UNIVERSITY of NORTHERN COLORADO

Exploring Gender Differences in Force Concept Inventory Results through Factor Analysis

by

Richard D. Dietz, Matthew R. Semak, Robert H. Pearson, Courtney W. Willis

University of Northern Colorado, Greeley, CO 80639

Abstract: It has been widely reported that males perform better than females on the Force Concept Inventory (FCI). The reasons for this situation remain obscure. We have determined that applying the technique of factor analysis to FCI responses by all students in introductory physics leads to the identification of several factors associated with distinct physical concepts. Here we apply the same technique to the analysis of FCI results but with the responses separated by gender in the hope of gaining some insight into the performance difference between males and females.

Introduction

In our earlier work, we studied patterns among our students' responses to the FCI to explore their conceptual approach to Newtonian mechanics. These patterns suggested that certain groups of the test's questions have well-defined physical concepts in common. The existence of these groups indicates that our students may recognize certain physical concepts during the process of solving the FCI problems. The analysis was accomplished using the data reduction technique of exploratory factor analysis.

Exploratory factor analysis is a statistical technique used to search for response patterns in large data sets. It investigates patterns of correlations among observed variables in order to identify and study constructs that are not directly observable but hypothesized to exist. Factor analysis does not identify what the unobserved constructs are, only the pattern of responses. The researchers must try to identify them from the questions which load on each factor. In our study, the variables are the FCI's questions, and the factors are the physical concepts common to certain sets of questions.

Hoping to gain insight into the widely reported gender gap in performance on the FCI, we performed the same analysis separately for female and male students. We found that the factors identified for both genders were the same as those for the entire population. However, in certain cases, the questions loading onto each factor differed for males and females.

Results of Past Work

In our previous work, we considered an FCI data set consisting of 406 pre-tests and 322 post-tests collected over a three-year period from students in introductory mechanics. Our exploratory factor analysis of the post-test data presented in the orange box to the right was done using the statistical package Mplus[1] with weighted least squares and a Promax rotation.

There is no universally accepted rule that prescribes how many meaningful factors may be extracted from a data set. The five factor solution was ultimately chosen because it proved to be the most interpretable. Our interpretations of the five factors are listed at the top of the orange box. The numbers listed for each are the "loadings", which are a measure of the degree of correlation between the question and the factor.

MPLUS EFA (WLS - Promax Rotation) of Post-Test (n=322)

	F1 Result of Applied Force	F2 Third Law	F3 Identify Forces	F4 Balanced Forces	F5 First Law
Q23	72				
Q27	66				
Q10	60				
Q8	59				
Q24	55				
Q9	52				
Q22	51				
Q19	51				
Q3	46				
Q14	*				
Q12	*				
Q20	*				
Q21	*				
Q2	*				
Q29	*				
Q4		84			
Q15		80			
Q28		64			
Q16		45			
Q18			78		
Q5			75		
Q11			65		
Q13			51		
Q30			*		
Q25				92	
Q17				54	
Q26				47	
Q6					78
Q7					66
Q1					55

Analysis by Gender

In this work, we use scree plots to compare the responses of females and males on the FCI post-test. To accomplish this, a factor analysis of the subset of questions relating to each of the five factors we identified was performed. A loading cut-off of 45 was used.

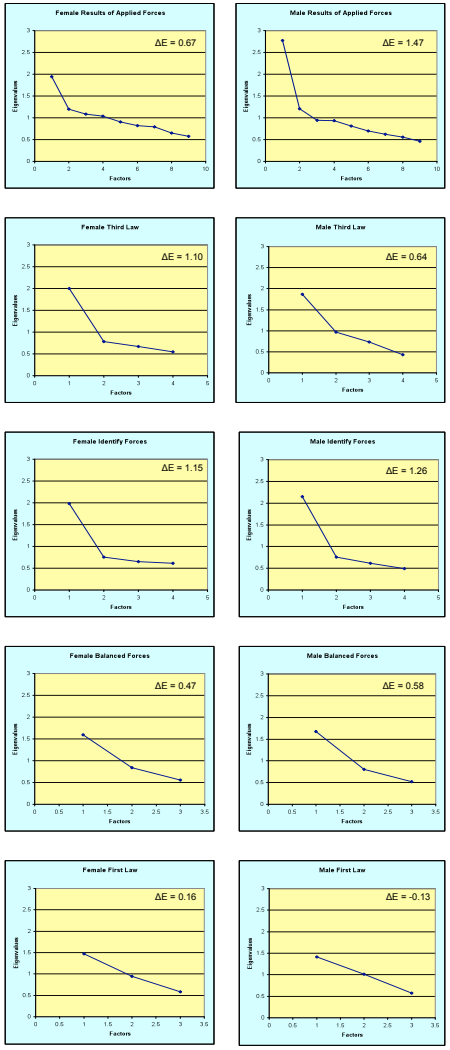
A scree plot is useful in determining how many meaningful factors are to be retained from a set of data. Each factor has a corresponding eigenvalue. The higher the eigenvalue, the better defined the underlying factor.

In performing the gender specific factor analysis, we test how strongly each of the five factors relates to its corresponding questions. If a factor is, in fact, "well-defined", the first eigenvalue will be much larger than the others, indicating it is primarily that one identified factor which the associated questions have in common.

The scree plot not only provides an effective visualization of a factor's robustness, it also suggests a way to quantify it. We compare the slope of the line joining the eigenvalues of factors 1 and 2 with the slope of a line fit to the eigenvalues of factors 2 and greater. Following Adams et al.[2], we define ΔE to be the slope of the second line minus the slope of the first line. The larger ΔE, the more pronounced is the scree plot's "elbow" and correspondingly the degree to which the first eigenvalue dominates.

From the shape of the scree plots it is evident that the first three factors are more well-defined than the last two. This suggests that there is room for improvement in our definition of factors.

The difference between the plots for males and females is not pronounced, but two may have some significance. In this analysis, the males seem to have formed a stronger factor structure in the case of applied forces. The same may be said of the females for the third law.



References

1. L. K. Muthén and B. O. Muthén, *Mplus User's Guide 5th Ed*. Los Angeles: Muthén & Muthén (1998-2009).
2. W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Weiman, "A new instrument for measuring student beliefs about physics and learning physics: the Colorado Learning Attitudes about Science Survey" *Phys. Rev. Special Topics: Physics Education Research*, **2**, 010101 (2006).