

Cultural and Linguistic Influences on the Force Concept Inventory: A Preliminary Study

Colm Mealy¹, Saleh Al Hashemi¹, Barbara M. Olds^{2*} and Ronald L. Miller²

¹Petroleum Institute in Abu Dhabi/ ²Colorado School of Mines

Introduction

The Force Concept Inventory (FCI) is one of the earliest and most widely used concept tests in physics.^{1,2} According to its authors, this multiple choice assessment tool is “designed to assess student understanding of the most basic concepts in Newtonian mechanics.” David Hestenes, developer of the FCI, and his colleagues have published widely about its use as have others, most notably Richard Hake, who has collected and published summaries of thousands of FCI data points from schools in the USA.³

As collaborators in a recent partnership between the Colorado School of Mines (CSM) and the Abu Dhabi National Oil Company (ADNOC) to establish the Petroleum Institute (PI) in Abu Dhabi, United Arab Emirates, the authors of this paper have worked together to:

- 1) develop an Arab-language version of the FCI
- 2) administer it, along with a modified version of the standard 1995 English version, to students at the Petroleum Institute
- 3) compare the results from the English and Arabic versions at the Petroleum Institute
- 4) compare the results from students at the PI to results from students taking Physics I at the Colorado School of Mines, and
- 5) analyze all of the results.

This paper includes a preliminary analysis of our results as well as suggestions for further study.

The CSM-PI Partnership

At the request of the Abu Dhabi National Oil Company (ADNOC), the Colorado School of Mines has been retained to provide academic leadership required to design, start up, and operate the Petroleum Institute (PI), a world-class educational institution dedicated to educating engineers for the oil and gas industry. When it is fully operational the PI will offer undergraduate and graduate degrees in five programs:

- chemical engineering
- petroleum engineering

* Contact author: Barbara M. Olds, Associate Vice President for Academic Affairs, Colorado School of Mines, Golden, CO 80401, bolds@mines.edu, 303-273-3827.

- petroleum geosciences engineering
- mechanical engineering
- electrical engineering (power; instrumentation and controls)

A Foundation Program is also offered to help incoming students make the transition from their high school preparation to the entry requirements for the baccalaureate curricula, especially in the area of English proficiency since all instruction at the PI is in English. The Petroleum Institute began its inaugural year in September 2001 with an initial class in the Foundation year of 150 students gleaned from an applicant pool of over 800. The second year of the program started in September 2002 with a class of 150 students admitted into the Foundation Program.

The Petroleum Institute is a unique enterprise that represents collaboration between academia in the United States and the private sector in the Middle East. The academic programs at the Petroleum Institute involve intensive interaction with industry and four industrial partners (BP-Amoco, JODCO, Shell, and TotalFinalElf) who have teamed with ADNOC to provide advice to and support for the PI.

The Force Concept Inventory

The Force Concept Inventory (FCI) was designed by David Hestenes and colleagues to probe students' understanding of Newtonian mechanics, specifically the central concept of force. The Inventory¹ "requires a forced choice between Newtonian concepts and commonsense alternatives." The FCI builds on work by Halloun and Hestenes in developing an earlier version, the Mechanics Baseline Test (MBT).⁴ The FCI does not require calculations and is designed to probe "in a way that is understandable to the novice who has never taken physics course, while at the same time rigorous enough for the initiate."³ Hestenes et al. argue that the FCI can be used for a variety of purposes: as a diagnostic tool, for evaluating instruction, and as a placement exam.

Halloun, Hake, and Mosca have revised the original version of the FCI to improve clarity and question sequencing.⁵ Their version contains 30 questions as opposed to 29 in the original FCI. A thorough evaluation of the MBT and FCI tests was completed by Hake, who analyzed results from over 6,000 students who had completed pre- and post-test versions of either the FCI or the MBT.³ Hake addressed a number of concerns that have been raised about the validity and reliability of the FCI and concluded that students in physics courses which employ interactive engagement instructional methods demonstrated significantly higher improvement between the FCI pre- and post exams than students in traditional lecture courses.

Although there have been criticisms of the FCI, most notably by Heller and Huffman, even critics conclude that "the FCI, which has reasonable face and content validity, is the best test currently available."^{2, 6, 7}

Our Experiment including Preliminary Results and Discussion

We have used the FCI (original version developed by Hestenes and colleagues, 1992) as a

*Proceedings of the 2003 American Society for Engineering Education Annual Conference & Exposition
Copyright © 2003, American Society for Engineering Education*

diagnostic and assessment tool at the Colorado School of Mines for several years. For this research we are using results of a pretest taken by 393 CSM first-year students enrolled in Physics I in the fall of 2001. For our study at the Petroleum Institute, Professor Saleh Al Hashemi and Instructor Khaled Al Wahedi translated into Arabic 23 questions from the modified version of the FCI developed by Halloun, Hake, and Mosca (available at <http://media4.physics.indiana.edu/~hake/>). This 23-question modified FCI was administered as a pretest to 57 freshman students at the Petroleum Institute in the fall of 2002. Twenty-eight took the exam in Arabic and 29 in English. These students had each completed a year of introductory science at the PI as part of their Foundation year studies, but some were not yet enrolled in the introductory physics course. For comparison purposes, each student was randomly assigned to take either the Arabic or English version of the test. To begin assessing the impact of the Foundation year program on PI students, an additional 112 students who entered the Foundation year in the fall of 2002 took the test in Arabic.

As shown in Table I, 21 questions on the PI version of the FCI are common to the Hestenes 1992 and Halloun 1995 versions. Responses to these 21 questions have been used to compare the performance of CSM and PI students from several perspectives.

Comparison of PI Freshmen (English) and PI Freshmen (Arabic). Our first analysis is a comparison of responses of the PI freshmen who took the test in English vs. those who took it in Arabic. As shown in Figure 1, there appears to have been substantial language effect on the results, since a much higher percentage of students (82.1%) who took the test in Arabic answered at least 25 percent of the questions correctly compared to only 48.3% of those who took the test in English (this difference is statistically significant using a two-tailed z-test for proportions with $p=0.007$). Although we have not yet had an opportunity to perform in-depth analysis of this result, it provides us with important information for designing the introductory physics curriculum at the Petroleum Institute. Instructors will have to plan their instruction in English in a way that ensures that their Arabic-speaking students are given the time needed to process concepts in English.

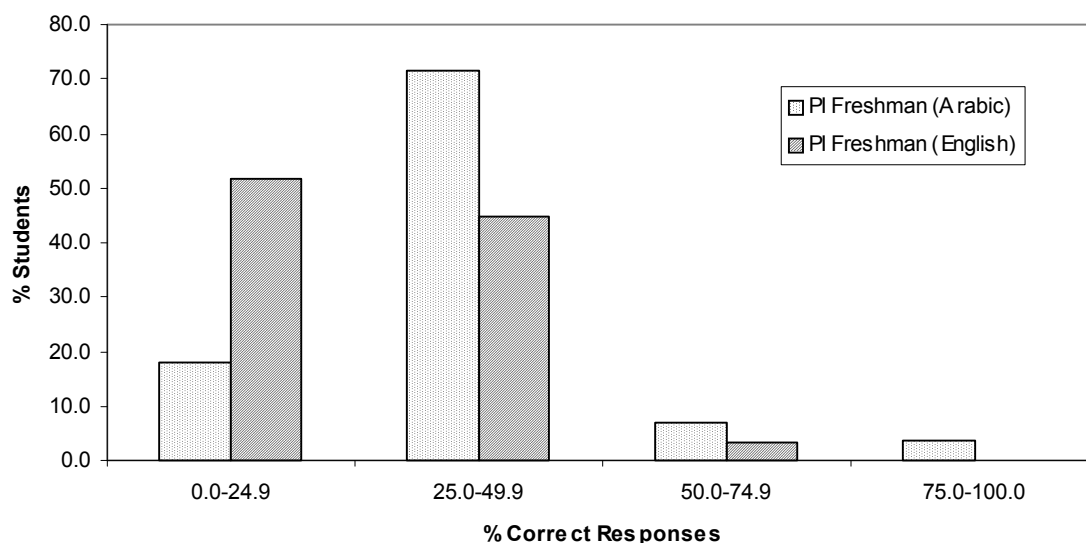


Figure 1 – Comparison of PI Freshmen Scores as a Function of Test Language

Table 1 – Comparison of Questions Included in the PI, Hakes, and Hestenes Versions of the Force Concept Inventory

PI version	Hakes version [Halloun et al, 1995]	Hestenes version [1992]
1	1	1
2	2	3
3	3	17
4	4	2
5	12	16
6	13	5
7	14	23
8	8	6
9	9	7
10	10	8
11	11	9
12	15	13
13	16	14
14	17	18
15	19	20
16	20	21
17	21	24
18	22	25
19	23	26
20	24	27
21	28	11
22	29	none
23	30	none

PI Freshmen Compared to PI Foundation Year Students. Our second objective was to compare PI freshmen student results with those of the Foundation year students. The freshman group had completed a year's worth of physical science in the Foundation year while the Foundation year students were entering the PI from high school with a broad range of preparation for college. Figure 2 shows our results for this comparison. To separate the effect of Foundation-year science instruction from the language effect, we compared only PI freshmen who took the FCI in Arabic with the Foundation year students, all of whom completed the assessment in Arabic. The data indicate that PI students who had completed the Foundation year answered more questions correctly than those who were just entering the Foundation year. For example, 82.1% of the freshmen answered at least 25% of the questions correctly while 56.3 % of the Foundation year students did. This difference is statistically significant using a two-tailed z-test

for proportion with $p = 0.012$. Since the freshman students have not completed their Physics I course and their overall performance is not yet known, the overall significance of these results is not apparent, but it is encouraging to see FCI score improvement as a result of completing the Foundation year curriculum.

Hestenes reports a mean pretest correct answer percentage of between 27 and 28 percent for high school students entering physics courses in the USA. The mean pretest correct answer percentages for the PI freshmen (Arabic version of FCI) and PI Foundation students were 35.4 and 27.6 percent, respectively. Overall, these results indicate that Arabic-speaking students entering the Petroleum Institute possess about the same level of Newtonian mechanics understanding as their peers in the USA and that additional physics instruction in the Foundation program improves conceptual understanding somewhat.

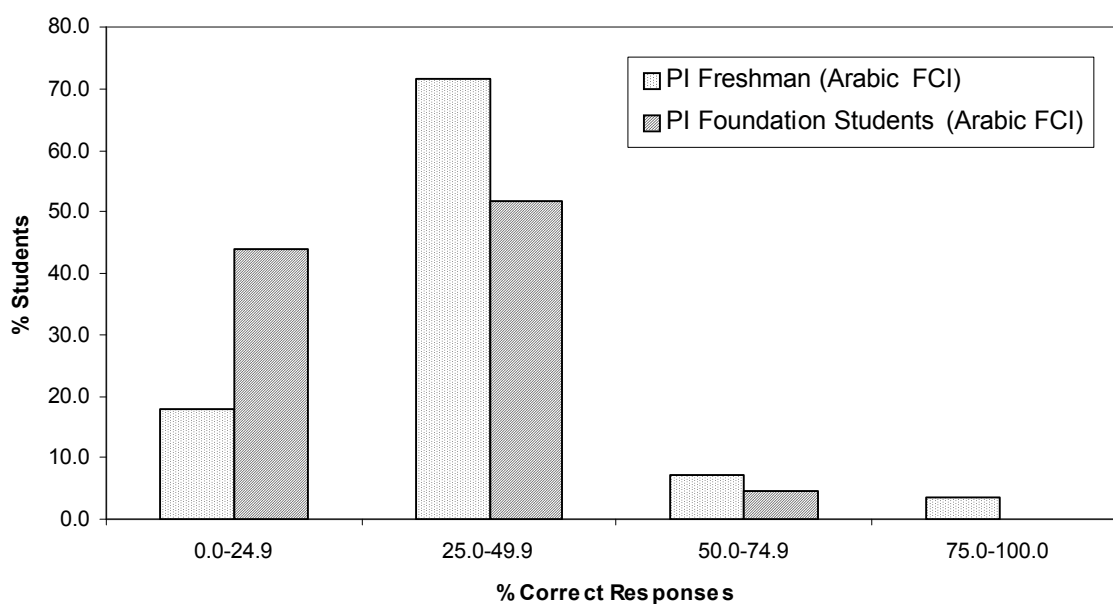


Figure 2 – Comparison of PI Freshman Scores vs. Foundation Student Scores

PI Freshmen Compared to CSM Freshmen in Physics I. We also compared scores for PI freshmen (both Arabic and English subgroups) with CSM freshmen at the beginning of their required Physics I course. As shown in Figure 3 and summarized in Table II, approximately the same number of PI freshmen (Arabic FCI) and CSM freshmen answered more than 25 percent of the FCI questions correctly (82.1% vs. 90.9%); the difference was not statistically different with $p = 0.132$. However, a large and statistically significant difference was noted when comparing the number of PI freshmen (Arabic FCI) and CSM freshmen who answered more than 50 percent of the questions correctly. This results suggests that the number of students who have sufficient conceptual understanding to be deemed Newtonian thinkers is larger in the CSM cohort and that additional work will be required in physics courses to increase scores among PI students. For example, additional class and study time will likely be required to help repair PI students misconceptions about basic Newtonian mechanics concepts.

All results for PI freshmen who completed the FCI in English are statistically lower than comparison scores for CSM freshmen. This result reinforces our earlier conclusion that FCI results for PI students contain a significant language bias and that this effect will need to be considered during curriculum development and assessment of student performance in PI science and engineering courses.

Table II – Statistical Comparison of Differences in FCI Performance Between PI and CSM Freshmen

	PI freshmen (Arabic FCI) (p ₁)	PI freshmen (English FCI) (p ₂)	CSM freshmen (p ₃)	Results of two-tailed z-test for proportions
% of students answering more than 25% of FCI questions correctly	82.1	48.3	90.9	p ₁ vs. p ₃ , p = 0.132 p ₂ vs. p ₃ , p=0.000
% of students answering more than 50% of FCI questions correctly	10.7	3.4	39.1	p ₁ vs. p ₃ , p=0.003 p ₂ vs. p ₃ , p=0.000

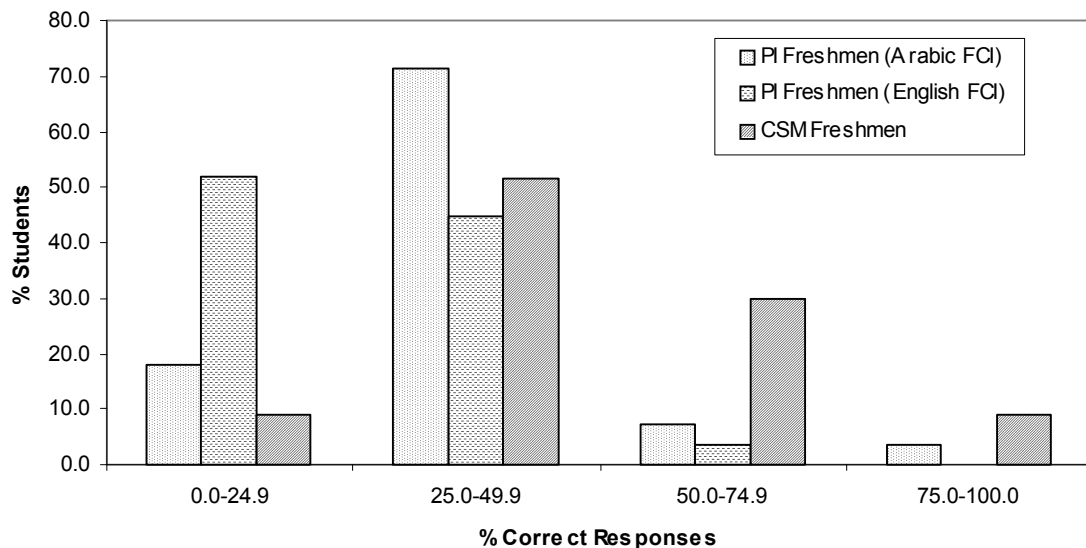


Figure 3 – Comparison of PI Freshmen (Arabic and English) vs. CSM Freshmen Scores

Comparison of Results from Individual Questions. Finally, we compared the results from individual questions for PI freshmen who took the FCI in Arabic and CSM students to see

whether there were significant differences in the responses of the two groups and, if there were, to analyze and discuss whether these differences might indicate cultural biases in the assessment. Figure 4 presents the results for the 21 common questions. The largest discrepancies appear to be on questions 1, 4, 7, 8, 9, 12, 19, and 20 (PI FCI numbering). Although our sample sizes are small and not necessarily representative of first-year students in either the USA or the UAE, speculation about these differences raises questions for further study.

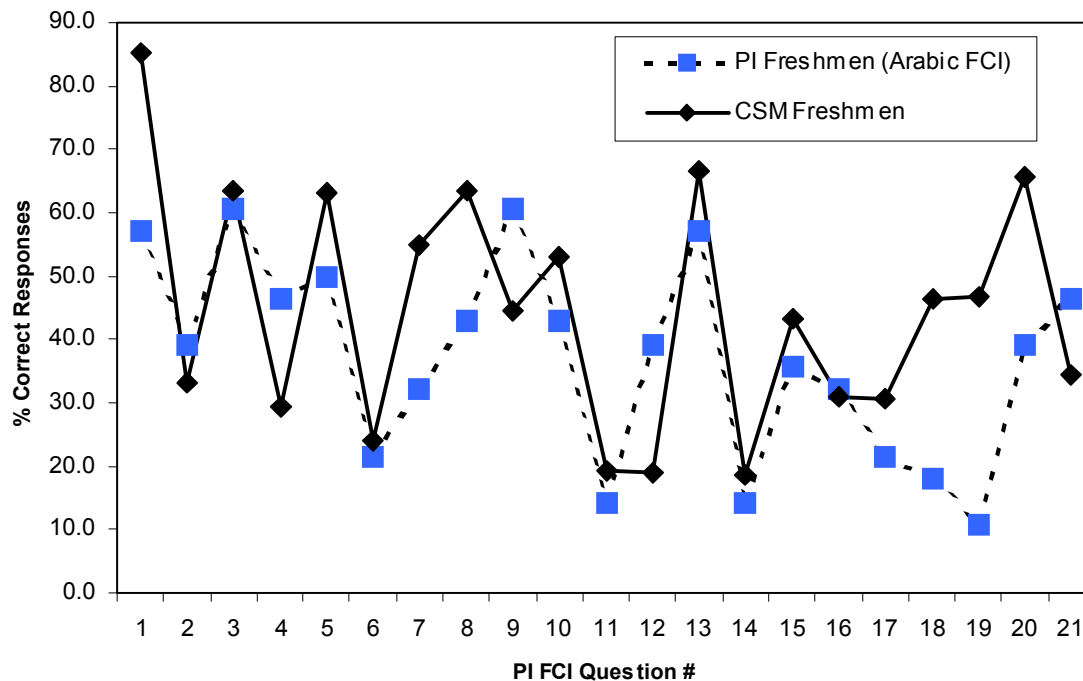


Figure 4 – PI Freshman (Arabic FCI) vs. CSM Freshmen Results by Question Number

The discrepancies between the CSM Freshmen and PI Students appear to be clustered around two concepts in the inventory--gravity and Newton's third law. Table III lists a selection of the questions probing these two concepts in the PI inventory. The table shows the distribution of Newtonian answers across the different versions (Arabic/English) together with the most frequent *alternative* choice for each question.

Table III – Answers to PI Version of FCI by PI Students Along with Most Common Alternative Choices

	"Newtonian" Answers			Most Common Alternative Choice		
	Freshman		Foundation	Freshman		Foundation
	English (n=29)	Arabic (n= 28)	Arabic (n=112)	English (n=29)	Arabic (n=28)	Arabic (n=112)
Q1	13	16	44	7 (a)	7 (a)	45 (a)
Q3	7	17	39	10 (b)	6 (a)	42 (b)
Q4	9	13	39	15 (a)	14 (a)	66 (a)
Q5	18	14	61	7 (c)	13 (c)	50 (c)
Q6	3	6	17	15 (b)	15 (c)	57 (c)
Q7	8	9	37	14 (a)	16 (a)	60 (a)
Q8	11	12	36	11 (b)	8 (a)	53 (a)
Q9	6	17	67	10 (c)	4 (c)	30 (c)
Q10	12	6	37	10 (d)	9 (d)	44 (d)
Q11	4	3	17	10(b) 10 (c)	10 (c)	48 (c)
Q12	9	11	24	12 (c)	11 (c)	64 (c)
Q13	10	13	57	8 (b)	9 (c)	43 (c)
Q14	3	4	9	8 (d)	16 (d)	52 (d)
Q17	10	6	34	8 (c)	8 (b)	42 (b)
Q18	7	5	26	11 (d)	12 (d)	58 (d)
Q19	5	3	13	10 (d)	9 (c) 9 (d)	49 (d)
Q20	13	11	51	9 (c)	7 (c)	33 (c)
Q21	10	13	29	13 (d)	12 (d)	56 (d)

The table shows that certain alternative choices (“misconceptions”) are extremely common among Foundation and Freshman students (in both Arabic and English versions). The student alternative choices to these inventory questions are possibly of greater benefit to physics instruction as they give a snapshot of the common sense notions of the incoming students. A brief analysis of student ideas relating to gravitation and Newton’s third law is given in the following paragraphs.

Gravitation

Questions 1, 3, 5, 6, 7, 11, 14 and 18 probe the concept of gravitation (acceleration independent of weight, parabolic trajectories). The most popular student alternatives to these questions indicate a prevalent view among PI students that heavier objects fall faster. Newton’s second law in relation to gravity (constant force produces constant acceleration) also appears to be poorly understood.

Newton’s Third Law

Questions 4, 12, 13, 17 and 21 probe Newton’s third law in relation to impulsive forces and continuous forces. The most common alternative responses would suggest that students equate greater mass with greater force. Other popular alternative choices indicate misconceptions in

relation to “action and reaction pairs”: (a) the most active agent produces the greater force, (b) greater mass implies greater force, (c) largest force determines motion.¹

Linguistic and Cultural Influences

The differences between CSM students and PI students on the inventory, and the differences between the two versions administered to PI students raise many questions about possible linguistic and/or cultural influences on students’ ideas about force.

The fact that PI students did better on the Arabic version of the inventory would primarily seem to suggest a mismatch between the language of instruction and the language of assessment. It would also seem to suggest a greater level of *scientific fluency* in the first language. The language used in the Arabic version of the inventory was similar to that used in schools and textbooks in the UAE at secondary school level. Students were therefore familiar with and comfortable with the Arabic scientific lexicon used in the inventory as opposed to the English lexicon in the English version. Even after a semester of science taught in English in the Foundation program, it would appear that students, although scientifically literate (i.e. they could recognize scientific words), had not reached a sufficient level of *fluency* in the second language to pick up the subtle differences in language in the distracters, some of which had quite complex syntax. This has a direct consequence for physics curriculum development, instruction and assessment at the PI. The ability to recognize the various exemplars of a given concept does require a certain level of fluency and a high degree of *interaction* with the language of science.

The subtle differences between the denotative and connotative meanings of scientific words in English and Arabic may also be contributing to the discrepancy on individual questions in the two versions. The Arabic language is based on a series of root letters where concepts, words and ideas are formed using various combinations and modifications of the root letters. The most commonly used Arabic word for “force” for example is “goowah.” The word has connotations of “strength” and “power” and is in fact used interchangeably to refer to “power” (e.g. the power of a storm). Intuitively, a student could mistakenly associate greater strength (power) with a greater force even though from a physics education point of view the terms power and force have completely different meanings.

Our results suggest that a rigorous study of possible language and cultural bias in using the FCI with native Arabic speakers would be worth undertaking. As in introduction to this inquiry, we will briefly discuss two questions.

Question 4 (PI FCI numbering) reads:

"A large truck collides head-on with a small compact car. During the collision:

- (A) the truck exerts a greater amount of force on the car than the car exerts on the truck.*
- (B) The car exerts a greater amount of force on the truck than the truck exerts on the car.*
- (C) Neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.*
- (D) The truck exerts a force on the car but the car does not exert a force on the truck.*

(E) The truck exerts the same amount of force on the car as the car exerts on the truck.

In this case, 39 percent of the PI students identified the correct answer, E, compared to 29.4 percent of the CSM freshmen. The overwhelming choice among CSM freshmen was A; it was also chosen by half of the PI students. Hestenes suggests that this question assesses students' understanding of Newton's Third Law for impulsive forces. In contrast, an incorrect response of (A) may indicate that a student has the mistaken concept that greater mass implies greater force.

Note that in the English version the distracters refer to the "amount of force" and "force," while the Arabic version simply uses the word "goowah" to refer to both. In this case Arabic is more economical: a single word can conjure up several concepts and sub-concepts. In the English version, the term "amount" (of force) may be quite distracting and imply a quantitative reply.

Question 1 (PI FCI numbering) reads:

"Two metal balls are the same size but one weighs twice as much as the other. The balls are dropped from the roof of a single story building at the same instant of time. The time it takes the balls to reach the ground below will be:

- (A) About half as long for the heavier ball as for the lighter one.*
- (B) About half as long for the lighter ball as for the heavier one.*
- (C) About the same for both balls.*
- (D) Considerably less for the heavier ball, but not necessarily half as long.*
- (E) Considerably less for the lighter ball, but not necessarily half as long.*

Eighty-five percent of the CSM students and 51% of the PI freshmen identified the correct answer (C). In both cases the most common incorrect answer was A, followed by D and then B. Hestenes suggests that a correct answer to this question indicates some level of understanding of the gravitation concept that acceleration is independent of weight.

Again, the Arabic language may have a subtle influence on the concept in question and contribute to the somewhat lower scores among the PI students. The word for gravity in Arabic, "jathibiyah," has unusual connotations. It implies a force of "attraction" or a "pull." The same word is often used interchangeably to refer to magnetism although magnetism is also translated as "goowah magnetisiah" or "magnetic power." Further research is needed to determine if there is a basic ontological difference between the concepts of gravity and magnetism in the minds of the students. This has implications for teaching, as instructors need to explore the connotative and denotative aspect of English words used to describe scientific concepts.

The impact of culture on the results is more difficult to assess. UAE students grow up in a society where traditional Islamic values permeate all aspects of life in a country with a rapidly developing modern economy. The country came into existence in 1971 and its GDP per capita is one of the highest in the world. The phenomenal pace of growth has led to a massive increase in population and a burgeoning education system. Female local teachers comprise the majority of teachers at the primary stages of education in the government sector. There are private secondary schools, some of which use western teaching methods, offering International Baccalaureate courses to

expatriate and local students. Government Secondary Schools tend to be staffed mainly by expatriate teachers from other Arab countries. Competition for third level places is keen and in recent years several private third level colleges have opened in the country offering western style curricula. While the scope of the secondary school curriculum is wide, it is nevertheless competitive and results-driven with little focus on higher order thinking skills such as synthesis, conjecture and evaluation. Consequently, rote learning and memorization is encouraged, mirroring traditional practices in the teaching of Islamic studies. This is in many ways in juxtaposition to the Islamic view of education which encourages “tafakir” or thinking and search for personal meaning (from the root word “fikrah” meaning “idea” or “thought”).

When the Force Concept Inventory was administered at the PI, it was one of the first “exams” taken by students where they were asked for their ideas about science. The whole procedure was quite novel and a break from the “right answer” syndrome of traditional school examinations that tend to focus on the applications of rules and the use of complicated algorithms. Although secondary school students are exposed to the basic Newtonian concepts as part of their Physics curriculum, instruction tends to be didactic and expository with very little time devoted to practical experimentation and the exploration of concepts on a deeper level. The paucity of practical courses and experimentation in the secondary curriculum reflects cultural expectations where administrative and managerial type jobs are more highly rated compared to vocational type work. Indeed much “practical” work, from gardening to maintenance of buildings and oilfields, is carried out by workers from the Indian subcontinent and other Arab countries. The higher percentage scores of the CSM freshmen on the FCI could be due to the more experiential and laboratory based background of US students and a more Socratic enquiry-centered curriculum.

Conclusions

Although it is premature to draw definitive conclusions from this study, we believe that our results raise important questions that should provide interesting avenues for further research. Among these questions are the following:

- To what extent does language affect the results of the assessment, even for students who have an acceptable (500+) TOEFL score?
- To what extent do cultural differences affect the outcomes of an inventory like the FCI, even when it is taken in the student’s native language?
- Are there other prevalent “misconceptions” held by students which the FCI does not measure that may be culturally significant?
- Do what extent do the learning styles of Emirati students entering the Freshman year at the PI differ from the learning styles of CSM Freshmen?
- To what extent do practical laboratory sessions aid in the amelioration of misconceptions in mechanics?

References

1. Hestenes, D., Wells, M., and Swackhamer, G., "Force Concept Inventory." *The Physics Teacher*, **30**, 414-158, 1992.
2. Hestenes, D. and Halloun, I., "Interpreting the FCI: A Response," *The Physics Teacher*, **33**, 502-506 (1995).
3. Hake, Richard R., "Interactive-engagement Versus Traditional Methods: A Six-Thousand Student Survey of Mechanics Test Data for Introductory Physics Courses," *American Journal of Physics*, **66** (1), January 1998, 64-74.
4. Halloun, I. and Hestenes, D., "The Initial Knowledge State of College Physics Students," *American Journal of Physics*, **53** (1043-1055 (1985).
5. Halloun, I, Hake, R., and Mosca, E., "Force Concept Inventory," revised August 1995. Available with permission at <http://psheldon.rmwc.edu/fci/fcipayes/fci.htm>
6. Heller, P. and Huffman, D., Interpreting the FCI: A Reply," *The Physics Teacher*, **33**, 503-511 (1995).
7. Huffman, D. and Heller, P., "What Does the Force Concept Inventory Actually Measure?" *Physics Teacher*, **33**, 138-143 (1995).

Biographical Information

COLM MEALY is lecturer in mathematics and one of the founding faculty members of the Petroleum Institute in Abu Dhabi. He has taught mathematics at High School and University Level for over 20 years, and has worked as an educational consultant to several multi-media companies in the area of e-learning. He served two tours of duty with UNOHCI in Baghdad as Education Observer reporting on the state of the Iraqi Education system. His research interests include learning styles and misconceptions in science and mathematics among high school students.

SALEH AL HASHEMI is an Assistant Professor in the Chemical Engineering Department at the Petroleum Institute in Abu Dhabi. He received his PhD from Tufts University, Medford, MA in May 2000 and subsequently worked in industry as a process engineer at Ruwais Refinery, Abu Dhabi. He joined the Petroleum Institute in September 2001. He is a recipient of the "Sheik Rashid Prize for Scientific Excellence."

RONALD L. MILLER is Professor of Chemical Engineering and Petroleum Refining at the Colorado School of Mines where he has taught chemical engineering and interdisciplinary courses and conducted research in educational methods for over seventeen years. He has received three university-wide teaching awards and has held a Jenni teaching fellowship at CSM. His paper entitled "Connections: A Longitudinal Study of an Integrated Freshman Program" (co-authored with Barbara Olds) won the award for best paper in the Educational and Research Methods Division of ASEE during the 2001 annual conference. He has received grant awards for educational research from the National Science Foundation, the U.S. Department of Education, the National Endowment for the Humanities, and the Colorado Commission on Higher Education.

BARBARA M. OLDS is Director of the Division of Research, Evaluation and Communication in the Education and Human Resources Directorate at NSF. She is a Professor of Liberal Arts and International Studies and former Associate Vice President for Academic Affairs at the Colorado School of Mines. She was awarded a Fulbright fellowship to teach and conduct research in Sweden during the 1998-99 academic year.