The Role Model Effect on Gender Equity:

How are Female College Students

Influenced by Female Teaching Assistants in Science?

by

Darilyn Ebert

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved November 2010 by the
Graduate Supervisory Committee:

Anton Lawson, Chair
Jane Maienschein
Julie Mustard

ARIZONA STATE UNIVERSITY

December 2010

ABSTRACT

The gender gap of women in science is an important and unresolved issue in higher education and occupational opportunities. The present study was motivated by the fact that there are typically fewer females than males advancing in science, and therefore fewer female science instructor role models. This observation inspired the questions: Are female college students influenced in a positive way by female science teaching assistants (TAs), and if so how can their influence be measured? The study tested the hypothesis that female TAs act as role models for female students and thereby encourage interest and increase overall performance. To test this "role model" hypothesis, the reasoning ability and self-efficacy of a sample of 724 introductory college biology students were assessed at the beginning and end of the Spring 2010 semester. Achievement was measured by exams and course work. Performance of four randomly formed groups was compared: 1) female students with female TAs, 2) male students with female TAs, 3) female students with male TAs, and 4) male students with male TAs. Based on the role model hypothesis, female students with female TAs were predicted to perform better than female students with male TAs. However, group comparisons revealed similar performances across all four groups in achievement, reasoning ability and self-efficacy. The slight differences found between the four groups in student exam and coursework scores were not statistically significant. Therefore, the results did not support the role model hypothesis. Given that both lecture professors in the present study were males, and given that professors

typically have more teaching experience, finer skills and knowledge of subject matter than do TAs, a future study that includes both female science professors and female TAs, may be more likely to find support for the hypothesis.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Introduction

This thesis was inspired by the under representation of women in the sciences. The Congress of the United States established a national policy ". . . that men and women have equal opportunity in education, training, and employment in scientific and technical fields" (p. 1190) by enactment of the Women In Science and Technology Equal Opportunity Act in 1980 (Handelsman, et al., 2005). Unfortunately, there are many published studies indicating that women have not yet achieved equity in these fields (Bettinger & Long, 2005; Buck, Clark, Leslie-Pelecky, Lu, & Cerda-Lizarraga, 2007; Butler & Christensen, 2003; Carrell, Page, & West, 2009; Ehrenberg, Goldhaber, & Brewer, 1995; Gilmartin, Denson, Li, Bryant, & Aschbacker, 2007; Handelsman, et al., 2005; Hanson, Schaub, & Baker, 1996; Taylor, Erwin, Ghose, & Perry-Thornton, 2001).

Women who do have careers in science have overcome a lot to get there. Scantlebury & Baker (2007) found that in the early twentieth century, a measure of progress was being made as women began to follow vocational education in large numbers. However, following World War I, that trend was replaced with a new interest in home-centered pursuits. Advanced education and burgeoning business interest became the typical male pathway while women's studies and activities were by far focused on domestic issues—even girls' chemistry classes adapted to cooking and nutrition topics. Through World War II, men prepared for work, partly through continuing education, while women prepared for family-based lives. These conditions remained static until the 1970's when the women's movement took hold, and it became not only acceptable, but desirable for women to work outside the home. In preparation for careers outside the home, female enrollment at universities increased steadily. Not only did women enter college campuses in high numbers, but they began to be interested in traditionally male-dominated fields of study. Xie and Shauman (2003) point out that as enrollments increased, the general body of college freshmen remained approximately balanced among males and females with similar aptitudes, educational backgrounds, and aspirations. However, even with increases in the number of women pursuing science majors, they still lag behind in the number of advanced degrees awarded, and they continue to be under represented in academic and scientific fields (Taylor et al., 2001). Why then do so few women choose and continue a field of study in science? What influences in college may affect these choices, and how do the choices affect potential career paths?

One possible answer may have to do with the number of females teaching science in college. In recognition of this, in 2003, the National Science Foundation (NSF) reported a deficiency in female science faculty at the college and university level, and attention has been given to the disparity of promotion and tenure rates for female science faculty as well (Scantlebury & Baker, 2007). Incentives by organizations such as the NSF have spurred academic institutions to create a more favorable environment for the promotion of women in science. According to Handelsman, et al. (2005), ". . . women scientists may not pursue academic careers simply because they are not encouraged to do so, question whether they have what it takes to be successful, or lack female role models who would help them envision themselves as faculty" (p. 1190). This position was previously espoused by Hackett, Esposito, & O'Halloran (1989) when they said, "The lack of female professorial and occupational role models has been identified as a significant barrier to women's career development while conversely, the availability of female role models has received support as an important positive influence" (p. 165). Several studies have supported the assertion that female instructors have positively motivated female students (e.g. Bettinger & Long, 2005; Butler & Christensen, 2003; Carrell, Page, & West, 2009; Dee (2007); Ehrenberg, Goldhaber & Brewer, 1995; Hoffman & Oreopoulos, 2009; Nixon & Robinson, 1999; Rask & Bailey, 2002).

Relevant Studies

Nixon and Robinson (1999) expanded upon earlier gender studies (Canes & Rosen, 1995; Ehrenberg, et al., 1995; Rothstein, 1995) by exploring how the numbers of high school faculty and professional staff, acting as role models, are able to affect the level of female high school student performance. They based their study on information obtained from the National Longitudinal Survey of Youth (NLSY) to trace student performance as it related to the proportion of their high school faculty. The participants included 12,000 youth from the ages of 14 to 22 in 1979. They narrowed the number of participants to female youth who completed the NLSY School Survey information and who completed ninth grade or above at a school that reported the percentage of female faculty and professional staff. Some of the variables used in their analysis included the total number of years of education, whether the respondent was a high school or college graduate, the amount of education they received beyond high school, the percentage of female faculty at their school, the percentage of faculty with advanced degrees, the faculty-student ratio, relevant faculty salaries, their mother's and father's educational level, whether or not both parents lived in the home, family income, etc. The total number of participants who met all the criteria was 2,378 females. As a control variable, Nixon and Robinson (1999) performed the same study for male youth, and narrowed the number of those participants meeting the same criteria as the female youth to 2,190. Four measures of educational attainment were used as dependent variables: the total number of years of school completed by age 26, whether the participant graduated from high school by age 26, whether the participant enrolled in college by age 26, and whether the participant graduated from college by age 26. Nixon and Robinson (1999) expected, ". . . that the resources put into teaching as measured by the faculty-student ratio, faculty's salaries, and the percentage of faculty with an M.A. or a Ph.D. will positively influence educational attainment" (p. 189). The results of the statistical analysis clearly show that the level of female students' educational attainment was positively and significantly influenced by the exposure to female faculty and professional staff in high school. Conversely, the study showed that female faculty and professional staff did not have a significant impact on the educational attainment for male youth. Thus, these study results support the female role model hypothesis (Nixon & Robinson, 1999).

There is evidence that students choose college majors whose faculty are populated by their own gender, and therefore Rask and Bailey (2002) conducted a study at Colgate University using data from student, transcript, and faculty records from the classes of 1988 – 2000 to determine if the persistent major gap in certain departments was due to the lack of female and minority faculty. Their research included 3,478 white female students, 912 minority students, and 3,779 white male students pursuing 22 different major courses of study. Rask and Bailey (2002) predicted that, ". . . female or minority faculty in a department could influence female and minority students to choose [a] major somewhat independent of human capital, precollege socialization, or postcollege expectations" (p. 100). In other words, the researchers predicted that students could choose a college major based on the role model effect rather than their personal attributes and backgrounds. The data gathered from 8,167 undergraduate student files included information from their precollege years (SAT, high school grades and student rank) together with their college transcripts (specific courses taken and when) as well as demographic records. The researchers attached race and gender information of the faculty who taught each course, along with the GPAs of each student prior to their selecting a major. They found that the number of classes taken with a matched instructor does positively affect the likelihood that a student will choose the same major represented by that teacher. They also found that collegiate academic success, quality of instruction, and socialization are also factors that determine a student's choice of major, so the choice of major is not exclusive to the role model effect. Rask and Bailey (2002) also, ". . . strongly suggest that the influence of students preferring professors like themselves is an important aspect of major choice" (p. 120). Therefore, they recommend that affirmative action focus on the under representation of female and minority faculty to help rectify the imbalance.

Bettinger and Long (2005) also addressed the problem of the under representation of women in mathematical, technical, and science-related fields by relating that, ". . . the health of the economy depends on the production of certain kinds of degrees, and the under representation of women in certain areas may contribute to shortages in critical fields" (p. 152). Therefore, they used a longitudinal database that included 54,000 students to assess if gender matched faculty members influenced student interest and enrollment in under represented subjects. Bettinger and Long (2005) expected that instructors would be influential in encouraging female students to succeed in these subjects. Participants in their study included first-time, full-time, and traditionally-aged freshmen who took the ACT, and attended a four-year college in Ohio in 1998 or 1999. Bettinger and Long (2005) examined three student outcomes: additional courses taken, total number of credit hours taken in the major, and choice of major. They used within course and student variation, and discovered that there was a small positive instructor effect on the choice of student major and course completion, thus supporting a possible role model effect. However, there were some male-dominated fields where there were no positive or significant effects found. Therefore they conclude that, ". . . female instructors have mixed effects on the interests of female students" (p. 156). Interestingly, on one hand, the data show that in the sciences, female students were less likely to take more than one course in biology or physics if the instructor was a female; but on the other hand, the data also show that female students were more likely to take an additional course from a female instructor in geology, math, and statistics. According to Bettinger and Long (2005), these results suggest that gender matched faculty members do have the potential to increase student interest and enrollment in some under represented subjects. They feel it is important to note that females were under represented as majors in math and geology. Therefore, the results did support the role model hypothesis. However, they say that in other male-dominated courses (e.g. engineering, physics, and computer science), no role model effect was found. The authors suggest this may be due to the lack of female faculty in these disciplines, and that future research outcomes might be different when more female instructors enter these academic areas.

Dee (2007) examined the influence of teacher gender on student achievement, student subject interest, and instructor judgment of student performance. He used the 1988 National Education Longitudinal Survey (NELS), which was a large sample of approximately 25,000 eighth grade students from 1,052 public and private schools, as well as within student comparisons to show if a gender matched instructor influenced educational outcomes. Dee (2007) expected that a gender matched instructor would positively affect the students. The researcher's survey included data from two different instructors for each student, and included each instructor's view of the student's performance. The subjects chosen were science, history, social studies, math and English. The results of the study showed significant positive effects from the gender influence of teachers on their students in test scores, student subject interest, and instructor judgment of student performance. Dee (2007) suggested that, ". . . the gender interactions between students and teachers are consequential and that it would be worthwhile to know more about why such student-teacher interactions matter" (p. 552). He also suggested that the areas of the study (e.g. student achievement, student subject interest, and instructor judgment of student performance) might be used as a guide for new policies that would close the gender gap in education.

Ehrenberg et al. (1995) also recognized the under representation of certain role models in secondary education. They analyzed how a teacher's race, gender, and ethnicity (RGE) influenced students with matched and non-matched RGE to determine if a future match of employers and employees by RGE would be important in building employment relationships. The researchers predicted that teacher-student matched RGE would positively affect student achievement and evaluations. They focused on how teachers related to their students and how they evaluated them. They also focused on the level of student performance by the use of standardized tests. Using the 1988 NELS, they regressed individual student test score gains between the eighth grade and tenth grade on teacher RGE. They found that high school student test scores were not significantly related to teacher-student matched RGE, but that teacher evaluations of students in math and science were significantly related to the criteria. The researchers concluded that if student achievement was most important, then the data do not support the need for a teacher-student RGE match. However, if encouragement was the desired attribute, then teacher RGE would matter. The researchers suggested that future studies into the resolution of which interpretation was correct would be important in promoting women and minorities in the labor market.

Butler and Christensen (2003) focused on an explanation for gender discrimination in a traditionally male-dominated course, and they suggested that the role model effect might be a possible response. They expected to find that female students would do better on assignments when they had a close interaction with a female teaching assistant, and they would not do as well without it. Butler and Christensen's 2003 three-year study included 600 students from a Political Inquiry course. The researchers related that the required introductory course was taught two days a week by a professor, and had an enrollment of approximately 150 students. Undergraduate teaching assistants (TAs) taught a weekly one-hour lab with an enrollment of between 10 and 20 students. Some of the responsibilities of the TA in this course were to explain and grade student writing assignments, administer and grade quizzes, review course content, and answer questions. Therefore, the TAs were considered appropriate representatives for a gender effects analysis. The course content was controlled, and other control variables included student GPA, year in school, age, number of credit hours enrolled, marital status and U.S. citizenship. The gender of the TA varied, and all TA assignments were random. Their study results did not show support for a role model effect for student achievement since the male students outperformed the females on the final exam as well as writing assignments when they had male TAs, and they did even better when they had female TAs—even though these results were not statistically significant. Butler and Christensen (2003) suggested that the content and nature of the Political Inquiry course might be biased in favor of males, and they also suggested that future research focus on methods that would better reach female students in order to eliminate teaching methods that would create unfair gender advantages.

The research of Hoffman and Oreopoulos (2009) was based on ". . . the impact of male and female undergraduates' exposure to same sex teachers and whether such exposure can affect student achievement and subject interest" (p. 481). They used a large sample of students (*n* = 34,352) and student and instructor administrative data from the University of Toronto's Arts and Science Faculty from 1996 to 2005. The students were between the ages of 17 and 20 at the beginning of their collegiate years. The study was confined to 88 large first year courses with at least 50 students enrolled in a class. Hoffman and Oreopoulos (2009) focused on large first year classes so that teacher and student interactions would be limited in order to minimize familiarity, differential treatment, instructor reputation, or other intangibles that might override the gender interaction effect. They also focused on college students to determine whether gender role model effects were present in higher education by using within student and within instructor variation. Hoffman and Oreopoulos (2009) found a slightly positive effect of the degree of influence that teacher gender has on student achievement and subject interest. Their data show that students with a gender matched instructor were only one percentage point more likely to drop a course than those without an instructor match; and students with a gender matched instructor showed a grade achievement of only 1 to 5 percent of a standard deviation higher than those without an instructor match. They point out that the results also seem to be affected more by the lower performance of male students when matched with female instructors, while female student performance remained the same. These results do not show much support for the role model hypothesis. The Hoffman and Oreopoulos (2009) study suggests that the gender of the instructor plays a very minor role when it comes to college student achievement, and that the role model effect might be greatest for younger students.

Carrell, Page and West (2009) also note that the under representation of women in science is a continuing problem. They state that even though women have penetrated many male-dominated careers, and have earned an equal number of graduate degrees in medicine, business and law, they still lag substantially behind males in the science, technology, engineering, and math (STEM) fields. The role of the professor's gender is the focus of their study, and they acknowledge other research that points out the importance of role model effects to educational outcomes. They suggest that gender bias may exist in the educational expectations of instructors, instruction methods, or the amount of instructor advice and encouragement offered. The researchers' data was gathered from the United States Air Force Academy (USAFA) since it has a large enrollment, but it also has many similarities to small liberal arts colleges in the United States in that student-teacher interactions are encouraged. Their study participants included 9,481 students, who were part of the graduating classes of 2000 – 2008. SAT scores measured student aptitude, and a composite of high school GPA, class rank, and quality of high school were recorded for each student. Students were randomly assigned to instructors over a wide selection of courses that were mandatory and standard, and course grades were not assigned by a single instructor. The results from the Carrell et al. (2009) study show that there is a significant role model benefit to female students and a minimal effect on male students, especially insofar as seeking an academic pathway in the STEM fields is concerned. They suggest that this outcome specifically emphasizes the benefit to female students who have a high aptitude for mathematics, and that these students are especially influenced by female role models who teach introductory classes in STEM. The researchers conclude that quality teachers, regardless of gender, are the primary asset to students, but that there may be specific traits to male or female teachers that enhance student performance. They suggest this as a subject for future study.

Hanson et al. (1996) summarized the under representation of females in science when they reported that the employment base in science fields has been wasting a quarter of its potential resources since working Americans have been divided nearly equally among men and women, but only 22% of those working in science fields are women. It is important to overcome the accumulated effects that may accompany female science students, and therefore further study into the causal conditions effecting equal opportunity collegiate retention in science may be necessary. Even though there have been studies showing a direct influence of role models upon elementary and secondary students (Dee, 2007; Ehrenberg et al., 1995; Gilmartin et al., 2007; Nixon & Robinson, 1999), little research has been conducted, and there are relatively few empirical studies concerning the particular impact of female role models for college science students, and the studies that have been performed present conflicting conclusions (Bettinger & Long, 2005; Carrell et al., 2009; Handelsman et al., 2005; Hoffman & Oreopoulos, 2007). Therefore, this thesis addresses these inconsistencies by using alternative instruments to those previously employed, and specifically focuses on the effect of female teaching assistants on the performance of female students in an introductory college biology course.

Overview of the Present Study

There are typically fewer females than males advancing in science, and therefore fewer female science instructor role models. This observation inspired two questions: Are female students positively influenced by female science teaching assistants? And if so, how can we measure that influence? Achievement, reasoning ability, and self-efficacy are widely accepted measures of student performance and cognitive development and have been measured in research for many years (Cavallo, 1996; Germann, 1994; Johnson & Lawson, 1998; Lent, Brown, & Gore, 1997; Shayer & Adey, 1993). Therefore, to test the hypothesis that female students are positively influenced by female science teacher role models, I predicted that this influence would manifest itself in superior achievement, and in superior gains in reasoning ability and science self-efficacy when female students are taught by female instructors as opposed to when they are taught by male instructors.

Methods

Participants

The participants in this study included 13 female TAs and six male TAs who taught an introductory biology lab at a major southwestern American university in the Spring 2010 semester. Their experience in science ranged from senior undergraduate science major to Ph.D. candidate. Also included in the study were 724 students enrolled in a college introductory biology course at the same university during the same semester.

Procedure

An empirical study was conducted using exams and assignment scores to measure student performance. A three-part evaluation included achievement measurements (lecture and lab examinations and laboratory coursework), reasoning ability tests in two phases, and self-efficacy surveys in two phases, and these instruments were the means to determine the progress of the students.

At the beginning of the semester, 640 students were tested to evaluate their science reasoning ability (reasoning ability pretest) prior to course instruction. Additionally, 331 students from the same enrollment were surveyed as to their level of confidence (self-efficacy pretest) regarding their ability to communicate scientific concepts. Following the semester of instruction, 637 students repeated the same reasoning ability test (posttest), and 318 students who took the first survey were administered the same survey instrument again (self-efficacy posttest). The pretest and posttest scores for reasoning ability and self-efficacy were used to determine student progress in these areas. At the conclusion of the semester, 724 students took the final examination for the lecture portion of the course. Of this group, 676 students completed the laboratory portion, including a lab final examination. The combination of lecture and laboratory scores allowed an evaluation of overall achievement. A number of students that took each component of the study were culled from the final results if they did not complete both phases (pre and posttests) of the reasoning and self-efficacy tests.

The instruments were used to compare and contrast student ability by gender as well as student ability within four types of group pairings (hereafter referred to as the four groups). The four groups were 1) female students who had female teaching assistants (TAs), 2) male students who had female TAs, 3) female students who had male TAs, and 4) male students who had male TAs. The students did not know who their teaching assistants were going to be in advance of enrolling in the course, therefore the four groups were randomly assigned.

Instruments

Lab final exam. The lab final exam was comprised of a short answer section worth 110 points, and a multiple-choice section with 25 possible points for a total of 135 points. A selection of example questions is cited below:

1. A 330 pound, 16-foot-long colossal squid was caught in the Ross Sea this year. A group of scientists notice that this squid has razor-sharp hooks on its tentacles. One scientist asks, “Are these razor-sharp hooks used for killing larger prey?” That scientist has

just: a) asked a causal question b) produced a hypothesis c) produced a prediction d) produced an observation e) produced a result.

2. Think about the Isopod lab and give a prediction based on the following hypothesis and experiment: IF the isopod beetles are attracted to moisture, and I add a moist cotton ball to one side of the isopod tray, then I would expect to find that the Isopods have: a) moved to the side of the tray where the moist cotton ball is. b) moved to the opposite side of the tray where there is no moist cotton ball. c) remained in their original positions at both ends of the tray.

Lab total score. The lab total score included 12 lab quizzes at 5 points each (60 points), 11 homework assignments at 10 points each (110 points), 15 participation points, a 40 point article summary assignment, a 40 point lab report, and the 135 point lab final examination. The laboratory portion of the course is the environment where students received maximum exposure to the teaching assistants. Therefore, laboratory achievement scores most directly relate to the focus of this study, although the influence of the teaching assistants could result in an improved performance in the lecture portion of the achievement analysis also by their tutoring of lecture material during TA office hours.

Lecture final exam. The lecture final exam, consisting of 60 multiple-choice questions, provided an additional possible 225 points toward the comprehensive achievement component of this three-part research. Two well-experienced male professors taught the lecture part of this introductory biology course.

Reasoning ability test. The reasoning ability test included 20 multiple-choice questions. The Classroom Test of Scientific Reasoning (Lawson, 1978) was the same instrument that was employed in both the pre and posttest administrations. A correct response to a question received a single point score and all incorrect answers were scored zero. Therefore, the student scores could range from 0 to 20 points. A selection of example questions is cited below and a full copy is found in Appendix A.

1. Suppose you are given two clay balls of equal size and shape. The two clay balls also weigh the same. One ball is flattened into a pancake-shaped piece. Which of these statements is correct? a) The pancake-shaped piece weighs more than the ball b) The two pieces still weigh the same c) The ball weighs more than the pancake-shaped piece

2. Farmer Brown was observing the mice that live in his field. He discovered that all of the mice were either fat or thin. Also, all of them had either black tails or white tails. This made him wonder if there might be a link between the size of the mice and the color of their tails. So he captured all of the mice in one part of his field and observed them. Below are the mice that he captured [drawings of

12 fat mice with black tails, 3 fat mice with white tails, 2 thin mice with black tails and 8 thin mice with white tails]. Do you think there is a link between the size of the mice and the color of their tails? a) appears to be a link b) appears not to be a link c) cannot make a reasonable guess

Reliability of measures in the study was measured for the posttest for reasoning ability. The composite posttest reasoning ability measure, including 20 items, had a KR-20 reliability coefficient of 0.78, indicating that the reasoning ability test measurements were consistent and reliable.

Validity of the reasoning ability test as a good measure of general reasoning skill has been established by previous research (e.g., Lawson, 1978, 1979, 1980, 1982, 1983, 1985, 1992; Lawson, Baker, DiDonato, Verdi, & Johnson, 1993; Lawson, Banks, & Logvin, 2007; Lawson & Thompson, 1988; Lawson & Weser, 1990; Lawson & Worsnop, 1992).

Self-efficacy. The self-efficacy survey (identical pre and posttest content) asked each student to identify their confidence level in their ability to describe specific science concepts on an escalating level of confidence from 1 – 5 (i.e. 1 = not at all confident, 2 = not very confident, 3 = average confidence, 4 = confident, 5 = very confident). There were a total of 15 survey items. The total score exhibiting the least confidence is 15 and the most confidence is represented by a score of 75. The composite posttest survey self-efficacy measure had a Cronbach's alpha coefficient of 0.95. A selection of example questions is cited below and a full copy is in Appendix B.

1. Look at a skull that you have not seen before and determine what the animal ate and where it lived.

2. Use the fossil record to construct an argument that supports evolution theory.

3. Test the hypothesis that water rises in a cylinder inverted over a burning candle due to consumed oxygen.

Analysis

A series of *t* tests and Analysis of Variance (ANOVA) tests were conducted. Pertaining to the analysis of the data obtained by the instruments, the results assume equal variances based on the non-significant Levene's test for the equality of variances. Caution needs to be taken when interpreting these results because there were fewer male teaching assistants ($n = 6$) than female teaching assistants ($n = 13$) in the sample, and a small sample size could lead to invalid conclusions.

Results and Discussion

Achievement

Lab total score. The means and standard deviations for the lab total score for each group are reported in Table 1. As predicted, the mean score of the female students with female TAs (mean = 337.35) was higher than the mean score of the female students with male TAs (mean = 331.32). Male students performed similarly with male TAs (mean = 329.38) and with females TAs (mean = 329.41).

Table 1

Lab Total Score Observed Statistics for Four Groups (400 Points)

Student Gender	TA Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	Male	117	329.38	42.35	3.92	[321.63, 337.14]
	Female	239	329.41	49.02	3.17	[323.16, 335.65]
Female	Male	57	331.32	44.28	5.87	[319.57, 343.06]
	Female	237	337.35	45.89	2.98	[331.48, 343.22]
Total		650	332.47	46.37	1.82	[328.89, 336.04]

Although the female students with female TAs scored slightly higher than the female students with male TAs, a one-way ANOVA comparing means across all four groups was not statistically significant, $F(3, 646) = 1.41$, $p = .24$ (see Table 2). This suggests that the groups performed similarly on the lab total score. Therefore, these results do not provide support for the role model hypothesis.

Table 2

One-Way ANOVA for Lab Total Score for Four Groups

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	9070.63	3	3023.54	1.41	.24
Within Groups	1386609.87	646	2146.46		
Total	1395681.00	649			

Lab final exam. The group means and standard deviations for the lab final exam are reported in Table 3. Unexpectedly, the mean score of the female students with female TAs (mean = 106.22) was lower than the mean score of the female students with male TAs (mean = 107.30). Further, regardless of the TA gender, the male students scored higher on the lab final exam than did the female students. The mean score of the male students with male TAs (mean = 108.29) was higher than the group mean for either female student group, while the mean score of the male students with female TAs (mean = 109.77) was the highest among all the groups.

Table 3

Lab Final Exam Scores Observed Statistics for Four Groups (135 Points)

Student Gender	TA Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	Male	90	108.29	15.13	1.60	[105.12, 111.46]
	Female	245	109.77	14.16	.91	[107.99, 111.55]
Female	Male	43	107.30	15.02	2.29	[102.68, 111.92]
	Female	245	106.22	14.25	.91	[104.43, 108.02]
Total		623	107.99	14.45	.58	[106.86, 109.13]

A one-way ANOVA found that the slight group differences were not statistically significant, $F(3, 619) = 2.52, p = .06$ (see Table 4). Thus again, there was no support found for the role model hypothesis.

Table 4

One-Way ANOVA for Lab Final Exam for Four Groups

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	1569.55	3	523.18	2.52	.06
Within Groups	128285.41	619	207.25		
Total	129854.96	622			

Lecture final exam. Table 5 shows the means and standard deviations for each of the four groups on the lecture final exam. All four mean scores were nearly identical with a range of 40.05 to 40.80. The one-way ANOVA was not statistically significant, $F(3, 635) = .40, p = .75$ (see Table 6).

Table 5

Lecture Final Exam Scores Observed Statistics for Four Groups (225 Points)

Student Gender	TA Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	Male	116	40.22	6.38	.59	[39.04, 41.39]
	Female	235	40.80	6.33	.41	[39.98, 41.61]
Female	Male	56	40.05	7.34	.98	[38.09, 42.02]
	Female	232	40.80	6.66	.44	[39.94, 41.66]
Total		639	40.63	6.55	.26	[40.12, 41.13]

Table 6

One-Way ANOVA for Lecture Final Exam for Four Groups

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	52.48	3	17.16	.40	.75
Within Groups	27284.13	635	42.97		
Total	27336.61	638			

The results from the three testing instruments for achievement showed that there was no significant difference found in performance among the four groups. If the role model effect were going to manifest itself, it was expected to be evident in the lab testing instruments. However, the same level of expectation was not held for the lecture final exam. Since the lectures were exclusively taught by male professors, and there was limited interaction between the male professors and the students, the lecture final exam scores were used as a transfer test of the possible role model effect. In other words, even though the lectures were taught by male professors, it was possible that improved achievement by females, due to the role model effect, might transfer to lecture final exam performance. Needless to say, this did not turn out to be the case.

Reasoning Ability

Reasoning ability pretest. Table 7 shows the mean scores on the reasoning ability pretest for male and female students. The mean score for males (mean = 13.42) was higher than for females (mean = 11.90). An independent *t* test was statistically significant, $t(673) = 4.87$, $p < .01$, indicating that male students performed significantly better than female students at the beginning of the semester.

Table 7

Reasoning Pretest Observed Statistics for Male and Female Students (20 Points)

Student Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	365	13.42	3.98	.21	[13.00, 13.83]
Female	310	11.90	4.08	.23	[11.45, 12.35]

Reasoning ability posttest. Examination of the reasoning ability scores at the end of the semester shows that the four groups performed similarly (see Table 8). Unexpectedly, the mean score of the female students with female TAs was lower than the mean score of the female students with male TAs (respective means of 15.41 and 16.46). Male students with male TAs had slightly higher means (mean = 15.81) than male students with female TAs (mean = 15.36).

Table 8

Reasoning Posttest Observed Statistics for Four Groups (20 Points)

Student Gender	TA Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	Male	113	15.81	3.55	.31	[15.20, 16.43]
	Female	229	15.36	3.39	.22	[14.93, 15.79]
Female	Male	56	16.46	2.64	.44	[15.60, 17.33]
	Female	242	15.41	3.24	.21	[15.00, 15.83]
Total		640	15.56	3.31		

A one-way ANOVA found that these small group differences were not statistically significant, $F(3, 636) = 2.05, p = .11$ (see Table 9). The finding that the female students with female TAs did not perform better than the female students with male TAs provides no support for the role model hypothesis. The fact that females did not perform as well as the males did on the pretest and similarly to the males on the posttest is of interest as it indicates that the apparent pretest female reasoning deficit can be overcome with inquiry instruction explicitly aimed at reasoning improvement (see Table 9).

Table 9

One-Way ANOVA for Reasoning Posttest for Four Groups

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	67.24	3	22.41	2.05	.11
Within Groups	6946.61	636	10.92		
Total	7013.86	639			

Reasoning ability gain scores. Gain scores (shown in Table 10) were computed by taking the difference between the students' pretest and posttest reasoning ability scores. The largest group gain was the female students with male TAs (mean gain = 4.11); while the smallest gain was the male students with female TAs (mean gain = 1.76). Group gains computed as percentages are shown in Figure 1.

Table 10

Reasoning Gain Scores for Four Groups

Student Gender	TA Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	Male	113	2.73	5.10	.48	[1.78, 3.68]
	Female	229	1.76	5.11	.34	[1.09, 2.43]
Female	Male	56	4.11	4.60	.62	[2.87, 5.34]
	Female	242	3.57	5.44	.35	[2.89, 4.26]
Total		640	2.82	5.26	.21	[2.41, 3.23]

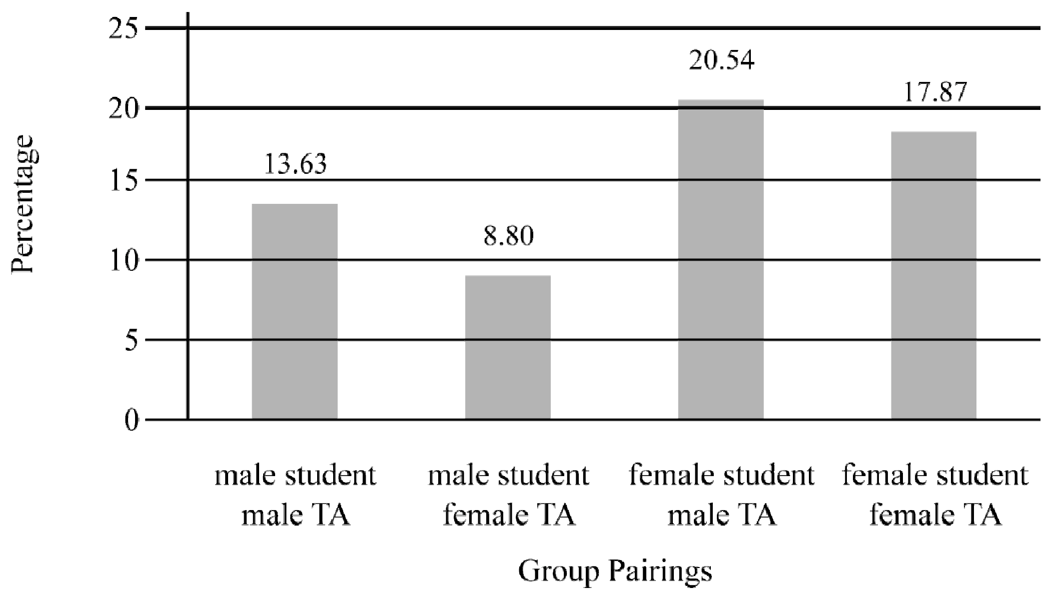


Figure 1. Percentage of Gains in Reasoning Ability by Groups (Pretest to Posttest).

An ANOVA found that gain differences were statistically significant, $F(3, 636) = 6.03$, $p < .01$ (see Table 11). Given this significant result, follow-up analyses were conducted to investigate where the differences among the four groups existed. The results yielded two significant comparisons. First, the female students with male TAs had significantly larger gains than the male students with female TAs (respective mean gains of 4.11 and 1.76, $p < .05$). Second, the female students with female TAs also had larger gains than the male students (mean = 1.76, $p < .01$). Of primary interest was the comparison between the female students with female TAs (mean = 3.57) and female students with male TAs (mean = 4.11). Clearly this is not the predicted result based on the role model hypothesis. Therefore once again, the hypothesis is not supported.

One more comparison is worth mentioning. Both female and male students made greater gains with male TAs than with female TAs (mean = 4.11 versus 3.57 and mean = 2.73 versus 1.76 respectively). Although these differences were not statistically significant, nor were they predicted, they are interesting as they suggest a possible male TA effect when it comes to reasoning gains. Although TA reasoning ability was not assessed in this study, it is possible that TA reasoning ability varied in the present TA sample. If male TAs did in fact have better reasoning ability than female TAs (similar to the differences found between the male and female students) they might have been in a better position to effect greater gains among their students. Of course this is merely speculation, but it might be worthwhile to assess TA reasoning ability in a future study and compare it to student reasoning gains as a test of this hypothesis.

Table 11

One-Way ANOVA for Reasoning Gain Scores

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	488.89	3	162.96	6.03	< .01
Within Groups	17178.80	636	27.01		
Total	17667.69	639			

Self-Efficacy

Self-efficacy pretest. Table 12 shows the mean pretest self-efficacy scores for female and male students. The mean score on the self-efficacy pretest was higher for female students (mean = 43.94) than for male students (mean = 41.09). An independent *t* test was statistically significant, $t(347) = -2.89$, $p < .01$, indicating that female students reported significantly higher levels of self-efficacy than male students at the beginning of the semester.

Table 12

Self-Efficacy Pretest Observed Statistics for Male and Female Students (75 Points)

Student Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	192	41.09	9.48	.68	[39.79, 42.39]
Female	157	43.94	8.75	.70	[42.50, 45.37]

Self-efficacy posttest. Table 13 includes posttest self-efficacy mean scores for the four groups. The highest mean was reported by female students with female TAs (mean = 51.08), while the lowest mean was reported by female students with male TAs (mean = 48.54). Although this result is consistent with the role model hypothesis, a subsequent statistical analysis found that the group differences were not statistically significant (F (3, 345) = .24, p = .87) (see Table 14). Male students with male TAs reported slightly lower self-efficacy (mean = 49.27) than with female TAs (mean = 50.94).

Table 13

Self-Efficacy Posttest Observed Statistics for Four Groups (75 Points)

Student Gender	TA Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	Male	37	49.27	15.29	2.51	[44.34, 54.20]
	Female	154	50.94	15.33	1.23	[48.52, 53.36]
Female	Male	13	48.54	16.70	4.23	[40.22, 56.86]
	Female	145	51.08	15.03	1.27	[48.59, 53.57]
Total		349	50.73	15.20		

Table 14

ANOVA on Self-Efficacy Posttest for Four Groups

Source	Sum of Squares	*df*	Mean Square	*F*	*p*
Corrected Model	165.51	3	55.17	.24	.87
Intercept	340296.88	1	340296.88	1462.83	< .01
Student/TA Gender	165.51	3	55.17	.24	.87
Error	80257.17	345	232.63		
Total	978609.00	349			
Corrected Total	80422.68	348			

Self-efficacy gain scores. Self-efficacy gain scores were computed by taking the difference between the students' pretest and posttest self-efficacy scores. Table 15 reports group means. The largest mean gain was reported by male students with female TAs (mean = 8.99), while the smallest gain was reported by male students with male TAs (mean = 5.65). Figure 2 presents the mean gain scores of the four groups in terms of percentages. As you can see, percentages ranged from a high of 11.98% to a low of 7.53%. Table 16 shows the results of an ANOVA indicating that the group differences were not statistically significant, $F(3, 345) = .38$, $p = .77$.

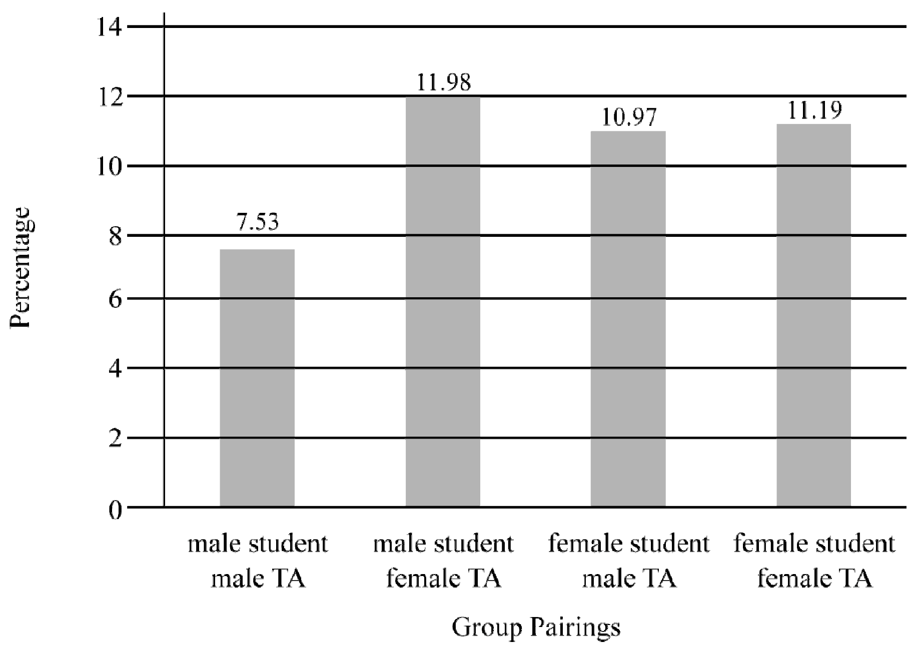


Figure 2. Percentage of Gains in Self-Efficacy by Groups (Pretest to Posttest).

Table 15

Observed Statistics of Gain Scores for Self-Efficacy

Student Gender	TA Gender	*n*	*M*	*SD*	*SE*	95% CI
Male	Male	37	5.65	17.15	2.82	[-.07, 11.37]
	Female	154	8.99	17.36	1.40	[6.22, 11.75]
Female	Male	13	8.23	18.60	5.16	[-3.01, 19.47]
	Female	145	8.39	16.62	1.38	[5.67, 11.12]
Total		349	8.36	17.03	.91	[6.57, 10.15]

Table 16

One-Way ANOVA for Self-Efficacy Gain Scores

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	332.92	3	110.97	0.38	.77
Within Groups	100619.31	345	291.65		
Total	100952.23	348			

The self-efficacy survey scores showed that female students performed statistically better than the males at the beginning of the semester. Interestingly, the posttest results showed that there were no significant differences between the four groups at the end of the semester. Therefore, compared to the pretest scores, the male student posttest scores improved, and were approximately the same as the female student scores. The question that was raised by these results centered on the cause of the student improvement, and in order to test the role model hypotheses, the self-efficacy survey gain scores were analyzed.

There were no significant role model outcomes found for the self-efficacy survey gain scores. Since the results of this measure do not support the role model hypothesis, an alternative explanation was that the course curriculum and methods assisted the male students in catching up with the females in self-efficacy over the semester.

Conclusion

The results of this study suggest that the role model effect may not be influential for increased student academic performance of college biology students. Given that there is evidence from earlier research that supports the role model effect as a positive influence for female science students, other aspects of the role model hypothesis, such as retention rate increases, increased female student interest in science, and advanced degree attainment should be studied (Bettinger, 2004; Butler & Christensen, 2003; Carrell et al., 2009; Dee, 2005; Hackett, 1989; Hoffman & Oreopoulos, 2007). Since college students and their instructors were the focus group for this research, a similar study using academic achievement, reasoning ability and self-efficacy could be conducted on high school, middle school or even elementary school students to see if these measures would show support for the role model effect for younger students.

As an overview of the accumulated data, there were limitations to the analyses. Among the study group of teaching assistants, the females were in the majority by a count of thirteen to six, and the total number of student instructors was small. However, within the student sample, the numbers of male and female students were approximately equal. Also, this study was conducted for only one semester. An expanded study covering more than one semester should be an important consideration for future research. Another consideration for future research would be to assess the teaching ability of the TAs. In this study, there was a combination of graduate and undergraduate student TAs, representing different levels of maturity as well as a wide range of ability and experience in education, teaching and research. Therefore, any strong assessment pertaining to the four groups should be made with caution. Additionally, among the students in the introductory biology course who were potential participants, many did not participate in either the pretest or the posttest for the reasoning ability test and self-efficacy survey. Therefore, they were excluded from the data. This accounts for the difference in sample size for each of the instruments in the research. However, in spite of the inequalities in the number of the study participants or length of the study the amount of data collected are sufficient for meaningful interpretation of the results.

Finally, a study that includes female science professors, in addition to the TAs, may show positive influence for female students. Female science professors typically have more teaching experience, finer skills and knowledge of subject matter than do the TAs. They are also important female representatives for those who have successfully completed science degrees, and attained successful science careers. These attributes are the building blocks for successful role models, and could make a difference in the increase of the number of females in science.

REFERENCES

Bettinger, E. P., & Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *The American Economic Review*, 95(2), 152-157.

Buck, G. A., Plano Clark, V. L., Leslie-Pelecky, D., Lu, Y., & Cerda-Lizarraga, P. (2007). Examining the cognitive processes used by adolescent girls and women scientists in identifying science role models: A feminist approach. *Wiley InterScience,* 688-707. doi:10.1002/sce.20257

Butler, D. M., & Christensen R. (2003). Mixing and matching: The effect on student performance of teaching assistants of the same gender. *Political Science and Politics*, 36(4), pp. 781-786.

Canes, B. J., & Rosen, H. S. (1995). Following in her footsteps? Faculty gender composition and women's choice of college majors. *Industrial and Labor Relations Review*, 48, 486-504.

Carrell, S. E., Page, M. E., & West, J. E., (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics*, 125(3), 1101-1144. Retrieved from http://ejournals.ebsco.com.ezproxy1.lib.asu.edu/direct.asp?ArticleID=470A89EE41BE4A45CD5F.

Cavallo, A. M. (1996). Meaningful learning, reasoning ability, and students' understanding and problem solving of topics in genetics. *Journal of Research in Science Teaching*, 33, 625-656.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3), 528-554.

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial and Labor Relations Review*, 48(3), 547-561.

Germann, P. J. (1994). Testing a model of science process skills acquisition: An interaction with parents' education, preferred language, gender, science attitude, cognitive development, academic ability, and biology knowledge. *Journal of Research in Science Teaching*, 31, 749-783.

Gilmartin, S., Denson, N., Li, E., Bryant, A., & Aschbacher, P. (2007). Gender ratios in high school science departments: The effect of percent female faculty of multiple dimensions of students' science identities. *Journal of Research in Science Teaching*. 44(7), 980-1009. doi:10.1002/tea.20179

Hackett, G., Esposito, D., & O'Haloran, M. S. (1989). The relationship of role model influences to the career salience and educational career plans of college women. *Journal of Vocational Behavior*, 35, 164-180.

Handelsman, J., Cantor, N., Carnes, M., Denton, D., Fine, E., Grosz, B., Hinshaw, V., Marrett, C., Rosser, S., Shalala, D., & Sheridan, J. (2005). More women in science. *Science*, 309, 1190-1191.

Hanson, S. L., Schaub, M., & Baker, D. P. (1996). Gender stratification in the science pipeline: A comparative analysis of seven countries. *Gender and Society*, 10(3), 271-290.

Hoffman, F., & Oreopoulos, P. (2009). A professor like me: The influence of instructor gender on college achievement. *Journal of Human Resources*, 44(2), 479-494.

Johnson, M.A., & Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? *Journal of Research in Science Teaching*, 35, 89-103.

Lawson, A.E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15, 11-24.

Lawson, A.E. (1979). Relationships among performances on group-administered items of formal reasoning. *Perceptual and Motor Skills*, 48, 71-78.

Lawson, A.E. (1980). The relationship among levels of intellectual development, cognitive style and grades in a college biology course. *Science Education*, 64, 95-102.

Lawson, A.E. (1982). The reality of general cognitive operations. *Science Education*, 66, 229-241.

Lawson, A.E. (1983). Predicting science achievement: The role of developmental level, disembedding ability, mental capacity, prior knowledge and beliefs. *Journal of Research in Science Teaching*, 20, 117-129.

Lawson, A.E. (1985). A review of research on formal reasoning and science teaching. *Journal of Research in Science Teaching*, 22, 569-618.

Lawson, A.E. (1992). The development of reasoning among college biology students. *Journal of College Science Teaching*, 21, 338-344.

Lawson, A.E., Baker, W.P., DiDonato, L., Verdi, M.P., & Johnson, M.A. (1993). The role of physical analogies of molecular interactions and hypothetico-deductive reasoning in conceptual change. *Journal of Research in Science Teaching*, 30, 1073-1086.

Lawson, A. E., Banks, D. L., & Logvin, M. (2006). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, 44, 706-724.

Lawson, A.E., & Thompson, L.D. (1988). Formal reasoning ability and misconceptions concerning genetics and natural selection. *Journal of Research in Science Teaching*, 25, 733-746.

Lawson, A.E., & Weser, J. (1990). The rejection of nonscientific beliefs about life: The effects of instruction and reasoning skills. *Journal of Research in Science Teaching*, 27, 589-606.

Lawson, A.E., & Worsnop, W.A. (1992). Learning about evolution and rejecting a belief in special creation: Effects of reflective reasoning skill, prior knowledge, prior beliefs and religious commitment. *Journal of Research in Science Teaching*, 29, 143-166.

Lent, R. W., Brown, S. D., & Gore, P. A. Jr. (1997). Discriminant and predictive validity of academic self-concept, academic self-efficacy, and mathematics-specific self-efficacy. *Journal of Counseling Psychology*, 44, 307-315.

Nixon, L. A., & Robinson, M. D. (1999). The educational attainment of young women: Role model effects of female high school faculty. *JSTOR*, 36(2), 185-194.

Rask, K. N., & Bailey, E. M. (2002). Are faculty role models? Evidence from major choice in an undergraduate institution. *The Journal of Economic Education*, 33(2), 99-124.

Rothstein, D. S. (1995). Do female faculty influence student's educational and labor market attainments? *Industrial and Labor Relations Review*, 48, 515-530.

Scantlebury, K., Baker, D. (2007). Gender issues in science education research: remembering where the difference lies. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 257-285). Mahwah, NJ: Lawrence Erlbaum Associates.

Shayer, M., & Adey, P. S. (1993). Accelerating the development of formal thinking in middle and high school students IV: Three years after a two-year intervention. *Journal of Research in Science Teaching*, 30, 351-366.

Taylor, V. S., Erwin, K. W., Ghose, M., & Perry-Thornton, E. (2001). Models to increase enrollment of minority females in science-based careers. *Journal of the National Medical Association*, 93(2), 74-77.

Xie, Y., & Shauman, K. A. (2003). Women in science: Career processes and outcomes (pp. 72-96). Cambridge, Massachusetts: Harvard University Press.

APPENDIX A

CLASSROOM TEST OF SCIENTIFIC REASONING

CLASSROOM TEST OF

SCIENTIFIC REASONING

Free Response Version

Directions to Students:

This is a test of your ability to apply aspects of scientific and mathematical reasoning to analyze a situation to make a prediction or solve a problem. For some items you will simply be asked to state your current belief. In others you will be asked to show your work and/or explain your answer. Try to answer as completely as you can in the spaces provided. On some items these explanations are more important than your actual answer. If you do not fully understand what is being asked in an item, please ask the test administrator for clarification.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Revised Edition: August 2000 by Anton E. Lawson, Arizona State University. Based on: Lawson, A.E. 1978. Development and validation of the classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15(1): 11-24.

1. Suppose you are given two clay balls of equal size and shape. The two clay balls also weigh the same. One ball is flattened into a pancake-shaped piece. *Which of these statements is correct?*

   a. The pancake-shaped piece weighs more than the ball
   b. The two pieces still weigh the same
   c. The ball weighs more than the pancake-shaped piece

2. *Please explain your selection.* ______________________________

____

____

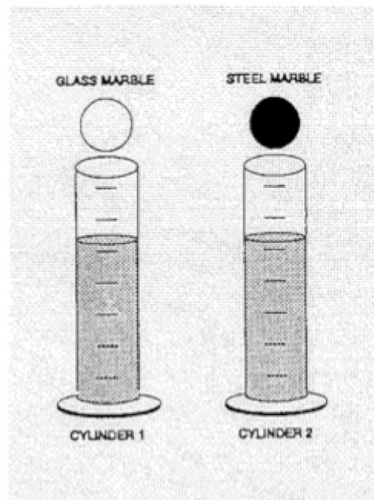
____

3. To the right are drawings of two cylinders filled to the same level with water. The cylinders are identical in size and shape.

   Also shown at the right are two marbles, one glass and one steel. The marbles are the same size but the steel one is much heavier than the glass one.

   When the glass marble is put into Cylinder 1 it sinks to the bottom and the water level rises to the 6th mark. *If we put the steel marble into Cylinder 2, the water will rise*

   a. to the same level as it did in Cylinder 1
   b. to a higher level than it did in Cylinder 1
   c. to a lower level than it did in Cylinder 1



4. *Please explain your selection.* ______________________________

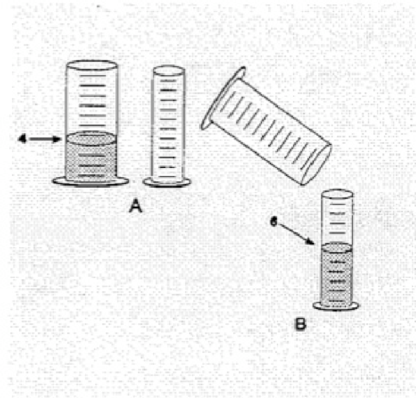
____

____

____

5. To the right are drawings of a wide and a narrow cylinder. The cylinders have equally spaced marks on them. Water is poured into the wide cylinder up to the 4th mark (see A). This water rises to the 6th mark when poured into the narrow cylinder (see B).

Both cylinders are emptied (not shown) and water is poured into the wide cylinder up to the 6th mark. *How high would this water rise if it were poured into the empty narrow cylinder?*



____

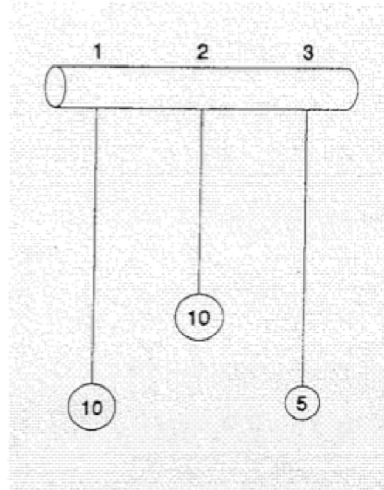
6. *Please show or explain how you arrived at your answer.* ____________________

7. Water is now poured into the narrow cylinder (described in Item 5 above) up to the 11th mark. *How high would this water rise if it were poured into the empty wide cylinder?*

____

8. *Please show or explain how you arrived at your answer.* ____________________

9. At the right are drawings of three strings hanging from a bar. The three strings have metal weights attached to their ends. String 1 and String 3 are the same length. String 2 is shorter. A 10 unit weight is attached to the end of String 1. A 10 unit weight is also attached to the end of String 2. A 5 unit weight is attached to the end of String 3. The strings (and attached weights) can be swung back and forth and the time it takes to make a swing can be timed.



Suppose you want to find out whether the length of the string has an effect on the time it takes to swing back and forth. *Which string(s) would you use to find out?*

____

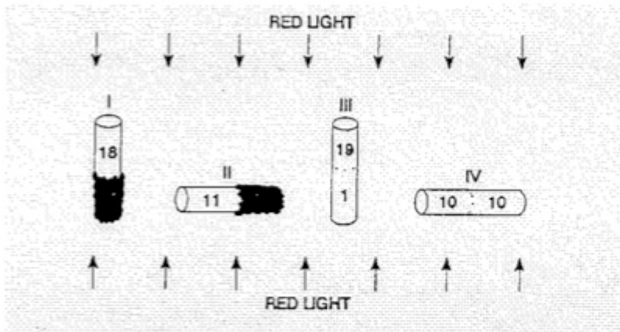
10. *Please explain why you chose the string(s) you did.* ____________________

____

____

____

11. Twenty fruit flies are placed in each of four glass tubes. The tubes are sealed. Tubes I and II are partially covered with black paper; Tubes III and IV are not covered. The tubes are placed as shown. Then they are exposed to red light for five minutes. The number of flies in the uncovered part of each tube is shown in the drawing.



This experiment shows that flies respond to (respond means move to or away from):

a. red light but not gravity
b. gravity but not red light
c. both red light and gravity
d. neither red light nor gravity

12. *Please explain your selection.* ______________________________

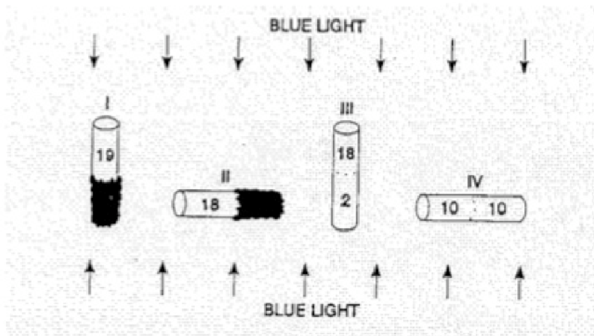
____

____

____

4

13. In a second experiment, a different kind of fly and blue light was used. The results are shown in the drawing.



These data show that these flies respond to (respond means move to or away from):

a. blue light but not gravity
b. gravity but not blue light
c. both blue light and gravity
d. neither blue light nor gravity

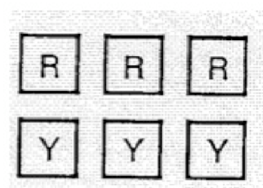
14. *Please explain your selection.* ______________________________________

____

____

____

15. Six square pieces of wood are put into a cloth bag and mixed about. The six pieces are identical in size and shape, however, three pieces are red and three are yellow. Suppose someone reaches into the bag (without looking) and pulls out one piece. *What are the chances that the piece is red?*



____

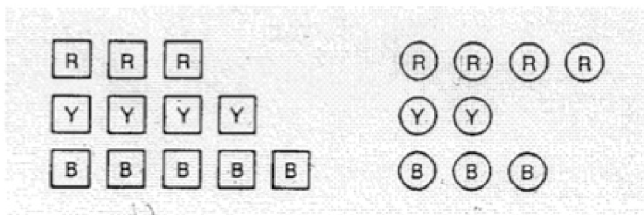
16. *Please show or explain how you arrived at your answer.* ____________________

____

____

____

17. Three red square pieces of wood, four yellow square pieces, and five blue square pieces are put into a cloth bag. Four red round pieces, two yellow round pieces, and three blue round pieces are also put into the bag. All the pieces are then mixed about. Suppose someone reaches into the bag (without looking and without feeling for a particular shape piece) and pulls out one piece.



What are the chances that the piece is a red round or blue round piece?

____

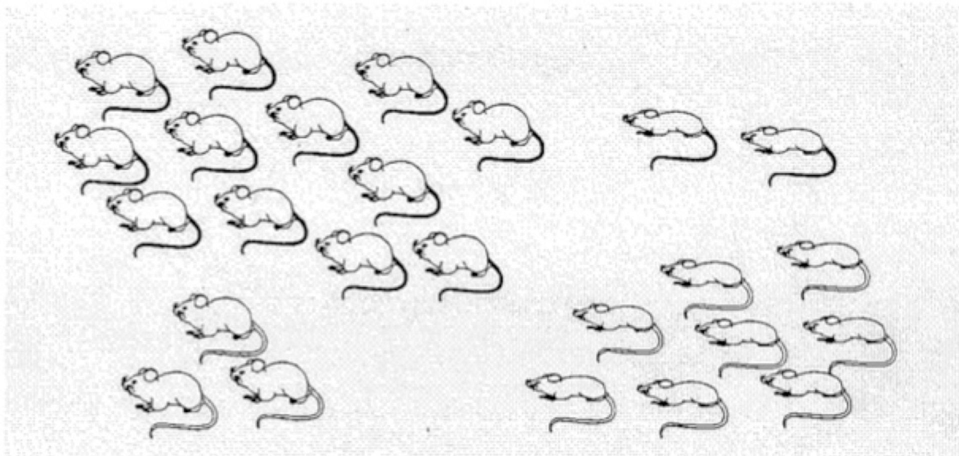
18. *Please show or explain how you arrived at your answer.* ____________________

____

____

____

19. Farmer Brown was observing the mice that live in his field. He discovered that all of them were either fat or thin. Also, all of them had either black tails or white tails. This made him wonder if there might be a link between the size of the mice and the color of their tails. So he captured all of the mice in one part of his field and observed them. Below are the mice that he captured.



Do you think there is a link between the size of the mice and the color of their tails?

____

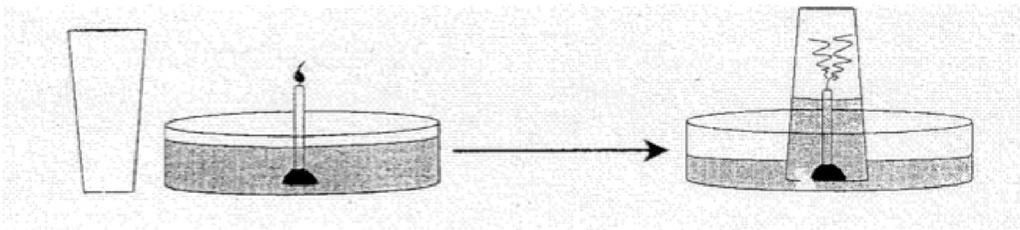
20. *Please show or explain how you arrived at your answer.* ________________

____

____

____

21. The figure below at the left shows a drinking glass and a burning birthday candle stuck in a small piece of clay standing in a pan of water. When the glass is turned upside down, put over the candle, and placed in the water, the candle quickly goes out and water rushes up into the glass (as shown at the right).



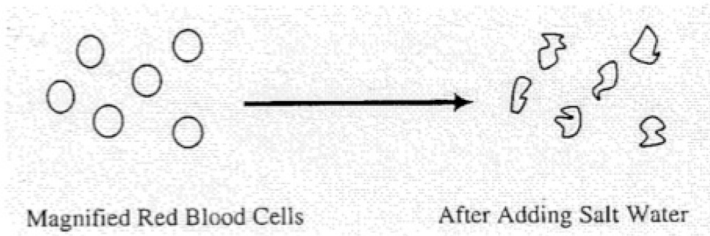
This observation raises an interesting question: Why does the water rush up into the glass?

Here is a possible explanation. The flame converts oxygen into carbon dioxide. Because oxygen does not dissolve rapidly into water but carbon dioxide does, the newly-formed carbon dioxide dissolves rapidly into the water, lowering the air pressure inside the glass.

Suppose you have the materials mentioned above plus some matches and some dry ice (dry ice is frozen carbon dioxide). *Using some or all of the materials, describe a way to test this possible explanation?*

22. What result of your test (mentioned in #21 above) would show that your explanation is probably wrong?

23. A student put a drop of blood on a microscope slide and then looked at the blood under a microscope. As you can see in the diagram below, the magnified red blood cells look like little round balls. After adding a few drops of salt water to the drop of blood, the student noticed that the cells appeared to become smaller.



This observation raises an interesting question: Why do the red blood cells appear smaller?

Here are two possible explanations: I. Salt ions (Na+ and Cl-) push on the cell membranes and make the cells appear smaller. II. Water molecules are attracted to the salt ions so the water molecules move out of the cells and leave the cells smaller.

Suppose you have some salt water, a very accurate weighing device, and some water-filled plastic bags, and suppose the plastic behaves just like red-blood-cell membranes.

Using some or all of the materials, describe an experiment to test these two explanations.

24. *What result of the experiment would best show that explanation I is probably wrong?*

What result of the experiment would best show that explanation II is probably wrong?

APPENDIX B

CLASSROOM TEST OF SCIENTIFIC REASONING

Use the following scale to respond to items 1-15. The items inquire into your confidence level in doing biology. For example, are you very confident, … confident, not at all confident?

a. very confident
b. confident
c. average
d. not very confident
e. not at all confident

1. Look at a skull that you have not seen before and determine what the animal ate and where it lived.

2. Explain to someone how a cow's skull is suited for eating plants.

3. Conduct a controlled experiment to find out whether isopods can sense light.

4. Conduct a controlled experiment to determine whether isopods can communicate.

5. Use the fossil record to construct an argument that supports evolution theory.

6. Based on observable similarities and differences, sort a sample of sea shells into groups that may represent different species.

7. Conduct an experiment to find out if the environment, genes, or both cause differences in shell coloration in a snail species.

8. Identify and analyze relevant data to test the hypothesis that cell phones cause brain tumors.

9. Gather and graph data that show the pattern of height variation in a sample of college students.

10. Design a test of the hypothesis that more grass grows on the north slope of "A" Mountain because its soil is wetter in the spring.

11. Conduct a controlled experiment to find out if a substance in a chemical reaction is broken apart or is a catalyst.

12. Explain to someone the difference between hypotheses and predictions including an example from the lab in which you mixed MnO_2 and H_2O_2 and tried to figure out which molecule is being broken apart.

13. Test the hypothesis that water rises in a cylinder inverted over a burning candle due to consumed oxygen.

14. Explain to someone why suction (defined as an attracting/pulling force) does not exist.

15. Identify similarities and differences in the vegetation growing at different elevations on a tall mountain.

APPENDIX C

IRB APROVAL

ARIZONA STATE UNIVERSITY
RESEARCH AND ECONOMIC AFFAIRS

Office of Research Integrity and Assurance

To:	Anton Lawson LSC
From:	for Mark Roosa, Chair SM JTF Soc Beh IRB
Date:	02/10/2010
Committee Action:	**Exemption Granted**
IRB Action Date:	02/10/2010
IRB Protocol #:	1002004799
Study Title:	Self-efficacy in science knowledge and learning

The above-referenced protocol is considered exempt after review by the Institutional Review Board pursuant to Federal regulations, 45 CFR Part 46.101(b)(2) .

This part of the federal regulations requires that the information be recorded by investigators in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. It is necessary that the information obtained not be such that if disclosed outside the research, it could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.

You should retain a copy of this letter for your records.