
	GrQM	17	.76	.44	.11	.54	.99
	ChemQM	14	.64	.50	.13	.36	.93
	Total	146	.54	.50	4.14E-02	.46	.62
Q21	ModPh	71	.32	.47	5.59E-02	.21	.44
	UgQM	44	.41	.50	7.50E-02	.26	.56
	GrQM	17	.65	.49	.12	.39	.90
	ChemQM	14	.36	.50	.13	7.00E-02	.64
	Total	146	.39	.49	4.05E-02	.31	.47
Q22	ModPh	71	.37	.49	5.76E-02	.25	.48
	UgQM	44	.36	.49	7.34E-02	.22	.51
	GrQM	17	.59	.51	.12	.33	.85
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.38	.49	4.02E-02	.30	.46
Q23	ModPh	71	.17	.38	4.48E-02	7.97E-02	.26
	UgQM	44	.32	.47	7.10E-02	.17	.46
	GrQM	17	.53	.51	.12	.26	.79
	ChemQM	14	.21	.43	.11	-3.16E-02	.46
	Total	146	.26	.44	3.64E-02	.19	.33
Q24	ModPh	71	8.45E-02	.28	3.32E-02	1.82E-02	.15
	UgQM	44	.14	.35	5.23E-02	3.08E-02	.24
	GrQM	17	.24	.44	.11	1.05E-02	.46
	ChemQM	14	7.14E-02	.27	7.14E-02	-8.29E-02	.23
	Total	146	.12	.32	2.66E-02	6.38E-02	.17

Table F.14 Result of ANOVA by Question

		Sum of Squares	df	Mean Square	F	Sig.
Q20	Between Groups	1.607	3	.536	2.196	.091
	Within Groups	34.646	142	.244		
	Total	36.253	145			
Q21	Between Groups	1.464	3	.488	2.082	.105
	Within Groups	33.282	142	.234		
	Total	34.747	145			
Q22	Between Groups	1.145	3	.382	1.636	.184
	Within Groups	33.135	142	.233		
	Total	34.281	145			
Q23	Between Groups	UgQM0	3	.667	3.625	.015
	Within Groups	26.110	142	.184		
	Total	28.110	145			
Q24	Between Groups	.358	3	.119	1.157	.328
	Within Groups	14.662	142	.103		
	Total	15.021	145			

Table 5-4.3 Follow-up Tukey HSD test

Question 20 no significant difference
Question 21 no significant difference
Question 22 no significant difference

Question 23	ModPh	UgQM	GrQM	ChemQM
ModPh			*	
UgQM				
GrQM	*			
ChemQM				

Question 24 no significant difference

VITA

Place of Birth: Aarau, Kanton Aargau, Switzerland

Date of Birth: April 20, 1968

Education:

1996	MS Science Education, The Middle East Technical University, Turkey
1992	BS Physics Education, The Middle East Technical University, Turkey
1987	Fatih Vatan High School, Istanbul, Turkey
1984	Oberentfelden Secundar Schule, Aargau, Switzerland
1980	Unterentfelden Elementary School, Aargau, Switzerland
1978	Meisterschwaden Elementary School, Aargau, Switzerland

Employment:

1996 – Present	Teaching Assistant, Science Education, Abant Izzet Baysal University, Bolu.
1993 -1996	Teaching Assistant, Physics Education, The Middle East Technical University, Ankara.
1992 -1993	Physics Teacher, Sener High School, Merter, Istanbul.

Membership in Professional Organization:

American Association of Physics Teacher
National Science Teacher Association
National Association for Research in Science Teaching