Physics education research (PER) has taught us much about how our students learn (or fail to learn) physics. Instructional strategies and curricula based on the findings of PER have been developed and implemented with great success. Given all that is known about how students learn in general, it is not surprising that the most successful reform strategies require significant interactions among students, and between students and instructor. It is therefore reasonable to wonder about the extent to which PER-based reform is possible in an environment where students speak many languages and come from many cultures. In this article we give a brief overview of PER and describe the implementation of highly interactive PER-based instructional strategies at City College of New York, a very diverse inner-city school.

Some Lessons from Physics Education Research

Physics education research involves the study of how students learn physics. The research is interpreted in terms of fundamental principles of how knowledge is developed. While experienced instructors are familiar with many of the findings, PER has detailed underlying difficulties that students have and the strategies that they use. PER has also provided specific strategies that can improve and evaluate instruction.

PER has informed us of persistent difficulties that students have with physics content. For example, student difficulties relating force and motion are well known. All physics instructors explain Newton’s second law and can easily get students to repeat \( F_{\text{net}} = ma \). However, away from the exact context in which the material was taught, many students will still answer...
Kofi Donnelly is a physics major at City College of New York who has worked on physics education for the past year and a half. He is currently applying to Ph.D. programs in physics.

City College of New York
135th St. at Convent Ave.
New York, NY 10031
kofi_donnelly@yahoo.com

Kofi Donnelly
tries and/or learned English as a second language (ESL). No single culture makes up a majority of the class. Students at CCNY are native to more than 90 countries representing Africa, Asia, Europe, Latin America, and many parts of New York City. About two-thirds of the students took high school physics, either in New York or in their native countries.

The instructional strategies and curricula described previously require significant interaction. In *Tutorials*, students work in groups of three or four, explaining what they are doing and interpreting and responding to others’ arguments. Students solve problems that require figuring out what the relevant physics is before engaging in mathematical manipulations. Lectures are interrupted by difficult conceptual questions where responses and justifications are shared. Given the premium on language and interpretation of context, can these interactive strategies be successful in a multicultural environment like CCNY?

The structure and content of the introductory calculus-based physics course here is fairly standard. For each of the two classes described in this article, students met three times per week in a lecture hall and twice per week in smaller groups ($N < 25$) for recitation. In addition, students met every other week for a standard laboratory ($N < 25$) and worked through traditional physics problems in workshops on the opposite weeks. During a recent semester, several PER-based instructional strategies were implemented in one lecture class, while a traditional lecture class was taught during the same semester. In the “reform” class, 84 students took the first midterm and 63 students took the final. In the traditional class, 74 students took the first midterm and 50 took the final. The difference between the numbers of students taking the midterm and the final were consistent with typical drop rates for this course, but the drop rate was less for the reform class than it was for the traditional class.

While the textbook, content, homework, and labs were essentially the same in the two classes, two significant differences were implemented in the reform class (reflecting a small fraction of the total time that students invest in the course). During one of the two recitations each week, students worked through the *Tutorials*. Also, during the lectures of the reform class students participated in interactive demonstrations where they had to make predictions and discuss them with a neighbor. These demonstrations accounted for about 20% of the lecture time in the reform class. Since no additional time was allotted for *Tutorials* and interactive lectures, students in the reform class had fewer opportunities to review standard homework problems in recitation and to see all of the demonstrations and derivations in lecture.

It is clear to us that differences in the two classes were primarily due to the *Tutorials* and interactive lecture demonstrations. Our results are very similar to numerous other institutions incorporating similar strategies.1,2,6,11 Also, it is worth noting that the lecturer in the traditional class was highly skilled and widely respected, and that at the time of registration students were unaware that reform strategies were to be implemented in one of the lecture classes.

**Results**

Students were active and on task in *Tutorials* and interactive demonstrations, which is comparable to what has been observed at other institutions using similar materials.15 (Fig. 1 shows students at work during tutorial.) Attendance was typically over 75%, and many would spontaneously remain after the class ended to continue working through the materials. The class was noisy and active, with the vast majority of students engaged in the material. Students demonstrated the same conceptual and reasoning difficulties that have been reported in the literature, and shortly into the semester, many students routinely insisted on a sound understanding of the basic ideas of physics, not just what the answers were. They questioned and challenged each other and the instructor.
Groups that the students formed were self-selected; all of the students spoke English in class and most of the groups were multicultural. While the majority of the groups were functional throughout the semester, a few were quiet with much less interaction among the members.

**Force Concept Inventory**

One frequently used tool in probing the development of student understanding of basic mechanics concepts is the Force Concept Inventory (FCI). The FCI asks about ideas that most instructors would hope their students understand after studying mechanics. The questions are worded in common language. Students are asked to describe the path of a sliding puck after being kicked or to identify the forces on a golf ball in flight. There are 29 questions in multiple-choice format. The questions and the distractors (incorrect choices) are both based on detailed investigations of student understanding of mechanics.

Tens of thousands of students from a great variety of classes have taken the FCI before and after instruction. To many instructors, student improvement after studying physics is surprisingly poor. A common measure of the improvement of a class from before to after instruction is the fraction of the possible gain:

\[
h = \frac{\text{post}\% - \text{pre}\%}{100 - \text{pre}\%}.
\]

Post\% and pre\% are the class FCI averages before and after instruction, so \( h \) represents the fractional improvement of the class as a whole. Across a huge spectrum of populations, student aptitudes, and instructional strategies, \( h \) is somewhat predictable. Hake has reported that for traditional instruction \( h = 0.23 \pm 0.04 \) and that for interactive classroom strategies, \( h = 0.48 \pm 0.14 \). These results correspond to a wide range of class initial scores on the FCI. At CCNY, the FCI was given before and after instruction in the two classes described. It was given without warning in recitation and did not count towards the student grades. Results were comparable to the numbers reported by Hake. For the traditional class, the average score before instruction was 40\% and after instruction was 53\% corresponding to \( h = 0.23 \). For the reform class, the average score went from 39\% to 65\% corresponding to \( h = 0.43 \). Only

---

**Fig. 2. FCI scores before and after instruction. The fraction of the possible gain for the traditional class is 0.23 and 0.43 for the reform class (0.46 for native English speakers and 0.42 for ESL students).**

---

A stone is thrown upward and hits the ground 72 m below the starting point. The total time between leaving the thrower’s hand and hitting the ground is 6 seconds.

a. With what speed was the stone thrown?

b. At some later time, the stone is moving down with a speed 12 m/s.
   i. How long a time is this after the stone was thrown?
   ii. What is the position of the stone at this time?

---

A stone is thrown upward and hits the ground 72 m below the starting point. The total time between leaving the thrower’s hand and hitting the ground is 6 seconds.

a. With what speed was the stone thrown?

b. At some later time, the stone is moving down with a speed 12 m/s.
   i. How long a time is this after the stone was thrown?
   ii. What is the position of the stone at this time?

---

--

**Fig. 3. Exam question given to both the traditional class and the reform class.**
the students who took the FCI both before and after instruction were included in the results. In the traditional class, there were 38 such students (12 native English speakers and 26 ESL students). In the reform class, there were 45 (16 native English speakers and 29 ESL students).16

Given the highly semantic components of both the modified instruction and the FCI, we were curious to see if the scores on the FCI were at all dependent on whether the students were native English speakers. Figure 2 shows the FCI scores before and after instruction in both the traditional and reform class. Scores for the entire class, for native English speakers, and for ESL students are included. In both classes, $b$ was only slightly higher for native English speakers than it was for ESL students, 0.26 versus 0.21 for the traditional class and 0.46 versus 0.42 for the reform class. These results suggest that native language is not a significant factor in the success of the PER-based reform.

Examinations

Students in the two classes were given four common examination problems, one on each of the first two midterms and two on the final. These problems were typical of the problems usually given in the physics department and for the most part did not address the types of changes that were implemented in the reform class. Other than the last part of the first midterm problem, the problems were mostly produced by the traditional instructor and not seen by the other instructor until after instruction.

On the first midterm, students in both classes were given the problem shown in Fig. 3. The two classes had finished instruction in one-dimensional kinematics, devoting about the same amount of time to the topic. The exam problem was initially written with just the quantitative parts a and b. The qualitative part c was added just before the exam was administered. Figure 4 shows the results of how students performed on the different parts of the exam. The quantitative score is the average percentage of correct answers for parts a and b. The qualitative score is the percentage of students who answered the entire part c correctly.

Even though the students in the reform class spent less time on quantitative problem solving during instruction, a larger portion of the students got correct answers to parts a and b. Apparently spending the extra time helping students to understand the underlying ideas of kinematics helped them succeed on this type of problem solving. (This result is consistent with what is known in general about how problem solving is learned.)3,4 The difference on the qualitative part of the exam is even more significant. Most of the students from the reform class were able to answer this correctly whereas most of the students in the traditional class were not.

Figure 4 also shows the breakdown of native English speakers versus ESL students. In the traditional class, there were 27 native English speakers, 30 ESL students, and 16 unknown. In the reform class the numbers were 42, 25, and 17. As seen in Fig. 4, it is clear that both the native English speakers and the ESL students met with better success in the reform class on both parts of the exam question.

Students in the reform class also had higher averages of correct responses on two of the other three exam problems that were common to the two classes. The problems covered a modified Atwood’s machine, racing cars, and conservation of energy. The averages in the reform class were 44%, 65%, and 37% respectively. In the traditional class, the averages were 66%, 41%, and 29%. Overall, there were not significant or consistent differences in performance between the native English speakers and the ESL students in either class.
**Student Attitudes**

In the reform class, there could be a concern that students would not appreciate the *Tutorials* and the interactive lectures because so much is demanded in terms of thinking during class and justifying answers. This could be a particular concern since quality of instruction is often judged by student opinion, even though the correlation of this measure and the real success of a course is dubious.

Throughout the semester, we were encouraged that attendance and participation in the reform class were extremely high, that morale was very positive, and that many students continued to work with each other outside of class hours. In addition, we gave a survey at the end of the semester. On a five-point scale, students were asked to evaluate how much the various components of the course helped them learn physics with one being “not at all helpful” and five being “extremely helpful.” The same survey was also given in the traditional class. Table I shows the average student responses for various components of the course in both classes. It is very encouraging that the results were extremely positive for *Tutorials* and the interactive lecture demonstrations. Not only were students succeeding at a higher rate in the reform class, it seems that they appreciated the value of the strategies used.

**Conclusions**

Even though this study was conducted with a single class of students within the specific context of CCNY, the results are very encouraging. Students who completed a class with PER-based instructional materials did better on conceptual questions and quantitative problems. They also reported favorable opinions of the reformed parts of the course and dropped out at a lower rate. Furthermore, there was success and participation for both native English speakers and students for whom English is a second language. While increased instructor enthusiasm and time associated with introducing a new teaching technique can improve student scores, the extreme similarities of the results reported here with those reported repeatedly elsewhere suggest that the improvements are robust.

### Table I. Results of a student survey about various components of the course.

<table>
<thead>
<tr>
<th>Survey Item</th>
<th>Traditional course</th>
<th>Reform course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lecture demonstrations</td>
<td>4.0</td>
<td>4.7</td>
</tr>
<tr>
<td>Laboratories</td>
<td>3.3</td>
<td>3.2</td>
</tr>
<tr>
<td>Recitations</td>
<td>4.3</td>
<td>4.0</td>
</tr>
<tr>
<td>Textbook</td>
<td>3.5</td>
<td>3.8</td>
</tr>
<tr>
<td>Tutorials</td>
<td>N/A</td>
<td>4.5</td>
</tr>
</tbody>
</table>

**Acknowledgments**

This work was supported by a Professional Staff Congress–City University of New York research award. We thank Joel Gersten, Fred Smith, and Martin Tiersten for their help with many phases of the work. We are also appreciative of Lillian McDermott, Joe Redish, Seth Rosenberg, and three anonymous reviewers for much constructive feedback on the paper.

**References**


15. One of us (RNS) has had extensive experience using *Tutorials* at both the University of Washington and the University of Maryland.

16. Although not the focus of this article, it is worth mentioning that asking students about race can have a negative impact on test scores of minority students. See C. Steele, “Thin ice: Stereotype threat and black college students,” *Atlantic Monthly* 284(4), 44, (1999).