FCI Normalized Gain, Scientific Reasoning Ability, Thinking in Physics, and Gender Effects

Vincent P. Coletta[1], Jeffrey A. Phillips[1], and Jeff Steinert[2]

[1]*Department of Physics, Loyola Marymount University, Los Angeles, CA 90045 USA*
[2]*Arizona School for the Arts, Phoenix, AZ 85004 USA*

Abstract. We observe no significant effect of gender on grades in our IE introductory mechanics courses at Loyola Marymount University, but we do observe a significant gender gap on FCI normalized gains, with males achieving higher gains than females. Over the past three years FCI gains have improved for both male and female students in IE classes taught with the Thinking in Physics (TIP) pedagogy. However, a gender gap on FCI gains remains, even when scientific reasoning abilities are taken into account. Indeed, the gap appears much greater for students with the strongest scientific reasoning skills and the highest FCI gains. Data from IE introductory physics courses using modeling at Edward Little High School in Maine show a similar result, with some additional data showing a reverse gender gap for those students with very weak scientific reasoning skills.

Keywords: Scientific reasoning, introductory mechanics
PACS: 01.40.Fk

INTRODUCTION

Interactive methods of teaching physics are more effective than traditional methods [1,2]. However, our research demonstrates that those students who have weak reasoning skills are likely to have limited success in physics even in interactive courses [3,4,5]. The Lawson Classroom Test of Scientific Reasoning Ability is a multiple-choice test that includes questions on conservation, proportional thinking, identification and control of variables, probabilistic thinking, and hypothetico-deductive reasoning [3,6]. We have found a strong correlation between individual students' normalized gains on the Force Concept Inventory (FCI) and their preinstruction scores on the Lawson Test in IE introductory mechanics classes at Loyola Marymount University (LMU) (N = 98, r = 0.54) and at Edward Little High School (ELHS) (N = 199, r = 0.53) [3,4]. The correlation between Lawson score and FCI *G* has been replicated by many others, including Dubson and Pollack at U. of Colorado [7], Diff and Tache at Santa Fe Community College [8], Pamela and Saul at U. of Central Florida [9], Savinainen at Kuopion Lyseo High School in Finland [10], and Pyper at BYU [11]. We have also found strong, positive correlations between individual students' pre-instruction SAT scores and their normalized FCI gains at ELHS (N= 335, r = 0.57) and at LMU (N = 292, r = 0.46) [5].

Our research reveals that many students come to IE physics classes unable to use certain scientific reasoning skills needed to learn physics, and for those students, achieving good conceptual understanding in mechanics is unlikely. Two of us created the Thinking in Physics (TIP) program at LMU to enhance those reasoning skills. TIP identifies students as being at risk in introductory physics, based on their pre-instruction performance on the Lawson test and/or the SAT. We have previously described some of our methods and the effect they have had on students' understanding of mechanics concepts, as measured by improved normalized gains on the FCI [12] and in improvements in problem solving skills [13].

A gender gap in performance on the FCI and FMCE has been widely observed [14,15,16], with males achieving consistently higher average normalized gains than females. Some have speculated that the test itself is gender biased [17]. Osborn Popp, Meltzer, and Megowan-Romanowicz have presented evidence that the test is not systematically biased in favor of males [18], suggesting that the observed gap represents a real difference between males and females in the conceptual understanding achieved in IE introductory mechanics. Recently researchers at the U. of Colorado have suggested that a cause of the

observed gender gap may be stereotype threat, and that it can be mitigated by reflective writing exercises [19,20].

We have examined various test data for our students in IE introductory mechanics courses and, in terms of class test averages and overall grades, we found no significant gender effects. However, males consistently achieved higher FCI normalized gains than females. We wondered whether TIP may have had an effect on this gap and whether the gap was consistent across all student reasoning levels. We found that, while TIP improved conceptual understanding for both males and females, a significant gap persists. We also found that the gap tends to be particularly large for students with the strongest reasoning skills and largest FCI gains.

THINKING IN PHYSICS AND THE GENDER GAP

TIP develops general thinking and problem solving skills. TIP requires students to keep thinking journals, encouraging metacognitive reflection. TIP utilizes a variety of games and puzzles, including Sudokus, to facilitate student problem solving development. Using software developed by the MIND Research Institute, TIP computer games build skills, using visual techniques that do not rely on language abilities, but which do make heavy demands on working memory, requiring planning steps in advance to achieve a favorable outcome. TIP uses the context of physics to teach thinking about variables and their relationships. TIP activities include student group work and lab experiments. A more complete description of TIP and its success in developing conceptual understanding and problem solving skills is provided in earlier papers [12,13].

Each of us at LMU has developed our own courses with somewhat different features, emphasizing different aspects of the TIP pedagogy and utilizing it in somewhat different ways. In both versions of TIP, TIPC (Coletta) and TIPP (Phillips), students show a significant gender gap on FCI scores, but show no significant gender gap on course grades. Students in TIPC have somewhat higher FCI gains than those in TIPP, and the gender gap is larger in TIPC.

Compared with students in earlier pre-TIP IE classes taught by the same instructor in 2000, 2002, and 2006, students in TIPC classes in 2007, 2009, and 2010 have shown improved conceptual understanding of introductory mechanics, as measured by improved FCI normalized gains G. However, the large pre-TIP gap of 0.18 (G = 0.27 for females and 0.45 for males) has not changed in TIPC classes (G = 0.40 for females and 0.58 for males) (Fig. 1). For TIPP, the instructor taught no comparable courses prior to the beginning of the TIP project. For TIPP classes taught in 2008, 2009, and 2010, with a total of 75 females and 71 males, the gender gap was 0.09 (G = 0.33 for females and 0.42 for males). The difference is statistically significant. A one-tailed t-test gives $p < 0.005$.

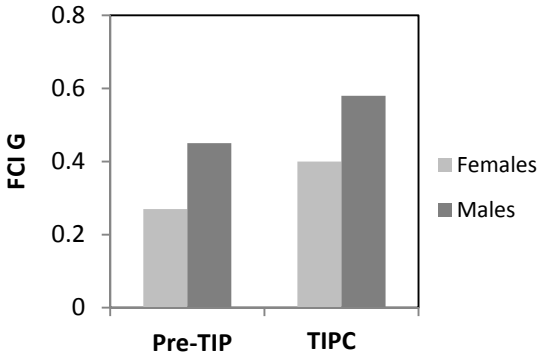

FIGURE 1. Thinking in Physics pedagogy has substantially improved FCI normalized gains for both females and males, but a large gender gap remains. There were 58 females and 25 males in the pre-TIP classes, and 59 females and 39 males in the TIPC classes. Standard errors for each of the four averages range from 0.02 to 0.04.

SCIENTIFIC REASONING ABILITY AND FCI GENDER GAP

Average scores for females on the Lawson test were significantly lower than for males in both TIP implementations (16.3 vs. 17.9 in TIPC and 15.9 vs. 17. 5 in TIPP). Given the correlation between Lawson scores and FCI gains, it is conceivable that the gender gap in Lawson scores might partially account for the gender gap in FCI gains. However, when we examined gender differences for students with similar Lawson scores, we found something more complex. The gender gap is larger than the overall gap for students with similar high Lawson scores, and smaller for students with similar low Lawson scores.

TABLE 1. Average FCI G and average Lawson scores for each quartile of Lawson scores for females and for males in TIPC. Standard errors for G within each quartile ranged from 0.03 to 0.07. Thus the male-female differences are significant for all but the lowest quartile. A one-tailed t-test on the highest male and female quartile gives $p < 10^{-4}$.

Females		Males	
Lawson	FCI G	Lawson	FCI G
9.8	0.33	13	0.36
15.2	0.32	17.2	0.48
18.8	0.44	19.3	0.67
21.4	**0.50**	22.4	**0.81**

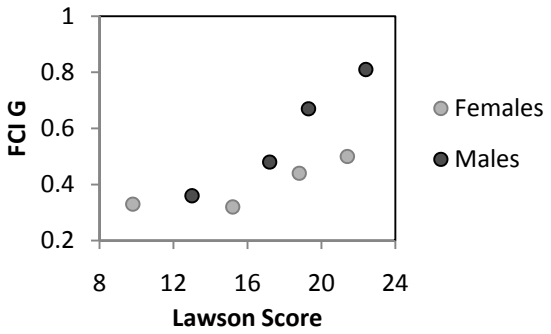

FIGURE 2. Average FCI G vs. average Lawson score for male and female Lawson quartiles in TIPC classes.

In TIPC classes there is little or no gender gap in FCI normalized gains for students with Lawson test scores in the lowest quartile, but in the highest quartile the gap is exaggerated: for males FCI G averaged 0.81, while for females FCI G averaged 0.50 (Table 1 and Fig. 2).

In TIPP classes there is little or no gender gap in FCI normalized gains for students with Lawson test scores in the lowest two quartiles. In the highest two quartiles the gaps are larger, but not statistically significant (Table 2). However, if we lump together Lawson scores of 17 or higher (the top two quartiles), providing a larger sample, the FCI *G* gap of 0.08 (0.50 for males and 0.42 for females) is statistically significant. A one-tailed t-test yields $p = 0.05$.

TABLE 2. Average FCI G and average Lawson scores for each quartile of Lawson scores for females and for males in TIPP. Standard errors for G within each quartile ranged from 0.03 to 0.07.

Females		Males	
Lawson	FCI G	Lawson	FCI G
11.9	0.23	10.6	0.28
15.0	0.29	15.2	0.28
17.3	0.37	18.0	0.47
20.8	0.45	21.6	0.53

Gender Gap at ELHS

From 2002 to 2005 one of us used Modeling to teach introductory mechanics at Edward Little High School to a total of 199 students, 98 female and 101 males. The females' average FCI normalized gain was 0.37 ± 0.02 (SE) and the males' average gain was 0.46 ± 0.03 (SE). A one-tailed t-test of these two populations shows statistical significance with $p < 0.005$. However, the males scored significantly higher than females on the Lawson test, 16.9 vs. 14.9, so this difference partially accounts for the difference in FCI gains. If we consider only data from students with Lawson scores of 19 or higher, a clearer picture emerges, consistent with the results from LMU. For the 20 females with scores in this highest range, the average FCI G was 0.53 ± 0.04 (SE), and for the 38 males in this Lawson range, the average FCI G was 0.66 ± 0.04 (SE), a highly significant difference. A one-tailed t-test yields $p < 0.01$. The average Lawson scores in this range were not significantly different for males and females.

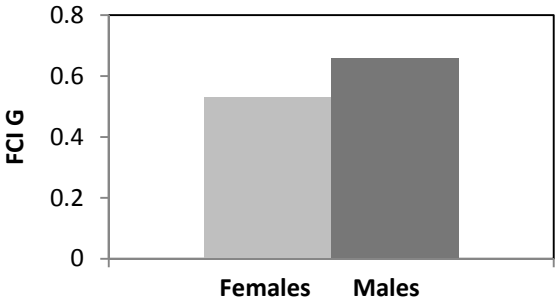

FIGURE 3. Gender gap at ELHS for Lawson scores of 19 or greater.

Reverse Gender Gap at ELHS

An interesting phenomenon is seen when we examine the subset of the ELHS data corresponding to Lawson scores of 10 or less, with 14 females and 7 males falling in this range. For those students with the weakest scientific reasoning skills, we observe a reverse gender gap on the FCI, with females achieving higher gains than males (Fig. 3).

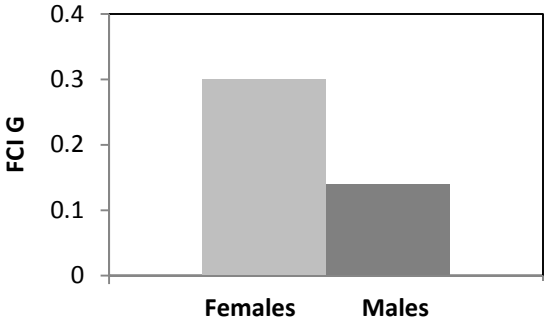

FIGURE 4. Reverse gender gap for lowest range of Lawson scores at ELHS, scores of 10 or less. The 14 females with scores in this range had an average FCI G of 0.30 ± 0.05 (SE) and the 7 males with scores in this range had an average FCI G of 0.14 ± 0.04 (SE). A two tailed t-test on the two groups yields $p = 0.02$.

A Conjecture

For IE physics classes in schools with high at-risk populations, classes in which scores on the Lawson Test are very low, we conjecture that the class average FCI normalized gains for females may be higher than the class average gains for males.

SUMMARY AND QUESTIONS

We find that, while there is no gender gap in course grades, there is a significant gender gap in FCI normalized gains. That gender gap is greatest for those students with the strongest reasoning skills and the highest FCI gains, with males having substantially higher gains than females. The gender gap is reversed for students with very low Lawson scores.

The size of the gender gap varies both with the general IE approach (TIP vs. Modeling) and with specific instructor and methods (TIPC vs. TIPP). These observations raise some important questions: Why the differences in the two versions of TIP? If stereotype threat is the source of much of the gender gap, why is there no gender gap in course grades, & why an enhanced gender effect for the strongest reasoners and a reverse gender effect for the weakest?

REFERENCES

1. R. R. Hake, *Am. J. Phys.*, **66**(1), p. 64-74 (1998).
2. E. F. Redish and R.N. Steinberg, *Phys. Today*, **52**(1): p.4-30 (1999).
3. V. P. Coletta and J. A. Phillips, *Am. J. Phys.*, **73**: p1172-1182 (2005).
4. V. P. Coletta, J. A. Phillips and J. J. Steinert, *Phys. Teach.*, **45**: p. 626-629 (2006).
5. V. P. Coletta, J. A. Phillips and J. J. Steinert, *Phys. Rev. ST: Phys. Educ.* **3** p. 010106-1– 010106-5 (2007)
6. A. E. Lawson, *Am. Biol. Teach.*, **62**(7):p. 482-95 (2000).
7. M. A. Dubson and S. J. Pollock. AAPT *Announcer* **36** (2): 90 (2006).
8. K. Diff and N. Tache, *AIP Conf Proc.* **951**: p. 85-87 (2007).
9. P. Pamela and J. Saul, AAPT Announcer **36**: p. 89 (2006).
10. A. Savinainen, private communication (2011).
11. 11. B. Pyper, 2011 invited PERC presentation.
12. V. P. Coletta and J. A. Phillips, *AIP Conf Proc.* **1179**: p. 117-120 (2009).
13. V. P. Coletta and J. A. Phillips, *AIP Conf Proc.* **1289**: p. 13- 16 (2010).
14. R.R. Hake, submitted to PERC; online at http://www.physics.indiana.edu/~hake/PERC2002h-Hake.pdf (2002)
15. M. Lorenzo, C. Crouch, and E. Mazur, *Am. J. Phys.*, **74 (2)**, p. 118-122 (2006).
16. L. E. Kost, S. J. Pollock, and N. D. Finkelstein, *Phys. Rev. ST: Phys. Educ.* **5** p. 010101 (2009)
17. L. McCullough, *J of International Women's Studies* **5** (4): p. 20-30.
18. S. E. Osborn Popp, D. E. Meltzer, and C. Megowan-Romanowicz, presentation at Am. Ed. Res. Assoc. (2011).
19. L. E. Kost-Smith *et al*, *PERC proceedings* **1289**: p. 197- 200 (2010).
20. A. Miyake *et al*, *Science*, **330,** p. 1234 -1237 (2010).