Predicting FCI Gain With A Nonverbal Intelligence Test

M.R. Semak, R.D. Dietz, R.H. Pearson and C.W. Willis

University of Northern Colorado, Greeley, CO 80639

Abstract. We have administered both a commercial, nonverbal intelligence test (the GAMA) and Lawson's *Classroom Test of Scientific Reasoning* to students in two introductory physics classes to determine if either test can successfully predict normalized gains on the *Force Concept Inventory.* Since gain on the FCI is known to be related to gender, we adopted a linear model with gain on the FCI as the dependent variable and gender and a test score as the independent variables. We found that the GAMA score did not predict a significant amount of variation beyond gender. Lawson's test, however, did predict a small but significant variation beyond gender. When simple linear regressions were run separately for males and females with the Lawson score as a predictor, we found that the Lawson score did not significantly predict gains for females but was a marginally significant predictor for males.

Keywords: Physics education research, student understanding, gender, bias, mechanics
PACS: 01.40.Fk, 01.40.gf, 01.75.+m

INTRODUCTION

The *Force Concept Inventory* (FCI) [1] is widely used to assess student understanding of some of the fundamental concepts of classical mechanics. When used as a pre- and post-test, the gain students achieve on the FCI is usually interpreted as a measure of instructional efficacy. If a reliable way were found to predict such gains, our understanding of the learning process would benefit. Coletta and Phillips [2] have advocated the use of Lawson's *Classroom Test of Scientific Reasoning* (found in the appendix of [2]) as such a predictor, and they found a significant correlation between scores on the Lawson test and FCI gain. The Lawson test is highly contextualized in that it deals with concepts frequently used in science. We investigated a test of logical reasoning that is abstract and devoid of any direct scientific or verbal context to determine how it would compare to the Lawson test as a predictor of FCI gain.

This test is the GAMA (General Ability Measure for Adults) by Naglieri and Bardos which is commercially available from Pearson Assessments [3]. These authors claim that the GAMA "evaluates an individual's overall general ability with items that require the application of reasoning and logic to solve problems that exclusively use abstract designs and shapes" [4]. It may be that such a nonverbal test can better serve the process of assessment for, in particular, a population with diverse language backgrounds. Indeed, we feel that the form of the GAMA allows us to minimize the issue of context as we test our students.

There are four types of questions posed on the GAMA examples of which are presented in Figure 1. The first kind is a *matching subtest* which requires one to identify the spatial orientation of a design consisting of yellow and blue shapes and find its match among a selection of such designs with variations in orientation and color. Second, there is the *analogies subtest* that presents two sample designs with some conceptual relationship that must be found between another design and a group of options. Third, the *sequences subtest* has one examine an incomplete sequential evolution of a pattern and identify the appropriate missing piece. Finally, fourth, the *construction subtest* asks one to recognize which design from a selection can be built from two shapes of specific color.

The GAMA itself consists of 66 items of increasing difficulty. Unlike the Lawson test, the GAMA is timed (25 minutes), and no one is expected to finish it. A complex scoring rubric leads to a single number, the GAMA IQ score, which if measured for a large random population will have a mean of 100 and a standard deviation of 15.

Overall, one can wonder if the proper interpretation of the results of a test of this kind can possibly help us develop insight as to the fundamental cognitive factors affecting a student's physical understanding. Before attempting to deal with such questions, however, we simply would like to know if, in fact, any significant correlation exists between our students' performance on the GAMA and the FCI, an assessment directly addressing their physical reasoning. That is the focus of this work.

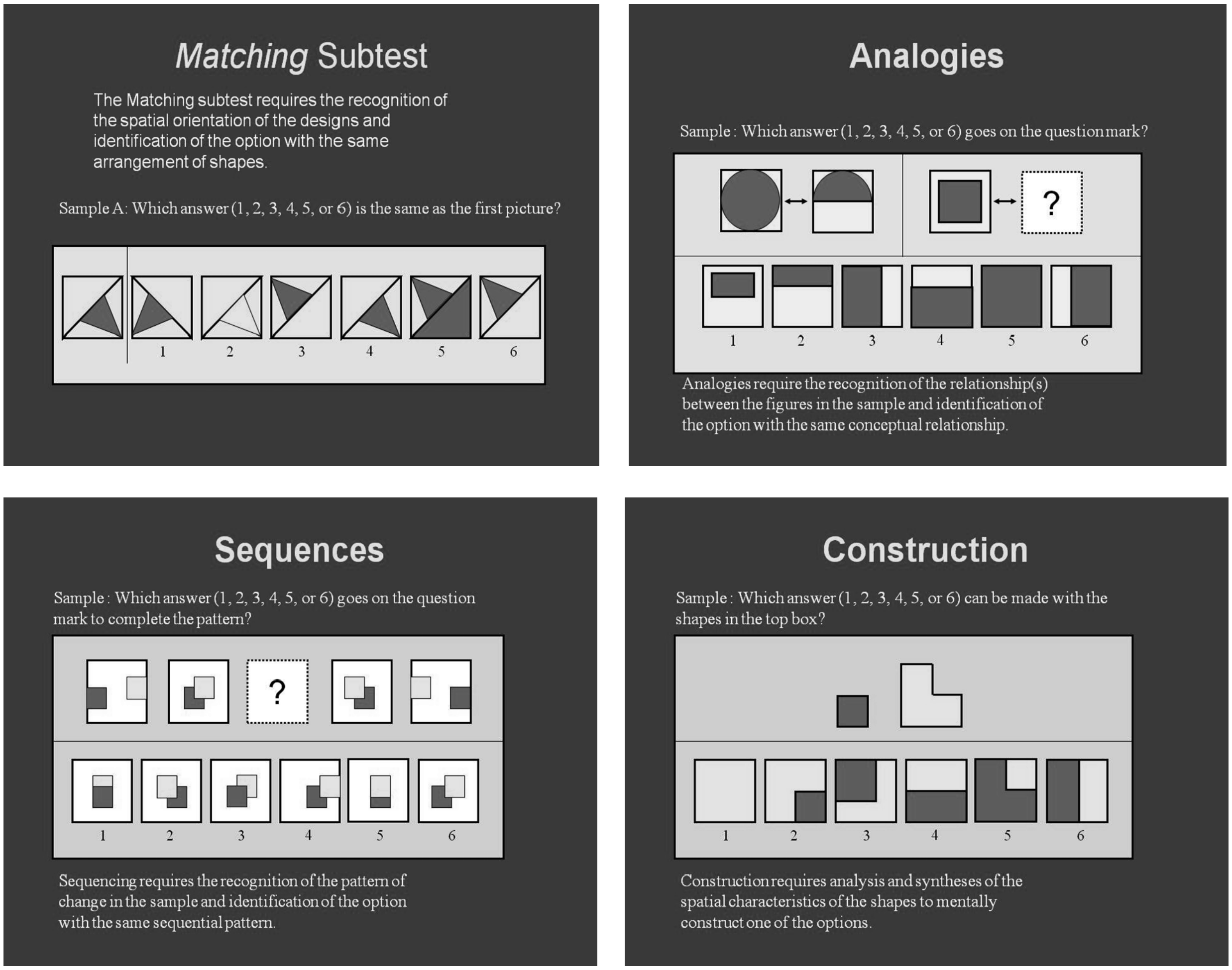

FIGURE 1. Simulated items similar to those found in the *General Ability Measure for Adults* (GAMA). Copyright © 2006 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "GAMA" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliate(s).

METHODOLOGY

The FCI, the GAMA and Lawson's *Classroom Test of Scientific Reasoning* were administered sequentially during one session to both an algebra-based and calculus-based class prior to instruction using the traditional lecture format. The FCI was then given to the same population after instruction. Seventy students participated in the study. The FCI (normalized) gain was calculated as the ratio between the actual change and the greatest possible change in one's score:

$$Gain = \frac{postscore\% - prescore\%}{100 - prescore\%}.$$

Next, the data were treated with a linear regression analysis as described below.

ANALYSIS

As stated above, we wished to investigate how well the FCI normalized gains over the course of a semester could be explained by two cognitive measures: the GAMA and Lawson's *Classroom Test of Scientific Reasoning*. Linear regression analysis was used to answer this question. In all of our regression models, FCI normalized gain of each student was used as the dependent variable. Since FCI scores are known to differ by gender [5], that factor was accounted for in all models by including the dummy variable MALE which was coded 1 for males and 0 for females.

First, we modeled the relationship between the GAMA IQ scores and FCI gains after adjusting for gender. Given that we expected gender to be related

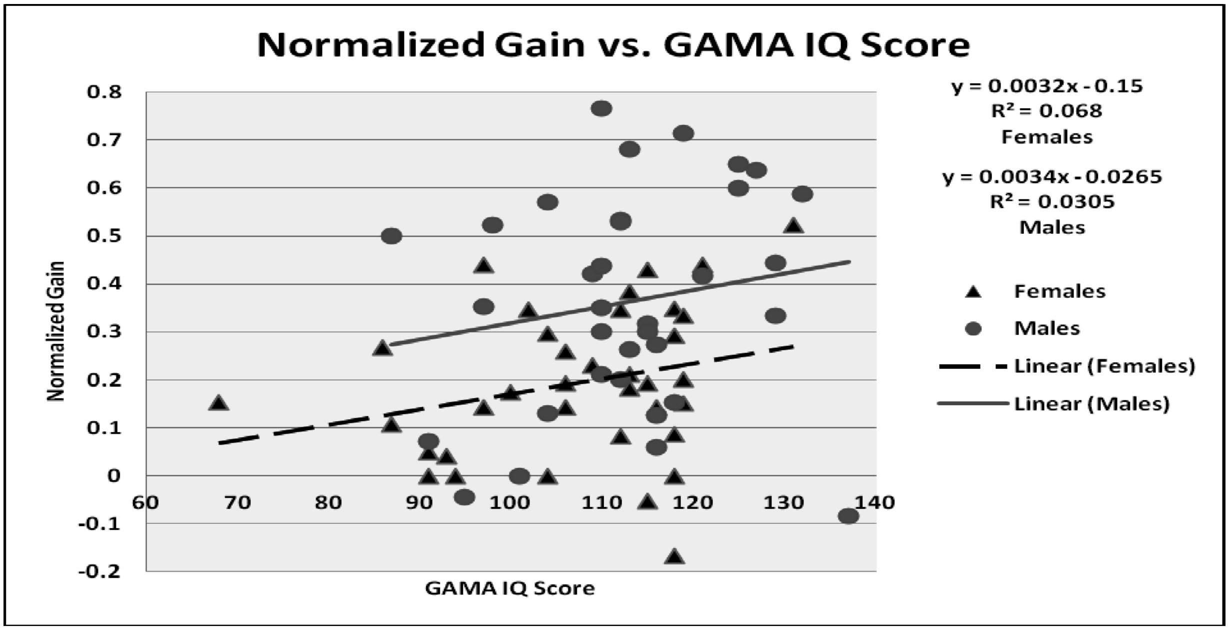

FIGURE 2. Simple linear regressions between the GAMA IQ scores and the FCI normalized gains done for the females and males. The small positive slopes are found to be statistically negligible.

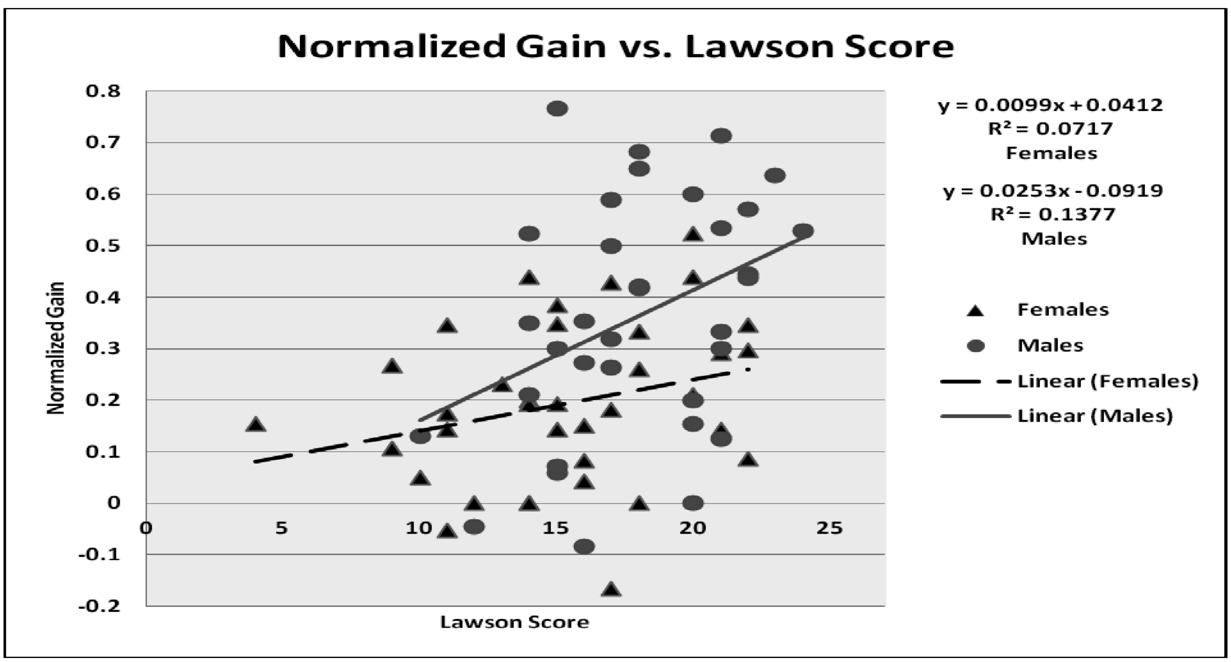

FIGURE 3. Simple linear regressions fitting gender data for Lawson's test scores and the FCI normalized gains. Note the marginally significant positive slope for the males' model whereas the slope for that of the females is statistically no different from zero.

to FCI gains, Type I (sequential) sums of squares were examined first for MALE and then for the GAMA IQ scores. Thus, the effect of MALE was unadjusted while the effect of the GAMA IQ scores corresponded to the variance it explained after accounting for MALE. This model was significant, $F(2, 67) = 8.24$, $p < 0.001$, $R^2 = 0.20$, but the GAMA did not explain a significant amount of additional variation beyond gender, $F(1, 67) = 3.01$, $p = 0.087$, $\omega^2 = 0.024$.

This model for the GAMA could be described as fitting separate simple regression lines for males and females which are constrained to be parallel. The test for MALE is of different intercepts, and the test for the GAMA is whether the slope is different from 0. Testing for slope differences requires computing a variable from the product of MALE and the GAMA IQ scores (denoted MxG). Since the GAMA did not significantly explain any more variance in FCI gains beyond gender, this interaction (MxG) could not be investigated as MxG and MALE are multicollinear (one couldn't tell how changes in these variables cause changes in the FCI gains). Moreover, since the GAMA is not significant, we did not expect useful explanatory information to be gained by examining separate simple regressions for genders. These regressions were considered here only to disconfirm the unlikely possibility that the two regression lines had slopes that were significantly different from 0 but in opposite directions as well as aid in visualizing the possible correlation between the FCI gains and the GAMA IQ scores for females and males (Figure 2).

Second, we examined a model of the relationship between Lawson's test scores and FCI gains after adjusting for gender, and, again, considered Type I sums of squares for individual effects. The overall model explained a significant (at the 0.05 level) amount of variance in gains, $R^2 = 0.24$, $F(2, 67) = 10.6$, $p < 0.0001$. After adjusting for MALE, Lawson's test explained a small but significant amount of additional variance, $F(1, 67) = 7.0$, $p = 0.010$, $\omega^2 = 0.067$.

This model for Lawson's test has the same general structure as that for the GAMA described above. Testing for differences in slope for females and males could not be done due to, again, multicollinearity issues. As an alternative, simple linear regressions were separately run for males and females using Lawson's test as a predictor (Figure 3). Lawson's test did not significantly predict gains for females, $R^2 = 0.07$, $F(1, 34) = 2.63$, $p = 0.114$ but was a marginally significant predictor (at 0.05) for males, $R^2 = 0.14$, $F(1, 32) = 5.11$, $p = 0.031$.

Also, note that residuals were examined after fitting both models and in neither case were there any suggestions of important assumption violations.

CONCLUSION

Interestingly, we found that the GAMA, the instrument lacking scientific and verbal context, does not help to predict FCI gains for our sample. We are left with the gender means which could be used to posit a guess as to a student's gain (in our case, 0.19 for females and 0.36 for males). Other studies [6, 7] have reported positive correlations between the GAMA and other cognitive measures. This raises the question of the role of reasoning ability with respect to the FCI.

As for Lawson's test, with its scientific context, it is found to have some possible significance after accounting for gender as a predictor for FCI gains. Separate linear regressions for females and males revealed marginal significance for males, yet none for females. This suggests that the scientific reasoning ability of males may help predict their FCI gains, but that would not be the case for females. This contradicts the findings of Coletta and Phillips [2] who claim a significant correlation between FCI gains and Lawson's test scores for their population of study. No distinction was made between genders in that study, so a direct comparison is not possible.

To extend this work, we would like to determine if significant (gender-specific) correlations exist between our students' FCI gains and their scores on other measures of cognitive ability. These would test various types of reasoning skills. This could be done in conjunction with conducting "think aloud" [8] studies concerning students' thought processes while taking the FCI. This approach may shed some light on, among other things, the relation between the types of reasoning used in answering FCI questions and gender.

REFERENCES

1. D. Hestenes, M Wells, and G. Swackhamer, "Force Concept Inventory," Phys. Teach. 30, 141–158 (1992).
2. V. P. Coletta and J. A. Phillips, "Interpreting FCI scores: normalized gain, preinstruction scores, and scientific reasoning ability," American Journal of Physics **12**, 1172 (2005).
3. J. A. Naglieri and A. N. Bardos, General Ability Measure for Adults (GAMA®), Copyright © 2012 Pearson Education, Inc. or its affiliate(s). http://psychcorp.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=PAg118
4. A.N. Bardos, "The General Ability Measure for Adults," in *Handbook of Nonverbal Assessment*, edited by R. S. McCallum, New York: Kluwer Academic/Plenum Publishers, 2003, pp. 163-174.
5. J. Docktor and K. Heller, "Gender Differences in Both Force Concept Inventory and Introductory Physics Performance," 2008 Physics Education Research Conference, AIP Conf. Proceedings 1064, AIP, Melville, NY, 2008, pp. 15-18.
6. K. S. Lassiter et al., "College Student Performance on the General Ability Measure for Adults and the Wechsler Intelligence Scale for Adults – Third Edition," Psychology in the Schools, **38(1)**, 1 (2001).
7. K .S. Lassiter et al., "Comparison of the General Ability Measure for Adults and the Kaufman Adolescent and Adult Intelligence Test with College Students," Psychology in the Schools, **39(5)**, 497 (2002).
8. M. W. van Someren, Y. F. Barnard and J. A. C. Sandberg, *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*, London: Academic Press Limited, 1994.