**CHAPTER 1:  INTRODUCTION**

The purpose of this study is to see if there are differences in the achievement and opinions of men and women in an introductory physics course.  There has been much work done on differences in math and science achievement and persistence and much debate over the causes of the differences.  This study will contribute to the ongoing debate on the causes of these differences.

This chapter provides a brief overview of previous research into these differences and into possible ways to overcome them.  Then some terms used in the thesis are defined and the research questions are presented.  Finally, some limitations of the study are briefly discussed.

**Sex Differences**

A word or two needs to be said here about the term "sex differences."  Most of the research on the differences between boys and girls or men and women in their performance on math and science tests, their persistence in the fields of science, math, and engineering, and their experiences in classrooms talks about "sex differences" rather than "gender differences."  Some people do not see a distinction between sex and gender, but many feminist scholars do (e.g., Levi-Strauss, 1971; Rubin, 1975).  When there is a difference between sex and gender, sex is biological and gender is social and cultural.  Men and women obviously have sex differences, which are physical and hormonal.  They also have gender differences, which arise from differences in the way children imitate their parents, the differences in what their are praised for doing or told not to do, and their self-socialization (Maccoby & Jacklin, 1974).  This is the subject of much debate, as

almost any observed difference between men and women is up for study. Much of the research discussed in the literature review of this thesis uses the "sex differences" terminology, but this author will make the distinction.

There has been considerable research on differences of ability in math and science, but that research does not show conclusive overall differences (Kahle & Meece, 1994). There are some differences in achievement on standardized tests in science and math, and there is also some inconclusive evidence of difference in understanding of physics.

There are differences in performance on the Scholastic Achievement Test (SAT). In 1991, 24% of boys scored at least 600 on the math SAT, while only 13% of girls scored as high (National Science Board, 1993). There are also differences in performance on both the math and physics National Assessment of Educational Progress (NAEP) exams. These differences grow as students progress through school, with the smallest differences in the fourth grade and the largest in the twelfth grade (National Science Board, 1993).

There is not much research that closely analyzes the differences in physics understanding. Boys outscore girls on the physics achievement test (Sadker & Sadker, 1994). In addition, a few studies indicate that male students have fewer alternative conceptions in physics than female students do. For example, David Maloney (1988) found that in one study of the rules college-age novices use in trying to solve projectile motion problems, women were less likely than men to ignore irrelevant information and therefore made more errors. This could not be completely explained by male/female differences in completion of high school physics courses, but might have been related to differential confidence levels (Maloney, 1988).

One study that was directly inspirational to the present study is this author's master's thesis (Blue, 1994). This was a study of the University of Minnesota's Physics 1041, an algebra-based introductory physics course for non-scientists in which students learned in cooperative groups and were taught an explicit problem solving strategy. Blue found that men and women in Physics 1041 learned several fundamental concepts equally well. The lack of difference could have been because the women in the course were slightly older or have slightly higher grade point averages than the men, so they may not have been as intimidated as most women are by a physics class (Kahle and Meece, 1994). Another set of reasons could be the design of the course, which was made to be friendly to women: women prefer cooperation to competition (Johnson and Johnson, 1989), they do well on free-response questions (Scantleburyz and Baker, 1992), and they liked being taught how to solve problems (Heller and Lin, 1992).

Several studies have shown that women who plan to major in science or engineering in college have high school backgrounds, SAT scores, and attitudes which are equivalent to those of men with the same plans (DeBoer, 1985; Seymour, 1992b, Whigham, 1988). Despite this, women drop out of science and engineering majors at a higher rate than men do (American Institute of Physics, 1996; National Science Board, 1993; Widnall, 1988). There are some differences in both the predictors for choosing a science major (Ware, Steckler, & Leserman, 1985) and in the reasons for dropping one (Seymour, 1992a; 1992b).

There have been several investigations into reasons for these differences. Some researchers who note the persistent differences in many populations believe that the differences are innate, caused by biological sex differences (Benbow, 1988; Benbow & Stanley, 1980; Maccoby & Jacklin, 1974). Others point to social and cultural differences

between the experiences of boys and girls as possible causes for the differences, such as cultural definitions of science and gender, and gendered classroom experiences. These are explored in some detail in Chapter 2.

One of the most critical times for students to re-evaluate their major and career plans is the first year of college (DeBoer, 1985). This is when science majors typically take large introductory courses. If students are to be retained in their physics majors, they need to experience some measure of success or enjoyment during their first physics course.

## Possible Solutions

Sheila Tobias has found that women, even more than men, are uncomfortable with the competitive "class culture" that exists in most introductory college science classes (Tobias, 1990). Research consistently shows that female students prefer to work in cooperative groups rather than to compete with other students (e.g., Johnson & Johnson, 1989, Kahle & Meece, 1994). Studies of precollege students also suggest that females might benefit from cooperative grouping. Boys are more active in traditional high school science classrooms, especially in whole-class interactions. They raise their hands more, dominate lab work, and are more often called on by teachers (Tobin & Garnett, 1987). These differences can translate into differences in achievement. A study of fourth grade girls in a math class found that teacher-initiated class participation correlated highly with the girls' achievement (Fennema & Peterson, 1986). Thus female students might benefit more from cooperative group work than from whole-class discussions from which they tend to be excluded.

Not all groups are cooperative groups. Students may have had negative experience in groups. Certain conditions need to be met to keep these things from happening. The experimental section of this study used structured cooperative learning based on the Johnson and Johnson model of cooperative learning (Johnson & Johnson, 1989, Johnson, Johnson, & Smith, 1989). This model is discussed in Chapter 2.

Another way to try to lessen differences in performance and persistence in a physics class is to teach an explicit problem solving strategy. One such strategy, based on research into the ways in which experts and novices solve problems, has been developed at the University of Minnesota (Heller & Hollabaugh, 1992, described in detail in Chapter 2). There are indications that being taught an explicit problem solving strategy could benefit women. Women are more likely than men to embrace the strategy (Heller & Lin, 1992), and their scores on both conceptual and problem solving tests are equivalent to the scores of men at the end of a course where students have used the strategy (Blue, 1994; Heller & Lin, 1992; Huffman, 1994).

**Definitions Of Terms**

There are several terms that, for convenience, will be defined here.

sex differences: For the purposes of this study, sex differences are those that can be attributed to biological differences; physical, hormonal, and perhaps genetic differences between boys and girls or men and women.

gender differences: For the purposes of this study, gender differences are those that can be attributed to the differential socialization of boys and girls. These can arise from

differences in the way children imitate their parents, the differences in what their are praised for doing or told not to do, and their self-socialization (Maccoby & Jacklin, 1974).

cooperative groups:  Students work in groups of three (or four, if the class is not divisible by three).  Group members sit together and work together in both problem sessions and laboratory sessions, they receive group grades on the problems they do in their problem sessions and on their laboratory reports, and they are assigned tasks and roles within their groups.

explicit problem solving strategy:  Students in one section, the experimental section, were taught a five-step problem solving strategy in lecture and problem sessions, are given answer sheets with spaces to do the five steps, and are graded on using the strategy on their exams.  The five steps are:   focus on the problem, describing the physics, plan the solution, execute the plan, and evaluate the solution.

opinions about the course:  At the end of the course, students filled out an evaluation of the course.  This evaluation included ratings of the lecture, laboratory, problem sessions, problem solving methods used, the teaching assistants, and the professors.  For the purposes of this study, the main differences between the sections were the cooperative grouping and the explicit problem solving strategy used in  the experimental section and not the traditional section.  For this reason only student ratings of the laboratory, problem sessions, and problem solving methods will be compared.

physics learning:  There are, of course, many facets to physics learning.  The two explored in this thesis are problem solving ability and conceptual understanding, both defined below.

problem solving ability:  This will be measured by performance on four problems on the final exam.  To solve these problems, students had to use principles of kinematics and Newton's laws.

conceptual understanding:  This will be measured in two ways in this study:  performance on the Force Concept Inventory (Hestenes, Wells, & Swackhamer, 1992), a 29 item multiple choice test, and on free-response questions.  The free-response questions concern acceleration, the nature of forces, and Newton's second and third laws (Blue, 1992, Huffman, 1994).

### Research Questions

There is considerable debate surrounding the causes of the observed differences in the physics performance of men and women.  Some think that the differences are so often observed that they must be biological, and some think that there is such a difference in the way that parents and teachers treat boys and girls that the differences in achievement must arise from these social and cultural influences.  To add to this debate, one of the research questions of this study is:

1. If there are minimal differences between men and women in their relevant physics background and initial performance when they start an introductory physics course which was designed to appeal to a broad

population, will there be differences in how much physics they learn by the end of the course?

The introductory course under study was one of the sections of Physics 1251 which was offered at the University of Minnesota in the fall of 1993. The problem solving sessions of one section incorporated cooperative group learning and an explicit problem solving strategy, both of which might help make it easier for women as well as men to experience some measure of success or enjoyment during their first physics course. This section is called the experimental section in this paper. Another section, called the traditional section here, had more traditional recitations for problem solving sessions. This study examines the experimental section and, as baseline to aid in the interpretation of the results, the traditional section.

About 22% of the students in the course were women. All the women on whom there is complete data (20 from the experimental section and 14 from the traditional section) were used in the sample for this study, and they were matched with an equal number of men. The matched sample was chosen to eliminate as many relevant, measurable differences as possible. Since men and women often have different high school backgrounds, the subjects were matched on whether they had taken physics in high school, whether they had taken calculus in high school, and their high school grade point average. Since men and women often are treated differently and learn different amounts in their math and science classes, their course background was not used as the sole measure of their initial physics knowledge. Subjects were matched on pretest Force Concept Inventory scores, on a pretest free-response conceptual test, and on how well they solved a pretest problem on density. Furthermore, subjects were matched on their year in school and on their locus of control over their own grades.

The first research question will be answered with respect to two main aspects of physics learning: problem solving ability and conceptual understanding. The first way in which physics learning was measured is through a careful scoring of the problem solving portions of the matched sample's final exams. Another way to look at physics learning is to look at students' conceptual understanding. For an initial look in this area, the matched sample's Force Concept Inventory post-test scores were compared. For a more detailed look at conceptual understanding, the matched sample's answers to some free response conceptual questions were analyzed.

The answer to the first research question will contribute to the ongoing debate on the causes of difference in math and science performance. If there are minimal differences between men and women in their background and pretest performance and there are still differences in their post-test performance, then that will suggest that the differences in performance will have been caused by the biological differences between men and women. Conversely, if there are minimal differences between men and women in their background and pretest performance and then there are no differences in post-test performance, that will suggest that the difference in performance that is usually seen is caused by societal and cultural gender differences rather than biological sex differences.

There is also concern that women drop out of their science, math, and engineering majors at a higher rate than men do. Past research has shown that women prefer working cooperatively to working competitively and that they are likely to embrace an explicit problem solving strategy. Perhaps the addition of cooperative group work and an explicit problem solving strategy to an introductory course would be welcome changes. The second research question of this study is:

2. Will there be differences in the opinions of men and women about an introductory physics course which was designed to appeal to a broad population?

This research question is much different than the first. Because the student evaluations at the end of the course were anonymous, it is not possible to study only the evaluations of the students in the matched samples. Therefore it will not be possible to distinguish between sex differences and gender differences in students' opinions about the course. Since the aspects of the experimental course that were designed to appeal to a broad population are the cooperative grouping and the explicit problem solving strategy, only the questions from the student evaluations that relate to the laboratory, the problem solving sessions, and the problem solving strategy are studied.

The answer to the second research question might contribute to the understanding of why even women who plan to major in science or engineering drop out of those majors at higher rates than men do. It could be that women have a lower opinion of some of the main aspects of their introductory physics courses. If so, some changes in those aspects of the course might help to retain women in their science majors.

**Limitations Of The Study**

The small sample size of 68 students is a limitation of this study. The sample size is necessarily small for two reasons: the low percentage of women in the course and the low attendance at lecture. Although the classes were large lecture courses (with 225 and

10

286 students finishing the quarter), the percentage of women was low (22%), and equal samples of women and men were sought. In addition, much of the data was taken during lecture, and sometimes only half of the students in a given section would attend on these days. Thus the number of women on whom complete data was collected was only 34, giving a total sample size of 68.

However, using matched samples is a powerful technique. The women and men have in the sample been matched on three pretest scores and on several demographic characteristics, so any differences in their post-test scores are attributable to their sex. The matched sample will be compared to the larger group in Chapter 4 to see whether the characteristics of students on whom complete data was collected are significantly different than the characteristics of the larger population.

The more significant limitation to the study is the need to rely on what students write in their problem solving solutions and the answers to the conceptual questions studied. At times student answers can be ambiguous, and often students do not include explanations of their reasoning even when they are asked to do so. In an attempt to interpret the answers in as straightforward a manner as possible, there was little second-guessing and little benefit of the doubt given. For example, when a student set acceleration to zero in a rotating system, it was assumed that the student was wrong rather than that the student was looking at the problem from a non-inertial reference frame. This was done consistently; the worst problem arising from this practice will be a uniform under-reporting of students' skills.

**Overview Of The Dissertation**

A review of the literature related to this study is provided in Chapter 2. Included in this review is the research on differences between men and women in preparation for, performance in, and persistence in college science, possible reasons for these differences, and possible solutions, which include cooperative grouping and the use of an explicit problem solving strategy.

A description of the research methods used in this study is provided in Chapter 3. Included in this chapter is a description of the setting, instruction, instruments and data analysis used in this study. Also included are descriptions of the student population and the selection of the matched sample.

The results of this study are presented in Chapter 4. Included are a comparison of the matched sample and the student population, the results of the problem solving test, the Force Concept Inventory, and the free-response conceptual test, and the course evaluations.

Finally, Chapter 5 provides a discussion of the results. Included in this chapter are comparisons of the course evaluations, the problem solving test, the Force Concept Inventory, and the free-response conceptual test, as well as discussions of the limitations of the study and implications for instruction.

# CHAPTER 2:  REVIEW OF THE LITERATURE

This study focuses on sex differences in the achievement and opinions of students in a physics course that has been designed to appeal to a large population.  There have been many previous studies focusing on sex differences in math and science and much discussion of the possible reasons for the differences.  Several ways of trying to make a course more gender fair have been explored, and two of them were used in the experimental section.  In addition, much is already known about student understanding of the concepts taught in the course.

Literature relevant to the study includes:

> Sex differences and possible reasons for them
> Possible solutions
> Conceptual understanding

## Sex Differences

There has been much interest in sex differences in preparation and performance of males and females in math and science as well as in their persistence in college majors and careers that relate to math and science.

**Preparation**

Several studies have shown that women who plan to major in science or engineering in college have high school backgrounds and SAT scores which are equivalent to those of men with the same plans (DeBoer, 1985; Seymour, 1992b).  For example, Myrna Whigham found few gender differences among engineering students at Iowa State.  Math and composite ACT scores were not different, although women had higher scores on the English and social studies sections.  Women were more likely than men to have a positive attitude toward success, as measured by their agreement with the phrases "happy to be regarded as an excellent math student," "being regarded as smart would be great," and disagreement with the phrases "would think I was some kind of grind if I got A" and "would like me less if I were really a good math student."  Women were more likely than men to say they were encouraged by their teachers.  They also disagreed even more strongly than the men with the idea that math is a male dominated field (Whigham, 1988).

**Performance**

There has been considerable research on sex differences of ability in math and science, but that research does not show conclusive overall differences (Kahle & Meece, 1994).  There are some sex differences in achievement on standardized tests in science and math, and there is also some inconclusive evidence of sex difference in understanding of physics.

The National Science Board reports scores on both the math and science National Assessment of Educational Progress (NAEP).  They have found that sex differences in

performance on these tests grow as students progress through school.  On the math NAEP, there is no difference in scores between boys and girls in fourth or eighth grade, but boys score higher than girls in twelfth grade.  On the science NAEP, the growth in difference in performance is more gradual.  On the 500 point test, boys outperform girls by three points in fourth grade, by seven points in eighth grade, and by eleven points in twelfth grade (National Science Board, 1993).

There are additional studies that have shown that males outperform females on mathematics achievement tests, especially those that emphasize problem solving, at the end of high school (e.g., Armstrong, 1985; Chipman & Thomas, 1985).  The most well known of these achievement tests is the Scholastic Aptitude Test (SAT).  In 1991, 24% of boys scored at least 600 on the math SAT, while only 13% of girls scored as high (National Science Board, 1993).

There are two caveats worth noting along with these statistics.  First, recent meta-analyses indicate that sex differences in math achievement are small and have been decreasing since 1975 (Friedman, 1989).  Second, the differences can appear larger or smaller depending on how people present the statistics.  In his meta-analyses, Alan Feingold (1992) found sex differences in the variability of achievement.  Males are more variable than females in their general knowledge, mechanical reasoning, quantitative ability, spatial reasoning.  For this reason, the sex differences in achievement in standard deviation units (difference between male and female means divided by the standard deviation within each sex) are smaller than sex differences between means (Feingold, 1992).

Although boys outscore girls on the SAT (both the math and the verbal sections) and on all science achievement tests, with the largest gap on the physics test (Sadker & Sadker, 1994), there is not much research that closely analyzes the sex differences in physics understanding. A few studies indicate that male students have fewer difficulties in solving physics problems than female students do. For example, David Maloney (1988) found that in one study of the rules college-age novices use in trying to solve projectile motion problems, women were less likely than men to ignore irrelevant information and therefore made more errors. This could not be completely explained by male/female differences in completion of high school physics courses, but might have been related to differential confidence levels (Maloney, 1988).

All of these results should be interpreted with caution, however, as some of these differences in performance might be due to the testing instruments rather than real differences in understanding. It has been found that, on average, female students have higher verbal skills than males and therefore have an advantage on extended written questions. Male students, on average, have an advantage on multiple choice questions and on timed tests (Sadker & Sadker, 1994; Scantleburyz and Baker, 1992).

One study that was directly inspirational to the present study is this author's master's thesis (Blue, 1994). This was a study of the University of Minnesota's Physics 1041, an algebra-based introductory physics course for non-scientists in which students learned in cooperative groups and were taught an explicit problem solving strategy. Blue found that Physics 1041 is gender-fair in the teaching of the concepts of acceleration, the nature of forces, and Newton's Second Law, as measured by students' responses to free-response questions. Possible reasons that women learn as successfully as men in Physics 1041 are that they are slightly older or have slightly higher grade point averages than the men, so

they may not have been as intimidated as most women are by a physics class (Kahle and Meece, 1994). Another set of reasons could be the design of the course, which was made to be friendly to women: women prefer cooperation to competition (Johnson and Johnson, 1989), they do well on free-response questions (Scantleburyz and Baker, 1992), and they liked being taught how to solve problems (Heller and Lin, 1992). For whatever reasons, it seems that the instruction in Physics 1041 is equally effective for men and women. Women and men both improve in their conceptual understanding of several basic physics concepts during the course. The same free-response questions were used in the present study, so that male and female understanding of the same physics concepts can be explored as a partial answer to the first research question: If there are minimal differences between men and women in their relevant physics background and initial performance when they start an introductory physics course which was designed to appeal to a broad population, will there be differences in how much physics they learn by the end of the course?

**Persistence**

Although women and men who declare majors in science and engineering have similar backgrounds, women drop out of science and engineering majors at a higher rate than men do. Given a population of 2000 males and 2000 females starting ninth grade together, on the average 140 males and 44 females will say that they plan to major in science when they enter college. Out of the students who declare this intent, 46 men and 20 women actually receive bachelor's degrees in science, and 5 men and 1 woman will receive a Ph.D. (Widnall, 1988). Statistics from the National Science Board show that in 1991, women earned 45% of bachelor's degrees in science and 16% of the bachelor's degrees in engineering (National Science Board, 1993). The most recent survey by the

17

American Institute of Physics found that in 1994, only 17% of physics bachelors degrees and 12% of physics doctorates went to women (American Institute of Physics, 1996). More specifically, a study of a group of Harvard and Radcliffe students has found that of those who declared an interest in a science major on their application, only 69% of the men and 50% of the women completed a science major (Ware, Steckler, & Leserman, 1985). This rate of defection is the highest of any college major (Seymour, 1992a). There were some sex differences in the predictors for choosing a science major. For men, the two most significant factors were high grades in their first college science course and being positive they would be science majors as early as the summer before college. For women, the two most significant predictors were having highly educated parents and high math SAT scores. These differences suggest that women, more than men, need both personal support and an external validation of their abilities in order to complete a major in science (Ware, Steckler, & Leserman, 1985).

These data are supported by a study on attrition from science, math, and engineering majors done at four different institutions in Colorado (Seymour, 1992b). Science, math, and engineering have the highest drop-out rates, of both men and women, of any major. People of both sexes who switch majors cite "structural and cultural sources" more than "personal inadequacy" (Seymour, 1992a). There are, however, some sex differences. More women than men said that they had chosen to major in science, math, or engineering because of the direct support and encouragement of a parent or mentor than because of their own interest in the fields. The reasons that women cited most often for dropping out of these majors were poor grades in introductory classes and a rejection of what they saw as the lifestyle of scientists and engineers. The reason that most men cited was poor teaching. Women also cited poor teaching, but they had very different definitions of good and poor teaching. Men thought a good teacher was someone who

was enthusiastic, explained things well, and told good stories, and women thought that a good teacher was someone who cared how they did in the class. Personal attention seems to have a lot to do with the persistence of women in science, math, and engineering majors (Seymour, 1992b).

One of the most critical times for students to re-evaluate their major and career plans is the first year of college (DeBoer, 1985). This is when science majors typically take large introductory courses. If students are to be retained in their physics majors, they need to experience some measure of success or enjoyment during their first physics course. This study will investigate whether being taught in cooperative groups and being exposed to an explicit problem solving strategy affect the opinions about the course and the physics learning of males and females differently.

**Possible Reasons for Sex Differences**

There are many possible reasons for the differences in performance and persistence between males and females with similar backgrounds. For example, Chipman and Wilson (1985) summarized seventeen studies on gender differences that were published between 1977 and 1981. One of their findings was that it was hard to separate variables in this kind of study: ability may affect attitude and vise versa. At the ninth grade level, they found no sex differences in cognitive variables, encouragement from parents and teachers, or in liking or interest in math. They did find that ninth grade boys were more likely than ninth grade girls to see math as useful, to expect a career in math, and to be confident in their math ability (Chipman & Wilson, 1985). Interestingly, girls were less likely than boys were to think that math was something for males and not for females (Chipman & Wilson, 1985). There have been several investigations into reasons for the

sex differences. Some suggested reasons are innate sex differences, cultural definitions of science and gender, and gendered classroom experiences. These are explored in the next section.

**Innate Sex Differences.** It has not been determined whether sex differences in achievement are due to innate biological differences in ability or due to cultural influences (e.g., Brophy, 1985; DeBoer, 1985; Jones & Wheately, 1990; Kahle & Lakes, 1993, Leder, 1990). For example, the Guest Comment, "Why So Few Women" by J. Button-Shafer (1990) in the American Journal of Physics precipitated a long and heated debate on these issues (Levin, 1990; Rushaki, 1990, Button-Shafer, 1991; Hawkins, 1991; Goldberg, 1991).

Some of the most outspoken believers in innate gender differences are Camilla Benbow and Julian Stanley, who work with The Johns Hopkins University's Center for Advanced Studies (Benbow, 1988; Benbow & Stanley, 1980). The Center for Advanced Studies conducts a search each year for gifted seventh graders to invite into their summer programs, using the SAT. Based on their observations of years of math SAT scores, Benbow and Stanley have concluded that sex differences in math achievement must be biological, since they are so often found (Benbow & Stanley, 1980). Similarly, a meta-analysis of research into sex differences in visual-spatial ability, which some link to mathematical ability, has prompted researchers to propose a biological or genetic cause of the difference in ability (Maccoby & Jacklin, 1974).

Many do not see the reasons for the sex differences so clearly, however. For example, Jane Brophy (1985) believes that sex differences exist, but they are not biological but socially determined. From infancy, adults and older children model behaviors that are

different for different sexes. They also expect different things from girls and from boys, and by these processes girls and boys learn to act differently. By the time girls and boys are in school, there are enough differences that teachers treat them differently. Boys are more likely to act out, either displaying their intelligence or disrupting the class, and they get more attention from teachers for doing these things.

In addition, Chipman and Wilson found evidence in their meta-analysis (1985) that spatial ability (such as measured on many standardized math tests) is not innate but can be taught. They also did not find much evidence that spatial ability is related to math performance, partly because it is not well defined (Chipman & Wilson, 1985).

Myra and David Sadker note that SATs overpredict school success for males and under predict it for females; a boy with an A+ average in high school will have total SAT scores an average of 83 points higher than a girl with an A+ average (Sadker & Sadker, 1994). Although total SAT scores for females are lower than those for males, women have higher grade point averages in college than men do (National Science Foundation, 1995). A possible reason for the discrepancy that echoes Brophy's findings is that girls might get good grades for good behavior; teachers say cooperative girls are the smartest (Sadker & Sadker, 1994).

**Cultural Definitions Of Science And Gender.** Science is often seen as a male domain, and that idea can keep women and girls from being interested in science (Belenky, et. al, 1986; Keller, 1985; Schiebinger, 1987). One of the fathers of modern science, Francis Bacon, was also one of the first to talk about the earth and nature as female. When he spoke of the practice of science, he used metaphors of marriage and rape as he talked about how men could discover the hidden secrets of (female) nature (Bacon, 1620/1989).

Whether or not one places much significance in the metaphors used by a philosopher of science three and a half centuries ago, it remains true that since then science has been practiced by men more often than by women. Perhaps because of that inbalance the qualities of a good scientist have come to be associated with men and not with women, in the same way that the qualities of a good nurturer have come to be associated with women

Table 2.1: Stereotypes Relating to Science and Gender Difference

| male | female |
|---|---|
| knower | known |
| mind | nature |
| reason | feeling |
| rational discourse | irrationality |
| scientific knowledge | unpredictability |
| "hard" science - physics, chemistry | "soft" science - education, psychology |

and not with men (Keller, 1985; Schiebinger, 1987). Some feminist theorists have suggested that the questions, practices, and answers of science might be different if scientists and decision makers were women rather than men (see, for example, Harding, 1986). As it is, though, many stereotypes abound in Western technological culture that relate to science and to sex differences, as can be seen in Table 2.1 (Keller, 1985; Schiebinger, 1987).

There are some women who feel so alienated from science, in part because they feel that it has too limited a definition of truth, that they reject science, logic, and rationality (Belenky, et. al., 1986). They are probably not represented in this study, however, as it is unlikely

that a woman who rejects these things would take a calculus-based physics course in college. Others have a view of science more compatible with that of current philosophers of science, that scientific theories and laws are created by people, taught to them by people, and might be changed in the future by other people (Belenky, et. al., 1986). It is

not often that students learn this in an introductory science class, however, since there is usually no time for theories to be presented to students by books and instructors as anything but timeless facts, somewhat mysterious in origin. It is usually not until the end of college or into graduate school that students learn that scientists, textbook authors, and professors are not omnipotent (Belenky, et. al., 1986).

**Gendered Classroom Experience.** Much research has found that girls and women get a poorer education in science than boys and men, even when they are in the same classrooms (Rosser, 1990, etc., etc.). This is true at all levels, in middle school, high school, and college.

Girls and boys enter school in kindergarten with the same measured math abilities. By the time they graduate from high school girls are behind boys in both math ability and self-esteem (American Association of University Women, 1992). Elizabeth Fennema and her colleagues have found that girls' confidence in math drops in middle school before their achievement also drops in middle school (American Association of University Women, 1992; Fennema & Peterson, 1986). Adolescent girls' self-esteem drops, especially as compared to boys' (American Association of University Women, 1992). Follow-up studies have shown that much of the change in self-esteem can be linked to girls' and boys' differential experiences in schools (Orenstein, 1994).

At all stages, girls get less attention from teachers than boys do (American Association of University Women, 1992, Sadker & Sadker, 1994). For example, Gail Jones and Jack Wheatley (1990) studied thirty chemistry classes and thirty physical science classes in high schools. They found that male students called out in class significantly more than females, they initiated private conversations with teachers about procedural matters

significantly more, and they received both praise and behavioral warnings significantly more often than females did (Jones and Wheatley, 1990).

Kahle and Lakes (1983) studied students ages nine, thirteen, and seventeen and found that girls had fewer science experiences than boys.  For the purposes of their study, they defined science experiences as use of scientific apparatus (meter stick, stopwatch, compass, scale, magnifying glass, balance, telescope, microscope); close observation of animals; fixing things; reading about or talking about science; doing science projects or hobbies.  The researchers speculated that possible reasons for the discrepancy in science experiences were (1) unequal training in the same classrooms as boys, where boys do the experiments and girls watch or record, (2) perception of science as a masculine thing, and (3) fewer extracurricular science activities (Kahle & Lakes, 1983).

High school girls are also likely to participate less in science than they do in  their other courses, saying science makes them feel stupid (Sadker & Sadker, 1994).  Counselors and teachers often excuse girls from taking science classes when the girls ask them to do so; most girls do not realize that this decision effectively shuts them out of many careers (Sadker & Sadker, 1994).

George deBoer surveyed men and women at a selective college and found that although the women had received higher grades in their high school science classes than the men did, they rated their high school science ability lower than the men (deBoer, 1986). Perhaps this is because their SAT scores under predict their success, as was noted above. The Sadkers have found that when girls see a discrepancy between their grades and their SAT scores, they are more likely to believe the scores, thinking they are more objective (Sadker & Sadker, 1994).  DeBoer also found that a high self-rating of high school

science ability was strongly related to choosing to take more than one or two science courses in college, and concluded that women's low self-ratings were affecting their persistence (deBoer 1986).

A study of students who had been the valedictorians of their high school found that the female valedictorians' self-evaluations lowered in college and men's did not, despite the fact that the women valedictorians get better grades in college then the men do (Sadker & Sadker, 1994). This drop in self-evaluation is caused at least in part by the continuing different experiences that men and women have in college. In their study of students in cooperative groups in a college physics class at the University of Minnesota, Heller and Hollabaugh (1992) found that even within small cooperative groups made up of two men and one woman, the men would dominate the woman, often ignoring her ideas, even if she were the highest achiever of the group and the one with the right ideas.

The answer to the first research question in this study will contribute to the ongoing debate on the causes of sex difference in math and science performance. Since there are only minimal gender differences between the men and women in the matched samples, whether or not there are sex differences in their performance on tests of physics knowledge at the end of the course will suggest whether there are innate sex differences in ability.

<div align="center">**Possible Solutions**</div>

**Cooperative Grouping**

Sheila Tobias has found that women, even more than men, are uncomfortable with the competitive "class culture" that exists in most introductory college science classes (Tobias, 1990). Research consistently shows that female students prefer to work in cooperative groups rather than to compete with other students (e.g., Johnson & Johnson, 1989, Kahle & Meece, 1994). Studies of precollege students also suggest that females might benefit from cooperative grouping. Boys are more active in traditional high school science classrooms, especially in whole-class interactions. They raise their hands more, dominate lab work, and are more often called on by teachers (Tobin & Garnett, 1987). These differences can translate into differences in achievement. A study of fourth grade girls in a math class found that this type of teacher-initiated class participation correlated highly with the girls' achievement (Fennema & Peterson, 1986). Thus female students might benefit more from cooperative group work, where teachers interact with groups rather than individuals, than from whole-class discussions from which they tend to be excluded.

Not all groups are cooperative groups. Students may have had negative experience in groups, running into things like the "free-rider effect," where students do less work because they expect other group members to pick up the slack (Johnson & Johnson, 1989). Or they could have seen the "sucker effect," where the more able group members also stop working so as not to become victim of the free rider effect (Johnson & Johnson, 1989). When the students with natural leadership skills take on leadership skills of the group, there is a "rich-get-richer effect" as their skills improve and no one else gets to practice their skills (Johnson & Johnson, 1989). Problem groups can also be characterized by dysfunctional divisions of labor, where some students do all the intellectual work or have all the contact with the laboratory equipment (Johnson & Johnson, 1989). Certain conditions need to be met to keep these things from happening.

The cooperative learning used by the experimental section in the laboratory and problem sessions was based on the Johnson and Johnson model of cooperative learning (Johnson & Johnson, 1989, Johnson, Johnson, & Smith, 1989). In this model there are five aspects of cooperative learning: a) positive interdependence, b) face-to-face interaction, c) individual accountability, d) interpersonal skills, and e) group processing.

Positive interdependence means that students are dependent upon each other such that if one does well, the other will do well or if one does poorly, the other will do poorly also. The second aspect of cooperative grouping is face-to-face interaction. When members of a group can see each other and look each other in the eye, they are more likely to work well together. The third aspect of cooperative learning is individual accountability. It is important that students not feel that, because they are working in a group, they can be lazy and let the other members of their group be the only responsible ones. The fourth aspect of cooperative learning is the teaching of interpersonal skills. Some people need to be reminded of skills like listening well to other group members, complimenting them on their ideas, and not insulting anyone. The fifth aspect of cooperative learning is group processing. Groups often need to take time to think about what they do well as a group, what they do poorly as a group, and how they could improve.

David and Roger Johnson of the Cooperative Learning Center at the University of Minnesota have done meta-analyses of over 500 studies (Johnson & Johnson, 1989) that compare cooperative group learning to learning in competitive environments. They have found that cooperative learning helps with both achievement and persistence. It has been consistently shown that cooperative groups can achieve more than even the best achieving

individual in the group (Heller, Keith, & Anderson, 1992; Johnson & Johnson, 1989; Johnson, Johnson, & Smith, 1991).

Perhaps more important to the issue of addressing sex differences is the finding that cooperative grouping can help with persistence in a physics major. The perception that physics is male dominated and competitive can intimidate female students (Kahle and Meece, 1994). Competitive learning makes students more anxious than cooperative learning (Johnson & Johnson, 1989). This effect is even more noticeable among students with above-average levels of anxiety, such as females in an introductory physics class. Research consistently shows that female students prefer to work in cooperative groups than to compete with other students (Johnson and Johnson, 1989; Kahle, 1990). In addition, in a competitive environment, only a few students can win and the rest must lose. The reaction of many of the "losers," both male and female, will be to drop out (Johnson & Johnson, 1989).

Working in cooperative groups has also been found to play a part in keeping students in college. Cooperative learning has been shown to improve students' attitudes towards the subject they are studying (Johnson & Johnson, 1989; Johnson, Johnson, & Smith, 1991), which might then lead to higher retention of majors in a subject. In fact, it has been shown that active learning can be instrumental in keeping students in college, especially for "withdrawal-prone" students (Johnson, Johnson, & Smith, 1991).

This can be explained by Tinto, an educator who studies college and college students. His theory of integration says that good encounters, both social and academic, get students more integrated into the attitudes and values of the college. The more integrated students become, the less they are likely to drop out of their college or their major

(Pascarella & Terenzini, 1991). There is evidence that lasting friendships and study groups form out of cooperative groups (Johnson & Johnson, 1989), leading students to become more integrated.

The effect of cooperative grouping on attitudes will be explored in this study by answering the second research question, concerning sex differences in opinions about the course.

**Teaching An Explicit Problem Solving Strategy**

There are indications that being taught an explicit problem solving strategy could benefit women. In a study of students in Physics 1041-1042, an algebra-based introductory physics course at the University of Minnesota, Heller and Lin (1992) found that a larger proportion of women than men adopted the explicit problem-solving strategy that was taught in the course. One reason could be that more female than male students in that course admitted to being intimidated by math or problem solving. Perhaps more of the men in that course were already comfortable with the way they solved problems. A recent study of high school students also found a gender difference in the reaction to a problem solving strategy. It was found that male high school students who are required to use a problem solving strategy actually make smaller gains in conceptual understanding than expected (Huffman, 1994). It is interesting to note that in both the Heller and Lin study and the Huffman study it was found that by using the explicit strategy, the female students achieved scores on problem-solving tests that were equivalent to those of the male students. It has also been found that students in Physics 1041 had no real differences in the responses of men and women to free-response conceptual questions (Blue, 1994).

In this study, students in the traditional section were encouraged to use the problem-solving strategy outlined in their textbook (Fishbane, Gasiorowicz, & Thornton, 1994) and students in the experimental section were required to use a strategy developed at the University of Minnesota. Outlines of the two strategies appear in Tables 3.1 and 3.2, on pages 74 and 76.

The explicit problem solving strategy taught in the experimental section was similar to those developed by Polya (1945) and by Reif, Larkin, & Brackett (1976). Polya's strategy outlined four basic steps: understand the problem, devise a plan to solve it, carry out the plan, and take a look back (Polya, 1945). Reif, Larkin, and Brackett adapted Polya's method into a strategy to teach to physics students at the University of California at Berkeley. They saw a clear improvement in the problem solving skills of their students when they taught them the four step strategy of describe, plan, implement, and check (Reif, Larkin, & Brackett, 1976).

The explicit problem solving strategy was developed by and is based on the expert novice research (e.g., Chi, Feltovich, and Glaser, 1981; Finegold & Mass, 1985; Heller, Keith & Anderson, 1992; Heller & Hollabaugh, 1992; Larkin, 1980; Larkin, 1981; Larkin & Reif, 1979).

For the purpose of this study, students' problem solving abilities were assessed. How can one tell that a physics student is a good problem solver? Some studies assume that individuals in certain categories, like professors or advanced graduate students of physics or students who always get As on physics tests, are good, expert problem solvers (e.g., Finegold & Mass, 1985; Larkin, 1980). Other studies try to score problem solving with

criteria such as errors per problem and length of time to solution (e.g., Tarmizi & Sweller, 1988). These methods are easy to use but do not have a strong relationship to expert/novice research.

The physics education group led by Patricia Heller at the University of Minnesota has refined a detailed coding scheme grounded in expert/novice research which scores problem solutions on their general approach, specific application of physics, logical progression, and appropriate mathematics. This coding scheme will be used in this study to analyze problem solving abilities.

1.      The general approach is what physics principles the student uses to solve a problem and how well the student understands those principles. Previous research has shown how experts are more likely than novices to see what physics principles are needed to solve a problem (Chi, Feltovich, & Glaser, 1981). Whether a student understands the principles or not is a result of how much physics he or she learned.

2.      The specific application of physics category assesses how well the student applies the general principles to the specific problem. Students might have difficulties applying the principles because they have failed to learn a small process that is part of solving many problems, for example resolving vectors into components. They might also have problems that result from not defining variables and coordinate systems well enough. Neglecting the descriptive stage of the solution is something novices are more likely to do than experts (Finegold & Mass, 1985; Heller & Hollabaugh, 1992; Larkin, 1980; Larkin, 1981; Larkin & Reif, 1979).

3.      Logical progression is a description of how organized and logical the solution is. One way a student might get a low score in this category is by failing to plan the solution before starting.  Novices are less likely than experts are to take this planning step (Finegold & Mass, 1985; Heller & Hollabaugh, 1992; Heller, Keith, & Anderson, 1992; Larkin, 1980; Larkin, McDermott, 1980).  Some students make illogical jumps or leave their solutions unfinished.  These mistakes could come from not having "chunked" enough information.  Most people can only store $7 \pm 2$ pieces of knowledge in short term memory (Chase & Simon, 1973).  Experts are more likely than novices to have their knowledge stored in large chunks (Chase & Simon, 1973; Hinsley, Hayes, & Simon, 1977; Larkin, 1979).  Novices sometimes literally forget what they are doing in the middle of a problem, so they do not remember what the next logical step should be.

4.      Appropriate mathematics is obviously important.  Students should not make mistakes in arithmetic, algebra, or calculus in their solution.  Furthermore, it is best for students to wait until the algebraic expression for the unknown is found before substituting numbers and units.

It is worth noting that this coding scheme, with scores for general approach, specific application of physics, logical progression, and appropriate mathematics, should not discriminate among students who have been taught different heuristics.  The scheme is based on expert/novice research and not on any specific heuristic.  Students who followed either the strategy from the textbook (Fishbane, Gasiorowicz, & Thornton, 1994), as was suggested in  the traditional section, or the explicit problem solving strategy taught in the experimental section would receive a high score.

In addition to problem solving, another fundamental aspect of physics learning is conceptual understanding. There has been a lot of work done on students' alternative conceptions (misconceptions) of physics; there are more than 300 studies on students' conceptual understanding in mechanics alone (Wandersee, Mintzes, and Novak, 1994).

## Conceptual Understanding

Alternative conceptions in physics, incorrect conceptions about physical ideas, have been found to be prevalent among both high school and college students (Ausubel, Novak and Hanesian, 1978). These conceptions are often quite resistant to change, even after instruction (Clement, 1982; Minstrell, 1982; McClosky, 1983; Gunstone, 1987). Although the existence of alternative conceptions is widely accepted, different researchers have different theories about the nature of students' alternative conceptions.

Some researchers think that students' conceptions hang together like coherent theories. These theories are sometimes similar to those of scientists like Aristotle and Galileo whose views have been replaced by modern theories. The major proponents of the belief that students' alternative conceptions are coherent, personal theories are John Clement and Michael McClosky (e.g., Clement, 1983; McClosky, 1983). Other researchers think that students' alternative conceptions are not coherent theories. Instead, students have many conceptions that are unconnected and dependent on the context in which they are used. This "knowledge-in-pieces" view of students' conceptions is held by Andrea diSessa, who calls the pieces "phenomenological primitives" (p-prims), and Jim Minstrell, who calls the pieces "facets" (diSessa, 1988; diSessa, 1993; Minstrell, 1991).

The difference in these views can be illustrated by an example. Many people have studied students' interpretations of the motion of a coin that is thrown up in the air with some initial velocity (e.g., Clement, 1983; McClosky, 1983; Champagne, Gunstone, and Klopfer, 1985; Halloun and Hestenes, 1985; diSessa, 1988). This motion is quite simple -- there is only one force acting on the coin (its weight), and there is only one constant acceleration. Many students, however, think that the motion is more complex. They

believe there is some sort of initial force given to the coin by the toss. This fades away as the coin rises, as evidenced by the coin's slowing down. After the coin reaches the top of its flight, there is a new force, gravity, that gets bigger and bigger as the coin falls towards the earth and speeds up.

Researchers who believe that students' alternative conceptions are in the form of coherent theories interpret these kinds of student responses as paralleling the theories of Aristotle and Galileo. Aristotle believed that a moving object is under the influence of a force that is directly proportional to their speed (Champagne, Gunstone, and Klopfer 1985). If students believe that, it is not surprising that they would invent forces to account for the changing speed of the coin. That is just what Galileo did. He wrote about an "impetus" theory, which is quite consistent with the belief that a toss imparts a force to a coin (McClosky, 1983).

Researchers who believe that students' knowledge is in pieces would not take student responses as evidence that they believe the same coherent theories of Aristotle or Galileo. diSessa (1988), for example, believes that students have knowledge in fragments that he calls phenomenological principles (or p-prims). These fragments are abstractions from experiences, and as such they are very context based and not necessarily consistent or coherent. diSessa would analyze the students' description of the tossed coin as follows: Initially, the action of the hand on the ball fits the student's "force as mover" p-prim. This principle comes from students' experience that pushing an object at rest causes it to move in the direction of the push. Like other p-prims, this is an over generalization that students make from their experiences in the world, where almost everything has friction. After the "force as mover" p-prim, students use the "dynamic balance" p-prim to explain the instantaneous zero velocity at the top of the trajectory, and the p-prims called "dying

away" and "overcoming" to explain why the ball comes down again. The initial force gets smaller and smaller ("dying away" p-prim) until at the top of the trajectory it balances gravity ("dynamic balance" p-prim). Gravity then "overcomes" the initial force to bring the ball down ("overcoming" p-prim).

No matter what form students' alternative conceptions take, whether they are coherent or in pieces, students have them. In this study, students' responses will be analyzed by judging the extent to which they deviate from the accepted theory, but no attempt will be made to decide whether students' conceptions are in the form of a coherent theory or not.

Some of the areas in which students' alternative conceptions have been studied are the differentiation of velocity and acceleration, the nature of forces, and Newton's Second and Third Laws. Each of these is reviewed briefly below.

**Velocity And Acceleration**

The major studies of students' conceptual understanding of kinematics have been done by Lillian McDermott and the Physics Education Group at the University of Washington. An understanding of kinematics is essential for the study of physics, especially the study of dynamics (Trowbridge and McDermott, 1980). Students who have never studied physics have only vague meanings of kinematics terms such as position, time, speed, and acceleration. In fact, everyday language often interchanges the terms speed and acceleration (Trowbridge and McDermott, 1980).

Trowbridge and McDermott (1980) explored the understanding that students taking an introductory physics class at the University of Washington had of velocity. Students were

interviewed as they did tasks involving comparison of the velocities of two balls rolling down inclines.  They found that students often confused speed with position, thinking that two balls that were in the same place had the same speed (Trowbridge and McDermott, 1980).  They also found that students had a great difficulty with the concept of instantaneous velocity.  When asked about speed "at an instant," some students replied that there is no such thing -- there needs to be a time interval in order for there to be a speed.  Trowbridge and McDermott interpreted this result to mean that students have learned the definition of speed as a finite distance divided by a finite time, but cannot understand instantaneous velocity.

In a second study, students were asked to compare accelerations in different situations (Trowbridge and McDermott, 1981).  One of the three interview tasks is especially worth noting, since it is used in written form as part of this study (see Figure 3.7, page 66).  A ball rolls up an incline with an initial velocity, reaches the top, and comes down the incline to its starting point.  Students were asked to describe the acceleration throughout the motion, and (if they did not bring it up themselves) they were also asked specifically about the acceleration at the top of the ramp.  Students demonstrated a great deal of confusion between acceleration and velocity, showing that they did not know the difference between them.  For example, most students said that the ball has no acceleration at the top of the ramp since it is not moving.  Overall, in Trowbridge and McDermott's sample of 39 students from a calculus based physics course, 36% were successful with the task before instruction and 64% were successful after instruction. This study will use a similar problem to look at any gender differences in students' understanding of acceleration.

**Nature Of Forces**

Students' knowledge of physics terms is so vague that many interchange the terms "velocity" and "acceleration" or "force," "energy," and "power" as though they cannot distinguish their meaning (Trowbridge and McDermott, 1980; 1981; Hestenes, Wells, and Swackhamer, 1992). Several studies have investigated students' understanding of the nature of forces.

Students tend not to know that forces arise only from interactions between two objects. That is, only objects can push and pull on each other. They often use phrases like "force of motion," "momentum force," or "inertia force" to describe forces on moving objects (Hestenes, Wells, and Swackhamer, 1992).

For example, Clement (1983) found that college students seemed attached to the idea of an active force. He presented students with the problem of a coin tossed in the air and had them draw a free body diagram of the coin. In a physics class of 34 engineering majors, 88% of them got this problem wrong before instruction, and 90% of those incorrect students had drawn a force pointing upwards on the coin (Clement 1983, p. 328). In interviews, various students identified this force as the "force of the throw," "the upward original force," the "applied force," "the force up from velocity," or the "force that I'm giving it" (Clement, 1983, p. 328). These are "pseudoforces;" students call them forces but they are not interactions between two objects.

In another study, college students were again found to use the term "force" for many different things. In a situation where college students were searching for the cause of motion for situations on the Mechanics Diagnostic Test, they cited such causes as "a force

of inertia," "a potential force," "the force of velocity," and "the force behind it . . . coming from the throw" (Halloun and Hestenes, 1985b, p. 1059).  None of these cited forces are described as interactions between objects -- they are also "pseudoforces."

One question in this study is whether students will identify "pseudoforces" when talking about  a different problem.  Possible gender differences in the understanding of the nature of forces are also explored.

**Newton's Second Law**

Newton's Second Law states that the vector sum of forces acting on an object will cause that object to accelerate in the direction of the sum of forces.  The numerical value of this acceleration is the vector sum of the forces divided by the mass of the object.  Students' lack of discrimination between velocity and acceleration logically leads to alternative conceptions of Newton's Second Law (see Trowbridge and McDermott, 1980; 1981).

Many physics students believe that when an unbalanced force is applied to an object, the velocity, not the acceleration, of that object will vary directly with the force (Clement, 1983; McClosky, 1983; Champagne, Gunstone, and Klopfer, 1985).  Some researchers believe that students, like Aristotle, have a coherent theory that force is directly proportional to speed (Champagne, Gunstone, and Klopfer, 1985).  Some believe that students are using the pieces-of-knowledge, or p-prims, of "force as a mover" and "continuous push."  Students know from their experience in a world full of friction that objects need continuous forces to keep them moving (diSessa, 1993).

There are no studies that specifically investigate students' knowledge that it is the sum of forces that is proportional to acceleration. However, there are situations where students start to articulate their beliefs about sums of forces. One of these came out of Halloun and Hestenes' interviews of college physics students about their answers to the Mechanics Diagnostic Test. Although no one quoted in their paper communicated the Second Law, some students did compare the magnitude of forces to the magnitudes of other things:

"The speed is equal to the force of pull."

"The initial velocity is greater than the force."

"The energy of blast has to be greater than the force."

(Halloun and Hestenes 1985b, p. 1059).

These students seem to have some idea of "overcoming," that something has to be bigger than something else for there to be motion. Some students also believe that one force in a Third Law pair can "overcome" another, as will be discussed below. Students with these ideas about "overcoming" seem to have some idea that something has to be bigger than something else for there to be motion. This is still several steps away from the Second Law, however.

This study looks explicitly at students' understanding of the idea that it is the sum of forces acting on an object that cause its acceleration. It is the first known to do so. Gender differences in the understanding of Newton's Second Law are explored, with special attention given to students' ability to communicate their knowledge that acceleration in any direction is caused by the *sum* of the forces in that direction.

**Newton's Third Law**

Newton's Third Law is hard to understand; in fact, some physics educators claim that it may be the hardest law of motion to learn (Brown, 1989; Maloney, 1994). The idea that interacting objects exert forces of equal magnitude on one another seems to contradict many common-sense beliefs. It can be hard to believe that stationary objects exert forces at all. Minstrell (1991) claims that one commonly held facet or idea is that "passive objects don't exert forces," and they use this facet to explain, for example, that there is no force from your chair on you when you are sitting in it.

For objects that are moving, the Third Law is still hard to believe. As was mentioned above, some students articulate beliefs about the Third Law when they are talking about what causes acceleration. Many students believe that when two objects interact one object exerts more force than the other (e.g., Hestenes, Wells, and Swackhamer, 1992; diSessa 1993). This seems to be common sense, learned from experience combined with a confusion between forces and the effects of forces. Some researchers explain these student responses by saying that students have a coherent theory based on something like a "dominance principle" -- when two objects are in "conflict," the more forceful object exerts the larger force. The more forceful object is the one that is heavier, larger, faster, or more active than the other (Hestenes, Wells, and Swackhamer, 1992). Others believe that students are applying an "overcoming" p-prim. From everyday life, students see that larger, faster objects can somehow "overcome" smaller, slower ones (diSessa, 1993).

**Force Concept Inventory**

The Force Concept Inventory was selected for this study because it is a well established, nationally recognized multiple choice test of students' conceptual understanding of mechanics. It is a 29 question multiple choice test designed to measure students' conceptual understanding of mechanics (Hestenes, Wells, & Swackhamer, 1992). Through repeated administrations of the test with both high school and college students, the authors of the test have demonstrated that the results of the Force Concept Inventory are consistent (Hestenes, Wells, & Swackhamer, 1992). While the authors of the test suggest that the Inventory measures student conceptions of several different aspects of force and mechanics, a recent debate in The Physics Teacher brings that into question (Heller & Huffman, 1995; Hestenes & Halloun, 1995; Huffman & Heller, 1995). For the purposes of this study, only the total score of the Inventory will be used. The test is designed in such a way that each question includes one Newtonian answer and several non-Newtonian answers that introductory physics students often believe are true. These alternatives were selected based on data from student interviews (Halloun & Hestenes, 1985; Hestenes, Wells, & Swackhamer, 1992). Furthermore, the authors of the test have also conducted follow-up interviews with students to confirm that responses to individual items on the Inventory closely match the responses in an interview situation (Hestenes, Wells, & Swackhamer, 1992, Halloun & Hestenes, 1985).

**CHAPTER 3: METHODS**

The specific research questions are as follows:

1. If there are minimal differences between men and women in their relevant physics background and initial performance when they start an introductory physics course which was designed to appeal to a broad population, will there be differences in how much physics they learn by the end of the course?

2. Will there be differences in the opinions of men and women about an introductory physics course which was designed to appeal to a broad population?

One section of Physics 1251 in the fall of 1993, called the experimental section in this study, was designed to appeal to a broad population. The students learned in cooperative groups in their laboratory and problem solving sessions, and they were exposed to an explicit problem solving strategy. This study also includes baseline data from a traditional section of the same course from the same quarter.

To answer the first research question a matched sample of students was studied. Men were matched to women on their high school math courses, science courses, and grade point average as well as on their locus of control over their grades and on three pretest physics scores. Then their scores on their problem solving final exam, the Force Concept

Inventory, and a free-response conceptual test given at the end of the course were compared.

It was not possible to use the same matched sample to answer the second research question, since the evaluations students filled out at the end of the course were anonymous. However, students did indicate their sex and section number on the evaluations, so it was possible to compare responses of men and women in each section. Questions about the laboratory, problem solving sessions, and problem solving methods were concentrated on, since they were the aspects of the experimental course that were changed so that the course would appeal to a broader population.

Descriptions of the research methods are organized in different sections in this chapter. Included are sections on the course and instruction, the instruments and analysis used, the student population, and the selection of the matched sample.

**The Course**

This study was conducted during the first quarter of a three quarter introductory, calculus based physics course for scientists and engineers at the University of Minnesota. There were three sections of the course taught. Each section was taught by a team of two professors and several graduate student teaching assistants. Students in all three sections had three hours of lecture a week from the professors. In addition, they enrolled in a one hour problem solving session and a two hour laboratory session from the teaching assistants. They also had a one hour exam each week.

The lectures were similar in the traditional and experimental sections, but they did have differences. The professors had different styles, plus one of the professors in the experimental section consistently modeled an explicit problem solving strategy. The main difference between the sections was in the learning environment of the sessions run by the teaching assistants. Students in the intermediate and the experimental sections worked in cooperative groups during the laboratory and problem sessions while students in the traditional did not. In addition, the students in the experimental section were exposed to an explicit problem solving strategy, which the students in the other two sections were not. For the purposes of this study, the traditional and the experimental sections, the most different from one another, will be compared.

**Lecture**

**Content.** The syllabuses for the two sections were roughly the same. The traditional section started week one of the ten week quarter going over the "Tooling Up" chapter of the textbook, Fishbane, Gasiorowicz and Thornton's <u>Physics for Scientists and Engineers</u> (1994). The remainder of week one and week two were spent on motion in a straight line. Week three and half of week four was spent on motion in a plane, and then discussion of Newton's laws started. Discussion of Newton's laws and their applications continued through the first part of week six, when discussion of kinetic energy and work started and continued through week eight. In week nine, potential energy and conservation of energy were discussed, and in week ten the traditional section discussed momentum, collisions, and conservation of momentum.

The experimental section also spent the first two weeks discussing straight line motion, and also discussed the problem solving strategy that they were to use in the course. Then,

like the traditional section, they went on to discuss motion in a plane during weeks three and four. Weeks five through eight were spent on Newton's laws and their applications. Instead of spending the last weeks of the course on conservation laws, weeks nine and ten were spent on rotational dynamics.

**Style.** Both sections had three fifty-minute lectures a week from one of the two professors teaching each section. The author has discussed lecturing styles with a professor from each course over electronic mail, and has discovered some similarities and some differences in the styles the professors say were used during the term being studied.

The professors in each section gave lectures that were related to but independent from the course text. They tried to do demonstrations nearly every day. In the traditional section, short problems were done in front of the class every day, and in the experimental section they were done almost every day. Each professor also did long problems in front of the class, taking most of a class period to do them, several times during the quarter. Professors in each section tried to liven up the period with demonstrations and funny stories.

There were some differences in the lecture styles between the traditional section and the experimental section. The professors in the experimental section had class discussion, asked students to talk to each other about questions raised in lecture, and had students turn in in-class problems more than once a week. The professors in the traditional section felt that "students don't have time" for these activities. They felt that the best use of their time was to lecture to and entertain the students, one saying that "I always try to be as funny and entertaining as possible to keep the students awake. Being a professor is 50% entertainment, 50% education, but the entertainment fosters the education."

**Description of Problem Solving Strategies**

The students in the experimental section were taught how to solve physics problems using an explicit problem solving strategy developed by Heller & Heller (1992). The five steps of the strategy are: (a) Focus the Problem, (b) Describe the Physics, (c) Plan the Solution, (d) Execute the Plan, and (e) Evaluate the Solution.

1.      *Focusing the Problem* involves translating it from the given written statement into a picture and a written description. This step has three parts: a sketch of the problem situation which includes the given information, a question that restates what the problem is asking, and a suggested approach for solving the problem.

2.      *Describing the Physics* involves translating the sketch into an idealized physics diagram of the situation. Where the sketch in Focusing the Problem might have stick figures and sketches of the environment of the problem situation, the idealized diagram will include only the most important problem elements, represented by simple dots or boxes. This step has three parts: a physics diagram with definitions of variables, a target variable for which the student will solve, and the suggested quantitative relationships to use to solve the problem.

3.      *Planning a Solution* involves translating the physics description into the mathematical equations which will be used to solve the problem. This step has three parts: constructing the specific equations, checking to see if there are as many equations as there are unknowns, and outlining the order in which the equations can be solved in order to isolate the target variable.

4.	*Executing the Plan* involves following the planned outline.  Students are to avoid substituting numbers for variables until the last step.  When they reach the last step, where they have an algebraic equation with the target variable on one side and known variables on the other side, they substitute numbers for the known variables and calculate a numerical answer.

5.	*Evaluating the Solution* is the last step.  Students must check their solution  and answer to see if it is complete, clear, and reasonable.  This includes checking the original question to see if it has been answered, checking to see if the original target variable has been found, checking to see if the answer is in appropriate units, and checking the numerical answer with anything they know from their experience to see if it is reasonable.  The explicit problem solving strategy is shown in Figure 3.1.

The students in the traditional section were not given an explicit problem solving strategy to use.  If they were unsure of how to solve a problem, they could looking at the suggestions in their textbook.  Most physics texts include some general problem solving tips.  The book being used by both sections of the course, Fishbane, Gasiorowicz, and Thornton's <u>Physics for Scientists and Engineers</u> (1994), has some suggestions which are listed in Figure 3.2.

These suggestions do have some resemblance to the explicit strategy used by the experimental section.  The first four suggestions, "read the problem, then read it again", "draw a sketch or diagram of the problem", "write down the given and known quantities", and "understand which quantities are to be found", correspond to roughly the first two steps of the explicit strategy, Focus the Problem and Describe the Physics.  Step three of the strategy, Plan a Solution, corresponds to the fifth and part of the sixth suggestion, "think about which principles link the quantities to be determined to those that are

known" and "find the equation or equations that contain the quantities in the problem."

The rest of

Figure 3.1:  Problem Solving Strategy for the Experimental Section

1 Focus the Problem

 • Picture & Given Information

 • Question(s)

 • Approach

2 Describe the Physics

 • Diagram & Define Variables

 • Target Variable(s)

 • Quantitative Relationships

3 Plan the Solution

 • Construct Specific Equations

 • Check for Sufficiency

 • Outline the Math Solution

4 Execute the Plan

 • Follow the Plan

 • Calculate Target Variable(s)

5 Evaluate the Answer

 • Is Answer Properly Stated?

 • Is Answer Reasonable?

 • Is Answer Complete?

(Heller & Heller, 1992)

the sixth suggestion, "the rest is mathematics!" and the seventh, "wait until the end to put in numbers and units', correspond to the fourth step of the explicit strategy, Execute the Plan.  And the final step of the explicit strategy, Evaluate the Solution, corresponds to the last two suggestions, "when you get to a number, think about it" and "use any checks you can find for your result" (Fishbane, Gasiorowicz, & Thornton, 1994; Heller & Heller, 1992, Heller, Keith, & Anderson, 1992).

Figure 3.2:  Problem Solving Tips from Fishbane, Gasiorowicz, and Thornton

1.  Read the problem, then read it again.  Failure to read the problem is perhaps the source of more false starts and wrong answers than is any other cause.

2.  Draw a sketch or diagram of the problem that will help you visualize the situation presented by the problem.

3.  Write down the given and known quantities.

4.  Make sure you understand which quantities are to be found.

5.  There are usually only a few principles applicable to the solution of a problem. Think about which principles link the quantities to be determined to those that are known.

6.  Use the principles that apply to the situation to guide you to the equation or equations that contain the quantities in the problem.  Pay attention to when certain equations apply and when they do not.  The rest is mathematics! Sometimes, several of the equations need to be manipulated together.  Count the number of equations available to see if there are enough equations to determine the unknowns.

7.  When you solve for an unknown in terms of the known quantities, use symbols, not numbers.  Wait until the end to put in numbers and units.  It is important to include units, both because the answer may require them and because the proper cancellation of units will provide a check.

8.  When you get to a number, think about it.  Does it make sense?  If you find that it takes about 3 minutes to drive from New York City to Los Angeles, you have probably made a mistake!

9.  Use any checks you can find for your result.

<div align="right">Fishbane, Gasiorowicz, & Thornton, 1994</div>

**Problem Solving (Recitation) Sessions**

The problem sessions in the traditional section were similar to traditional recitations. Every week, the students were given a practice quiz problem, which they solved individually and without help from the teaching assistant. For the remainder of the fifty minute period, the students asked questions about the homework they had just completed and the teaching assistants did problems on the blackboard.

The experimental section of the course used cooperative groups in the laboratory and problem solving sessions. Cooperative groups differ from traditional groups in that they are carefully structured and managed to maximize the active and appropriate participation of all group members (Johnson & Johnson, 1989; Johnson, Johnson, & Smith, 1991; Heller & Hollabaugh, 1992). In addition, having students work in cooperative group helps the teaching assistants (Heller & Heller 1993). It can be easier for the teaching assistant to interact with 5 to 7 groups than with 15 to 21 individual students. When students work together in cooperative groups and discuss what they are doing with one another, teaching assistants have a chance to eavesdrop on them and learn what their misconceptions are.

The cooperative learning used by the experimental section in the laboratory and problem sessions was based on the Johnson and Johnson model of cooperative learning (Johnson, Johnson, & Smith, 1989). In this model there are five aspects of cooperative learning: a) positive interdependence, b) face-to-face interaction, c) individual accountability, d) interpersonal skills, and e) group processing. Three of these aspects were consistently

represented in the laboratory and problem solving session of the experimental section positive interdependence, face-to-face interaction, and individual accountability.

Positive interdependence means that students are dependent upon each other such that if one does well, the other will do well or if one does poorly, the other will do poorly also. The main way in which this was represented in the experimental section is by the giving of group grades. One problem on each of the four problem solving tests given during the quarter was done as a group during the problem solving session. When combined with joint laboratory reports, in all about fifteen percent of each student's grade was based on work they had done in a group and received a group grade on. The other way in which positive interdependence was represented in the experimental section was by the assigning of group roles. At the beginning of the quarter, teaching assistants would assign members of a three person group the roles of manager, checker/recorder, and skeptic when they were working together. If there was a fourth member in a group, he or she was assigned the role of energizer. When students were assigned roles they knew that they could count on each other to perform specific tasks within the groups, and when all tasks were performed the group would function better. After a few weeks of the course, when the teaching assistants judged that the students were working in groups fairly well, they stopped assigning roles to students.

The second aspect of cooperative grouping is face-to-face interaction. When members of a group can see each other and look each other in the eye, they are more likely to work well together. This was implemented in the experimental section partly by keeping the groups small: the teaching assistant assigned three students to each group (there would be one group of four if the class was not divisible by three). The other way in which this was implemented was by having the students face one another. Most of the classrooms

where the problem sessions were held had moveable chairs with attached desks, and these could be easily moved into clusters of three or four. In the laboratory sessions, students sat around tables designed for four students.

The third aspect of cooperative learning is individual accountability. It is important that students not feel that, because they are working in a group, they can be lazy and let the other members of their group be the only responsible ones. To enforce individual accountability, eighty-five percent of a student's course grade in the experimental section was based on work and exams where he or she had worked individually.

The fourth aspect of cooperative learning is the teaching of interpersonal skills. Some people need to be reminded of skills like listening well to other group members, complimenting them on their ideas, and not insulting anyone. This training in interpersonal skills was never done formally in the experimental section. When a teaching assistant saw extreme problems with a group or an individual in a group, he or she would speak to the person or group with the problem.

The fifth aspect of cooperative learning is group processing. Groups often need to take time to think about what they do well as a group, what they do poorly as a group, and how they could improve. Most teaching assistants in the experimental section were reluctant to take time from their laboratory or problem solving sessions for group processing and only did it once in the quarter, usually in the second week.

**Laboratory Sessions**

In the traditional section, students worked together and shared laboratory equipment, but they were not cooperatively grouped. The students in each section had four sets of laboratory exercises in the ten week quarter, with each set of exercises lasting two or three weeks. In the traditional section, the four sets dealt with the following topics: Motion in One Dimension, Projectile Motion and Free Fall, Forces, and Conservation of Energy and Momentum. The teaching assistant chose which exercises the students did. Students could complete two exercises in a typical two hour laboratory session. At the end of the set of exercises, the teaching assistant would assign each student one exercise to write up as a lab report for an individual grade.

In the experimental section, the students worked in the same cooperative groups they had in the problem sessions. They had one different lab from the students in the traditional section. Like the students in the traditional section, their first three labs were Motion in One Dimension, Projectile Motion and Free Fall, and Forces. Since the last weeks of the course dealt with circular motion and not conservation of energy and momentum, correspondingly the last lab of the course was Circular Motion. As in the traditional section, the teaching assistant chose which exercises the students did and students could complete two exercises in a typical two hour laboratory session. But in the experimental section, at the end of the set of exercises the teaching assistant would assign group lab reports which were given a group grade.

**Preparation of the Teaching Assistants**

The teaching assistants for both sections went through an orientation during the week before classes started in September. For the first few days the teaching assistants for both the experimental and the traditional sections were in their orientation together and then,

since they were to teach in different ways, the last parts of their orientations were separate.

There were three aspects of the first part of the orientation, when all the teaching assistants were together.  First, they were taught about the goals of the course and their responsibilities that came with the job of being a teaching assistant.  Second, they learned about common difficulties students might have, including common alternative conceptions.  And third, in preparation for teaching their laboratory sessions, they were taught about the structure of the lab, they did two sample exercises and wrote up reports in the style that they would expect from their students.

The teaching assistants preparing to teach the experimental section went on to learn about cooperative group work, including the rationale behind it, how to form cooperative groups, and what sample lesson plans of their problem solving sessions might look like.  In addition, they learned the explicit problem solving strategy that they were to teach and grade for.

The teaching assistants preparing to teach the traditional section went on to learn presentation skills as preparation for teaching their problem solving session.  They practiced presenting problem solutions in front of their peers.  The teaching assistants from the experimental section did not learn presentation skills, and the teaching assistants from the traditional section did not learn about cooperative grouping or about the experimental section's problem solving strategy.

**Instruments and Analysis**

The students were assessed in several ways. To test their initial problem solving abilities, they took an ungraded pretest problem. The final exam for all sections contained four common problems, so that problem solving abilities could be compared among all students. Students' conceptual understanding was tested by two means before and after instruction. The first was a multiple choice test on forces, the Force Concept Inventory (Hestenes, Wells, & Swackhamer, 1992). The second measure was a collection of three free-response questions on acceleration and forces. The Force Concept Inventory was given again at the
end of the course, and the free-response questions were given again at the beginning of the next course in the sequence. All of these conceptual tests were ungraded.

**Problem Solving Tests**

Pretest problem solving ability was measured by the asteroid problem (see Figure 3.3). Students were to find the size of an asteroid. Students who had a sensible method for finding the size of the asteroid, whether in terms of its radius, diameter, volume, or surface area, earned a score of "high." Students who had a method for finding the mass, not size, of the asteroid earned a score of "middle." And students who had neither earned a score of "low."

The final exams for both the traditional and the experimental sections had four common problems on them, representing the common topics of the syllabuses. Both sections had spent considerable time on motion in one and two dimensions and on Newton's laws and their applications. Therefore, the four common problems on the final exam were as follows: the Balloon Problem, where students use their knowledge of motion in two dimensions to solve where a package will fall when dropped from a moving balloon; the

Modified Atwood Problem, where students use Newton's laws to find tension in a rope and the speed of the system at a certain time; the Enterprise Problem, where students apply Newton's laws to find the force a woman exerts on her seat in an amusement park ride;  and the Ski Problem, where students use both Newton's laws and kinematics to find how

Figure 3.3:  The Asteroid Problem

The following problem is for diagnostic purposes only and will not be graded.  It will give you an idea of what you will have to do in this course, and will help us tailor the course to your problem-solving needs.

You have been given a job working on a team which is examining telescope pictures for asteroids which might collide with the Earth.  In your orientation, your team was told how important these observations might be.  Current theories indicate that dinosaurs and may other organisms became extinct when the Earth was struck by a large asteroid.  65 million years ago dust from an asteroid impact was lofted into the upper atmosphere all around the globe, where it lingered for at least several months and blocked the sunlight reaching the Earth's surface.  On the dark and cold Earth that temporarily resulted, many forms of life became extinct.  In has been suggested that such an asteroid collision is likely to happen again, perhaps causing the collision is likely to happen again, perhaps causing the extinction of the current dominant life form on Earth, namely us.  As you scan space for danger, how large an asteroid should you be watching for if the dangerous asteroid size is roughly the same as the one that wiped out the dinosaurs?  Available evidence suggests that about 20% of that asteroid's mass ended up as dust spread uniformly over Earth after eventually settling out of the upper atmosphere.  About 0.020 $g/cm^2$ of dust, which is chemically different than the  Earth rock, covered the Earth's surface.  Typical asteroids have a density of about 2.0 $g/cm^3$.

Possibly Useful Information:

Earth Radius = 6380 km

Volume of sphere = $\frac{4}{3}\pi R^3$

1 km = $10^5$ cm

Surface area of sphere = $4\pi R^2$

Volume density = $\frac{mass}{volume}$

Surface density = $\frac{mass}{area}$

Please write down all of your reasoning as you solve this problem.

long it takes for a lost ski to catch up to a skier on the slopes. These problems are shown in Figure 3.4.

Written problem solutions were scored with a coding scheme developed at the University of Minnesota, based on research into how experts and novices solve problems. Solutions were coded on four measures of problem solving. First, their general approach, which involves how well students choose the principles to solve a problem. Second, their specific application of physics, which is how well those principles are applied. Third, their logical progression, or how organized and logical the solution is. And fourth, the use of appropriate mathematics. These four measures are discussed in more detail in Chapter 2, pages 31 - 32, and the different errors students made on each are shown in Appendix A.

After each response was categorized, the categories were assigned numerical values between 0 and 10, where "a" was always worth 0, the category representing correct responses was always worth 10, and the intermediate categories were assigned normalized ranks. The scores for General Approach, Specific Application of Physics, Logical Progression, and Applied Mathematics were added across each problem, giving each student four scores out of 40. The intra-rater reliability of this scoring scheme is 0.95. These total rank scores were analyzed using the Wilcoxin Matched Pairs Signed-Rank Test. They were analyzed separately for the traditional section and the experimental section since only differences within the sections were of interest. This statistical test was to determine whether there were differences in the post-test problem solving ability of men and women in matched samples.

This coding scheme, with scores for general approach, specific application of physics, logical progression, appropriate mathematics, and clear communication, is not biased toward students who have been taught one heuristic or another.  The scheme is based on

Figure 3.4: Final Exam Problems

1. The Enterprise Problem

The "Enterprise," a ride at Valley Fair, consists of a vertical wheel of radius 9 m rotating about a fixed horizontal axis with seats for the occupants around its outer edge. The wheel rotates so that the occupants are moving at 11 m/s. The seats pivot so that the occupants' heads are towards the center of the wheel. When a 56 kg woman is upside down at the top of the wheel, what is the force she exerts on the seat?

2. The Atwood Problem

In the diagram shown below, block 1 of mass 1.5 kg and block 2 of mass 4 kg are connected by a light taut rope that passes over a massless frictionless pulley. Block 2 is just over the edge of a ramp inclined at an angle of 30_, and the blocks have a coefficient of sliding friction of 0.21 with the surface. At t=0, the system is given an initial speed of 11 m/s that starts block 2 down the ramp.
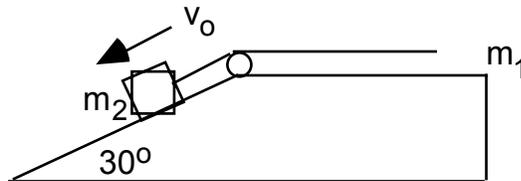


a. Find the tension in the rope.

b. Find the speed of the two masses at t = 2 s.

3. The Balloon Problem

As you sit in a park on a fine summer day, you notice a hot-air balloon approaching in the sky above you. It is losing altitude (its height above ground) at a constant rate of 1.5 m/s, and drifting toward your van with a breeze at a constant rate of 1.2 m/s. When the balloon is 30 m above the ground over the center of your van, someone in the balloon releases a bag of sand. The bag of sand falls toward your van. Your van is 4.0 m long and 2.5 m high. Perform a calculation to determine whether the bag strikes the rood of the van. You may ignore the effect of air resistance on the bag's motion.

4. The Ski Problem

Pat starts to ski down a smooth slope, which is inclined at 20_ to the horizontal and is 1.5 km long. Almost immediately, near the top, Pat falls and stops. He loses one ski and starts sliding down the slope. 4 seconds after his fall, the lost ski starts sliding without friction on its smooth bottom. If the coefficient of sliding friction between Pat and the snow is 0.12, then how long does it take after Pat's fall for the ski to catch up to him?

expert/novice research and not on any specific heuristic. Students who followed either of the heuristics discussed in this paper would receive a high score. Examples of good solutions from students from each section, along with their scores, are shown in Figure 3.5.

**Conceptual Tests**

**Force Concept Inventory.** The Force Concept Inventory is a 29 question multiple choice test designed to measure students' conceptual understanding of mechanics (Hestenes, Wells, & Swackhamer, 1992). For the purposes of this study, only the total score of the Inventory was used. The test is designed in such a way that each question includes one Newtonian answer and several non-Newtonian answers that introductory physics students often believe are true. These alternatives were selected based on data from student interviews (Halloun & Hestenes, 1985; Hestenes, Wells, & Swackhamer, 1992). A few of the questions are shown in Figure 3.6, and the full inventory is included in Appendix B.

Each of the questions was worth one point. Questions were scored as correct if the Newtonian response was selected, and incorrect if one of the non-Newtonian alternatives was selected. A complete copy of the Force Concept Inventory, with correct answers marked, is included in Appendix B. A matched-sample t-test was used to compare women and men in the traditional and experimental sections.

**Free Response Conceptual Questions.** Students' conceptual understanding of a few major topics in mechanics was further measured by three free response conceptual questions. Students' understanding of acceleration was addressed by student answers to a

Figure 3.5a: An Actual Student Solution to the Enterprise problem, Experimental Section
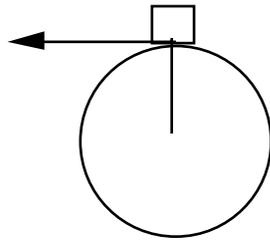
$r = 9$ m

$v = 11 \; \dfrac{m}{s}$

$m = 56$ kg

Question: What is the force on the seat?

Approach: Newton's laws

Free body diagram:

N

W      a
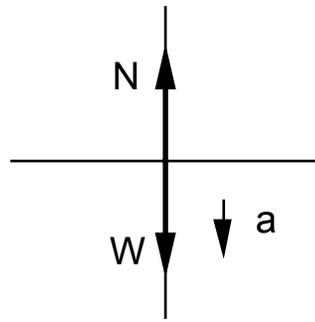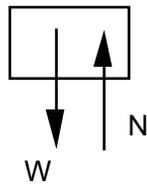
N = force of woman on the seat

Quantitative Relationships

$\Sigma F = ma, \quad a = \dfrac{v^2}{r}$

Plan:

|  | unknowns | Execution: |
|---|---|---|
| $\Sigma F = ma$ | a | $N - mg = ma$ |
| $\Sigma F = N - W$ | N, W | $N = ma + gm = m\,(a + g)$ |
| $W = mg$ |  | $N = m\left(\dfrac{v^2}{r} + g\right)$ |

$$a = \frac{v^2}{r}$$

$$N = 56 \text{ kg} \left( \frac{(11\text{m}/\text{s})^2}{9 \text{ m}} + 9.8 \frac{\text{m}}{\text{s}^2} \right) = 1302 \text{ N}$$

Figure 3.5b: An Actual Student Solution to the Enterprise Problem, Traditional Section
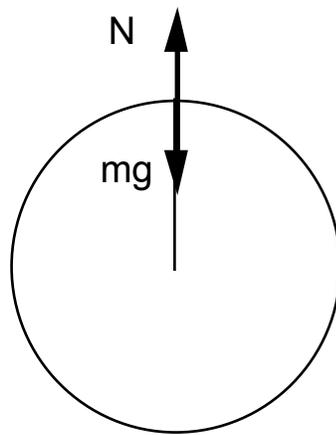
$$m = 56 \text{ kg}$$

$$r = 9 \text{ m}$$

$$v = 11 \; \frac{\text{m}}{\text{s}}$$

$$\Sigma F_y = N - mg = \frac{mv^2}{r}$$

$$N = \frac{mv^2}{r} + mg$$

$$N = \frac{(56 \text{ kg})(11 \text{ m / s})^2}{9 \text{ m}} + (56 \text{ kg}) \, (9.8 \; \frac{\text{m}}{\text{s}^2})$$

$$N = 1301.7$$

$$\boxed{N = 1300 \text{ N}}$$

Figure 3.5c:  Scoring for Both Examples


General Approach:  **category f**  (8.33 points)  The solution approach is mostly correct but a serious error is made about certain features of the physical events.  In this case, either students are confused about the direction of the normal force or they are confused about the distinction between the force the seat exerts on the woman and the force she exerts on the seat.


Specific Application of Physics:  **category h**  (10 points)  Specific equations do not exhibit clear inconsistencies with student's general physics approach and solution seems quite complete in its identification of quantities and their relative directions.


Logical Progression:  **category g**  (10 points)  Solution progresses from general principles to answer.  (Solution proceeds in a straightforward manner toward solution.)  Solution is successful in isolating desired unknown.


Appropriate Mathematics:  **category g**  (10 points)  Mathematics is correct.


The total score for both students, using the different problem solving techniques as taught in the experimental and traditional sections, would be 38.33 points out of 40.

Figure 3.6  Some Questions from the Force Concept Inventory
(correct answers appear in bold face)

2.  Imagine a head-on collision between a large truck and a small compact car.  During the collision,

   (A)  the truck exerts a greater amount of force on the car than the car exerts on the truck.
   (B)  the car exerts a greater amount of force on the truck than the truck exerts on the car.
   (C)  neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.
   (D)  the truck exerts a force on the car but the car doesn't exert a force on the truck.
   **(E) the truck exerts the same amount of force on the car as the car exerts on the truck.**

3.  Two steel balls, one of which weighs twice as much as the other, roll off of a horizontal Table with the same speeds.  In this situation:

   **(A) both balls impact the floor at approximately the same horizontal distance from the base of the Table.**
   (B)  the heavier ball impacts the floor at about half the horizontal distance from the base of the Table than does the lighter.
   (C)  the lighter ball impacts the floor at about half the horizontal distance from the base of the Table than does the heavier.
   (D)  the heavier ball hits considerably closer to the base of the Table than the lighter, but not necessarily half the horizontal distance.
   (E)  the lighter ball hits considerably closer to the base of the Table than the heavier, but not necessarily half the horizontal distance.

5.   A boy throws a steel ball straight up. **Disregarding any effects of air resistance** the force(s) acting on the ball until it returns to the ground is (are):

   (A)  its weight vertically downward along with a steadily decreasing upward force.
   (B)  a steadily decreasing upward force from the moment it leaves the hand until it reaches its highest point beyond which there is a steadily increasing downward force of gravity as the object gets closer to the earth.
   (C)  a constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point after which there is only the constant downward force of gravity.
   **(D) a constant downward force of gravity only.**
   (E)  none of the above -- the ball falls back down to the earth simply because that is its natural action.

question first posed by the physics education group at the University of Washington (Trowbridge and McDermott, 1981). In this question, known as the Ramp Question, a ball is sent up a ramp with an initial velocity, rolls to the top of the ramp, and rolls down. Students were asked to compare the accelerations on the way up the ramp and on the way down the ramp and to tell if there was an acceleration at the top of the ramp (see Figure 3.7). Correct understanding of acceleration is suggested by responses where a student knows that acceleration is the same throughout the problem and can explain why.

The second question was the Car and Bug Question. In this question, a car traveling at 55 miles per hour collides with a bug. Students were asked to draw and label all of the forces acting on the car and on the bug at the moment of impact, taking care to make the length of the arrows representing the forces correspond to the strength of the forces (see Figure 3.8). This question tests student understanding of Newton's third law. Correct understanding of the third law is suggested by responses with the force of the bug on the car equal to the force of the car on the bug.

The third question was the Car and Passenger Question. This problem, designed by Patricia Heller, is shown in Figure 3.9. In this question, a car is accelerating from 30 to 55 miles per hour. Students were asked to draw and label all of the forces acting on a passenger within the car and on the car itself and to explain what force or forces cause the car and the passenger to accelerate. The first part of the question, drawing and labeling the forces, served as an indirect measure of students' understanding of the nature of forces. A correct understanding of the nature of forces is suggested by forces that are all really acting on either the car or the passenger and that arise from an interaction between two objects. The second part of the question, explaining why the car and passenger

accelerate, is a more direct measure of students' understanding of Newton's second law.

A correct

Figure 3.7: Ramp Problem

Motion Up the Ramp

1    →    2

Motion Down the Ramp

3    ←    2
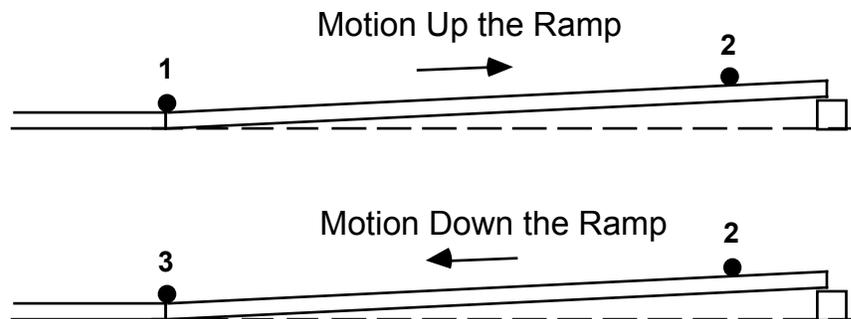
A steel ball is launched with some initial velocity, slows down as it travels up a gentle incline, reverses direction, and then speeds up as it returns to its starting point. Assume friction is negligible.

(a) Suppose we calculated the acceleration of the ball as it's moving *up* the ramp (from 1 to 2), and the acceleration as it's moving *down* the ramp (from 2 to 3). How would these two accelerations compare? (i.e., Are the accelerations the same size? The same direction?) Explain your reasoning.

(b)    Does the ball have an acceleration at its highest point on the incline (at position 2)? Explain your reasoning.

# Figure 3.8: Car and Bug Problem

A car traveling at 55 mph strikes a hapless bug and splatters it on the windshield.  On the left side of the table below, use arrows to draw all of the forces acting on the bug and the car at the moment of impact.  The length of the arrows should indicate the relative sizes of the forces  (i.e. a larger force should be represented by a longer arrow, equal forces should be represented by arrows of equal length).

To the right of the table, describe each force in words. (i.e. What kind of force is it? Is the force a push or pull? .What object, if any, is exerting the force? What object is being affected by the force?).
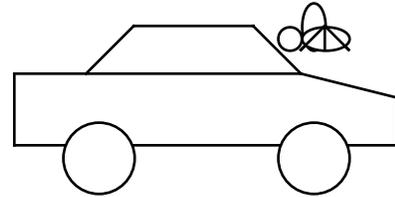
Note:  Bug is not drawn to scale.

| Bug | Description of Each Force |
|---|---|
| | |
| **Car** | **Description of Each Force** |
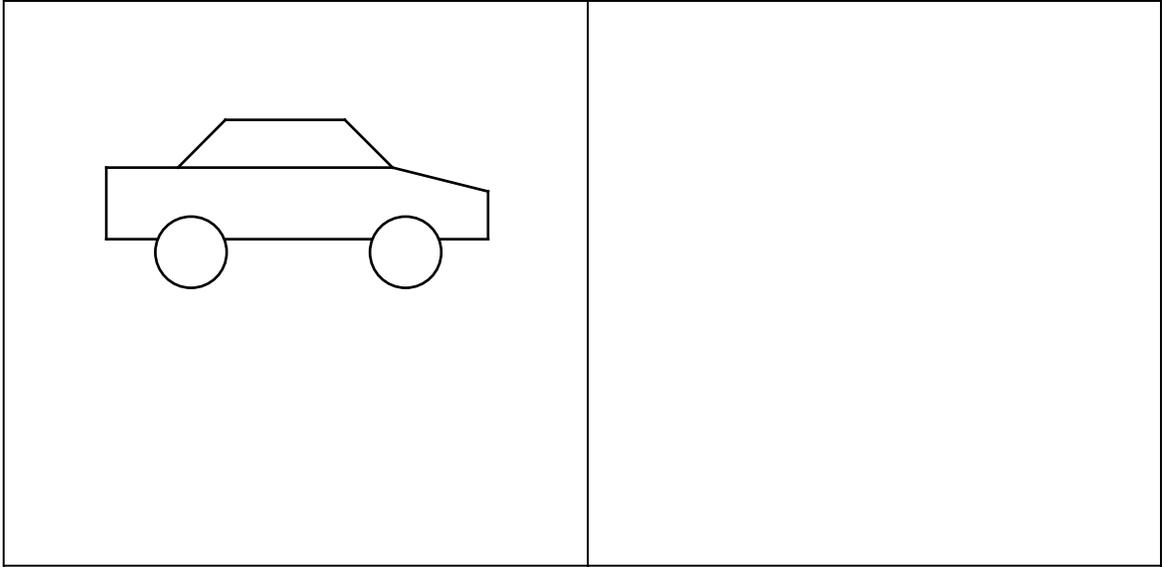
## Figure 3.9: Car and Passenger Problem

You are a passenger in a car which is traveling on a straight road while it is increasing speed from 30 mph to 55 mph. You wonder what forces cause you and the car to accelerate. When you pull over to eat, you decide to figure it out.

(a)  On the left side of the table below, draw and label arrows representing all the forces acting on *you* (passenger) while the car is accelerating. The length of the arrows should indicate the relative sizes of the forces (i.e., a larger force should be represented by a clearly longer arrow, equal forces by arrows of equal length). On the right side of the table, describe each force in words (i.e., What kind of force is it? Is the force a push or a pull? What object, if any, is exerting the force? What object is being affected by the force?).

| You (Passenger) | Description of Each Force |
|---|---|
|  |  |

(b)  Which force(s) cause **you** (passenger) to accelerate? Explain your reasoning.

# Figure 3.9, page 2

(c) On the left side of the table below, draw and label arrows representing all the forces acting on the *car* while it is accelerating. The length of the arrows should indicate the relative sizes of the forces (i.e., a larger force should be represented by a clearly longer arrow, equal forces by arrows of equal length). On the right side of the table, describe each force in words (i.e., What kind of force is it? Is the force a push or a pull? What object, if any, is exerting the force? What object is being affected by the force?).

| Car | Description of Each Force |
|---|---|
|  | |

(d) Which force(s) cause the **car** to accelerate? Explain your reasoning.

understanding of Newton's second law is suggested by students summing forces in the direction of motion and in the opposite direction.

In this look at conceptual understanding, the student answers to the free response conceptual questions were analyzed. This analysis was done in a similar manner as has been used by Rosiland Driver and Lillian McDermott in their research, and by the author in a masters degree thesis (Blue, 1992). It has several steps. First, student names were covered on the test papers to hide the identity and sex of the students. Then all student responses were read, and a list was made of all the different kinds of student responses. Then, for ease of comparison between questions, the responses were grouped into four broad categories, which are as follows:

1. Responses which include the accepted idea and a clear explanation.
2. Responses which are close to the accepted idea, but are either vague, unclear, partially incorrect, or include the accepted idea with no reasoning or explanation.
3. Responses which include alternative conceptions (this category will have many subcategories, which will probably be different for the different questions).
4. Uncodeable responses: responses that are either blank, incomprehensible, or otherwise cannot be categorized above.

The Car and Passenger problem yielded information about student understanding of two concepts. From their drawings and descriptions of forces on the passenger and on the car, inferences could be made about their understanding of the nature of forces. And from the questions, "Which force(s) cause **you** (passenger) to accelerate? Explain your reasoning," and "Which force(s) cause the car to accelerate? Explain your reasoning,"[1] student

---
[1]Emphasis in the original.

understanding of Newton's Second Law was gathered. The specific codes for the ramp problem, the car and passenger problem, and the car and bug problem, showing what kinds of alternative conceptions the students in this study held are listed in Appendix C.

A preliminary analysis of these questions showed that student understanding of both the nature of forces and Newton's Second Law was context-dependent: they showed different proficiencies on the passenger question and the car question. For these reasons the passenger and the car were analyzed separately.

It is important to note that student understanding of Newton's Second Law was analyzed separately from student understanding of the nature of forces. Because of this, a student could get quite a high score on their understanding of the Second Law even though they had a low score on their understanding of the nature of forces.

For each of the concepts, responses in the first category were given two points, those in the second category were given one point, and those in the third and fourth category were given no points. Since students were asked six questions all together (ramp, car and bug, nature of forces on both the passenger and the car, and Second Law on both the passenger and the car), their total possible score on free response conceptual question was 12 points. The intra-rater reliability of this scoring scheme was 0.93.

The total scores were analyzed using the Wilcoxin Matched Pairs Signed-Rank Test. They were analyzed separately for the experimental section and the traditional section since only differences within the sections were of interest. This test was to determine whether there was a difference in conceptual understanding between men and women in matched samples.

**Course Evaluations**

The course evaluations were given at the end of the quarter, in December 1993. Students were asked to evaluate their own performance, the lectures, problem solving sessions, and laboratory sessions, the problem solving methods, their teaching assistants, their professors, and the course as a whole. Students in the experimental section were cooperatively grouped in their laboratory and problem solving sessions and were exposed to an explicit problem solving strategy, and students in the traditional section were not. Therefore for the purposes of this study, only the student ratings of the laboratory, the problem solving sessions, and the problem solving methods of the course were compared. The questions to be analyzed in the study are shown in Figure 3.10 and the full course evaluation is included in Appendix D. Each of these questions was answered on a scale from -2 to +2, where -2=Strongly Disagree, 0=Neutral, and +2=Strongly Agree.

For uniformity of analysis, student answers to the "waste of time" questions were reversed so that all responses greater than 0 reflected positive answers. For example, a student who strongly disagreed (a score of -2) with the statement "The laboratory activities were a waste of time" would be expected to strongly agree (a score of +2) with the inverse statement. Then the answers to the eight questions were averaged, giving one score per student, between -2 and +2, reflecting how positively they evaluated the laboratory sessions, problem solving sessions, and problem solving methods of their course. Then t-tests were done on the averages to see if there were significant differences in how the males and females of each section evaluated these portions of their course.

Figure 3.10:  The Analyzed Questions from the Course Evaluation

**1=Strongly Disagree   2=Disagree   3=Neutral   4=Agree   5=Strongly Agree**

11.  The laboratory activities were interesting.

13.  The laboratory activities were a waste of time.

14.  The laboratory activities helped me understand how to solve physics problems.

15.  The recitation sessions were interesting.

17.  The recitation sessions were a waste of time.

18.  The recitation sessions helped me understand how to solve physics problems.

**The problem solving methods used in this course**

20.      were a waste of time.

21. helped me solve physics problems.

**Student Population**

The traditional section started out with 293 students and had 225 finish the course, and

the experimental section started out with 333 and had 268 finish.  There were no

significant differences in the background of the students in the two sections, as shown in

Table 3.1.

Table 3.1:  Demographic Data for Populations

| | Experimental Section | | Traditional Section | |
|---|---|---|---|---|
| | Men (n = 211) | Women (n = 57) | Men (n = 175) | Women (n = 50) |
| <ins>Last math course in high school</ins> | | | | |
| algebra, geometry, or trigonometry | 12% | 9% | 10% | 12% |
| precalculus, functions, or analysis | 31 | 40 | 35 | 34 |
| calculus | 57 | 51 | 55 | 54 |
| <ins>Physics in high school?</ins> | | | | |
| yes | 88 | 82 | 91 | 78 |
| no | 12 | 18 | 8 | 22 |
| <ins>High school GPA</ins> | | | | |
| 3.5-4.0 | 56 | 74 | 59 | 78 |
| 3.0-3.4 | 37 | 23 | 33 | 16 |
| 2.5-2.9 | 5 | 3 | 4 | 4 |
| 2.0-2.4 | 1 | 0 | 2 | 0 |
| below 2.0 | 1 | 0 | 0 | 2 |
| <ins>Grade level</ins> | | | | |
| freshman | 70 | 54 | 64 | 48 |
| sophomore | 16 | 26 | 25 | 22 |
| junior | 9 | 14 | 8 | 20 |
| senior | 4 | 6 | 3 | 8 |

Most of the students, both men and women, had completed a high school physics course, and about half had also completed a high school calculus course.  Most of the students had a high school GPA between 3.0 and 4.0 on a 4.0 scale, with the women's GPAs a bit higher than those of the men.  Most students in Physics 1251 were either freshmen or sophomores.

The pretest and post-test scores of the population of the students who took Physics 1251 in the fall of 1993 are shown in Table 3.2.  These data show that the population was fairly typical of past research, with women having lower pretest and post-test scores than men on measures of physics learning:  the Force Concept Inventory, the free-response conceptual tests, and problem solving tests.

In the experimental section, there is a significant difference between the pretest Force Concept Inventory scores of men and women (t (df = 219) = 2.82, p = 0.004).  This is also true in the traditional section (t (df = 188) = 2.01, p = 0.023).  In both sections the pretest scores of the men were higher than those of the women.  Men also had higher average post-test Force Concept Inventory scores than the women in both sections.  Although the differences were not as large as they were on the pretest, they were significant in both the experimental section (t (df = 168) = 1.23, p = 0.120) and the traditional section (t (df = 139) = 1.78, p = 0.037).

Since the number of men and women in the population who took the test is small, the Force Concept Inventory results were checked with those from 1995.  This course was taught in the same way as the experimental course was in 1993, with cooperative grouping.  The 85 women in the course had an average pretest score of 10.9 and an average post-test score of 17.1, while the 269 men in the course had an average pretest

score of 14.3 and an average

post-test score of 20.6. In this course, as in the 1993 course, there was a large and significant difference in both the pretest scores (t (df = 352) = 6.32, 0.001 > p) of men and women and their post-test scores (t (df = 352) = 5.98, 0.001 > p).

There were also differences in the scores of men and women on the free response conceptual question, both pretest and post-test, although not all differences were statistically significant. The median pretest scores in both sections were 0 out of 12, and although the men in both sections had greater ranges of scores, and greater percentages of men than women in both sections had scores above 0, the difference was not statistically significant (0.20 > p > 0.10 for each section). The post-test scores were a bit higher. In the experimental section, the median was 4 out of 12, and a significantly higher proportion of men than women scored above 4 points (p < 0.001). In the traditional section the median was 2, and although a higher proportion of men than women scored above 2 points the difference was not significantly significant (0.30 > p > 0.20).

The scores on the asteroid problem, the pretest measure of problem solving ability, were also statistically significantly higher for men than for women in both the experimental ($\chi^2$ = 8.96, 0.02 > p > 0.01) and traditional sections ($\chi^2$ = 11.87, 0.01 > p > 0.001). On the final exam, the scores of the men were slightly but not significantly higher than those of the women in either the experimental section ($\chi^2$ = 2.53, 0.20 > p > 0.10) or the traditional section ($\chi^2$ = 0.33, 0.70 > p > 0.50).

These results indicate that the population of students in both the experimental and the traditional sections is not unusual. As expected from previous research, men outperform women on both pretest and post-test measures of both conceptual understanding and problem solving ability, although not all these differences are statistically significant.

## Selection Of Matched Sample

In the population of students taking Physics 1251 the men have higher pretest scores on all three pretest measures of physics knowledge. The first research question of this study asks how post-test scores of men and women would compare if there were minimal differences between them at the start of the course, so it is necessary to match students to get a sample for study. There were far fewer women than men in each section of the course. All women who had taken the three pretests and the three post-tests and filled out a demographic questionnaire were chosen. Then, men on whom there was also complete data were matched to these women on seven measures; three pretest scores, three high school background characteristics, their year in college, and their locus of control over their grades. With so many variables, it was impossible to pair identical men to the women. The matching criteria are discussed below.

First, men and women were matched on their pretest scores. As shown on Table 3.3a, 90% of the pairs have Force Concept Inventory scores within 2 points of each other, and about 30% had the same scores. In addition, 90% have scores within two points of each other on the free response conceptual questions, and again 30% had the same scores. The asteroid problem was scored on a scale of high, medium, and low, and about 75% of the pairs were men and women with the same score, with only one pair of a high scorer with a low scorer.

Students were also matched on three aspects of their high school background. It was possible to always match students who had taken high school physics together and those students who had not with each other. As seen on Tables 3.3b and c, more than half the

Table 3.3a:  Pretest Data on Matched Pairs, Experimental and Traditional Sections

| Experimental Section | | | | | Traditional Section | | | | |
|---|---|---|---|---|---|---|---|---|---|
| pair | sex | FCI | free response | asteroid | pair | sex | FCI | free response | asteroid |
| 1 | M | 7 | 0 | medium | 21 | M | 10 | 0 | high |
| 1 | F | 9 | 0 | medium | 21 | F | 8 | 2 | high |
| 2 | M | 9 | 2 | low | 22 | M | 9 | 1 | high |
| 2 | F | 10 | 1 | medium | 22 | F | 9 | 0 | high |
| 3 | M | 9 | 0 | low | 23 | M | 12 | 2 | high |
| 3 | F | 11 | 1 | medium | 23 | F | 9 | 0 | high |
| 4 | M | 12 | 2 | medium | 24 | M | 15 | 0 | high |
| 4 | F | 12 | 0 | medium | 24 | F | 12 | 3 | high |
| 5 | M | 8 | 1 | high | 25 | M | 10 | 1 | high |
| 5 | F | 8 | 3 | high | 25 | F | 13 | 0 | high |
| 6 | M | 13 | 2 | high | 26 | M | 15 | 2 | low |
| 6 | F | 11 | 1 | high | 26 | F | 12 | 0 | low |
| 7 | M | 13 | 3 | high | 27 | M | 9 | 2 | medium |
| 7 | F | 11 | 2 | high | 27 | F | 12 | 4 | low |
| 8 | M | 12 | 0 | high | 28 | M | 12 | 1 | low |
| 8 | F | 12 | 0 | high | 28 | F | 14 | 1 | low |
| 9 | M | 14 | 0 | high | 29 | M | 15 | 0 | low |
| 9 | F | 14 | 1 | high | 29 | F | 14 | 1 | low |
| 10 | M | 14 | 2 | high | 30 | M | 9 | 1 | low |
| 10 | F | 14 | 2 | high | 30 | F | 4 | 1 | low |
| 11 | M | 19 | 1 | high | 31 | M | 9 | 3 | low |
| 11 | F | 16 | 3 | high | 31 | F | 10 | 0 | low |
| 12 | M | 20 | 1 | high | 32 | M | 11 | 0 | low |
| 12 | F | 20 | 0 | high | 32 | F | 12 | 0 | low |
| 13 | M | 6 | 0 | medium | 33 | M | 16 | 1 | high |
| 13 | F | 8 | 0 | low | 33 | F | 13 | 0 | medium |
| 14 | M | 9 | 0 | low | 34 | M | 10 | 2 | high |
| 14 | F | 10 | 0 | low | 34 | F | 16 | 3 | medium |
| 15 | M | 15 | 1 | low | | | | | |
| 15 | F | 12 | 1 | low | | | | | |
| 16 | M | 12 | 2 | low | | | | | |
| 16 | F | 12 | 0 | low | | | | | |
| 17 | M | 16 | 3 | low | | | | | |
| 17 | F | 16 | 4 | low | | | | | |

| 18 | M | 9 | 4 | high |
|----|---|----|---|--------|
| 18 | F | 5 | 1 | low |
| 19 | M | 9 | 1 | medium |
| 19 | F | 9 | 1 | high |
| 20 | M | 13 | 0 | high |
| 20 | F | 11 | 0 | high |

Table 3.3b: Demographic Data on Matched Pairs, Experimental Section

| pair | sex | last hs math | physics? | GPA | year | locus of control |
|------|-----|--------------|----------|-----|------|------------------|
| 1 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 1 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 2 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 2 | F | calculus | yes | 3.5-4.0 | sophomore | low |
| 3 | M | precalc | yes | 3.5-4.0 | sophomore | average |
| 3 | F | precalc | yes | 3.5-4.0 | sophomore | average |
| 4 | M | precalc | yes | 3.0-3.4 | freshman | average |
| 4 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 5 | M | calculus | yes | 3.0-3.4 | freshman | average |
| 5 | F | calculus | yes | 3.5-4.0 | freshman | high |
| 6 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 6 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 7 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 7 | F | precalc | yes | 3.0-3.4 | freshman | average |
| 8 | M | precalc | yes | 3.0-3.4 | sophomore | average |
| 8 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 9 | M | calculus | yes | 3.0-3.4 | freshman | low |
| 9 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 10 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 10 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 11 | M | precalc | yes | 3.0-3.4 | sophomore | average |
| 11 | F | calculus | yes | 3.5-4.0 | junior | average |
| 12 | M | calculus | yes | 3.0-3.4 | freshman | average |
| 12 | F | calculus | yes | 3.0-3.4 | freshman | average |
| 13 | M | precalc | yes | 3.5-4.0 | freshman | average |
| 13 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 14 | M | precalc | yes | 3.5-4.0 | freshman | average |
| 14 | F | precalc | yes | 3.5-4.0 | sophomore | average |
| 15 | M | precalc | yes | 3.5-4.0 | junior | average |
| 15 | F | precalc | yes | 3.5-4.0 | junior | low |
| 16 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 16 | F | precalc | yes | 3.5-4.0 | freshman | average |
| 17 | M | calculus | yes | 3.5-4.0 | freshman | high |
| 17 | F | calculus | yes | 3.0-3.4 | freshman | average |
| 18 | M | trigonometry | no | 3.0-3.4 | junior | average |
| 18 | F | precalc | no | 3.0-3.4 | senior | average |
| 19 | M | precalc | no | 3.0-3.4 | freshman | average |

| 19 | F | calculus | no | 3.5-4.0 | sophomore | average |
|----|---|----------|-----|---------|-----------|---------|
| 20 | M | calculus | no | 3.5-4.0 | sophomore | average |
| 20 | F | precalc | no | 3.0-3.4 | sophomore | average |

Table 3.3c: Demographic Data on Matched Pairs, Traditional Section

| pair | sex | last hs math | physics? | GPA | year | locus of control |
|------|-----|--------------|----------|-----|------|------------------|
| 21 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 21 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 22 | M | precalc | yes | 3.5-4.0 | junior | average |
| 22 | F | calculus | yes | 3.5-4.0 | sophomore | average |
| 23 | M | trigonometry | no | 2.5-2.9 | senior | high |
| 23 | F | trigonometry | no | 3.0-3.4 | senior | average |
| 24 | M | precalc | no | 3.5-4.0 | freshman | average |
| 24 | F | trigonometry | no | 3.0-3.4 | junior | average |
| 25 | M | calculus | yes | 3.0-3.4 | senior | average |
| 25 | F | calculus | yes | 3.5-4.0 | sophomore | average |
| 26 | M | calculus | no | 3.0-3.4 | sophomore | high |
| 26 | F | calculus | no | 3.5-4.0 | freshman | average |
| 27 | M | precalc | yes | 3.5-4.0 | freshman | average |
| 27 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 28 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 28 | F | precalc | yes | 3.5-4.0 | freshman | average |
| 29 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 29 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 30 | M | calculus | yes | 3.5-4.0 | freshman | low |
| 30 | F | calculus | yes | 3.0-3.4 | junior | average |
| 31 | M | precalc | no | 3.0-3.4 | sophomore | average |
| 31 | F | calculus | no | 3.5-4.0 | junior | average |
| 32 | M | calculus | yes | 3.5-4.0 | freshman | average |
| 32 | F | calculus | yes | 3.5-4.0 | freshman | average |
| 33 | M | geometry | no | 2.5-2.9 | junior | average |
| 33 | F | algebra | no | 3.5-4.0 | junior | average |
| 34 | M | calculus | yes | 3.5-4.0 | sophomore | average |
| 34 | F | calculus | yes | 3.5-4.0 | freshman | average |

Figure 3.11: Locus of Control Questions

1.   Luck often plays a role in the grades I receive.

2.   Most instructors are fair to students.

3.   Most students don't realize the extent to which their grades are influenced by accidental happenings.

4.   If a student is well prepared, there is rarely such a thing as an unfair test.

5.   There is a direct connection between how hard I study and the grades I get.

6.   The amount I learn in a course depends upon the instructor.

7.   I feel I have control over the grade I receive in a course.

8.   Many times exam questions tend to be so unrelated to course work that studying is really useless.

9.   Sometimes I can't understand how instructors arrive at the grades they give.

10.  The amount I learn in a course depends upon how hard I study.

pairs were of students who had taken the same amount of math in high school, and most of the rest of the pairs were of a student who had taken calculus with one who had taken precalculus.  Nearly all of the students in the sample had good high school grade point averages (GPAs).  Nearly half of the pairs were of two students who both had GPAs in the 3.5 - 4.0 range, and another 40% were of one student with a GPA in the 3.5 - 4.0 range and another with a GPA in the 3.0 - 3.4 range. More than 60% of the pairs were of students in the same year in college, and another 35% were of students that were only one year apart in school.  There was only one pair that was separated by two years.

Subjects were also matched on their locus of control over their own grades, which has been shown to be nearly as good a predictor of physics course grades as GPA and pretest scores (Lawrenz et. al, 1991). Locus of control was measured by student responses to ten questions, shown in Figure 3.11. Students answered each question on a five point scale, from strongly disagree (a score of -2) to strongly agree (a score of +2). For purposes of analysis, the scores on statements indicating a low locus of control such as "Sometimes I can't understand how instructors arrive at the grades they give" were reversed so that all positive scores indicated a high locus of control. Then the scores were averaged, and the top third scores of the students on whom there was complete data were assigned the rank "high," with the other thirds labeled average and low. About 75% of the pairs had the same rank, and all the rest were only separated by one rank; there was no pair between a student with a high locus of control and a student with a low locus of control.

# CHAPTER 4:  RESULTS

This chapter is in three main sections.  First, the matched sample is compared to the population, both in terms of their demographic characteristics and in terms of their performance on the pretests.  Then the results of the matched sample's physics learning are given; physics learning here includes both problem solving ability and conceptual understanding.  Finally, the opinions of men and women about the course are given.

## Comparison Of Matched Sample To Remaining Population

Demographic data for the matched sample and the remaining population in both the experimental and the traditional sections are shown in Table 4.1.  Binomial tests were done to compare whether these groups were equivalent in whether they had taken calculus and physics in high school, whether their high school grade point average was 3.5 or higher, and how many of them were freshmen, with a p of 0.05 or less taken as significant.

The results of the binomial tests are shown in Table 4.2.  The only significant difference in the demographics of the matched sample and the remaining population is between the groups of men in the traditional section.  A much lower percentage of men in the matched sample than in the remaining population had taken physics.

The matched sample was also compared to the population on three pretest measures of physics knowledge:  the Force Concept Inventory, the free-response conceptual questions, and the asteroid problem (see Figure 3.1, page 49).  The pretest scores for the sample and

the population from both the experimental and the traditional sections are shown on Table 4.3.

Table 4.2a:  Results of Binomial Tests on Demographic Data, Experimental Section

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | matched sample | remaining population | p | matched sample | remaining population | p |
| High school calculus | 55% | 57% | 0.17 | 65% | 51% | 0.08 |
| High school physics | 85 | 88 | 0.22 | 85 | 82 | 0.22 |
| H. S. GPA over 3.5 | 60 | 56 | 0.17 | 75 | 74 | 0.20 |
| College freshmen | 70 | 70 | 0.19 | 60 | 54 | 0.16 |

Table 4.2b:  Results of Binomial Tests on Demographic Data, Traditional Section

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | matched sample | remaining population | p | matched sample | remaining population | p |
| High school calculus | 57% | 55% | 0.21 | 72% | 54% | 0.09 |
| High school physics | 64 | 91 | 0.002 | 64 | 78 | 0.11 |
| H. S. GPA over 3.5 | 64 | 59 | 0.20 | 79 | 78 | 0.25 |
| College freshmen | 50 | 64 | 0.12 | 50 | 48 | 0.21 |

There was no difference in the Force Concept Inventory scores for the women in the experimental section (t (df = 19) = -0.04, p = 0.590) or the traditional section (t (df = 13) = 1.53, p = 0.153). On the other hand, the men in the matched sample had significantly lower pretest Force Concept Inventory scores than the men in the remaining population, both in the experimental section (t (df = 19) = 3.74, p = 0.002) and the traditional section (t (df = 13) = 3.08, p = 0.009). It is interesting to note that the average Force Concept Inventory scores for the men in the population were higher than the average scores for the women in the population, for both sections. Since the men were matched to the women to make the matched samples, the men in the samples had lower scores than the men in the populations.

The difference between the free response conceptual test scores of the matched sample and the remaining population was determined by finding the combined median and then using the $\chi^2$ test to compare numbers of responses above and below that combined median. The median for men in each section was 1 out of 12, and the median for men in each section was 0. As expected with such low scores, there were no differences in the pretest scores of the matched sample and the remaining population for men in the experimental section ($\chi^2 = 1.07$, p = 0.30), women in the experimental section ($\chi^2 = 0$, p > 0.99), men in the traditional section ($\chi^2 = 0.45$, 0.70 > p > 0.50), or women in the traditional section ($\chi^2 = 0$, p > 0.99).

Pretest problem solving ability was measured by the asteroid problem (see Figure 3.1). Students were to find the size of an asteroid. Students who had a sensible method for finding the size of the asteroid, whether in terms of its radius, diameter, volume, or surface area, earned a score of "high." Students who had a method for finding the mass, not size, of the asteroid earned a score of "middle." And students who had neither earned

a score of "low." Differences between the matched sample and the remaining population were tested using the $\chi^2$ test. There were no differences in the pretest scores of the matched sample and the remaining population for men in the experimental section ($\chi^2 = 3.69$, $0.20 > p > 0.10$), women in the experimental section ($\chi^2 = 0$, $p > 0.99$), men in the traditional section ($\chi^2 = 3.94$, $0.20 > p > 0.10$), or women in the traditional section ($\chi^2 = 1.37$, $0.70 > p > 0.40$).

Overall, there were two differences between the matched samples and the populations. A smaller proportion of men in the matched sample than in the population of the traditional section had taken high school physics, and men in both matched samples had lower average Force Concept Inventory scores than the men in both populations. Since the men in the matched sample were chosen since they matched to the women and the women in the sample are not significantly different than the women in the populations, these differences reflect differences between the men and women in the populations.

### Problem Solving Tests

The first research question concerned whether there would be differences in physics learning between men and women who had been matched on several variables in order to remove much of the differences between them. The first way in which this question was answered was by comparing the performance of the matched sample on four questions on their final exam, which are shown in Figure 3.4 on page 59.

The problem solutions were coded on their general approach, specific application of physics, logical progression, appropriate mathematics, and clear communication. After each response was categorized, the categories were assigned numerical values between 0

and 10, where "a" was always worth 0, the category representing correct responses was always worth 10, and the intermediate categories were assigned normalized ranks. The scores for General Approach, Specific Application of Physics, Logical Progression, and Applied Mathematics were added across each problem, giving each student four scores out of 40. The scores for the four problems were then added together, giving each student a total problem solving score out of 160. These scores are arranged by pair in Table 4.4. Although this study was not designed to compare the two sections, it is interesting to note that both men and women in the traditional section scored higher than both men and women in the experimental section on these final exam problems. This will be discussed further in Chapter 5.

These scores were analyzed using Wilcoxin's Matched Pairs Signed-Rank Test, which looks for differences between members of matched pairs. They were analyzed separately for the traditional section and the experimental section since only differences within the sections were of interest. There was no difference in overall problem solving in either the experimental section (W (n=19) = 67, p > 0.10) or the traditional section (W (n=14) = 50, p > 0.10).

Although there are no overall differences in problem solving ability, it is possible that there are differences in the understanding of specific questions which are masked by adding the four questions together. Therefore, it is interesting to examine the responses to the four questions individually. Problem solving scores by problem are shown in Table 4.5.

In the experimental section, there was no differences in the overall performance on the Enterprise problem ($W = (n=17)$ 56, $p > 0.10$), the force portion of the modified Atwood machine problem ($W (n=18) = 79.5$, $p > 0.10$), the balloon problem ($W (n=17)= 67$, $p >$

Table 4.4:  Total Problem Solving Scores

| | Experimental Section | | | | Traditional Section | | |
|---|---|---|---|---|---|---|---|
| pair | man | woman | Δ | pair | man | woman | Δ |
| 1 | 120 | 124.29 | -4.29 | 21 | 147.14 | 139.99 | 7.15 |
| 2 | 61.67 | 75.23 | -13.56 | 22 | 156.66 | 134.05 | 22.61 |
| 3 | 106.9 | 87.62 | 19.28 | 23 | 158.57 | 139.03 | 19.54 |
| 4 | 145.71 | 70 | 75.71 | 24 | 129.76 | 153.33 | -23.57 |
| 5 | 145.46 | 131.68 | 13.78 | 25 | 149.99 | 126.66 | 23.33 |
| 6 | 109.58 | 114.76 | -5.18 | 26 | 91.2 | 115.95 | -24.75 |
| 7 | 133.32 | 143.34 | -10.02 | 27 | 107.15 | 155.23 | -48.08 |
| 8 | 132.14 | 45.24 | 86.9 | 28 | 130.71 | 116.18 | 14.53 |
| 9 | 114.05 | 102.87 | 11.18 | 29 | 90.5 | 139.04 | -48.54 |
| 10 | 95.25 | 37.14 | 58.11 | 30 | 144.75 | 108.33 | 36.42 |
| 11 | 104.76 | 100.01 | 4.75 | 31 | 151.66 | 115 | 36.66 |
| 12 | 151.66 | 151.66 | 0 | 32 | 152.37 | 122.14 | 30.23 |
| 13 | 113.57 | 105.95 | 7.62 | 33 | 96.9 | 150.46 | -53.56 |
| 14 | 32.63 | 75.23 | -42.6 | 34 | 120 | 103.33 | 16.67 |
| 15 | 120.01 | 123.33 | -3.32 | median | 130.71 | 126.66 | 14.53 |
| 16 | 103.34 | 119.52 | -16.18 | | | | |
| 17 | 94.05 | 92.6 | 1.45 | | | | |
| 18 | 101.43 | 58.81 | 42.62 | | | | |
| 19 | 120.96 | 148.8 | -27.84 | | | | |
| 20 | 143.34 | 129.27 | 14.07 | | | | |
| median | 113.57 | 102.87 | 1.45 | | | | |

0.10), or the ski problem (W (n=14) = 30, p > 0.10).  For comparison, the traditional section was also studied.  In this section, there were also no differences in the overall performance on the Enterprise problem (W (n=12) = 30.5, p > 0.10), the force portion of the modified Atwood machine problem (W = (n=14) 46, p > 0.10), the balloon problem (W (n=13) = 24.5, p > 0.10), or the ski problem (W (n=13) = 43.5, p > 0.10).

Although there are no differences in the overall performance on any of the final exam problems, it is possible that there are differences in the performance on the different measures, general approach, specific application of physics, logical progression, and applied mathematics, which are masked by adding the four measures together.  Therefore, it is interesting to examine the performance on each measure within each problem. Scores by measure on the four problems are shown in Tables 4.6 through 4.9.  Appendix E provides another summary of the performance on each measure across the four problems.  The differences in responses are analyzed with $\chi^2$ tests between responses with no error ("g" for General Approach, Logical Progression, and Applied Mathematics, and "h" for Specific Application of Physics) and responses with an error.

The scores to the Enterprise problem, by measure, are shown in Table 4.6.  In the experimental section, there was no difference in the general approach ($\chi^2 = 0$, p > 0.99), the specific application of physics ($\chi^2 = 0.69$, $0.50 > p > 0.30$), the logical progression ($\chi^2 = 0$, p > 0.99), or the appropriate mathematics ($\chi^2 = 1.20$, $0.30 > p > 0.20$) in the Enterprise problem.  And in the traditional section, there was also no difference in the general approach ($\chi^2 = 0$, p > 0.99), the specific application of physics ($\chi^2 = 0$, p > 0.99), the logical progression  ($\chi^2 = 0.29$, $0.70$ p > 0.50), or the appropriate mathematics ($\chi^2 = 0$, p > 0.99) in the Enterprise problem.

Table 4.6a: "Enterprise" Problem Scores by Measure, Experimental Section

| pair | General Approach | | Specific Application | | Logical Progression | | Appropriate Mathematics | |
|---|---|---|---|---|---|---|---|---|
| | man | woman | man | woman | man | woman | man | woman |
| 1 | f | b | h | h | g | g | e | g |
| 2 | b | a | h | a | b | a | g | a |
| 3 | e | b | h | h | g | b | g | g |
| 4 | e | f | h | h | g | g | g | e |
| 5 | f | b | h | h | g | b | g | g |
| 6 | b | b | h | h | b | b | g | g |
| 7 | f | e | f | h | g | g | g | g |
| 8 | g | e | h | a | g | a | g | a |
| 9 | b | b | h | h | b | b | g | g |
| 10 | b | a | h | a | b | a | g | a |
| 11 | b | b | h | h | b | b | g | g |
| 12 | c | f | h | h | g | g | g | f |
| 13 | d | g | h | h | g | g | g | g |
| 14 | b | f | a | h | a | d | a | f |
| 15 | b | f | h | a | b | a | g | a |
| 16 | e | e | h | h | g | g | f | g |
| 17 | b | c | h | h | b | g | g | g |
| 18 | g | b | h | e | g | g | g | g |
| 29 | b | f | h | h | b | g | g | g |
| 20 | d | f | h | h | g | g | g | g |
| median | c | c | h | h | g | d | g | g |

Table 4.6b:  "Enterprise" Problem Scores by Measure, Traditional Section

| pair | General Approach | | Specific Application | | Logical Progression | | Appropriate Mathematics | |
|------|-----|-------|-----|-------|-----|-------|-----|-------|
| | man | woman | man | woman | man | woman | man | woman |
| 21 | d | g | h | h | g | g | g | g |
| 22 | f | c | h | h | g | g | g | g |
| 23 | g | g | h | h | g | g | g | g |
| 24 | d | g | h | h | g | g | g | g |
| 25 | c | c | h | h | g | g | g | g |
| 26 | c | b | h | h | g | b | g | g |
| 27 | b | f | h | h | b | g | g | g |
| 28 | b | c | h | h | b | g | g | g |
| 29 | b | d | b | h | b | g | g | g |
| 30 | g | f | h | h | g | g | g | e |
| 31 | f | c | h | h | g | g | g | g |
| 32 | g | e | h | e | g | g | g | g |
| 33 | c | g | h | h | g | g | g | g |
| 34 | g | f | h | h | g | g | g | g |
| median | d | e | h | h | g | g | g | g |

The scores to the Atwood machine problem, by measure, are shown in Table 4.7. In the experimental section, there was no difference in the general approach ($\chi^2 = 0.57$, $0.50$ p > $0.30$), the specific application of physics ($\chi^2 = 0$, p > $0.99$), the logical progression ($\chi^2 = 0$, p > $0.99$), or the appropriate mathematics ($\chi^2 = 0.94$, $0.50 > $ p > $0.30$) in the Atwood machine problem. And in the traditional section, there was also no difference in the general approach ($\chi^2 = 0$, p > $0.99$), the specific application of physics ($\chi^2 = 0$, p > $0.99$), the logical progression ($\chi^2 = 0$, p > $0.99$), or the appropriate mathematics ($\chi^2 = 0$, p > $0.99$) in the Atwood machine problem.

The scores to the balloon problem, by measure, are shown in Table 4.8. In the experimental section, there was no difference in the general approach ($\chi^2 = 1.82$, $0.20 > $ p > $0.10$), the specific application of physics ($\chi^2 = 0$, p > $0.99$), the logical progression ($\chi^2 = 0.91$, $0.50 > $ p > $0.30$), or the appropriate mathematics ($\chi^2 = 0.12$, $0.80 > $ p > $0.70$) in the balloon problem. And in the traditional section, there was also no difference in the general approach ($\chi^2 = 1.31$, $0.30 > $ p > $0.20$), the specific application of physics ($\chi^2 = 2.30$, $0.20 > $ p > $0.10$), the logical progression ($\chi^2 = 1.58$, $0.30 > $ p > $0.20$), or the appropriate mathematics ($\chi^2 = 0.21$, $0.70 > $ p > $0.50$) in the balloon problem.

The scores to the ski problem, by measure, are shown in Table 4.9. In the experimental section, there was no difference in the general approach ($\chi^2 = 0.04$, $0.95 > $ p > $0.90$), the specific application of physics ($\chi^2 = 0.12$, $0.80 > $ p > $0.70$), the logical progression ($\chi^2 = 0$, p > $0.99$), or the appropriate mathematics ($\chi^2 = 0.46$, p = $0.50$) in the ski problem. And in the traditional section, there was also no difference in the general approach ($\chi^2 = 0.16$, $0.70 > $ p > $0.50$), the specific application of physics ($\chi^2 = 0.97$, $0.50 > $ p > $0.30$), the logical progression ($\chi^2 = 0$, p > $0.99$), or the appropriate mathematics ($\chi^2 = 014$, $0.80 > $ p > $0.70$) in the ski problem.

Table 4.7a:  Atwood Machine Problem Scores by Measure, Experimental Section

| pair | General Approach | | Specific Application | | Logical Progression | | Appropriate Mathematics | |
|---|---|---|---|---|---|---|---|---|
| | man | woman | man | woman | man | woman | man | woman |
| 1 | d | e | d | d | e | g | d | g |
| 2 | a | f | a | h | a | g | a | g |
| 3 | g | e | h | f | g | g | c | g |
| 4 | d | a | f | a | g | a | g | a |
| 5 | f | g | f | h | g | g | f | e |
| 6 | d | f | b | f | b | g | g | e |
| 7 | e | e | h | h | b | b | g | g |
| 8 | c | a | b | a | b | a | g | a |
| 9 | c | c | b | b | b | b | g | g |
| 10 | e | a | h | a | g | a | g | a |
| 11 | c | e | b | h | b | g | g | g |
| 12 | g | e | h | h | g | g | g | g |
| 13 | e | g | h | f | g | e | g | b |
| 14 | e | c | d | b | g | b | g | g |
| 15 | e | g | h | h | g | g | g | g |
| 16 | c | c | b | c | b | g | g | g |
| 17 | d | f | e | e | g | f | g | f |
| 18 | e | g | d | h | g | d | e | c |
| 29 | g | g | d | f | g | g | f | f |
| 20 | e | g | h | e | b | g | g | c |
| median | e | e | d | f | g | g | g | f |

115

Table 4.7b: Atwood Machine Problem Scores by Measure, Traditional Section

| pair | General Approach man | General Approach woman | Specific Application man | Specific Application woman | Logical Progression man | Logical Progression woman | Appropriate Mathematics man | Appropriate Mathematics woman |
|---|---|---|---|---|---|---|---|---|
| 21 | g | g | f | h | d | g | g | g |
| 22 | g | g | h | d | g | g | g | g |
| 23 | g | f | h | f | g | g | g | g |
| 24 | c | g | b | h | b | g | g | g |
| 25 | g | c | h | b | g | b | g | g |
| 26 | e | c | d | e | b | b | g | g |
| 27 | e | g | f | h | e | g | g | g |
| 28 | d | g | e | f | f | e | g | b |
| 29 | e | d | d | f | d | g | g | g |
| 30 | g | g | f | h | g | g | f | g |
| 31 | g | e | h | h | g | d | g | f |
| 32 | f | c | f | b | g | b | g | g |
| 33 | e | f | h | f | b | f | g | g |
| 34 | a | g | a | h | a | e | a | c |
| median | e | f | f | f | e | f | g | g |

Table 4.8a: Balloon Problem Scores by Measure, Experimental Section

| pair | General Approach man | General Approach woman | Specific Application man | Specific Application woman | Logical Progression man | Logical Progression woman | Appropriate Mathematics man | Appropriate Mathematics woman |
|---|---|---|---|---|---|---|---|---|
| 1 | g | f | e | g | g | g | d | g |
| 2 | g | g | h | g | g | f | f | g |
| 3 | d | g | g | f | g | g | g | c |
| 4 | g | e | g | h | g | f | g | g |
| 5 | g | g | h | h | g | g | g | g |
| 6 | g | d | g | g | g | f | f | g |
| 7 | g | g | f | h | g | g | g | g |
| 8 | g | g | h | g | g | g | g | g |
| 9 | g | g | g | h | g | g | g | g |
| 10 | g | g | g | f | g | g | e | g |
| 11 | c | g | h | h | g | g | g | g |
| 12 | g | g | h | h | g | g | g | g |
| 13 | g | g | g | f | g | g | g | g |
| 14 | a | c | a | f | a | g | a | e |
| 15 | g | d | g | h | g | g | g | g |
| 16 | g | d | f | g | g | g | g | f |
| 17 | g | d | h | g | g | g | g | d |
| 18 | g | b | f | b | g | a | c | a |
| 29 | d | d | h | h | g | g | d | g |
| 20 | g | g | h | h | g | g | g | g |
| median | g | g | g | g | g | g | g | g |

117

Table 4.8b: Balloon Problem Scores by Measure, Traditional Section

| pair | General Approach | | Specific Application | | Logical Progression | | Appropriate Mathematics | |
|---|---|---|---|---|---|---|---|---|
| | man | woman | man | woman | man | woman | man | woman |
| 21 | g | c | h | h | g | f | g | g |
| 22 | g | d | h | b | g | g | f | g |
| 23 | g | d | g | g | g | f | g | g |
| 24 | g | g | h | h | g | g | g | g |
| 25 | g | g | h | g | g | f | g | g |
| 26 | e | d | g | g | g | g | g | g |
| 27 | b | g | h | g | c | f | e | g |
| 28 | g | b | h | e | g | g | g | f |
| 29 | b | g | h | g | g | g | d | g |
| 30 | e | a | g | a | g | a | g | a |
| 31 | g | c | h | g | g | g | g | g |
| 32 | g | g | g | h | g | g | g | g |
| 33 | g | g | g | h | f | g | f | g |
| 34 | g | a | h | a | g | a | g | a |
| median | g | d | h | g | g | g | g | g |

Table 4.9a:  Ski Problem Scores by Measure, Experimental Section

| pair | General Approach | | Specific Application | | Logical Progression | | Appropriate Mathematics | |
| | man | woman | man | woman | man | woman | man | woman |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | g | c | h | b | b | b | c | g |
| 2 | a | a | a | a | a | a | a | a |
| 3 | c | a | a | a | a | a | a | a |
| 4 | g | a | h | a | b | a | f | a |
| 5 | f | g | h | h | b | b | d | b |
| 6 | d | b | g | e | b | b | e | g |
| 7 | f | f | e | h | b | b | f | g |
| 8 | g | a | e | a | b | a | g | a |
| 9 | g | b | e | b | b | b | g | g |
| 10 | a | a | a | a | a | a | a | a |
| 11 | f | a | h | a | b | a | c | a |
| 12 | f | f | h | h | b | b | g | g |
| 13 | c | c | a | a | a | a | a | a |
| 14 | a | a | a | a | a | a | a | a |
| 15 | b | g | b | h | b | b | f | g |
| 16 | b | c | b | b | b | b | b | g |
| 17 | a | a | a | a | a | a | a | a |
| 18 | c | a | a | a | a | a | a | a |
| 29 | d | g | h | h | b | b | g | g |
| 20 | g | c | h | g | b | b | g | a |
| median | d | b | e | a | b | a | c | a |

Table 4.9b: Ski Problem Scores by Measure, Traditional Section

| pair | General Approach | | Specific Application | | Logical Progression | | Appropriate Mathematics | |
|---|---|---|---|---|---|---|---|---|
| | man | woman | man | woman | man | woman | man | woman |
| 21 | g | d | h | h | b | b | g | c |
| 22 | g | g | h | h | b | b | g | g |
| 23 | g | f | h | h | b | b | g | c |
| 24 | f | f | h | h | b | b | g | d |
| 25 | f | g | h | h | b | b | f | g |
| 26 | a | f | a | h | a | b | a | g |
| 27 | c | g | h | h | b | b | f | g |
| 28 | f | f | h | h | b | b | g | c |
| 29 | b | g | b | h | b | b | g | c |
| 30 | g | c | e | h | b | b | f | g |
| 31 | c | b | h | b | b | b | g | e |
| 32 | g | f | h | h | b | b | f | d |
| 33 | a | f | a | h | a | b | a | f |
| 34 | g | d | h | h | b | b | g | g |
| median | f | f | h | h | b | b | g | e |

120

There is no difference in the problem solving ability of the men and women in either matched sample, as measured by their performance on their final exam problems.

## Conceptual Tests

### Force Concept Inventory

Another way in which students' physics learning was measured was through their scores on the Force Concept Inventory, the 29 point multiple choice test. Post-test scores for the matched pairs of both sections are shown in Table 4.10. Although this study was not designed to compare the two sections, it is interesting to note that both men and women in the experimental section scored higher than both men and women in the traditional section on the Force Concept Inventory. This will be discussed further in Chapter 5. A matched-sample t-test indicates that there is no overall significant difference between the scores of males and females in the experimental section $(t \ (df = 19) = -0.80, p = 0.430)$ or in the traditional section $(t \ (df = 13) = -1.06, p = 0.312)$.

### Free-Response Conceptual Questions

Another way to assess the conceptual understanding of students is to look at their responses on conceptual questions. In this study students were given the ramp question to assess their understanding of velocity and acceleration, the car and passenger question to assess their understanding of both the nature of forces and Newton's Second Law, and the car and bug question to assess their understanding of Newton's Third Law. These questions are shown in Figures 3.7, 3.8, and 3.9. As was discussed in Chapter 3, responses for each question were scored on a scale of 0 to 2, with a total possible score of

Table 4.10:  Scores on the Force Concept Inventory Post-Test

| | Experimental Section | | | | Traditional Section | | |
|---|---|---|---|---|---|---|---|
| pair | man | woman | Δ | pair | man | woman | Δ |
| 1 | 21 | 16 | -5 | 21 | 17 | 17 | 0 |
| 2 | 13 | 20 | 7 | 22 | 16 | 12 | -4 |
| 3 | 17 | 14 | -3 | 23 | 27 | 15 | -12 |
| 4 | 24 | 15 | -9 | 24 | 20 | 18 | -2 |
| 5 | 14 | 24 | 10 | 25 | 21 | 16 | -5 |
| 6 | 22 | 22 | 0 | 26 | 13 | 15 | 2 |
| 7 | 25 | 20 | -5 | 27 | 14 | 9 | -5 |
| 8 | 24 | 21 | -3 | 28 | 14 | 14 | 0 |
| 9 | 19 | 24 | 5 | 29 | 16 | 20 | 4 |
| 10 | 19 | 25 | 6 | 30 | 13 | 11 | -2 |
| 11 | 22 | 25 | 3 | 31 | 9 | 15 | 6 |
| 12 | 27 | 25 | -2 | 32 | 11 | 13 | 2 |
| 13 | 21 | 16 | -5 | 33 | 22 | 19 | -3 |
| 14 | 17 | 15 | -2 | 34 | 20 | 21 | 1 |
| 15 | 22 | 22 | 0 | average | 16.64 | 15.36 | -1.00 |
| 16 | 22 | 18 | -4 | | | | |
| 17 | 23 | 14 | -9 | | | | |
| 18 | 22 | 15 | -7 | | | | |
| 19 | 17 | 24 | 7 | | | | |
| 20 | 26 | 22 | -4 | | | | |
| average | 20.85 | 19.85 | -1.29 | | | | |

12 points on all questions. The total scores earned are shown in Table 4.11. Although this study was not designed to compare the two sections, it is interesting to note that both men and women in the experimental section scored higher than both men and women in the traditional section on these questions. This will be discussed further in Chapter 5.

These total scores were analyzed using Wilcoxin's Matched Pairs Signed-Rank Test. They were analyzed separately for the traditional section and the experimental section since only differences within the sections were of interest. For the experimental section, W (n = 18) = 68, which is not significant even at the 10 percent level. For the traditional section, W (n = 11) = 32.5, which is also not significant even at the 10 percent level.

Although there are no overall differences in conceptual understanding, it is possible that there are differences in the understanding of specific questions which are masked by adding the six questions together. Therefore, it is interesting to examine the responses to the six questions individually. First, to determine whether there are differences in student understanding of the difference between velocity and acceleration, $\chi^2$ tests were done to compare the frequency of responses in categories 1 and 2, correct or close to correct, and the frequency of responses in categories 3 and 4, incorrect, incomprehensible, or blank.

The ramp responses are shown in Table 4.12. There were no significant differences in the responses of the women and men in the experimental section ($\chi^2 = 0.43$, $0.70 > p > 0.50$) or the traditional section ($\chi^2 = 0.29$, $0.70 > p > 0.50$).

Next, to determine whether there are differences in student understanding of the nature of forces, $\chi^2$ tests were done to compare the frequency of responses in categories 1 and 2 and the frequency of responses in categories 3 and 4. Student responses relating to the

nature

Table 4.11:  Total Scores on Free Response Conceptual Questions

| Experimental Section | | | | Traditional Section | | | |
|---|---|---|---|---|---|---|---|
| pair | man | woman | Δ | pair | man | woman | Δ |
| 1 | 4 | 12 | -8 | 21 | 2 | 2 | 0 |
| 2 | 0 | 5 | -5 | 22 | 0 | 1 | -1 |
| 3 | 4 | 7 | -3 | 23 | 5 | 1 | 4 |
| 4 | 4 | 6 | -2 | 24 | 3 | 4 | -1 |
| 5 | 7 | 9 | -2 | 25 | 4 | 0 | 4 |
| 6 | 2 | 3 | -1 | 26 | 4 | 1 | 3 |
| 7 | 1 | 2 | -1 | 27 | 0 | 1 | -1 |
| 8 | 7 | 8 | -1 | 28 | 1 | 4 | -3 |
| 9 | 2 | 3 | -1 | 29 | 1 | 3 | -2 |
| 10 | 3 | 3 | 0 | 30 | 1 | 0 | 1 |
| 11 | 2 | 2 | 0 | 31 | 1 | 2 | -1 |
| 12 | 6 | 5 | 1 | 32 | 0 | 0 | 0 |
| 13 | 5 | 3 | 2 | 33 | 0 | 2 | -2 |
| 14 | 6 | 3 | 3 | 34 | 4 | 4 | 0 |
| 15 | 11 | 8 | 3 | median | 1 | 1 | -1 |
| 16 | 6 | 2 | 4 | | | | |
| 17 | 6 | 2 | 4 | | | | |
| 18 | 10 | 5 | 5 | | | | |
| 19 | 10 | 4 | 6 | | | | |
| 20 | 11 | 3 | 8 | | | | |
| median | 5 | 3 | 0 | | | | |

Table 4.12: Frequencies of Response Categories on the Ramp Problem

| Type of Response | Experimental | | Traditional | |
|---|---|---|---|---|
| | Men (n=20) | Women (n=20) | Men (n=14) | Women (n=14) |
| 1. <u>Correct response with explanation</u> -- the acceleration of the ball is the same everywhere because the force is the same everywhere | **7** | **5** | **2** | **1** |
| 2. <u>Response close to accepted idea</u>, but vague, unclear, partially incorrect, or correct with no explanation | **2** | **1** | **1** | **0** |
| 3. <u>Response in terms of alternative ideas about velocity and acceleration</u> | **11** | **14** | **11** | **13** |
|    3a. Velocity and acceleration not completely discriminated: | **8** | **13** | **11** | **12** |
|       • acceleration the same up and down the ramp, but zero at top; | 0 | 0 | 0 | 2 |
|       • accelerations up and down the ramp are same size but opposite directions -- acceleration at top; | 4 | 6 | 3 | 0 |
|       • accelerations up and down the ramp are same size but opposite directions -- no acceleration at top; | 2 | 2 | 5 | 10 |
|       • accelerations up the ramp is larger or smaller than acceleration down the ramp; | 1 | 2 | 3 | 0 |
|       • "rates" of acceleration up and down ramp are the same - acceleration at top; | 1 | 2 | 0 | 0 |
|       • "rates" of acceleration up and down ramp are different - acceleration at top; | 0 | 1 | 0 | 0 |
|    3b. Indicate complete confusion between acceleration and velocity | **3** | **1** | **0** | **1** |
|       • "rates" of acceleration up and down ramp are the same - no acceleration at top; | 2 | 1 | 0 | 1 |

| | | | | |
|---|---|---|---|---|
| • "rates" of acceleration up and down ramp are different - no acceleration at top; | 1 | 0 | 0 | 0 |
| 4.  Idiosyncratic, incomprehensible, or blank | **0** | **0** | **0** | **0** |

of forces on the passenger are shown in Table 4.13. There were no significant differences in the responses of the women and men in the experimental section ($\chi^2 = 0.90$, $0.50 > p > 0.30$) or in the traditional section ($\chi^2 = 2.49$, $0.20 > p > 0.10$). Student responses relating to the nature of forces on the car are shown in Table 4.14. There were no significant differences in the responses of women and men in the experimental section ($\chi^2 = 0.10$, $0.80 > p > 0.70$) or in the traditional section ($\chi^2 = 0.97$, $0.80 > p > 0.70$).

Next, to determine whether there are differences in student understanding of the Second Law, $\chi^2$ tests were done to compare the frequency of responses in categories 1 and 2 and the frequency of responses in categories 3 and 4. Student responses relating to why the passenger accelerates are shown in Table 4.15. There were no significant differences in the responses of women and men in the experimental section ($\chi^2 = 0.10$, $0.80 > p > 0.70$) or in the traditional section ($\chi^2 = 2.62$, $0.20 > p > 0.10$). Student responses relating to why the passenger accelerates are shown in Table 4.16. There were no significant differences in the responses of women and men in the experimental section ($\chi^2 = 0.10$, $0.80 > p > 0.70$) or in the traditional section ($\chi^2 = 0.37$, $0.70 > p > 0.50$).

Next, to determine whether there are differences in student understanding of the Third Law, $\chi^2$ tests were done to compare the frequency of responses in categories 1 and 2 and the frequency of responses in categories 3 and 4. Student responses relating to the car and bug problem are shown in Table 4.17. There were no significant differences in the responses of women and men in the experimental section ($\chi^2 = 0.94$, $0.80 > p > 0.70$) or in the traditional section ($\chi^2 = 0$, $p > 0.99$).

There is no difference in conceptual understanding in either the experimental or the traditional section, as evidenced by scores on the Force Concept Inventory, overall scores

Table 4.13:  Frequencies of Response Categories on the Nature of Forces on the Passenger

| Type of Response | Experimental | | Traditional | |
|---|---|---|---|---|
| | Men (n=20) | Women (n=20) | Men (n=14) | Women (n=14) |
| 1.  Correct response with explanation -- All forces shown are interactions between two objects and are drawn pointing to the point of contact | **10** | **7** | **2** | **0** |
| 2.  Response close to accepted idea, but vague, unclear, partially incorrect, or correct with no explanation | **2** | **1** | **0** | **4** |
| 3.  Response in terms of alternative ideas about the nature of forces | **7** | **9** | **10** | **10** |
| 3a.  At least one Third Law force is drawn, a force from and not on the passenger | 2 | 0 | 1 | 3 |
| 3b.  At least one force is drawn that is not an interaction between two objects | 3 | 8 | 1 | 0 |
| 3c.  More than one type of error is made | 2 | 1 | 8 | 7 |
| 4.  Idiosyncratic, incomprehensible, or blank | **1** | **3** | **2** | **0** |

Table 4.14:  Frequencies of Response Categories on the Nature of Forces on the Car

| Type of Response | Experimental | | Traditional | |
|---|---|---|---|---|
| | Men (n=20) | Women (n=20) | Men (n=14) | Women (n=14) |
| 1.  Correct response with explanation -- All forces shown are interactions between two objects and are drawn pointing to the point of contact | **3** | **1** | **0** | **0** |
| 2.  Response close to accepted idea, but vague, unclear, partially incorrect, or correct with no explanation | **7** | **11** | **1** | **4** |
| 3.  Response in terms of alternative ideas about the nature of forces | **9** | **7** | **12** | **10** |
| 3a.  At least one Third Law force is drawn, a force from and not on the car | 3 | 1 | 0 | 0 |
| 3b.  At least one force is drawn that is not an interaction between two objects | 2 | 1 | 0 | 1 |
| 3c.  More than one type of error is made | 4 | 5 | 12 | 9 |
| 4.  Idiosyncratic, incomprehensible, or blank | **1** | **1** | **1** | **0** |

Table 4.15:  Frequencies of Response Categories on Why the Passenger Accelerates

| Type of Response | Experimental | | Traditional | |
|---|---|---|---|---|
| | Men (n=20) | Women (n=20) | Men (n=14) | Women (n=14) |
| 1. <u>Correct response with explanation</u> -- shows correct understanding of the Second Law | **11** | **6** | **1** | **1** |
| 2. <u>Response close to accepted idea</u>, but vague, unclear, partially incorrect, or correct with no explanation | **2** | **5** | **6** | **1** |
| 3. <u>Response in terms of alternative ideas about the Second Law</u> | **6** | **6** | **6** | **10** |
| 3a. Attributes acceleration to more than one force but does not talk about summing of net forces | 0 | 2 | 1 | 2 |
| 3b. Attributes acceleration to only one force when there are other forces on the diagram that should have been taken into account. | 1 | 0 | 1 | 5 |
| 3c. Attributes acceleration to a force that is in the opposite direction of acceleration | 2 | 0 | 0 | 0 |
| 3d. Attributes acceleration to something other than a force on the passenger | 3 | 4 | 4 | 3 |
| 4. <u>Idiosyncratic, incomprehensible, or blank</u> | **1** | **3** | **1** | **2** |

Table 4.16:  Frequencies of Response Categories on Why the Car Accelerates

| Type of Response | Experimental | | Traditional | |
|---|---|---|---|---|
| | Men (n=20) | Women (n=20) | Men (n=14) | Women (n=14) |
| 1.  <u>Correct response with explanation</u> -- shows correct understanding of the Second Law | **4** | **4** | **0** | **1** |
| 2.  <u>Response close to accepted idea</u>, but vague, unclear, partially incorrect, or correct with no explanation | **5** | **3** | **4** | **6** |
| 3.  <u>Response in terms of alternative ideas about the Second Law</u> | **10** | **12** | **9** | **7** |
|    3a.  Attributes acceleration to more than one force but does not talk about summing of net forces | 2 | 4 | 2 | 1 |
|    3b.  Attributes acceleration to only one force when there are other forces on the diagram that should have been taken into account. | 4 | 2 | 5 | 3 |
|    3c.  Attributes acceleration to a force that is in the opposite direction of acceleration | 3 | 6 | 1 | 0 |
|    3d.  Attributes acceleration to something other than a force on the car | 1 | 0 | 1 | 3 |
| 4.  <u>Idiosyncratic, incomprehensible, or blank</u> | **1** | **1** | **1** | **0** |

Table 4.17:

Frequencies of Response Categories on the Car and Bug Problem

| Type of Response | Experimental | | Traditional | |
|---|---|---|---|---|
| | Men (n=20) | Women (n=20) | Men (n=14) | Women (n=14) |
| 1.  Correct response with explanation -- the force of the car on the bug equals the force of the bug on the car | **10** | **14** | **2** | **2** |
| 2.  Response close to accepted idea, but vague, unclear, partially incorrect, or correct with no explanation | **0** | **0** | **0** | **0** |
| 3.  Response in terms of alternative ideas about the Third Law | **9** | **6** | **11** | **12** |
| 3a.  Force of car on bug > force of bug on car | 8 | 2 | 8 | 10 |
| 3b.  There is no force of bug on car although there is a force of car on bug | 1 | 2 | 2 | 1 |
| 3c.  There are no interactions between bug and car (although papers are not blank) | 0 | 2 | 1 | 1 |
| 4.  Idiosyncratic, incomprehensible, or blank | **1** | **0** | **1** | **0** |

on free-response conceptual questions, and scores on individual free-response conceptual questions. Since there is also no difference in the problem solving ability of either matched sample, as measured by their performance on their final exam problems, the first research question has been answered. When the relevant measurable differences are removed from a sample of men and women in an introductory physics course designed to appeal to a broad population, there are no differences in how much physics they learn by the end of the course. There are also no differences in how much physics a similar sample from a more traditional course learns.

**Course Evaluations**

The second research question concerned differences in the opinions about the course. To answer this question, some questions from the course evaluation were studied. As was discussed earlier, only those questions pertaining to the laboratory, the problem solving sessions, and the problem solving methods were selected for study, since these were the areas which were changed in the experimental section in order to make the course appeal to a broader population. The specific questions are shown in Table 4.18. Each of these questions was answered on a scale from -2 to +2, where -2=Strongly Disagree, 0=Neutral, and +2=Strongly Agree. For the purpose of analysis student answers to the "waste of time" questions were reversed so that all responses greater than 0 reflected positive answers. T-tests were done on the averages to see if there were significant differences in how the males and females of each section evaluated these portions of their course.

To ensure anonymity, students' names and university identification numbers were not on their course evaluations, so it was impossible to compare the evaluations belonging to

members of the matched sample. However, students' sections and sex were on the evaluations, so it was possible to compare the evaluations by sex within sections.

The answers to the eight evaluation questions studied are shown in Table 4.18. The mean responses for each question are shown on the far right of the table. To see whether there was a difference in the overall opinion about the relevant aspects of the course, the means for the eight relevant questions were averaged. In the experimental section, the average opinion of the men was a +0.01, and the average opinion of the women was a +0.34. These opinions were significantly different (t (df = 162) = 2.67, p = 0.010).

To find out more about where the differences between men and women were, the answers to the three questions about the laboratory, the three questions about the problem solving sessions, and the two questions about the problem solving methods were averaged individually. Again, the answers to the "waste of time" questions were reversed so that every average above a 0 would be positive.

In the experimental section, the main difference between the men and the women in the experimental section was in their opinion of the problem solving methods, which the women liked better than the men did. The average opinion of the men about the problem solving methods was a +0.05, while the average opinion of the women about the problem solving methods was a +0.60. These were significantly different (t (df=162) = 2.97, p = 0.005). The average opinion of the men about the laboratory was a -0.10, and the average opinion of the women was a +0.14, and the average opinion of the men about the problem solving sessions was a +0.10, and the average opinion of the women was a +0.35. Neither of these opinions were significantly different (t (df = 162) = 1.47, p = 0.147 and t (df = 162) = 1.53, p = 0.153).

Responses to the Evaluation Questions from the Experimental Section

| | SA† | A | N | D | SD | Mean |
|---|---|---|---|---|---|---|
| **Laboratory** | | | | | | |
| 11. The laboratory activities were interesting. | 5* | 28 | 25 | 19 | 23 | -0.27 |
| | 7 | 36 | 29 | 14 | 14 | +0.09 |
| 13. The laboratory activities were a waste of | 11 | 20 | 23 | 30 | 16 | -0.19 |
| time. | 5 | 7 | 25 | 52 | 11 | -0.59 |
| 14. The laboratory activities helped me | 3 | 23 | 30 | 29 | 15 | -0.29 |
| understand how to solve physics problems. | 7 | 23 | 25 | 29 | 16 | -0.25 |
| **Problem Solving Sessions** | | | | | | |
| 15. The problem solving sessions were | 6 | 22 | 24 | 29 | 19 | -0.33 |
| interesting. | 9 | 27 | 25 | 27 | 12 | -0.05 |
| 17. The problem solving sessions were a waste | 9 | 12 | 34 | 34 | 11 | -0.26 |
| of time. | 5 | 20 | 14 | 43 | 18 | -0.50 |
| 18. The problem solving sessions helped me | 8 | 46 | 26 | 12 | 8 | -0.33 |
| understand how to solve physics problems. | 18 | 43 | 21 | 16 | 2 | +0.59 |
| **Problem Solving Methods** | | | | | | |
| 20. The problem solving methods used in this | 19 | 20 | 23 | 29 | 8 | +0.13 |
| course were a waste of time. | 2 | 18 | 16 | 46 | 18 | -0.59 |
| 21. The problem solving methods used in this | 10 | 43 | 19 | 18 | 10 | +0.23 |
| course helped me solve physics problems. | 14 | 52 | 20 | 9 | 5 | +0.61 |

† The rating scale was 2 = strongly agree, 1 = agree, 0 = neutral, -1 = disagree, and -2 = strongly disagree.

* First row percentages are for men in the experimental section (n = 120); second row percentages are for women in the experimental section (n = 44).

Table 4.18b:

Responses to the Evaluation Questions from the Traditional Section

| | SA† | A | N | D | SD | Mean |
|---|---|---|---|---|---|---|
| **Laboratory** | | | | | | |
| 11. The laboratory activities were interesting. | 3* | 26 | 30 | 26 | 14 | -0.23 |
| | 5 | 44 | 29 | 17 | 5 | +0.27 |
| 13. The laboratory activities were a waste of | 10 | 25 | 22 | 30 | 13 | -0.12 |
| time. | 5 | 15 | 19.5 | 41 | 19.5 | -0.56 |
| 14. The laboratory activities helped me | 3 | 14 | 24 | 46 | 13 | -0.53 |
| understand how to solve physics problems. | 0 | 22 | 32 | 41 | 5 | -0.29 |
| **Problem Solving Sessions** | | | | | | |
| 15. The problem solving sessions were | 6 | 36 | 32 | 15 | 11 | +0.09 |
| interesting. | 10 | 20 | 24 | 41 | 5 | +0.32 |
| 17. The problem solving sessions were a waste | 10 | 12 | 16 | 37 | 25 | -0.54 |
| of time. | 5 | 7 | 7 | 44 | 37 | -1.00 |
| 18. The problem solving sessions helped me | 16 | 46 | 24 | 6 | 8 | +0.55 |
| understand how to solve physics problems. | 20 | 56 | 9.5 | 9.5 | 5 | +0.76 |
| **Problem Solving Methods** | | | | | | |
| 20. The problem solving methods used in this | 1 | 4 | 22 | 51 | 22 | -0.88 |
| course were a waste of time. | 0 | 0 | 20 | 51 | 29 | -1.10 |
| 21. The problem solving methods used in this | 10 | 55 | 29 | 5 | 1 | +0.68 |
| course helped me solve physics problems. | 15 | 59 | 19 | 7 | 0 | +0.80 |

† The rating scale was 2 = strongly agree, 1 = agree, 0 = neutral, -1 = disagree, and -2 = strongly disagree.

* First row percentages are for men in the traditional section (n = 138); second row percentages are for women in the traditional section (n = 41).

In the traditional section, the average opinion of all three aspects of the course was a +0.27 for the men and the average opinion of the women was a +0.56. These opinions were significantly different (t (df = 177) = 2.54, p = 0.014). Again, the answers to the three questions about the laboratory, the three questions about the problem solving sessions, and the two questions about the problem solving methods were averaged individually in order to find out more about where the differences between men and women were.

In the traditional section, the main difference between the men and the women in the experimental section was in their opinion of the laboratory, which the women liked better than the men did. The average opinion of the men about the laboratory was a -0.21, and the average opinion of the women was a +0.18. These opinions were significantly different (t (df = 177) = 2.41, p = 0.018). The average opinion of the men about the problem solving sessions was a +0.40, and the average opinion of the women was a +0.69. And the average opinion of the men about the problem solving methods was a +0.79, while the average opinion of the women about the problem solving methods was a +0.95. Neither of these opinions were significantly different (t (df = 177) = 1.60, p = 0.115 and t (df=177) = 1.33, p = 0.187).

Women had higher opinions than men of both the experimental section and the traditional section. The experimental section was designed to appeal to a broad population, with the use of cooperative grouping in the laboratory and problem solving sessions and the introduction to an explicit problem solving strategy. When questions about all three of these aspects of the course are grouped together, women seem to like these aspects of both experimental and the traditional section better than the men did. When a closer look

is taken at which of the aspects is influencing the total, women in the experimental section rate the problem solving strategy higher than the men, and women in the traditional section rate the laboratory higher than the men.

**Summary**

This study answered two research questions:

1. If there are minimal differences between men and women in their relevant physics background and initial performance when they start an introductory physics course which was designed to appeal to a broad population, will there be differences in how much physics they learn by the end of the course?

2. Will there be differences in the opinions of men and women about an introductory physics course which was designed to appeal to a broad population?

The first question was answered in two ways, by examining students' performance in problem solving and in conceptual understanding. First there was a careful coding and scoring of four problems from the final exams. No differences were found between men and women in any aspect of problem solving in either the experimental or the traditional section. Then the conceptual understanding of the students was explored, both with the Force Concept Inventory and free-response conceptual questions. No differences were found between men and women in any of the explored conceptual areas in either the experimental or the traditional section.

The second question was answered by looking at the opinions of the men and women regarding the aspects of the courses that had been changed in the experimental section to make it appeal to a broad population. It was found that women in both the experimental and traditional sections had higher opinions of these aspects of the course than the men did, with the women in the experimental section especially liking the problem solving strategy better than the men and the women in the traditional section especially liking the laboratory better than the men.

**CHAPTER 5: CONCLUSIONS**

There is evidence that, on a national level, boys score higher than girls on science and math tests (National Science Board, 1993). Not surprisingly, there were overall differences between the women and men taking Physics 1251 at the University of Minnesota in the fall of 1993 in their high school physics background, their pretest Force Concept Inventory scores, and their initial problem solving ability, as is shown in Tables 4.1 and 4.3 on pages 86 and 88. There have been several investigations into reasons for the differences in performance between men and women. Some researchers who note the persistent differences in many populations believe that they are innate, caused by biological sex differences (Benbow, 1988; Benbow & Stanley, 1980; Maccoby & Jacklin, 1974). Others point to social and cultural differences between the experiences of boys and girls ("gender differences") as possible causes for the observed differences, such as cultural definitions of science and gender and gendered classroom experiences. There have been some indications that a course with cooperative grouping and an explicit problem solving strategy would appeal to a broad population of students, so that women and men could both learn physics and enjoy their course (e.g., Heller and Lin, 1992; Johnson & Johnson, 1989, Kahle & Meece, 1994).

The first research question is:

> 1. If there are minimal differences between men and women in their relevant physics background and initial performance when they start an introductory physics course which was designed to appeal to a broad population, will there be differences in how much physics they learn by the end of the course?

Answering the first research question would contribute to the ongoing debate on the causes of the difference in math and science performance between men and women. If the initial differences between men and women can be minimized, and there are no differences in their final performance, this finding would provide supporting evidence for the hypothesis that the usual difference in performance seen in most studies is caused by social and cultural gender differences rather than biological sex difference.

There has also been research showing that even women who plan to major in science or engineering drop out of the majors at higher rates than men do (e.g., Ware, Steckler, & Leserman, 1985). Some of the reasons that they drop out include bad experiences in their introductory courses and teachers who do not care about them (Seymour, 1992b). The experimental section of Physics 1251 in the fall of 1993 was designed in part to appeal more to women, with the inclusion of cooperative group learning and an explicit problem solving strategy. Past research has shown that women prefer cooperative learning to the more common competitive atmosphere of science courses (Johnson & Johnson, 1989) and that both men and women are less likely to drop majors or even drop out of college if they have been working in cooperative groups (Johnson, Johnson, & Smith, 1991). It has also been seen that girls and women will appreciate and use an explicit problem solving strategy even more than men will (Heller & Lin, 1992; Huffman, 1994), perhaps because they are not as well prepared in math and physics as men are (National Science Board, 1993). If women like the cooperative grouping and the explicit problem solving strategy, they might be more happy with the course, which might in turn lead to higher retention of women in the science and engineering majors that they have declared. The second research question is therefore:

2. Will there be differences in the opinions of men and women about an introductory physics course which was designed to appeal to a broad population?

The second research question was designed to contribute to the understanding of why there would be differences. It was thought that women might have a lower opinion of some of the main aspects of their introductory physics courses.

The remainder of this chapter discusses the results of the study, including its limitations and implications for further study and for instruction.

## Research Question 1

**Summary of results**

The first research question was answered by assessing the performance of matched samples of men and women on their final exams, the Force Concept Inventory, and on free response conceptual questions. A summary of these results is shown in Table 5.1.

The problem solutions from the final exams were scored with a coding scheme developed at the University of Minnesota, based on research into how experts and novices solve problems. Solutions were coded on their general approach, specific application of physics, logical progression, appropriate mathematics, and clear communication. These scores were analyzed using Wilcoxin's Matched Pairs Signed-Rank Test.

Table 5.1:  Summary of Results for Research Question 1

| | Experimental | | | Traditional | | |
|---|---|---|---|---|---|---|
| | men | women | p | men | women | p |
| problem solving final median out of 160 (range) | 114 (33-146) | 103 (37-152) | p > 0.10 | 131 (90-159) | 127 (103-155) | p > 0.10 |
| Force Concept Inventory average out of 29 (standard error) | 20.9 (0.85) | 19.9 (0.96) | p = 0.43 | 16.6 (1.70) | 15.4 (0.80) | p = 0.31 |
| free response qns. median  out of 12 (range) | 5 (0-11) | 3 (2-12) | p > 0.10 | 1 (0-4) | 1 (0-4) | p > 0.10 |

In the experimental section, there was no differences in the overall performance of matched men and women on the four problems.  There was also no sex difference in performance on any of the individual problems:  the Enterprise problem the Atwood machine problem, the balloon problem, or the ski problem.  Furthermore, there was no difference in performance on any one of the four measures of problem solving ability: general approach, specific application of physics, logical progression, or appropriate mathematics.  To provide baseline data, the traditional section was also studied.  In this section, there were also no differences in the overall performance on the four problems. There was also no difference in performance on any of the individual problems or on any one of the four measures of problem solving ability.

Students' conceptual understanding was assessed both with the Force Concept Inventory and with free-response conceptual questions.  On the Force Concept Inventory, a 29 point

multiple choice test, there was no difference in the scores of the matched sample for either the experimental section or the traditional section. Further probing into students' conceptual understanding was done through several free-response conceptual questions. The responses to these questions were coded and combined, and there was no difference in the total scores of the matched sample for either the experimental or the traditional section. There was also no difference in performance on any of the individual questions for either section.

**Discussion**

Although the student population taking Physics 1251 in the fall of 1993 followed the usual trends, with the men scoring higher than the women on both pretest and post-test measures (see Table 3.3, page 79), this was not the case with the matched sample. When men and women are matched on high school backgrounds and pretest scores, there is no difference in post-test physics performance. If there were biological sex differences in physics ability between the men and women in the matched sample, there would have been differences in their post-test scores. The fact that there were none provides supporting evidence for the hypothesis that the usual difference in performance seen in most studies is caused by social and cultural **gender** differences rather than biological **sex** difference.

There is a limit to how much this result can be generalized, which is that the matched sample in this study is different from the population of students in the course and the population at large. The matched sample has lower scores on all three pretests than the remaining population (see Table 4.3, page 88). The reason for the lower scores in the sample is that the men were matched to the women, since there were far fewer women

145

than men in the course, and the average scores of the women in the population were lower than the average scores of the men. The Force Concept Inventory scores of the women were statistically significantly different than the men's scores. The scores on the asteroid problem were also significantly lower for women than for men. And although the scores on the free response conceptual questions were so low that there was no significant difference, the range of women's scores is smaller than the range of men's scores. In both sections, men's scores range from 0-9 out of 12 possible, while women's scores range from 0-4 in the experimental section and 0-6 in the traditional section.

Why were there no women in the population of students taking Physics 1251 in the fall of 1993 with pretest scores as high as those of the highest scoring men? Unfortunately, there are no data from this study to answer this question. And it is an even larger question because this was a course for students who planned to become scientists and engineers. One might expect that the women who were not good at science would not be in the course, especially given that the course was only 22% women. In addition, previous research has shown that women who study science and engineering in college have similar backgrounds as the men in their courses do (DeBoer, 1985; Seymour, 1992b; Whigham, 1988), and this is also true for the population from the fall of 1993, as shown in Table 4.1, page 86. These studies did not look at initial understanding of physics, however. So while the results of the present study suggest that there is no biological sex difference in physics ability between the men and women in the matched sample, the question of why there are differences in performance between men and women in the population has not been answered. There is need for further study in this area, to see whether these differences already exist at the start of high school physics courses or even much earlier.

**Implications for Instruction**

Past research has shown that men often outperform women on tests of science and math ability and achievement (e.g., Benbow & Stanley, 1980; Chipman & Thomas, 1985; National Science Board, 1993). There is a growing interest in learning how to manage instruction so that courses are gender fair. Toward this end, two ideas that were implemented in the experimental section of Physics 1251 in the fall of 1993 were cooperative grouping and teaching an explicit problem solving strategy.

At first glance, it seems that neither the traditional section or the experimental section were gender fair. As can be seen on Table 3.2 on page 76, the post-test scores of men in the populations of both sections were higher than the scores of women. However, there is some discussion of what "gender fair" would mean. One model of gender-fairness is that men and women should have equivalent post-test scores, to come out of a course with the same amount of knowledge. Another model is that men and women gain the same amount of knowledge in the course, and under that model it might be expected that women have lower post-test scores, since they often have lower pretest scores. In fact, some who use this model would argue that a course where women start with lower pretest scores and end up with equivalent post-test scores is unfairly favoring women.

As can be seen on Table 3.2, the women in the populations of both sections had lower pretest scores on all three measures than the men did. Therefore, although the courses were not gender fair according to the equivalent post-test model, it is not possible to rule out the idea that the courses are gender fair according to the equivalent gain model just by looking at the scores of the population. Instead it is useful to look at the scores of the matched samples.

The result that there were no differences in physics performance of the matched samples at the end of either the experimental or the traditional section of Physics 1251 has significant implications for instruction. If either section were not gender fair, then men and women who had equivalent pretest scores would not have had equivalent post-test scores. Since there was no difference in the scores of men and women at the end of the course, both the experimental and traditional sections are gender fair and the usual difference in post-test scores can be attributed to the difference in men and women's pretest scores at the start of college courses. Therefore it is argued that both the experimental section of the course, which was specifically designed to appeal to a large population, and the traditional section, which was not, are gender fair according to the equivalent gain model.

In the Physics 1251 course from the fall of 1993, not only did the men have higher pretest scores than the women, they had larger and higher ranges of scores on both the Force Concept Inventory and the free response conceptual questions, as can be seen on Table 3.2. There are men in the course whose pretest scores are higher than those of any women in the course, and these men were necessarily not in the matched sample. The significance of the difference in ranges can be seen when the percent correct on the post-test Force Concept Inventory is plotted against the percent correct on the pretest, as in Figure 5.1.[2] Men have higher average post-test scores than the women do because they have higher pretest scores than the women do.

---

[2]The author wishes to thank Laura McCullough of the University of Minnesota for this plot.

The implications for instruction are clear. Both the experimental and traditional sections of Physics 1251 had differences in performance between the men and women in the population, but that does not mean that the courses were not fair. The sample chosen to

Figure 5.1:  Average Force Concept Inventory Post-test Score vs. Pretest Score
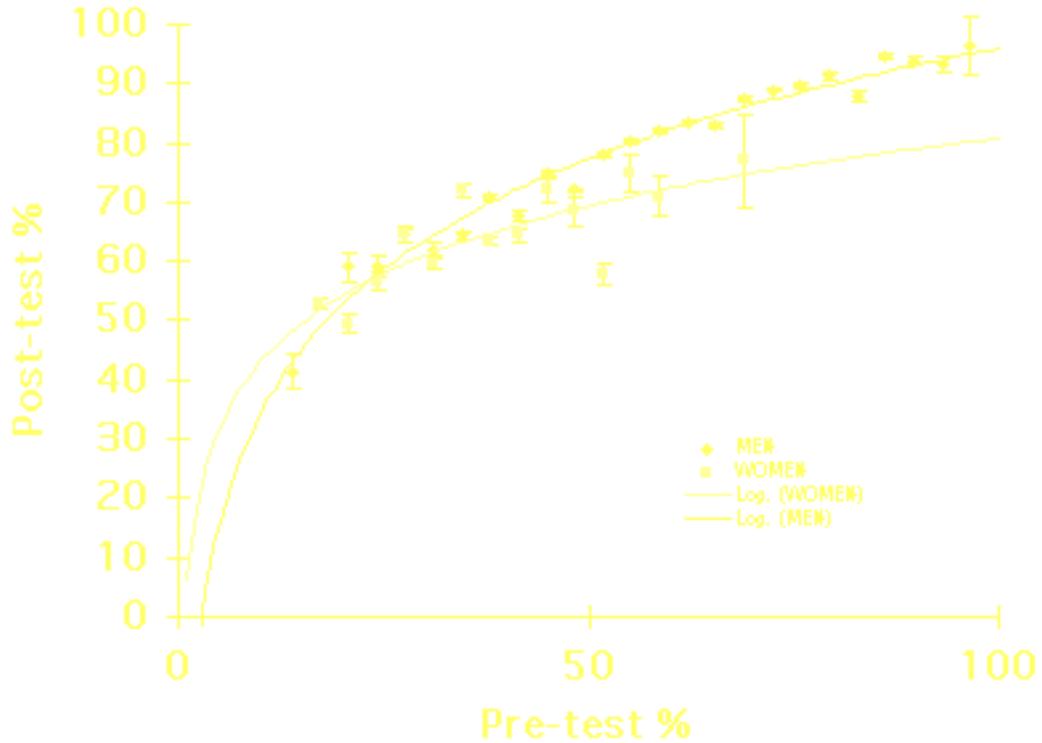


have few differences between men and women, both in their high school background and

their pretest scores, had no differences in their post-test scores.  If the courses had been

unfair, there would have been a difference.  The differences in the population post-test

scores were not due to the curriculum or instruction of the course but to the differences in

the pretest scores.

**Differences Between Sections**

Although this study was not designed to compare the experimental and traditional sections to each other, it is apparent from Tables 4.4, 4.10, and 4.11 on pages 92, 108, and 109 that the sections are not the same. It would appear that working in cooperative groups and being introduced to an explicit problem solving strategy affects the physics learning of both men and women. Although men and women did not score differently from each other within sections, both men and women in the traditional section scored better on the problem solving final exam than men and women in the experimental section, while men and women in the experimental section scored better on both tests of conceptual understanding.

There is no data from this study to conclusively say why this is so. In this study the variables between the sections were not well controlled, since comparing the sections was not the purpose of the study. The professors for each section were different people with different personalities and lecture styles. And although four problems on the final exam were the same for both sections, there was no attempt to control the form or content of earlier exams.

Consequently, there could be several reasons that the students performed differently in the two sections. For example, with regard to problem solving, one reason could be that the students may have received mixed messages from their instructors about the explicit problem solving strategy. One of the two professors consistently modeled the strategy every time he did a problem during the lecture, while the other usually chose to do more problems in less detail. It is also possible that the ten week quarter was not long enough for the students to learn to use the strategy effectively. Most of the students in this class,

Physics 1251, had taken physics before, where they had presumably solved many problems using a different strategy. It takes time to adapt to a new method. People who have been accustomed to using a novice strategy in any area, from problem solving to sports, often seem to regress in skill when they start using a more expert strategy. It is interesting to note that previous research into the success of the explicit problem solving strategy taught at the University of Minnesota had been on an algebra based course, Physics 1041, where it was found that problems solved by students who had worked in cooperative groups using the explicit strategy scored higher than problems solved by students from a traditional section (Heller, Keith, & Anderson, 1992). A smaller percentage of students in that course had taken an earlier physics course, so it could be that they had more success with the explicit strategy than the 1251 students did because they had less to unlearn. Even more significantly, that study compared problem solving performance after two quarters, not one quarter as was done in the present study. More research is clearly necessary to see how long it takes Physics 1251 students to move from a novice problem solving strategy to successful use of the explicit problem solving strategy.

While students in the traditional section scored higher on the problem solving final exam than the students in the experimental section did, they scored lower in both measures of conceptual understanding, as is shown on Tables 4.10 and 4.11 on pages 108 and 109. One possible reason for this could be that students in the experimental section were cooperatively grouped. Students working in cooperative groups are forced to talk to one another about the physics they are doing, which might well lead to a deeper conceptual understanding.

There may in fact be an interaction between attaining conceptual understanding and problem solving skills during a short course. In one study of high school students, one group of whom were taught the explicit strategy and the other group the strategy from their textbook, it was found that male high school students who are required to use a problem solving strategy actually make smaller gains in conceptual understanding than the male students who did not use the strategy, although their problem solving gains were higher (Huffman, 1994). Similarly, in the present study, the two sections seemed to learn different things differently. It is surprising to note that students in the experimental section, with their superior scores on conceptual understanding, consistently earned lower scores on their General Approaches to the problems on the final exam, as can be seen in Appendix E, pages 173 - 176. More research is necessary to see why this would be so, since the General Approach would seem to be where conceptual understanding is most applied to problem solving.

**Limitations**

The small sample size of 68 students is a limitation of this study. As was discussed in Chapter 1, the sample size is necessarily small for two reasons: the low percentage of women in the course and the low attendance at lecture. However, using matched samples is a powerful technique. The women and men have in the sample been matched on three pretest scores and on several demographic characteristics, so any differences in their post-test scores are very likely attributable to their sex. Also, it has been shown that the matched sample was quite different from the remaining population in each section. The resulting limits to how much this study can be generalized of the study have been discussed above.

The more significant limitation to the study was the need to rely on what students write in their problem solving solutions and the answers to the conceptual questions studied. At times student answers can be ambiguous, and often students do not include explanations of their reasoning even when they are asked to do so. In an attempt to interpret the answers in as straightforward a manner as possible, there was little second-guessing and little benefit of the doubt given. This was justified in part by the extensive research into student conceptions during and after introductory course, including studies which involved student interviews (e.g., Clement, 1982, 1983; diSessa, 1988, 1993; Halloun and Hestenes, 1985b; Trowbridge & McDermott, 1980, 1981).

There has been less research into the kind of assessment of student problem solving that was done in this study. There has certainly been research into novice problem solving (e.g., Chi, Feltovich, and Glaser, 1981; Finegold & Mass, 1985; Heller, Keith & Anderson, 1992; Heller & Hollabaugh, 1992; Larkin, 1980; Larkin, 1981; Larkin & Reif, 1979). However, these studies did not always address exactly when and why students make their mistakes. For example, consider a student who sets the acceleration of a system to zero incorrectly, as several did in the modified Atwood machine problem (see table in Chapter 4). In some cases, it is difficult to assess why the student did that. Is it because he or she has a conceptual physics problem, and really thinks that the acceleration is zero? Or is it done later in the problem, when a student not confident in his or her math ability either consciously or subconsciously sets acceleration to zero in order to reduce the algebra necessary? (When acceleration is non zero, the problem involves two equations and two unknowns, which can cause problems for some students.) Or is the difficulty something in between, perhaps a difficulty not with physics concepts or with algebra but with problem solving itself, where a student is unable to keep track of the physical meaning of variables for the length of time it takes to complete a problem.

There is certainly room for more research in this area, perhaps having students think aloud as they solve problems or perhaps interviewing them afterwards.

It has been shown that both the experimental and traditional sections of Physics 1251 were gender fair, since the post-test scores of men and women in the matched samples were equivalent. Another issue is whether men and women had equivalent opinions of the course.

**Research Question 2**

The rate of defection of both men and women from science, math, and engineering majors is higher than for other college majors (Seymour, 1992a), and this rate is higher for women than for men (Ware, Steckler, & Leserman, 1985; Widnall, 1988). One of the most critical times for students to re-evaluate their major and career plans is the first year of college, when many students are taking large, introductory courses (DeBoer, 1985).

The experimental section of the course was designed to appeal to women as well as to men. Previous research has shown that women are often at a disadvantage in traditional math and science classes (Fennema & Peterson, 1986; Tobias, 1990; Tobin & Garnett, 1987) and that many women prefer to work cooperatively rather than competitively or individually (Johnson & Johnson, 1989, Kahle & Meece, 1994). Students in the experimental section of the course worked in cooperative groups in their laboratory and problem solving sessions.

There is also evidence that women would prefer to be taught an explicit problem solving strategy. Heller and Lin (1992) and Huffman (1994) found that larger proportions of

women than men adopted and liked an explicit strategy. The students in the experimental

section were exposed to an explicit strategy through one of their two professors, and they

Table 5.2: Summary of Results for Research Question 2

|  | Experimental | | | Traditional | | |
|---|---|---|---|---|---|---|
|  | men | women | p | men | women | p |
| opinions of all three aspects<br><br>(on a scale of -2 to +2)<br>(standard error) | +0.01<br>(0.07) | +0.34<br>(0.10) | 0.010 | +0.27<br>(0.06) | +0.56<br>(0.10) | 0.014 |
| . . . of the laboratory<br>(standard error) | -0.10<br>(0.09) | +0.14<br>(0.13) | 0.147 | -0.21<br>(0.08) | +0.18<br>(0.13) | 0.018 |
| . . . of the p. s. sessions<br>(standard error) | +0.10<br>(0.08) | +0.35<br>(0.14) | 0.135 | +0.40<br>(0.09) | +0.69<br>(0.16) | 0.115 |
| . . . of the p. s. strategy<br>(standard error) | +0.05<br>(0.10) | +0.60<br>(0.13) | 0.005 | +0.79<br>(0.06) | +0.95<br>(0.10) | 0.187 |

were expected to use the strategy on the problems they did in their problem sessions and on their exams.

**Summary of results**

To answer the second research question, student responses to selected items from the course evaluations were compared. The eight questions selected for study were those which directly dealt with the laboratory activities, the problem solving sessions, and the problem solving methods used in the course. The rationale for this selection was that the main differences between the experimental and the traditional sections were manifested in the problem solving and laboratory sessions, where the students in the experimental

section were cooperatively grouped, and in the problem solving methods, which were different for the two sections.

To ensure anonymity, students' names and university identification numbers were not on their course evaluations, so it was impossible to compare the evaluations belonging to members of the matched sample. However, students' sections and sex were on the evaluations, so it was possible to compare the evaluations by sex within sections.

The women in the experimental section had higher opinions of the selected aspects of the course than the men did. When questions about the laboratory, the problem solving sessions, and the problem solving strategy were analyzed by themselves, it was found that the main difference was in the opinion of the problem solving strategy, which the women liked better than the men did. The women in the traditional section also liked the corresponding aspects of their course better than the men did, with further probing showing the main difference to be that they liked the laboratory better than the men did.

**Discussion**

**Overall.** Since the cooperative grouping and explicit problem solving strategy were included in the experimental section in part to appeal to women, it is not surprising that women rated those aspects of the course higher than men did. What is surprising, in light of past research, is that women also rated those aspects of the traditional section higher than men in that section did. Past research would suggest that it would be easier for men to dominate discussions in the problem solving sessions and equipment in the laboratory sessions of the traditional section (e.g., Tobin & Garnett, 1987), which would presumably be a source of frustration for the women. Perhaps that was not the case for these women,

however, since they had already chosen a major that most women avoid. There has been research showing that female engineering majors are as likely to be self-confident as the male majors (Whigham, 1988). It is quite possible that they could hold their own in a competitive environment. It is also possible that men did dominate the discussions and equipment and that the women either did not notice or did not mind. Differences in classroom experiences between men and women have been documented as early as in elementary school (e.g., Fennema & Peterson, 1986), so by college they might not be noticeable.

Past research would also suggest that the women might have preferred a more explicit problem strategy. In the traditional section, a textbook strategy was suggested but not required. While equivalent percentages of men and women in each section had completed a calculus course, as is shown in Table 3.1 on page 74, women in each section scored statistically significantly lower on the problem solving pretest than the men did, as is shown on Table 3.2 on page 76. If women were aware that they had poor problem solving skills, they might have appreciated being shown and required to use an explicit strategy. It might be true that they did lean heavily on the textbook strategy. It certainly is true that the course evaluation should not be interpreted as a vote between the strategies taught in the two sections, since the evaluation was only of the strategy the students had used.

**Opinions about problem solving in the experimental section.** In the experimental section, the difference of opinion between the men and women was about the problem solving methods used in the course. The course evaluation had two questions about the problem solving methods, which are shown on Table 4.18, page 118. When asked if the problem solving methods were a waste of time, the average reply of the women in the

experimental section indicated disagreement while men agreed. When asked if the problem solving methods helped them to solve physics problems, the average reply of both women and men indicated agreement, although a higher percentage of women agreed than men did. When the two questions were combined for analysis, it was found that the women's opinion of the problem solving methods was statistically significantly higher than the opinion of the men.

This result is consistent with previous research. A study at the University of Minnesota found that a larger proportion of women than men in an algebra-based introductory physics course adopted the explicit strategy (Heller & Lin, 1992), and a study of high school students found that not only did the girls like the strategy better than the boys did, the boys who had to use the strategy had lower conceptual test scores than was expected (Huffman, 1994). Girls and women often start off a physics class more worried about their ability to solve problems than boys and men are, which has been suggested as a reason that they would welcome a very explicit strategy for solving problems even when boys and men dislike it and think it is time consuming and unnecessary. What is more surprising is that the women in the traditional section had a more positive opinion of their textbook strategy than the women in the experimental section had about their strategy, as is shown on Table 5.2. It could be that the explicit strategy is not suited for this population of students as well as it is for high school students or students taking an algebra-based course. Most of the students, both women and men, in the population of 1251 students had taken physics already, and over half had also taken calculus, as is shown on Table 4.1, page 86. Perhaps, although the women are more receptive to it than the men are, none of these students need an explicit problem solving strategy.

It is interesting that, although only aspects of the experimental section of the course were specifically designed to appeal to a broad population, women had higher opinion of those aspects of both courses than men did. In the experimental section, this difference seems to be concentrated in opinions about the problem solving strategy, and in the traditional section, the difference seems to be concentrated in opinions about the laboratory.

**Opinions about the laboratory in the traditional section.** The course evaluation had three questions about the laboratory of the course, which are shown on Table 4.18, page 118. When asked if the laboratory activities were interesting, the average reply of the women in the traditional section indicated agreement while men disagreed. When asked if the laboratory activities were a waste of time, the average reply of both men and women indicated disagreement, although a higher percentage of women than men disagreed. And when asked if the laboratory activities helped them to understand physics problems, the average reply of both men and women indicated disagreement again, though this time a higher percentage of men disagreed. When the three questions were combined for analysis, it was found that the women's overall opinion of the laboratory was positive, while the men's opinion was negative. This difference was statistically significant.

More research would be needed to find out where this difference came from. It is interesting that women liked lab better than men did, since previous research has shown that this is one of the places in traditional courses that can discriminate against women the most, with students falling into sex roles where the men use the equipment and the women recording the results (e.g., Tobin & Garnett, 1987). Also, note that another interpretation could be that the women disliked the lab less than the men did. The

laboratory received lower ratings than the problem solving sessions or problem solving

strategies did.

## Implications for Instruction

Past research has shown that men tend to dominate traditional classrooms (Fennema & Peterson, 1986; Tobin & Garnett, 1987) and that women prefer to work in cooperative groups (Johnson, Johnson, 1989). It has also been shown that women enter college with lower abilities in math and science than men do (National Science Board, 1993) and that they are more likely to embrace an explicit problem solving strategy (Heller & Lin, 1992; Huffman, 1994). The experimental section of Physics 1251 that was taught in the fall of 1993 included both cooperative group work and an explicit problem solving strategy in order to appeal to women.

It was found that women in the experimental section had higher opinions of these aspects of the course than men did, which was not surprising. The surprise was that women in the traditional section also rated the corresponding aspects of their section higher than the men did. The implications for instruction are not clear. As was discussed above, it is not clear from these results whether or not the men in the traditional section were dominating the laboratory and problem solving sessions. If they were, and women still rated the sessions higher than the men did, that might mean that it does not matter, that women are just as happy when they do not participate very much. Then instructors might not need to monitor themselves for giving different amounts of attention to men and women or to monitor the students for levels of participation.

It is also not clear what the relationship is between high course ratings and eventual persistence in a major. Longer studies are needed, with students who volunteer to put their names on their evaluations, in order to make this connection.

A word needs to be said about the low ratings of some of these aspects of the course. The laboratory was not rated highly by men or women; in fact the men in both sections rated it negatively. These low ratings suggest that instructors who want a laboratory with their course should think about why they do and then tell their students. Similarly, if the explicit problem solving is to be taught, the results of this study suggest that instructors will need to sell the strategy to the students. The explicit strategy was rated lower by both men and women in the experimental section than the textbook strategy was by men and women in the traditional section. If courses are being taught in a team, then all professors as well as all teaching assistants need to be enthusiastic about the strategy, modeling it themselves and showing the students how it can help them.

**Limitations**

The data from the course evaluation tell what student opinions were, but not on what they were based. This data could possibly have been gathered from further written surveys, or more ideally from interviewing students. Unfortunately, the course population was so big that it was not feasible to conduct student interviews at the end of the course. Perhaps further research could include interviews about those areas of the course where the opinions of the men and women were different.

The inspiration for studying the opinions of men and women was the differing persistence of men and women in science and engineering majors. However, because of the anonymity of the course evaluations it was not possible to track men and women with more and less positive opinions to see whether or not they persisted through to the end of the third quarter of their introductory sequence, let alone whether they finished their declared majors.

## Summary

This thesis asked two questions about an introductory, calculus-based college physics course that included cooperative grouping and an explicit problem solving strategy in order to appeal to a both men and women. First, it was asked whether a matched sample of men and women would have different post-test scores, and it was found that there was no difference in the post-test scores in either the experimental section or a traditional section. This is a significant finding, considering that in the larger population, where men had higher pretest scores than women, men had higher post-test scores. Both the experimental and traditional sections seem to be gender fair. More research is needed, at the high school level or with still younger children, to find out why the pretest scores of the women in the population were so much lower than those of the men.

Second, it was asked whether opinions about the course would be different, and it was found that in the experimental section women had a higher opinion of the problem solving strategy than the men did, while in the traditional section women had a higher opinion of the laboratory than the men did. Perhaps additional research will be able to further illuminate the causes of these differences between men and women.

# REFERENCES

American Association of University Women (1992).  <u>Shortchanging Girls, Shortchanging America:  Executive Summary</u>.  Washington, DC:  American Association of University Women.

American Institute of Physics (1996, January).  Enrollments and Degrees Report.  (AIP Pub No. R-151.32).  College Park, MD:  P.J. Mulvey & E.Dodge.

Armstrong, J.M. (1985).  A national assessment of participation and achievement in mathematics.  In S.F. Chipman, L.R. Brush, & D.M. Wilson (Eds.), <u>Women and Mathematics:  Balancing the Equation.</u>  (pp. 59-94).  Hillsdale, NJ:  Erlbaum.

Ausubel, D.P., Novak, J.D., & Hanesian, H. (1978).  Educational Psychology:  A Cognitive View.  New York:  Holt, Reinhart, and Winston.

Bacon, F. (1989).  <u>The Great Instauration</u>.  (J. Weinberger, Ed.).   Arlington Heights, IL:  Harlan Davidson.  (Original work published 1620).

Belenky, M.F., Clinchy, B.M., Goldberger, N.R., & Tarule, J.M. (1986).  <u>Women's Ways of Knowing</u>.  Basic Books:  USA.

Benbow, C.P. (1988).  Sex differences in mathematical reasoning ability in intellectually talented preadolescents:  Their nature, effects, and probable causes.  <u>Behavior and Brain Science, 11</u>, 169-232.

Benbow, C. P. & Stanley, J. C. (1980).  Sex differences in mathematics ability:  fact or artifact?  <u>Science, 210</u>, 1262-1264.

Blue, J. (1994).  Gender Differences in Conceptual Understanding.  Unpublished Master's Degree Thesis.

Brophy, J. (1985). Interactions of male and female students with male and female teachers. In L.C. Wilkinson & C.B. Marrett (Eds.), <u>Gender Influences in Classroom Interaction</u> (pp. 115-142), New York: Academic Press.

Brown, D. (1989). Students' concept of force: The importance of understanding Newton's third law. <u>Physics Education, 24</u>(6), 353-358.

Button-Shafer, J. (1990). Guest Comment: Why So Few Women? <u>American Journal of Physics, 58</u>(1), 13-14.

Button-Shafer, J. (1991). Letter to editor: Response to 'Women - why so few' by M. Levin. <u>American Journal of Phyiscs, 59</u>(3), 199.

Champagne, A.B., Gunstone, R.F., & Klopfer, L.E. (1985). Instructional consequences of students' knowledge about physical phenomena. In L.H.T. West and L. Pines (Eds.), Cognitive Stucture and Conceptual Change (pp. 61-90). Orlando, FL: Academic Press.

Chase, W.G. & Simon, H.A. (1973). Perception in Chess. <u>Cognitive Psychology, 4</u>, 55-81.

Chi, Feltovich, and Glaser (1981). Categorization and representation of physics problems by experts and novices. <u>Cognitive Science, 5</u>, 121-152.

Chipman, S.F. & Thomas, V.G. (1985). Women's participation in mathematics: Outlining the problem. In S.F. Chipman, L.R. Brush, & D.M. Wilson (Eds.), <u>Women and Mathematics: Balancing the Equation.</u> (pp. 1-24). Hillsdale, NJ: Erlbaum.

Chipman, S.F. & Wilson, D.M. (1985). Understanding mathematics course enrollment and mathematics achievement: A synthesis of the research. In S.F. Chipman, L.R. Brush, & D.M. Wilson (Eds.), <u>Women and Mathematics: Balancing the Equation.</u> (pp. 275-328). Hillsdale, NJ: Erlbaum.

Clement, J. (1982). Students' preconceptions in introductory mechanics. <u>American Journal of Physics, 50</u>(1), 66-71.

Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In D. Gennes & A. Stevens (Eds.), <u>Mental Models</u> (pp. 325-329). Hillsdale, NJ: Erlbaum.

Collins, A., Brown, J.S., & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L.B. Resnick (Ed.), <u>Knowing, Teaching and Instruction: Essays in Honor of Robert Glaser</u> (pp. 453-494). Hillsdale, NJ: Erlbaum.

DeBoer, G.E. (1985). Characteristics of male and female students who experienced success or failure in their first science course. <u>Journal of Research in Science Teaching, 22</u>(2), 153-162.

DeBoer, G.E. (1986). Percieved science ability as a factor in the college selections of men and women in college. <u>Journal of Research in Science Teaching, 23</u>(4), 343-352.

diSessa, A. (1988). Knowledge in pieces. In G. Forman and P. Pufall (Eds.), <u>Constructivism in the Computer Age</u> (pp. 49-70). Hillsdale, NJ: Erlbaum.

diSessa, A. (1993). Toward an epistemology of physics. <u>Cognition and Instruction, 10</u>(2&3), 105-225.

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversey. <u>Review of Educational Research, 62</u>(1), 61-84.

Fennema, E. & Peterson, P. L. (1986). Teacher-student interactions and sex-related differences in learning mathematics. <u>Teaching and Teacher Education, 2</u>(1), 19-42.

Finegold, M. & Mass, R. (1985). Differences in the processes of solving physics problems between good physics problem solvers and poor physics problem solvers. <u>Research in Science & Technological Education, 3</u>(1), 59-67.

Fishbane, Gasiorowicz, & Thornton. (1994). <u>Physics for Scientists and Engineers</u>.

Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks. Review of Educational Research, 59(2), 185-213.

Goldberg, S. (1991). Letter to editor: Are males better at mathematical reasoning and if so why? A reply to Rushaki. American Journal of Phyiscs, 59(9), 775.

Gunstone, R.F. (1987). Student understanding in mechanics: A large population survey. American Journal of Physics, 55(8), 691-696.

Halloun, I.A. & Hestenes, D. (1985a). The initial knowledge state of college physics students. American Journal of Physics, 53(11), 1043-1055.

Halloun, I.A. & Hestenes, D. (1985b). Common sense conceptions about motion. American Journal of Physics, 53(11), 1056-1065.

Harding, S. (1986). The Science Question in Feminism. Cornell University Press: Ithaca, NY.

Hawkins, B. (1991). Letter to editor: Fewer women than there could be. American Journal of Phyiscs, 59(4), 295.

Heller, K. & Heller, P. (1992). The Competent Problem Solver: A Strategy for Solving Problems in Phyiscs. Unpublished.

Heller, K. & Heller, P. (1993). Breaking the Circle: A Teaching Apprenticeship Program for Physics Graduate Students. Unpublished (grant application).

Heller, P. & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping, part 2: Designing problems and structuring groups. American Journal of Physics, 60(7), 637-644.

Heller, P. & Huffman, D. (1995). Interpreting the Force Concept Inventory: A reply to Hestenes and Halloun. The Physics Teacher, 33(8), 503, 507-511.

Heller, P. & Lin, H. (1992). Teaching physics problem solving through cooperative learning: Do men perform better than women? Paper presented at the annual meeting of the National Association of Research in Science Teaching, Boston.

Heller, P., Keith, R., & Anderson, S. (1992). Teaching physics problem solving through cooperative grouping. Part 1: Group versus individual problem solving. American Journal of Physics, 60, 627-636.

Hestenes, D. & Halloun, I. (1995). Interpreting the Force Concept Inventory: A response to March 1995 critique by Huffman and Heller. The Physics Teacher, 33(8), 502, 504-506.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. Physics Teacher, 30 (March), 141-158.

Hinsley, D.A., Hayes, J.R., & Simon, H.A. (1977). From words to equations: Meaning and representation in algebra word problems. In M.A. Just & P.A. Carpenter (Eds.), Cognitive Processes in Comprehension (pp. 89-106). Hillsdale, NJ: Erlbaum.

Howell, D.C. (1992). Statistical Methods for Psychology, Third Ed. Boston: PWS-Kent.

Huffman, D. & Heller, P. (1995). What does the Force Concept Inventory actually measure? The Physics Teacher, 33(3), 138-143.

Huffman, D. W. (1994). The effect of explicit problem solving instruction on students' conceptual understanding of Newton's laws. Unpublished Thesis.

Johnson, D. W. & Johnson, R. T. (1989). Cooperation and Competition: Theory and Research. Edina, MN: Interaction Book Company.

Johnson, D.W., Johnson, R.T., & Smith, K.A. (1991). Active Learning: Cooperation in the College Classroom. Interaction Book Company: Edina, MN.

Jones, M.G. & Wheately, J. (1990). Gender differences in teacher-student interactions in science classrooms. Journal of Research in Science Teaching, 27(9), 861-874.

Kahle, J.B. & Lakes, M.K. (1993). The myth of equality in science classrooms. Journal of Research in Science Teaching, 20(2), 131-140.

Kahle, J.B. (1990). Real students take chemistry and physics: Gender issues. In K. Tobin, J.B. Kahle, and B.J. Frazer (Eds.), Windows into Science Classrooms: Problems Associated with Higher-Level Cognitive Learning (p. 92-134). New York: Falmer.

Kahle, J.B. and Meece, J. (1994). Research on gender issues in the classroom. In D. Gabel (Ed.), Handbook of Research on Science Teaching and Learning. Washington, DC: National Science Teachers Association.

Keller, E.F. (1985). Reflections on Gender and Science. Yale University Press: New Haven, CT.

Larkin, J. (1981) Cognition of learning physics. American Journal of Physics, 49(6), 534-541.

Larkin, J., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.

Larkin, J.H. & Reif, F. (1979). Understanding and teaching problem solving in physics. European Journal of Science Education, 1, 191-203.

Larkin, J.H. (1979). Information processing models and science instruction. In J. Lochhead & J. Clement (Eds.), Cognitive Process Instruction. Philadelphia: Franklin Institute.

Larkin, J.H. (1980). Skilled problem solving in physics: A hierarchical planning model. Journal of Structural Learning, 6, 271-297.

Larkin, J.H. (1983).  The role of problem representation in physics.  In D. Gentner & A.L. Stevens (Eds.), Mental Models.  Hillsdale, NJ:  Erlbaum.

Lawrenz, F., Heller, K., Johnson, W., and Heller, P. (1991?).  Evaluation of an Innovative, Introductory Physics Course.  Unpublished.

Leder, G.C. (1990).  Gender differences in mathematics:  An overview.  In E. Fennema & G. C. Leder (Eds.), Mathematics and Gender.  New York:  Teachers College Press.

Levin, M. (1990).  Women - why so few?  American Journal of Phyiscs, 58(10), 905-906.

Levi-Strauss, C. (1971).  "The Family" in Man, Culture, and Society, edited by H. Shapiro.  London:  Oxford U Press.

Maccoby, E.E. & Jacklin, C.N. (1974).  The Psychology of Sex Differences. Stanford, California:  Stanford University Press.

Maloney, D.P. (1988).  Novice rules for projectile motion.  Science Education, 72, 501-513.

Maloney, D.P. (1994).  Research on problem solving:  Physics.  In D. Gabel (Ed.), Handbook of Research on Science Teaching and Learning.  Washington, DC:  National Science Teachers Association.

McClosky, M. (1983).  Naive theories of motion.  In D. Gentner and A. Stevens (Eds.), Mental Models.  Hillsdale, NJ:  Erlbaum.

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

Minstrell, J. (1991).  Facets of students' knowledge and relevant instruction.  In R. Duit, F. Goldberg, & H. Niedderer (Eds.), Research in Physics Learning:  Theoretical Issues and Empirical Studies.  Proceedings of an International Workshop, University of Bremen.

National Science Board (1993). <u>Science and Engineering Indicators</u> (NSB 91-1). Washington, D.C. Government Printing Office.

National Science Foundation (1995). <u>1994 Indicators of Science and Mathematics Education</u> (Draft Copy).

Orenstein, P. (1994). <u>Schoolgirls: Young Women, Self-Esteem, and the Confidence Gap</u>. Doubleday: NY.

Pascarella, E.T. & Terenzini, P.T. (1991). <u>How College Affects Students</u>. San Fransisco: Jossey-Bass.

Polya, G. (1945). <u>How to Solve It.</u> Garden City, NY: Doubleday.

Reif, F., Larkin, J.H., & Brackett, G.C. (1976). Teaching general learning and problem-solving skills. <u>American Journal of Physics, 44</u>(3), 212-217.

Rosser, S. (1990). <u>Female Friendly Science</u>. New York: Pergamon Press.

Rubin, G. (1975). The traffic in women: Notes on the "political economy" of sex. In R.R. Reiter (Ed.), <u>Toward an Anthropology of Women</u> (pp. 157-210), New York: Monthly Review Press.

Rushaki, M.B. (1990). Guest comment: Are there innate cognitive gender differences? Some comments on the evidence in response to a letter by M. Levin. <u>American Journal of Physics, 59</u>(1), 11-14.

Sadker, M. & Sadker, D. (1994) <u>Failing at Fairness: How America's Schools Cheat Girls</u>. Scribners: NY.

Scantleburyz, K. & Baker, D. (1992). Achieving a gender equitable classroom. In F. Lawrenz, K. Cochran, J. Krajick and P. Simpson (Eds.), <u>Research Matters to the Science Teacher</u>. Manhattan, KS: National Association for Research in Science Teaching.

Schiebinger, L. (1987). The history and philosophy of women in science: A review essay. In <u>Sex and Scientific Inquiry</u>, S. Harding & J. O'Barr, Eds. (pp. 7-34). Chicago: University of Chicago.

Seymour, E. (1992a). "The problem iceberg" in science, math, and engineering education: Student explanations for high attrition rates. <u>Journal of College Science Teaching, 21</u>(4), 230-238.

Seymour, E. (1992b). Undergraduate problems with teaching and advising in SME majors -- explaining gender differences in attrition rates. <u>Journal of College Science Teaching, 21</u>(5), 284-292.

Siegel, S. & Castellan, N.J., Jr. (1988). <u>Nonparametric Statistics for the Behavioral Sciences, Second Edition.</u> New York: McGraw Hill.

Tarmizi, R.A. & Sweller, J. (1988). Guidance during mathematical problem solving. <u>Journal of Educational Psychology, 80</u>(4), 424-436.

Tobias, S. (1990). <u>They're Not Dumb, They're Different</u>. Tuscon, AZ: Research Corporation.

Tobin, K. & Garnett, P. (1987). Gender related differences in science activities. <u>Science Education, 71</u>(1), 91-103.

Trowbridge, D.E. & McDermott, L.C. (1980). Investigation of student understanding of the concept of velocity in one dimension. American Journal of Phyiscs, 48(12), 1020-1028.

Trowbridge, D.E. & McDermott, L.C. (1981). Investigation of student understanding of the concept of acceleration in one dimension. <u>American Journal of Physics, 49</u>(1), 242-253.

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. Washington, DC: National Science Teachers Association.

Ware, N. C., Steckler, N. A., & Leserman, J. (1985). Undergraduate women: Who chooses a science major? Journal of Higher Education, 56(1), 73-84.

Whigham, M.A. (1988). Gender-related differences in engineering students. NACADA Journal, 8(1), 35-45.

Widnall, S. (1988). AAAS presidential lecture: Voices from the pipeline. Science, 241, 1740-1745.

# APPENDIX A:  CODING WRITTEN PROBLEM SOLUTIONS

## 1.   General Approach

a   Nothing written.

b   Invalid or inappropriate principles (general formulas) are used.

c   The solution indicates a clear misunderstanding of how the central principle(s) are systematically applied in general to physical events.

d   The solution indicates an absurd assumption or interpretation regarding certain information needed for solution of the problem.  The assumption/interpretation contradicts the assumption/interpretation that the instructor feels it reasonable to expect from any student who has been actively enrolled in class up to that point in the course.

e   The solution approach is partially correct.  The solution includes correct identification of the central principle; but another concept important to the solution is either omitted, or there is indication of a serious misunderstanding of this concept.

f   The solution approach is mostly correct but a serious error is made about *certain features of the physical events*.

g   The solution correctly uses all of the required principles.  Errors in the solution are in the details of application to the specific problem, rather than in the general application of concepts and principles to physical events.

## 2. Specific Application of Physics

a   Nothing written.

b   Difficult to assess because the individual's use of principles is fundamentally flawed.
Because it is difficult to characterize the nature of the individual's approach, it is
impossible to determine whether or not the individual applied the ideas in a
consistent manner.

c   Specific equations are incomplete.  Not all of the equations needed for a correct
solution are presented.

d   Confusion regarding resolution of vectors into components.

e   Wrong variable substitution:  The specific equations exhibit an incorrect variable
substitution.

f   Careless use of coordinate axes or inconsistent attention to direction of vector
quantities:  The specific equations exhibit inconsistencies with regard to the signs
associated with variable quantities (e.g. In a problem where the v and a of an object
are in the same direction, the equation assigns different signs to the v and a
variables).

g   Careless substitution of given information:  Incorrect given information is substituted
into equation for specified variable.

h   Specific equations do not exhibit clear inconsistencies with student's general physics
approach and solution seems quite complete in its identification of quantities and
their relative directions.

## 3. Logical Progression

a   Nothing written.

b   Not applicable.  Solution is essentially a one-step problem, i.e. individual's solution involves given information substituted into a single principle relationship.

c   Solution does not show a logical progression in the use of equations.  The use of equations appears haphazard.

d   Solution is logical to a point, then one or more illogical or unnecessary jump is made.  Student may not understand how to combine equations to isolate variables.  In solution it may appear that earlier physics claims are abandoned in an attempt to reach an mathematical solution.

e   Solution is logical but unfinished.

f   Solution involves occasional unnecessary calculations but there is a logical progression of equations that leads to an answer.

g   Solution progresses from general principles to answer.  (Solution proceeds in a straightforward manner toward solution.)  Solution is successful in isolating desired unknown.

## 4. Appropriate Mathematics

a   Nothing written.

b   Solution is terminated for no apparent reason.

c   When an obstacle to mathematical solution (e.g. incorrect occurrence of $\sqrt{-1}$) is encountered, either "math magic" or additional (non-justified) relationships are introduced in order to get an answer or the solution is terminated.

d   Solution violates rules of algebra, arithmetic, or calculus (e.g. $\frac{x}{a+b} = \frac{x}{a} + \frac{x}{b}$).

Students apparently does not have mastery of basic mathematical operations or of transitive, commutative, or distributive properties of numbers.

e   Mistakes from line to line, like sign changes.

f   Mathematics is nearly correct , with only minor mistake such as a calculator error  or neglected factor of 2 and complete.

g   Mathematics is correct.

**APPENDIX B: FORCE CONCEPT INVENTORY**
**(correct answers appear in bold face)**

1. Two metal balls are the same size, but one weighs twice as much as the other. The balls are dropped from the top of a two story building at the same instant of time. The time it takes the balls to reach the ground below will be:

   (A) about half as long for the heavier ball.
   (B) about half as long for the lighter ball.
   **(C) about the same time for both balls.**
   (D) considerably less for the heavier ball, but not necessarily half as long.
   (E) considerably less for the lighter ball, but not necessarily half as long.

2. Imagine a head-on collision between a large truck and a small compact car. During the collision,

   (A) the truck exerts a greater amount of force on the car than the car exerts on the truck.
   (B) the car exerts a greater amount of force on the truck than the truck exerts on the car.
   (C) neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.
   (D) the truck exerts a force on the car but the car doesn't exert a force on the truck.
   **(E) the truck exerts the same amount of force on the car as the car exerts on the truck.**

3. Two steel balls, one of which weighs twice as much as the other, roll off of a horizontal table with the same speeds. In this situation:

   **(A) both balls impact the floor at approximately the same horizontal distance from the base of the table.**
   (B) the heavier ball impacts the floor at about half the horizontal distance from the base of the table than does the lighter.
   (C) the lighter ball impacts the floor at about half the horizontal distance from the base of the table than does the heavier.
   (D) the heavier ball hits considerably closer to the base of the table than the lighter, but not necessarily half the horizontal distance.

(E) the lighter ball hits considerably closer to the base of the table than the heavier, but not necessarily half the horizontal distance.

4.  A heavy ball is attached to a string and swung in a circular path in a horizontal plane as illustrated in the diagram to the right. At the point indicated in the diagram, the string suddenly breaks at the ball. If these events were observed from directly above, indicate the path of the ball after the string breaks.
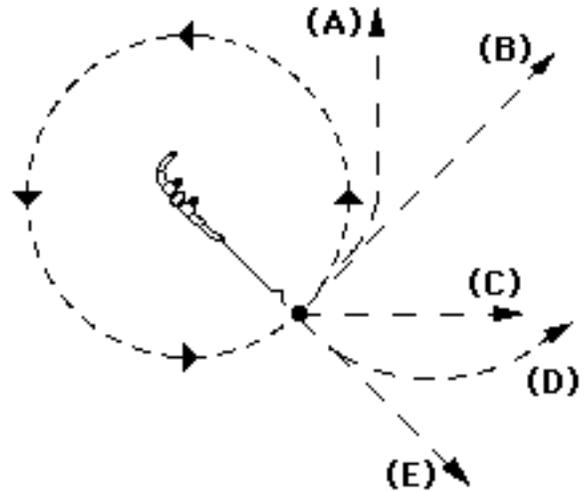
    **(correct answer:  B)**

5.  A boy throws a steel ball straight up. **Disregarding any effects of air resistance** the force(s) acting on the ball until it returns to the ground is (are):

    (A)  its weight vertically downward along with a steadily decreasing upward force.
    (B)  a steadily decreasing upward force from the moment it leaves the hand until it reaches its highest point beyond which there is a steadily increasing downward force of gravity as the object gets closer to the earth.
    (C)  a constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point after which there is only the constant downward force of gravity.
    **(D) a constant downward force of gravity only.**
    (E)  none of the above -- the ball falls back down to the earth simply because that is its natural action.

**Use the statement and diagram below to answer the next four questions:**
**The diagram depicts a hockey puck sliding with a <u>constant velocity</u> from point "a" to point "b" along a <u>frictionless horizontal surface.</u> When the puck reaches point "b" it receives an instantaneous horizontal "kick" in the direction of the heavy print arrow.**

182
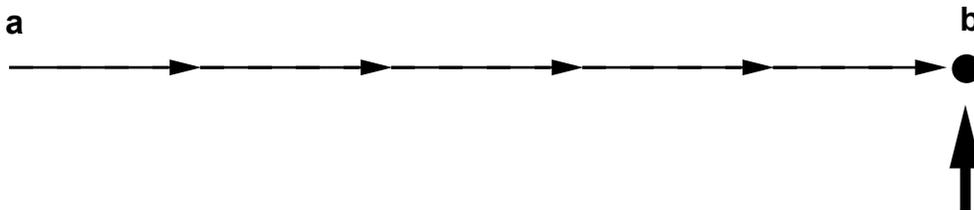
6. Along which of the paths below will the hockey puck move **after** receiving the "kick"?



(correct answer: B)
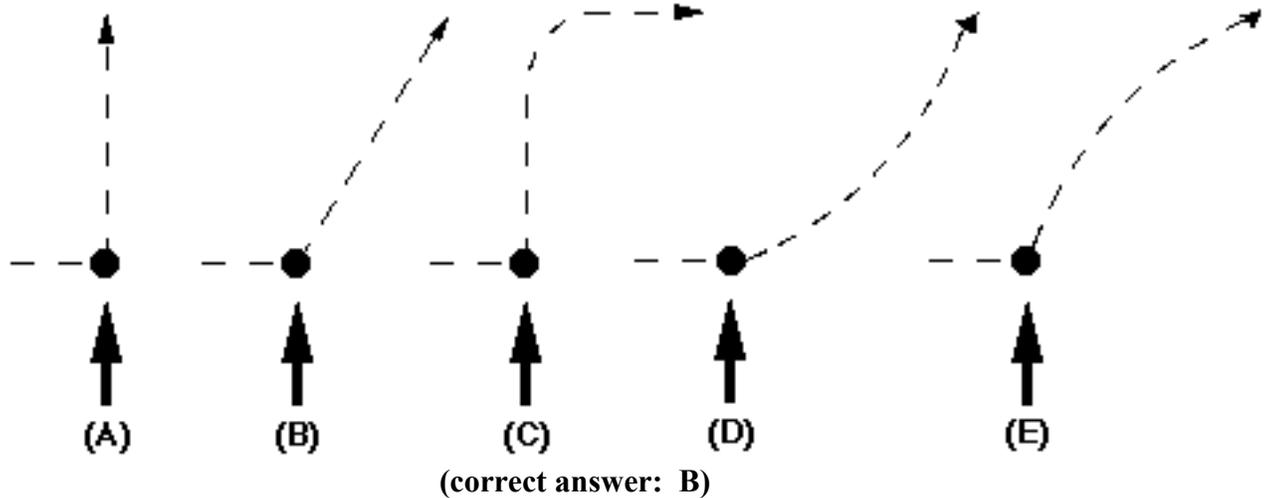
7. The speed of the puck just **after** it receives the "kick"?

    (A) Equal to the speed "$v_0$" it had before it received the "kick".
    (B) Equal to the speed "$v$" it acquires from the "kick", and independent of the speed "$v_0$".
    (C) Equal to the arithmetic sum of speeds "$v_0$" and "$v$".
    (D) Smaller than either of speeds "$v_0$" or "$v$".
    **(E) Greater than either of speeds "$v_0$" or "$v$", but smaller than the arithmetic sum of these two speeds.**

8. Along the **frictionless** path you have chosen, how does the speed of the puck vary **after** receiving the "kick"?

    **(A) No change.**
    (B) Continuously increasing.
    (C) Continuously decreasing.
    (D) Increasing for a while, and decreasing thereafter.
    (E) Constant for a while, and decreasing thereafter.

9. The main forces acting, **after** the "kick", on the puck along the path you have chosen are:

    (A) the downward force due to gravity and the effect of air pressure.
    (B) the downward force of gravity and the horizontal force of momentum **in the direction of motion.**
    (C) the downward force of gravity, the upward force exerted by the table, and a horizontal force acting on the puck **in the direction of motion.**

**(D) the downward force of gravity and an upward force exerted on the puck by the table.**

(E)  gravity does not exert a force on the puck, it falls because of the intrinsic tendency of the object to fall to its natural place.

10.  The accompanying diagram depicts a semicircular channel that has been securely anchored, in a **horizontal plane**, to a table top.  A ball enters the channel at "1" and exits at "2".  Which of the path representations would most nearly correspond to the path of the ball as it exits the channel at "2" and rolls across the table top.
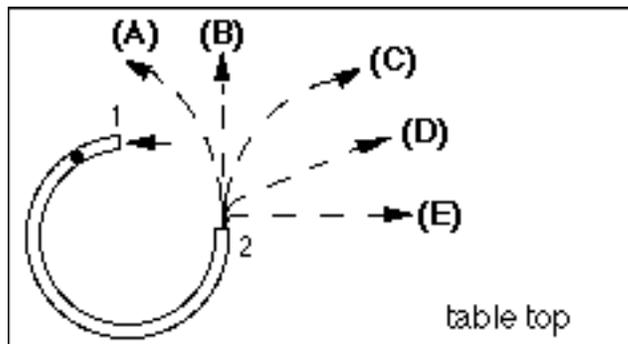


**(correct answer:  B)**

**Two students, student "a" who has a mass of 95 kg and student "b" who has a mass of 77 kg sit in identical office chairs facing each other.  Student "a" places his bare feet on student "b's" knees, as shown below.  Student "a" then suddenly pushes outward with his feet, causing both chairs to move.**

11.  In this situation,

(A)  neither student exerts a force on the other.
(B)  student "a" exerts a force on "b", but "b" doesn't exert any force on "a".
(C)  each student exerts a force on the other but "b" exerts the larger force.
(D)  each student exerts a force on the other but "a" exerts the larger force.
**(E) each student exerts the same amount of force on the other.**
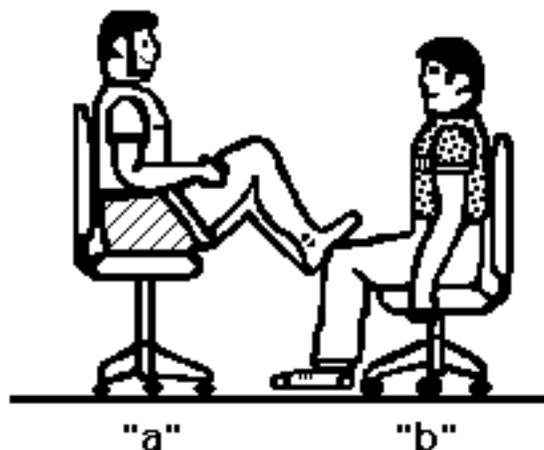


12.  A book is at rest on a table top. Which of the following force(s) is(are) acting on the book?

1. A downward force due to gravity.
2. The upward force by the table.
3. A net downward force due to air pressure.
4. A net upward force due to air pressure.

(A)  1 only
**(B) 1 and 2**
(C)  1, 2, and 3
**(D) 1, 2, and 4**
(E)  none of these, since the book is at rest there are no forces acting on it.

**Refer to the following statement and diagram while answering the next two questions.**

**A large truck breaks down out on the road and receives a push back into town by a small compact car.**



13. While the car, still pushing the truck, is **<u>speeding up</u>** to get up to cruising speed;

    **(A) the amount of force of the car pushing against the truck is equal to that of the truck pushing back against the car.**
    (B) the amount of force of the car pushing against the truck is less than that of the truck pushing back against the car.
    (C) the amount of force of the car pushing against the truck is greater than that of the truck pushing against the car.
    (D) the car's engine is running so it applies a force as it pushes against the truck but the trucks engine is not running so it can't push back against the car, the truck is pushed forward simply because it is in the way of the car.
    (E) neither the car nor the truck exert any force on the other, the truck is pushed forward simply because it is in the way of the car.

14. After the person in the car, while pushing the truck, reaches the cruising speed at which he/she wishes to continue to travel at a constant speed;

    **(A) the amount of force of the car pushing against the truck is equal to that of the truck pushing back against the car.**
    (B) the amount of force of the car pushing against the truck is less than that of the truck pushing back against the car.
    (C) the amount of force of the car pushing against the truck is greater than that of the truck pushing against the car.
    (D) the car's engine is running so it applies a force as it pushes against the truck but the trucks engine is not running so it can't push back against the car, the truck is pushed forward simply because it is in the way of the car.
    (E) neither the car nor the truck exert any force on the other, the truck is pushed forward simply because it is in the way of the car.

15. When a rubber ball dropped from rest bounces off the floor, its direction of motion is reversed because;

    (A) the energy of the ball is conserved.
    (B) the momentum of the ball is conserved.

**(C) the floor exerts a force on the ball that stops its fall and then drives it upward.**

(D) the floor is in the way and the ball has to keep moving.

(E) none of the above.

16. Which of the paths in the diagram below best represents the path of the cannon ball?
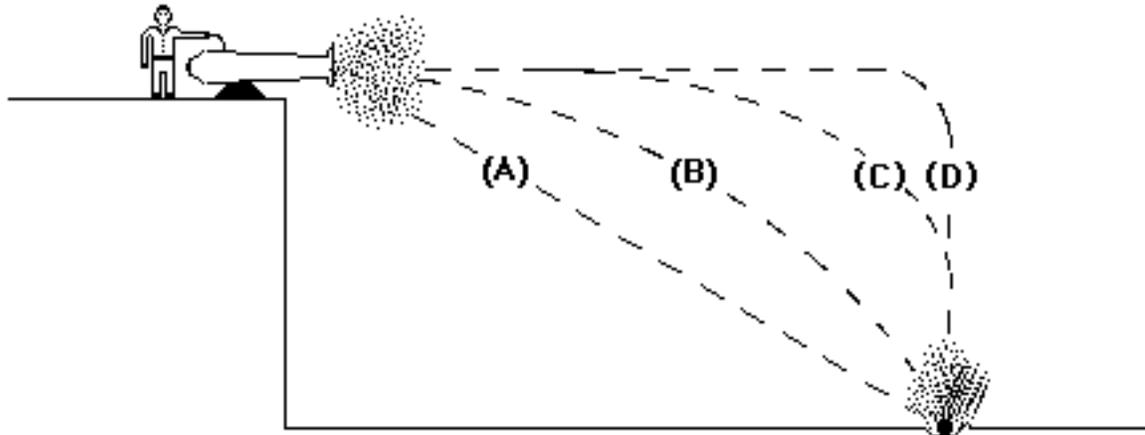


**(correct answer:  B)**

17. A stone falling from the roof of a single story building to the surface of the earth;

    (A) reaches its maximum speed quite soon after release and then falls at a
    constant speed thereafter.
    (B) speeds up as it falls, primarily because the closer the stone gets to the earth,
    the stronger the gravitational attraction.
    **(C) speeds up because of the constant gravitational force acting on it.**
    (D) falls because of the intrinsic tendency of all objects to fall toward the earth.
    (E) falls because of a combination of the force of gravity and the air pressure
    pushing it downward.

**When responding to the following question, assume that any frictional forces due to
air resistance are so small that they can be ignored.**

18. An elevator, as illustrated, is being lifted up an elevator shaft by a steel cable. When
    the elevator is moving up the shaft at a **constant velocity;**

(A) the upward force on the elevator by the cable is greater than the downward force of gravity.

**(B) the amount of upward force on the elevator by the cables equal to that of the downward force of gravity.**

(C) the upward force on the elevator by the cable is less than the down ward force of gravity.

(D) it goes up because the cable is being shortened, not because of the force being exerted on the elevator by the cable.

(E) the upward force on the elevator by the cable is greater than the downward force due to the combined effects of air pressure and the force of gravity.
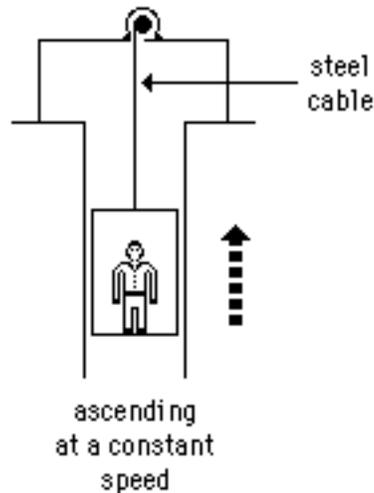
19. Two people, a large man and a boy, are pulling as hard as they can on two ropes attached to a crate as illustrated in the diagram to the right. Which of the indicated paths (A-E) would most likely correspond to the path of the crate as they pull it along?
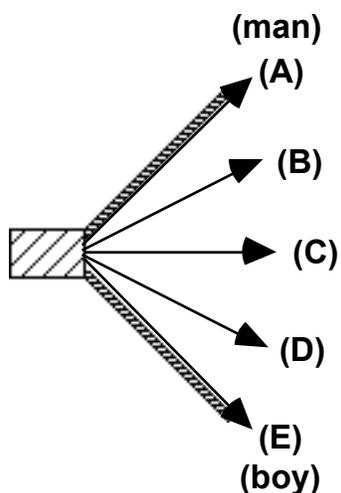
**(correct answer: B)**

**The positions of two blocks at successive 0.20 second time intervals are represented by the numbered squares in the diagram below. The blocks are moving toward the right.**
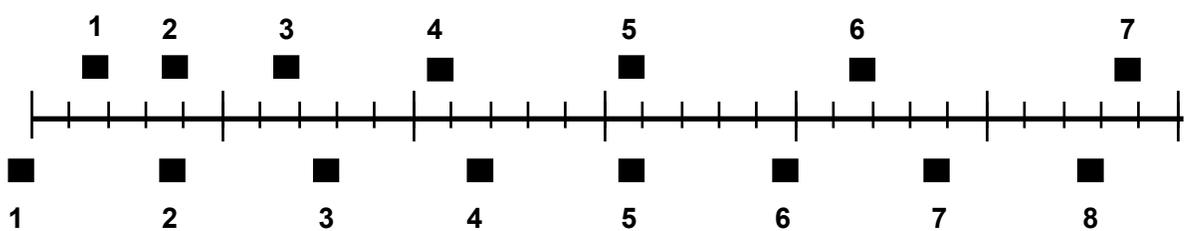
20. Do the blocks ever have the same speed?

(A) No.
(B) Yes, at instant 2.
(C) Yes, at instant 5.
(D) Yes at instant 2 and 5
**(E) Yes at some time during interval 3 to 4.**

**The positions of two blocks at successive equal time intervals are represented by numbered squares in the diagram below. The blocks are moving toward the right.**
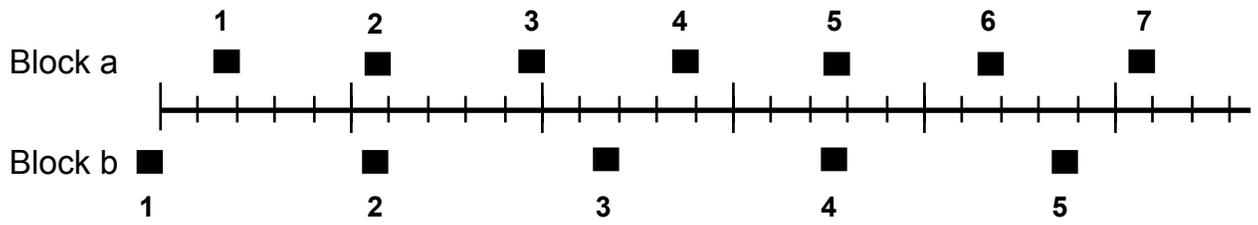
21.    The acceleration of the blocks are related as follows:

(A) acceleration of "a" > acceleration of "b"
(B) acceleration of "a" = acceleration "b" > 0
(C) acceleration of "b" > acceleration "a"
**(D) acceleration of "a" = acceleration of "b" = 0**
(E) not enough information to answer.


22. A golf ball driven down a fairway is observed to travel through the air with a trajectory (flight path) similar to that in the depiction below.
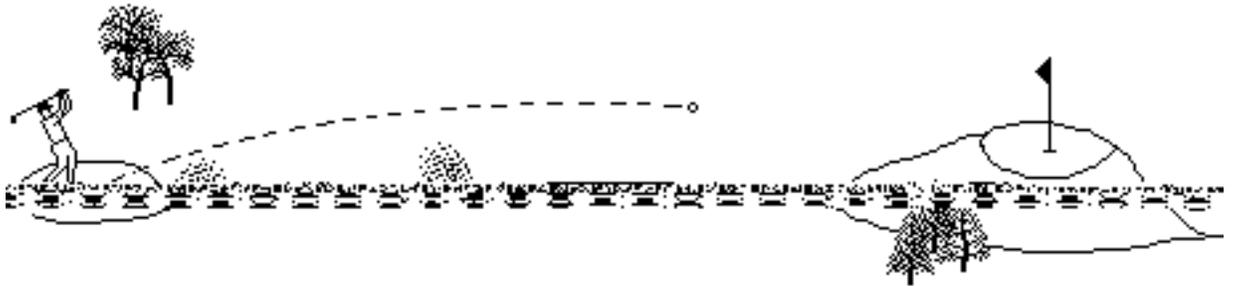
Which following force(s) is(are) acting on the golf ball during its entire flight?

1. the force of gravity.
2. the force of the "hit".
3. the force of air resistance.

(A) 1 only
(B) 1 and 2
(C) 1, 2, and 3
**(D) 1 and 3**
(E) 2 and 3

23. A bowling ball accidentally falls out of the cargo bay of an airliner as it flies along in a horizontal direction. As seen from the ground, which path would the bowling ball most closely follow after leaving the airplane?

(A)   (B)   (C)   (D)   (E)
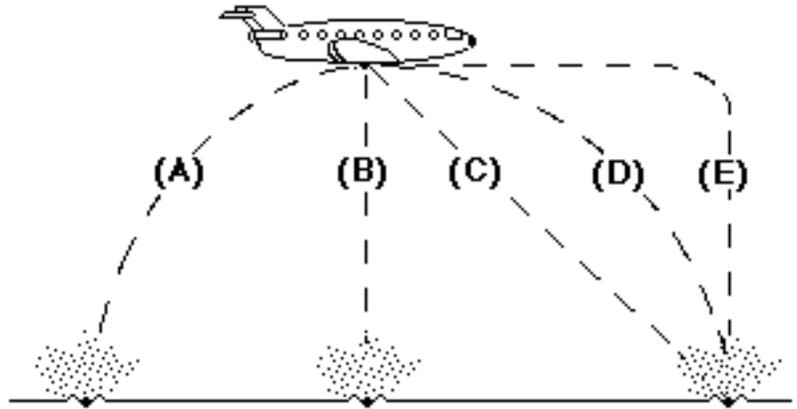
**(correct answer: D)**

**When answering the next four questions, refer to the following statement and diagram.**

A rocket, drifting sideways in outer space from position "a" to position "b" is subject to no outside forces. At "b", the rocket's engine starts to produce a constant thrust at right angles to line "ab". The engine turns off again as the rocket reaches some point "c".
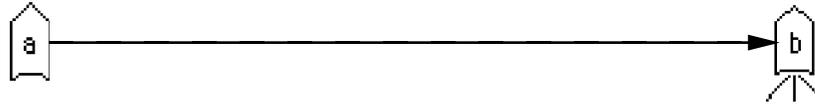
24. Which path below best represents the path of the rocket between "b" and "c"?
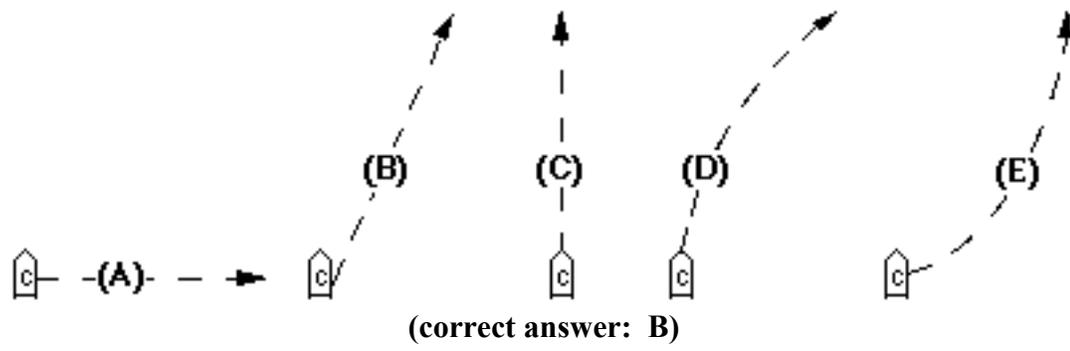
(A)        (B)        (C)        (D)        (E)
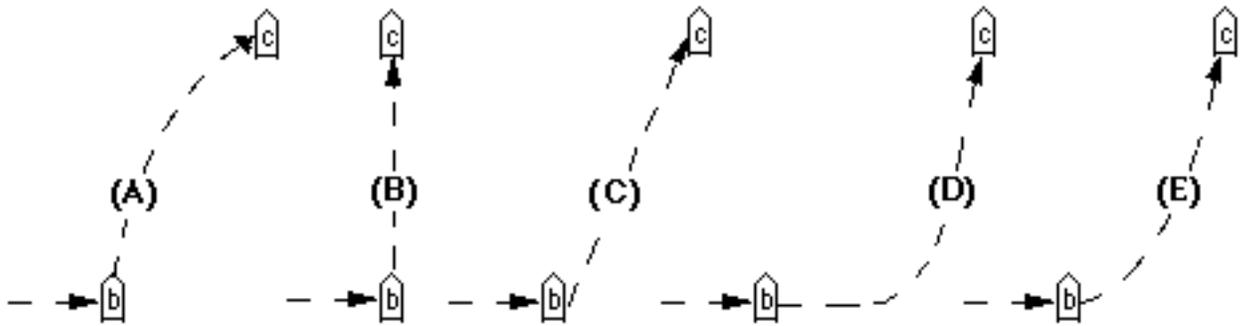
(correct answer:  E)

25. As the rocket moves from "b" to "c", its **speed** is

   (A)  constant.
   **(B) continuously increasing.**
   (C)  continuously decreasing.
   (D)  increasing for a while and constant thereafter.
   (E)  constant for a while and decreasing thereafter.

26. At "c" the rocket's engine is turned off. Which of the paths below will the rocket follow beyond "c"?

(B)    (C)    (D)    (E)

C— -(A)- — ➤    C    C    C    C

**(correct answer:  B)**

27. Beyond "c", the speed of the rocket is;

   **(A) constant.**
   (B) continuously increasing.
   (C) continuously decreasing.
   (D) increasing for a while and constant thereafter.
   (E) constant for a while and decreasing thereafter.

28. A large box is being pushed across the floor at a **constant speed** of 4.0 m/s. What can you conclude about the forces acting on the box?

   (A) If the force applied to the box is doubled, the constant speed of the box will increase to 8.0 m/s.
   (B) The amount of force applied to move the box at a constant speed must be more than its weight.
   **(C) The amount of force applied to move the box at a constant speed must be equal to the amount of the frictional forces that resist its motion.**
   (D) The amount of force applied to move the box at a constant speed must be more than the amount of the frictional forces that resist its motion.
   (E) There is a force being applied to the box to make it move but the external forces such as friction are not "real" forces they just resist motion.

29. If the force being applied to the box in the preceding problem is suddenly discontinued, the box will;

   (A) stop immediately.
   (B) continue at a constant speed for a very short period of time and then slow to a stop.
   **(C) immediately start slowing to a stop.**
   (D) continue at a constant velocity.
   (E) increase its speed for a very short period of time, then start slowing to a stop.

# APPENDIX C: CODING FOR FREE RESPONSE CONCEPTUAL
# QUESTIONS

### Codes for Ramp Problem (Velocity and Acceleration)

1. Correct: the acceleration is the same everywhere, with explanation: the forces on the ball are the same everywhere

2a. Correct without explanation

2b. Mostly correct with a small problem

3a. Indicate some confusion between acceleration and velocity

3b. Indicate complete confusion between acceleration and velocity

4a. Idiosyncratic and incomprehensible

4b. Blank

### Codes for Car and Bug Problem (Third Law)

1. Correct: The force of the car on the bug equals the force of the bug on the car

2. There is no response in this category

3a. The force of the car on the bug is not equal to the force of the bug on the car

3b. There is no force of the bug on the car

3c. There are no interactions between the bug and the car (although papers are not blank)

4a. Can't tell whether there are interactions between the bug and the car

4b. Blank

### Codes for Car and Passenger Problem (Nature of Forces)

1. Correct: All forces shown are interactions between two objects and are drawn pointing to the point of contact.

2. At least one of the forces drawn is not pointing in the right place (examples are normal forces pulling up on the passenger's head and friction pointing to the car's bumper), although all forces drawn are interactions.

3a. At least one Third Law force is drawn, a force from and not on the car or passenger (examples are passenger pushing on seat or tires pushing on road).

3b. At least one force is drawn that is not an interaction between two objects (examples include car engine, acceleration, momentum).

3c. More than one of the mistakes cited in (2), (3a) and (3b) is made.

4a. Idiosyncratic and incomprehensible

4b. Blank

### Codes for Car and Passenger Problem (Newton's Second Law)

1. Correct: Shows correct understanding of the Second Law, either talking about summing real forces or an unbalanced real force.

2. Not correct, but talks about summing things the student treats as forces (or combinations of real forces and "pseudoforces") or unbalanced "pseudoforces".

3a. Attributes acceleration to more than one force but does not talk about summing of net forces.

3b. Attributes acceleration to only one force when there are other forces on the diagram that should have been taken into account.

3c. Attributes acceleration to a force that is in the opposite direction of acceleration.

3d. Attributes acceleration to something other than a force on the passenger or car.

4a. Idiosyncratic and incomprehensible

4b. Blank

**APPENDIX D: FULL COURSE EVALUATION**

**STUDENT QUESTIONNAIRE**
**PHYSICS 1251**

**BACKGROUND INFORMATION:**

1.  What is your gender?
    (a)  Male
    (b)  Female

2.  What is your intended major?
    (a)  Physics
    (b)  Mathematics
    (c)  Engineering
    (d)  Biological Science
    (e)  Chemistry

3.  How many hours per week are you employed?
    (a)  None
    (b)  1-10 hours/week
    (c)  11-20 hours/week
    (d)  21-30 hours/week
    (e)  31 or more hours/week

4.  On average, how many hours per week did you spend studying physics outside of class?
    (a)  Less that 3 hours/week
    (b)  4-6 hours/week
    (c)  7-9 hours/week
    (d)  10-12 hours/week
    (e)  13 or more hours/week

5.  Overall, how would you rate your own performance in Physics 1251?
    (a)  Poor
    (b)  Marginal
    (c)  Adequate
    (d)  Good
    (e)  Excellent

6.  Overall, how much have you learned in this course?
    (a)  none
    (b)  a little
    (c)  some
    (d)  quite a bit

(e)    a great deal

## LECTURE, LABORATORY & RECITATION:
**1=Strongly Disagree   2=Disagree   3=Neutral   4=Agree   5=Strongly Agree**

7. The lectures were interesting.
8. The lectures helped me understand the concepts and principles of physics.

9. The lectures were a waste of time.
10. The lectures helped me understand how to solve physics problems.

11. The laboratory activities were interesting.
12. The laboratory activities helped me understand the concepts and principles of physics.

13. The laboratory activities were a waste of time.
14. The laboratory activities helped me understand how to solve physics problems.

15. The recitation sessions were interesting.
16. The recitation sessions helped me understand the concepts and principles of physics.

17. The recitation sessions were a waste of time.
18. The recitation sessions helped me understand how to solve physics problems.

**The problem solving methods used in this course**

19. were understandable.
20. were a waste of time.
21. helped me solve physics problems.
22. helped me understand the concepts and principles of physics.

## INSTRUCTOR RATINGS:

**How would you rate your:**
23. TA's attitude toward teaching.
24. TA's knowledge of the subject matter.
25. TA's level of enthusiasm.
26. TA's level of confidence.
27. TA's organizational skills.
28. TA's ability to manage lab activities.
29. TA's ability to explain physics concepts.
30. TA's ability to lead a discussion.
31. TA's general ability to relate to students.
32. TA's ability to work with individual students.
33. TA's ability to work with groups of students.
34. Overall, how would you rate your TA's teaching ability?
35. Professors' attitude toward teaching.

36. Professors' level of enthusiasm.
37. Professors' organizational skills.
38. Professors' ability to explain physics concepts.
39. Professors' general ability to relate to students.
40. Professors' knowledge of the subject matter.

41. Overall, how would you rate your professors' teaching ability?
42. Overall, how would you rate this course?