UNIVERSITY of NORTHERN COLORADO

Factor Analysis and Question Categorization in the Force Concept Inventory

by

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

University of Northern Colorado, Greeley, CO 80639

Abstract: The application of factor analysis to the Force Concept Inventory (FCI) has proven to be problematic. Some studies have suggested that factor analysis of test results serves as a helpful tool in assessing the recognition of Newtonian concepts by students. Other work has produced at best ambiguous results. One explanation for this may lie in the classification scheme of questions to which the results of the factor analysis are compared. We investigate some different categorizations of the questions on the FCI to see if we obtain better agreement between those categories and factors identified through factor analysis of student responses.

Introduction

The exchange of views between Huffman, Heller, Hestenes, and Halloun[1,2,3] in the pages of *The Physics Teacher* back in 1995 brought the concept of factor analysis to the attention of the physics education research community. In the ensuing fifteen years little if anything has been published by the AAPT on this topic. Our work here is an attempt to continue and extend the conversation on factor analysis.

Factor analysis is a statistical technique used to search for response patterns in large data sets. It does this by investigating 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.

Technique

Our FCI data set consists of 406 pre-tests and 322 post-tests collected over a three-year period from students in introductory physics. Newtonian mechanics was the subject of study over most of the semester for all students. The application of factor analysis to such a data set is often more of an art than a science since there are a variety of mathematical approaches that differ in significant ways.

The leftmost pair of panels compares pre- and post-test results analyzed in exactly the same way using SPSS[4]. Only three factors are identified in the pre-test results indicating weak correlations between questions that could easily arise from guessing.

Two different software packages were used to analyze the post-test results (rightmost pair of panels). Our original analysis was done using an older version of SPSS[4]. Our final analysis was done using a new and more sophisticated package titled Mplus[5]. The results differ in detail, but the conclusions seem to be robust in that the same five factors were identified.

Justification

There is no universally accepted rule that prescribes how many meaningful factors may be extracted from a data set. We employed a variety of techniques to determine the number of factors to retain for the post-test data. The eigenvalues-greater-than-one criterion suggested eight factors but is known to overestimate. Parallel analysis suggested four factors. The five factor solution was ultimately chosen for two reasons. First and foremost it proved to be the most interpretable. Also, Mplus provides a chi-square lack-of-fit statistic associated with the weighted least squares procedure. The associated p-value tests the null hypothesis that m factors are adequate vs. the alternative that more factors are needed. In the present study the five factor model was the smallest not to be rejected.

The two leftmost analyses below were done using SPSS with principal axis factoring and Quartimax rotation. The rightmost analysis was done using Mplus[5] with weighted least squares and a Promax rotation.

SPSS (PAF - Quartimax Rotation) Pre-Test (n=406)

	F1	F2	F3	F4	F5
Q10	61				
Q23	55				
Q14	51				
Q24	50				
Q12	47				
Q6	44				
Q20	42				
Q8	*				
Q1	*				
Q27	*				
Q7	*				
Q19	*				
Q22	*				
Q16	*				
Q3	*				
Q21	*				
Q13		54			
Q30		51			
Q11		47			
Q18		47			
Q26		40			
Q17		*			
Q5		*			
Q25		*			
Q2		*			
Q29		*			
Q15			62		
Q28			60		
Q4			49		
Q9				*	

SPSS (PAF - Quartimax Rotation) Post-Test (n=306)

	Result of Applied Force	Third Law	Identify Forces	Balanced Forces	First Law
	F1	F2	F3	F4	F5
Q23	56				
Q10	54				
Q13	52		43		
Q19	50				
Q27	46				
Q22	45				
Q20	45				
Q14	43				
Q3	43				
Q24	43				
Q26	43				
Q30	41				
Q8	41				
Q9	*				
Q12	*				
Q21	*				
Q2	*				
Q29	*				
Q4		69			
Q15		59			
Q28		51			
Q16		*			
Q18			56		
Q5			52		
Q11	44		45		
Q25				77	
Q17				42	
Q6					47
Q7					47
Q1					*

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

	Result of Applied Force	Third Law	Identify Forces	Balanced Forces	First Law
	F1	F2	F3	F4	F5
Q23	72				
Q27	66				
Q10	60				
Q8	59				
Q24	55				
Q9	52				
Q22	51				
Q19	51				
Q3	46				
Q14	44				
Q12	42				
Q20	40				
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

Pre and post data analyzed using the exact same statistics package, SPSS.

Post data analyzed using two different software packages (SPSS and Mplus)

Analysis

Hestenes and Halloun[2] maintained that factor analysis was an inappropriate tool to use in the study of FCI results. Our work does not support that contention nor does it support the contention made by Huffman and Heller[1] that the FCI does not measure student's understanding of well-defined aspects of Newtonian mechanics.

We found that the exploratory factor analysis of the pre-test data indicates no clear-cut factors that can easily be identified with introductory physics concepts, However, analysis of the post-test results does show that the students' responses have changed considerably and produces five factors that can be associated with specific physics concepts. Two of those factors appear in the list of six conceptual dimensions proposed by the authors of the FCI[6].

References

1. D. Huffman and P. Heller, "What does the force concept inventory actually measure?" *Phys. Teach.* **33**, 138-143 (1995).
2. D. Hestenes and I. Halloun, "Interpreting the Force Concept Inventory" *Phys. Teach.* **33**, 502-506 (1995).
3. P. Heller and D. Huffman, "Interpreting the Force Concept Inventory" *Phys. Teach.* **33**, 503-511 (1995).
4. SPSS for Windows, Rel. 6.1.4. Chicago: SPSS Inc. (2001)
5. L.K. Muthén and B.O. Muthén, *Mplus User's Guide 5th Ed*. Los Angeles: Muthén & Muthén (1998-2009).
6. D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory" *Phys. Teach.* **30,** 141-158 (1992).