Can a Nonverbal Intelligence Test Predict FCI Performance?

Matthew R. Semak, Cynthia S. Galovich, Richard D. Dietz
Department of Physics and Astronomy , University of Northern Colorado, Greeley, CO 80639

Abstract: We have administered a commercial, nonverbal intelligence test (GAMA) to students in introductory physics. We examined scores on that test to determine if it or any of its parts can predict student achievement on the *Force Concept Inventory* before and/or after instruction. In addition, given that a gender gap in performance is often seen to exist, gender is also considered as a predictor.

Introduction: The *Force Concept Inventory* (FCI) is widely used to assess student understanding of some of the fundamental concepts of classical mechanics. We investigated a test of logical reasoning that is abstract and devoid of any context, scientific or otherwise, to determine how it would serve as a predictor of FCI performance. We also investigate how the four different types of questions on the test may serve as individual predictors. Moreover, issues related to gender (gender gap) are often encountered when the FCI is administered. As gender may correlate with students' performance on the FCI, we check its role as another predictor along with the test of logical reasoning.

The GAMA: The test we used is the GAMA (General Ability Measure for Adults) by Naglieri and Bardos which is commercially available from Pearson Assessments. The GAMA uses four types of questions, and we consider the collection of each type of question to be a separate subtest. Examples of the four types of questions posed on the GAMA are presented in the next column. The test itself consists of 66 similar items of increasing difficulty. 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. The results for our population of 83 students show a mean of 112 and a standard deviation of 11.

Multiple Regression: The stepwise multiple regression process selects the best predictors of a dependent variable (such as FCI score) from among the set of independent variables which are gender and the individual subtest scores. This method also detects any interactions and/ or correlations among the independent variables which might preclude them from being independent predictors. No such interactions/correlations were indicated. Finally, residuals were examined for any patterns that would indicate that our predictors are missing some explanatory information (any error in the fit should not contain predictive information). No such patterns were found. Statistical significances for the regression process were set at the 0.05 level.

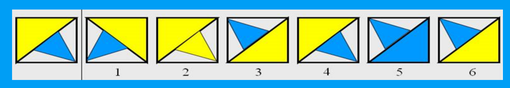
Analysis: We considered three cases for our study. First, the score on the entire FCI was used as the dependent variable. Then, in an effort to better understand our students' reasoning process, we split the FCI questions into two parts: one part consists of questions we feel would require strong reasoning ability and the other part consists of questions that could be answered based on past experience. This division was informed by many years of instruction of introductory physics students. Then, with the scores on these two question sets as dependent variables, a separate regression was performed for each.

For all three cases, an F – test was performed (at the 0.05 level) to assess the significance of the regression models; i.e., to test whether the coefficients for the models are significantly different from zero, and thus, the relationships found between independent and dependent variables are statistically significant. Also, an R – squared value was calculated to determine how much of the variance of a data set is explained by the relationship found through the regression process.

Matching subtest

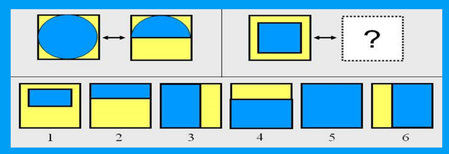
Requires the recognition of the spatial orientation of the designs and identification of the option with the same arrangement of shapes.

Sample : Which answer (1, 2, 3, 4, 5, or 6) is the same as the first picture?



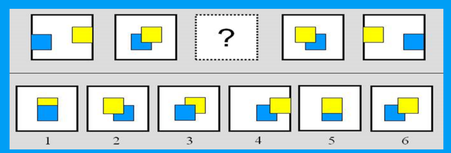
Analogies subtest

Requires the recognition of the relationship(s) between the figures in the sample and identification of the option with the same conceptual relationships.



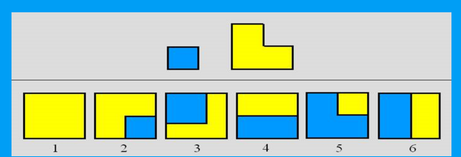
Sequences subtest

Requires the recognition of the pattern of change in the sample and identification of the option with the same sequential pattern.



Construction subtest

Requires analysis and syntheses of the spatial characteristics of the given shapes to mentally construct one of the options.



Conclusions:

- Only two parts of the GAMA test can be used as predictors: the matching and sequences subtests.
- The model which includes gender and the two subtests as predictors is more useful for the post-test.
- Gender is always a significant predictor.
- We could not identify a useful model to explain gain or normalized gain.
- In case 2, it is not surprising that GAMA subtests didn't serve as predictors for the pre-test performance, as it is difficult to reason about something about which you have limited knowledge. This is in contrast to case 3, in which questions can be answered based on personal experience.

Results:

Case 1: All FCI Questions Used in Regression

For both the pre- and post-test performance, gender and the matching and sequences subtests were found to be significant predictors.

- For the pre-test scores, a model with gender alone accounts for 27% of the explained variance whereas including the two subtests increases this value to 35%.
- For the post-test scores, a model with gender alone accounts for 34% of the explained variance whereas including the two subtests increases this value to 50%.

Relevant Statistics

FCI Test	Predictors	P-Value for F-test	R-Squared
Pre-test	Gender and Matching and Sequences Subtests	0.014	0.35
Post-test	Gender and Matching and Sequences Subtests	0.000	0.50

Case 2: FCI Questions Requiring Strong Reasoning Ability

For the pre-test performance, only gender is a significant predictor. For the post-test performance, gender and the matching and sequences subtests were found to be significant predictors.

- For the pre-test scores, a model with gender alone accounts for 23% of the explained variance.
- For the post-test scores, a model with gender alone accounts for 27% of the explained variance whereas including the two subtests increases this value to 39%.

Relevant Statistics

FCI Test	Predictors	P-Value for F-test	R-Squared
Pre-test	Gender	0.000	0.23
Post-test	Gender and Matching and Sequences Subtests	0.001	0.39

Case 3: Remaining FCI Questions

As in case 1, for both the pre- and post-test performance, gender and the matching and sequences subtests were found to be significant predictors.

- For the pre-test scores, a model with gender alone accounts for 20% of the explained variance whereas including the two subtests increases this value to 29%.
- For the post-test scores, a model with gender alone accounts for 31% of the explained variance whereas including the two subtests increases this value to 48%.

Relevant Statistics

FCI Test	Predictors	P-Value for F - test	R-Squared
Pre-test	Gender and Matching and Sequences Subtests	0.009	0.29
Post-test	Gender and Matching and Sequences Subtests	0.000	0.48

Acknowledgment: We thank Susan Hutchinson, Professor of Statistics, Applied Statistics and Research at UNC.