Clickers in the classroom:

Implementing the Harvard *Peer Instruction* approach in Cegep

Peer Instruction (PI) is a student-centered, information technology driven, instructional approach developed at Harvard by physicist Eric Mazur (1997). In *PI*, students use wireless handheld devices -colloquially called 'clickers'- to provide real-time feedback to the instructor. This feedback is then used to shape the instruction. The method has been warmly welcomed by the science community and adopted by a large number of American colleges and universities[1], due among other reasons to its common sense approach and its documented effectiveness (Fagen *et al*, 2002; Crouch & Mazur, 2001, Mazur, 1997).

The purpose of this paper is fourfold. The first purpose is to present the *PI* approach to Cegep instructors. The second purpose is to determine whether the *PI* approach can be implemented at the Cegep level. Indeed, although this method has been effectively used for 15 years in American colleges and universities, this is the first study documenting its applicability and effectiveness in Quebec Cegep institutions[2]. The third purpose is to find out whether *PI* is more effective than traditional instruction in Cegep. The final purpose is to isolate the specific contribution of wireless 'clickers' to student learning. This paper will address these issues as concisely as possible. A full description of the PAREA study and its results can be found online through the Centre de Documentation Collegiale website (www.cdc.qc.ca).

What is PI?

Brief History

As recounted in his book on *PI*, Eric Mazur (1997) developed the approach when in the early 1990s he became aware of a non-numerical, conceptual inventory of Newtonian physics concepts called the Force Concept Inventory (Halloun *et al.*,1995; Hestenes, Wells & Swackhamer, 1992; Halloun & Hestenes, 1985). The authors of the FCI devised the test to quantitatively gauge the

[1] Fagen *et al* (2002) reports survey data of 384 instructors –outside of Harvard- having used Peer Instruction. Note that of these only 6% were 2-year colleges that would bear some resemblance to Cegeps.
[2] Searches of ERIC and Google Scholar yield not entries for 'Peer Instruction' and 'Quebec' or 'cegep'

extent of students' preconceived –often "Aristotelian" (DiSessa, 1982)- views of the world, despite formal physics training. The FCI, a multiple choice instrument, is unique in that it asks conceptual physics questions in simple terms and proposes distractors[3] that are compiled from the most prevalent misconceptions given by students in interviews (Halloun & Hestenes, 1985a,b). Thus, to answer FCI questions, students do not resort to computations or memorized algorithms but have to identify the accurate concept from a number of "d*istractors*". In putting forward these misconceptions, the FCI reemphasizes that physics is often counter-intuitive and that students enter physics classrooms not as blank slates but rather with many pre-conceptions. To experts, the correct answers to these questions are straightforward, at times bordering triviality.

Mazur decided to give the test to his students and the end of the semester. He presented it to students and downplayed its importance, worried that students would scoff at such a basic test. Yet, his students were uneasy with the test as best exemplified by one who asked:

> "*Professor Mazur, how should I answer these questions? According to what* **you taught us** *or according to the way* **I think** *about these things?*"

In fact, to Mazur's great surprise, not only did the students not grasp the fundamental concepts after 1 or 2 years of seemingly successful high school physics training (which after all got them into Harvard…) but a large number of misconceptions remained even after a semester of *his* instruction! Even some of the high performing students did not fully grasp the basic concepts (Mazur, 1997). In fact, this turns out to be one of the most revealing finding of large scale FCI data studies. Indeed, a meta-analysis of more than 6500 respondents (Hake, 1998) has shown that a semester of traditional instruction changes only marginally students' conceptual understanding[4] of physics. Furthermore, this gap between what instructors think their students understand and what the FCI shows has made the FCI "*the most widely used and thoroughly tested assessment instrument*" in physics (McDermott & Redish, 1999) and has rendered the FCI into the central role in driving the physics reform efforts of the past decade (Crouch *et al*, 2001).

[3] "Distractors" are defined as incorrect choices of the FCI which were compiled from most prevalent wrong answers given by students in interviews (Halloun & Hestenes, 1985a).

[4] Data suggests that traditional instruction yields "normalized gains" <g> of approx 20%. This implies that 80% of missing basic concepts on entry are still not acquired after a semester of traditional instruction.

Note that <g> is defined as: <g> = (Post-test score% - Pre-test score%)/(100%- Pre-test score%)
Which is the ratio of the actual gain (<%post> - <%pre>) by the maximum possible gain (100% - <%pre>).

Description of the Method

Mazur developed *PI* to explicitly address his students' misunderstanding of basic concepts. This required making some modifications to the instruction format. Students are presented with a brief lecture (7-10 minutes, within limits of average adult attention span), the content of these lectures being similar to traditional curriculum differing only by an increased emphasis on concepts. After the brief lecture, students are presented with a *ConcepTest*: a multiple choice conceptual question having misconceptions available as possible answers. To gauge what all students are thinking, each student was initially given five flashcards each with a letter (A,B,C,D,E) corresponding to the five available choice of answers. When presented with a ConcepTest, students would raise the flashcard corresponding to their preferred choice.

Figure 1.1 Students involved in *PI* using Flashcards.
Reproduced with consent from author from: *PI: A user's manual*

This provided the instructor with real-time feedback of the approximate proportion of correct answers as well as the distribution of misconceptions. To get a more accurate picture of the distribution of answers, Mazur later replaced flashcards with "clickers", that is, one-way infra-red wireless keypad devices which bear resemblance to TV remote controls.

Figure 1.2 An infrared 'clicker'

To state their choice of answer when presented with a conceptual question, students simply press the corresponding choice number on the clicker and the data is transmitted to the instructor's computer. Clickers also allow students to state their level of confidence (Hi, 0, Lo) for each question answered. The instructor then has *instant feedback* on how the students in his classroom have grasped the concept by assessing in real time the exact percentage of the class having the correct answer as well as the percentage of students holding each misconception.

The ability to assess student comprehension in real-time allows instructors to decide on the spot whether to build on newly acquired knowledge or if more time is required to consolidate previously presented concepts. Indeed, if the concept is poorly understood (< 35% of correct answers on ConcepTest), the instructor will *revisit the concept* and explain further before resubmitting the ConcepTest to the group. However, if the correct response rate is very high (>85%), students have well understood the concept, and the instructor may simply address the remaining misconceptions that 15% of the class believes before proceeding to the next concept. Most frequently, the rates of correct response are neither very high nor very low. When moderate response rates (35%-85%) are obtained, students are asked to turn to their neighbour and try to convince them of their choice. This leads to 2-3min of discussion between students: the Peer Intruction *per se*.

Fig 1.3 Students involved in a *PI* discussion.

Reproduced with consent from author from: *PI: A user's manual*

This discussion forces students to formulate their thoughts clearly and better represent the concept. Furthermore, a discussion of concepts between students withdraws the authoritative nature that a discussion with an expert instructor can have. Indeed, students may take an instructors' explanation as 'fact' and not pursue a line of reasoning as elaborate as would be done with a peer. Beyond having a more evenly balanced debate of conceptions, students also discuss from perspectives that are often foreign to the expert-instructor. Thus, students may be better equipped than instructors at understanding their peers' misconceptions and conceptual change may thus be facilitated. After discussion, students are presented with the same ConcepTest and are asked to revote. The instructor then acknowledges the correct response and explains why the remaining misconceptions are wrong. The method can thus be schematized as follows:

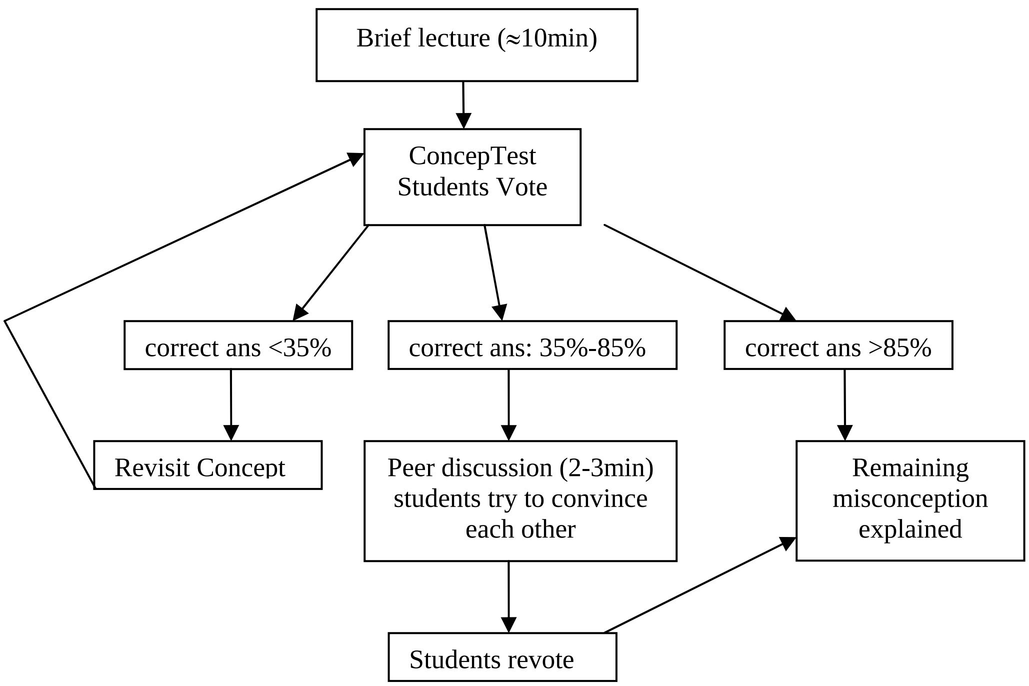


Figure 1.4 *PI* Algorithm

Replicated findings in American colleges and universities show not only that after the discussion between peers, rates of right answers increase significantly[5] but that the acknowledged levels of confidence for the correct answer also increase (Fagen *et al*, 2002; Crouch & Mazur, 2001; Mazur, 1997).

[5] Harvard, 10-year data shows rates of wrong-to-right answers of 32% compared to right-to-wrong rates of 6%, with overall 72% correct answers on the second vote and significant confidence level increases (Crouch & Mazur, 2001). Data of a large number (384) of non-Harvard users (Fagen *et al*,2002) indicates that significant conceptual knowledge gains (normalized gain = 0.39) occur with Peer Instruction.

Empirical Research Questions

This study focused on the following three empirical research questions.

1) Can the Harvard *PI* approach be implemented in a Cegep context?
    a. Does the approach fit within institutional constraints?
    b. What modifications to course structures must be made?
    c. Are the required modifications easily feasible?
    d. How is the approach received by other instructors?
    e. How is the approach received by students?

2) Is *PI* more effective than traditional didactic lecturing approaches?
    a. Does *PI* ***increase*** conceptual change?
    b. Does *PI* ***reduce*** traditional problem solving abilities?
    c. Does *PI* work better for students of ***higher proficiency***?

3) Is *PI* with clickers more effective than with flashcards?
    a. Does the use of clickers ***increase*** conceptual change?
    b. Does using clickers affect students' traditional ***problem solving abilities***?

Study Description and Experimental Design

To address the first empirical question, the reception of *PI* by administrators, colleagues and students will be described. Also described is the feasibility of the required modifications to implement the approach.

To address the second and third empirical questions, the following quasi-experimental study design was used. Students were pseudo-randomly assigned by the registrar to one of three groups consisting of two *PI* treatment conditions and one control section. Of the two *PI* groups, one used clickers (n=41) while the other used flashcards (n=42) to respond to in-class ConcepTests. Both *PI* groups were taught by the primary investigator. The third group consisted of a control section (n=38) where students were taught through traditional lecturing. The instructor for the control group was chosen as a match to the primary investigator by gender (M), age (+/-3yrs), teaching experience (+/- 1yr) and was anecdotally reported by students to be of similar teaching style.

To isolate the contribution of the technology to the approach, the *PI group* with clickers was compared to the *PI* group with flashcards. To compare the effectiveness of *PI* with respect to traditional didactic lecturing, both *PI* groups were pooled and compared to the control section. Comparison measures are presented below in the '*Instruments*' section.

Instruments

Three different quantitative variables were assessed in this study: quantitative problem solving skills (Exam), conceptual learning (FCI) and concept-confidence (Conf).

Exam

Physics ability is traditionally measured through quantitative problem solving. These skills were assessed using the John Abbott physics department's comprehensive final examination. This exam was constructed by a committee of physics professors (none of which were involved with this study) and had to be approved unanimously by all those teaching the course (10-12 instructors). Each instructor marked a single exam question for the entire cohort (not just for his or her students). This insured that no group had an exam of a differing difficulty, or a corrector of different generosity. Furthermore, the correctors of the exam questions were unaware of which students belonged to which treatment condition.

Conceptual Knowledge: FCI

In physics, students may know how to solve problems without having a complete conceptual understanding of the physics involved (Kim & Pak, 2002). Therefore, conceptual understanding was also measured the first and last week of the semester with the Force Concept Inventory (Halloun et al., 1995; Hestenes et al., 1992). To avoid ceiling and floor effects, ***normalized gains*** in the FCI are compared. Normalized gains are defined as:

$$g = (\text{Post T} - \text{Pre T}) / (\text{max T} - \text{Pre T}) \qquad \text{Eq.1}$$

When the post-test score is greater or equal to the pre-test score, normalized gains yield a value between 0 and 1 representing the fraction of the concepts learned to the total concepts initially left to learn. For instance, a student scoring 40% before instruction has 60% of concepts left to learn. If she scores 70% after instruction, then she gained 30% of the total 60% possible left to gain, thus g=0.50. Among compelling arguments given for using normalized gains (g), is the reported finding that g is ***uncorrelated*** to pre-test scores (Hake, 1998, 2001, 2002) and therefore gives a better description of the conceptual gain due to instruction. In contrast, post-test scores are highly correlated with pre-test scores which would be expected if no instruction were present.

Results Part 1: Can PI be implemented in Cegeps?

Implementing an extensive and somewhat costly instructional approach is often problematic in public institutions. Yet, John Abbott College's administration was very supportive of this project. From the physics department chair to the Academic Dean by way of the Dean of Science, each actor manifested great interest in *PI* and provided more than adequate support to implement it.

Since the first implementation of the approach, more than half of the full-time physics department (8/14) members currently use some form of *PI* in their classrooms. Instructors in other departments have learned about the method from presentations given at the college and from word of mouth. In the chemistry department, one professor has successfully used the clickers in his introductory course, and is looking forward to repeating the experience. A nursing instructor is currently looking into using the method in her courses next semester. Numerous other instructors have inquired about the hardware and may opt to use it in their classrooms. From the reception at the different levels of administration to instructors in diverse fields, it is fair to say that *PI* was warmly welcomed by our Cegep community.

Modifications to course structure

Using *PI* with clickers in the classroom requires a minimum amount of changes, as would any new technology. To present students with ConcepTests that will allow clicker votes, one simply needs to write or import conceptual questions into PowerPoint. Many ConcepTests can be found either online though the rich Project Galileo website at Harvard (http://galileo.harvard.edu/) or through textbook publishers that now package 'clicker questions' with their textbooks. Thus, there are currently sufficient resources available to make the use of ConcepTests quite feasible. Other changes related to the clicker technology include familiarization with the clicker hardware and software. It is strongly recommended that all interested instructors setup the clickers and receivers and try them a few times before attempting to use them in class. Certain hardware and software issues have been documented and can be accessed online (see full PAREA report: www.cdc.qc.ca).

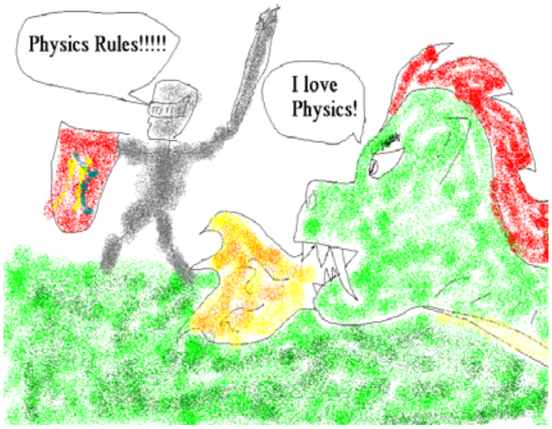
Reception by students

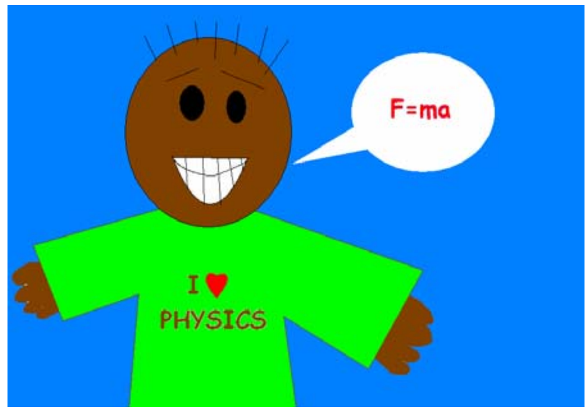
Students warmly welcomed using clickers in the classroom. Students in the flashcard section were also quite content with using flashcards. However this contentment tapered when these students realized the other section was using clickers. To gauge the appreciation of the method in both the clicker and flashcard section, students were asked to rate their level of agreement (5 = completely agree to 1 = completely disagree) with each of the seven statements below. Note that both the clicker and flashcard groups responded using clickers.

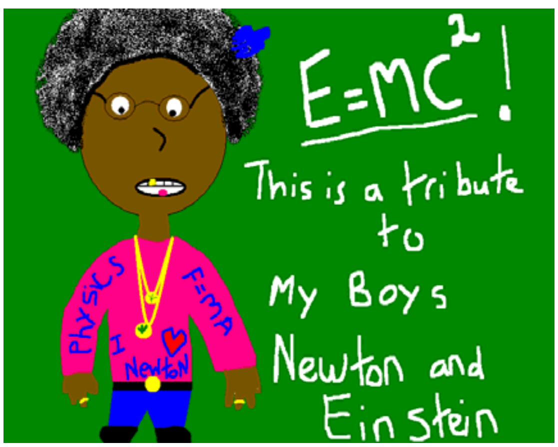
1) *PI* helped me recognize what I misunderstood
2) *PI* showed me that other students had misconceptions similar to mine
3) I actively discuss problems with my classmates
4) Convincing other students helps me to understand concepts
5) The mini-lectures help to clarify the concept for me
6) *PI* helps to learn better then traditional lectures
7) If I had the choice between a PI course and a traditional course I would choose PI

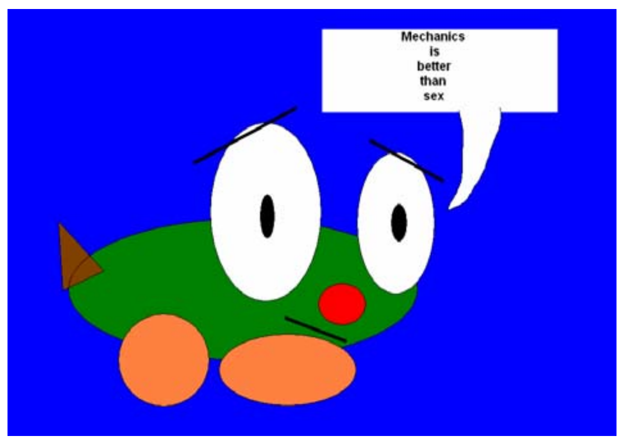
Answers were collapse from a 5-point Likert scale onto 3 categories: agree/strongly agree; neutral; disagree/strongly disagree. To determine whether students ***agreed*** with a statement more than would be expected by chance (2/5 or 40%), a binomial probability (agree $p=0.4$; not $q=0.6$; $n=30$) was calculated. Results show that students in the flashcard and clicker sections responded positively to *PI* by significantly ($p<0.05$) acknowledging its advantages as an instructional approach (Q1-5) and by preferring it to traditional instruction (Q6,7). Furthermore, 61% of students in the Flashcard section - using clickers for the first time- agreed that they would have participated more if they had clickers instead of flashcards.

Interesting unsolicited student feedback was also found in the form of computer doodles made in Microsoft Paint and placed on physics lab computers as screen savers. The pictures below were found on the physics lab computers screens after *PI* students had left. These pictures were not present before students entered the lab.









In summary, implementing *PI* seems quite feasible at the Cegep level. Indeed, the modifications to course structures required are minor and quite feasible and it has been very well received by administrators, teachers and students.

Effectiveness of PI vs. Traditional lecturing

Conceptual Learning

In this part of the study both *PI* groups (clicker group and flashcard group) were merged and learning measures were compared to the control section. The following table displays the FCI

Pre-test score, Post-test scores, normalized gains for both *PI* groups and the control group. Also shown below are p-values obtained using t-tests to determine whether the difference in averages between sections is significant.

Table 4.1
FCI data for *PI and* traditional control group

	<Pre-test> (%)	**<Post-test> (%)**	**g** (norm. gain)
PI **(n= 69)**	42.6	68.6%	0.50
Control (n=22)	46.0	63.3%	0.33
t-test (2-tailed) p	0.427	0.283	***0.008***

These results show that although no significant difference exists between groups before instruction (p=0.427) the *PI* group gained significantly more conceptual knowledge after instruction (p=0.008) as measured by the FCI. This result shows unequivocally that *PI* enables more conceptual learning than traditional instruction. Note that the results found here replicate results found by Mazur (1997) on conceptual learning and by Hake (1998) in the difference between non-traditional active engagement methods (including *PI*) and traditional instruction.

Traditional problem solving

Physics instructors may hesitate to use non-traditional methods such as *PI.* A frequent concern is that time spent on concepts implicitly takes away time spent on problem solving skills that students are expected to have and display on exams. The following table shows the average grades on the Fall 2005 common final exam as well as the p-values found by using a t-test to compare the exam averages of students between groups.

Table 4.2
Common final exam data for *PI* and traditional control group

	Exam Avg (%)
PI **(n= 79)**	68.0
Control (n=35)	63.0
t-test (2-tailed) p	0.21

These results show that *PI* students show non-significantly better results (p=0.21) on the exam. Therefore, although more time is spent on conceptual learning and less time is spent on algorithmic problem solving, students in *PI* groups do not have lesser problem solving skills. This may be due to the positive contribution of conceptual knowledge in traditional problem solving. That is, one must spend more time to learn many algorithms by rote than is required with solid conceptual knowledge.

The effect of clickers on learning

To determine the specific contribution of clickers on learning, the FCI pre-test, FCI post-test, FCI normalized gain and exam data are compared for both *PI* groups below:

Table 4.4
The effect of clickers: difference in learning data between flashcard and clicker groups

	PreFCI /30	**PostFCI /30**	***g***	***Exam (%)***
Clickers (n= 35)	**11.9**	**19.9**	**0.486**	**69.8**
Flashcards (n=34)	**13.6**	**21.3**	**0.520**	**71.6**
t-test (2-tailed) p	**0.209**	**0.351**	**0.745**	**0.630**

These results shows that both groups did not differ significantly in FCI score at the beginning of the semester (p=0.209) or at its end (0.351). Therefore, the use of clickers does not seem to add to the amount of conceptual learning or the problem solving skills. Indeed, although clickers have been reported to have a motivating influence, over the course of a semester no significant differences were found in conceptual learning (p = 0.745) nor in problem solving skills (0.630). This implies that *PI* is an effective instructional approach which is independent from the use of technology such as clickers.

Effectiveness of Peer Instruction: the role of proficiency

One may contend that what works at Harvard may not necessarily work in a public college setting. The question addressed in this section is whether student aptitudes in physics, or equivalently their proficiency level, contribute to the effectiveness of *PI*. To this effect, the initial proficiency level of all students was associated to their FCI score before instruction. Students from all groups were pooled and a median FCI score before instruction of 12/30 was found. Two groups were then constructed by taking all students at the median pre-test FCI score or below in one group, and all those above the median in another. Normalized gains for high and low proficiency students were then compared and differences in average normalized gains between groups were sought using a t-test. The following table illustrates the results.

Table 4.3
Effect of student proficiency on learning in *Peer Instruction* and control

	PreFCI ≤ Median ***G***	**PreFCI > Median** ***g***	***t-test (2-tailed) p***
Peer Instruction **(n= 69)**	**0.387**	**0.672**	< ***0.00001***
Control (n=22)	**0.264**	**0.383**	0.337
t-test (1-tailed) p	0.07	***0.00022***	

A difference is found between low proficiency students and high proficiency students in both sections. In the *PI* group, the difference between proficiency groups is very large (0.387 vs 0.672) and quite statistically significant ($p < 0.00001$). Furthermore, *PI* students with higher proficiency levels achieve significantly more conceptual learning (0.672 vs 0.373; p=0.00022) than high proficiency students in the control section. But low proficiency students also benefit from PI. Indeed, low proficiency *PI* students perform non-significantly better (p=0.07) then their counterparts in the control group. The lack of robust significance ($p<0.05$) is possibly due to a lack of statistical power since there were only 9 low proficiency students in the control section.

Discussion of Results

Effectiveness over Traditional instruction in Cegep context

As expected from studies in American colleges and universities, *PI* in Cegep enabled significantly more (p=0.008) conceptual learning than didactic lecturing. Such a result may not be sufficient to convince certain instructors from adopting the method. Indeed, some claim that given student difficulties with quantitative problem solving, time spent on basic concepts takes time away from in-class problem solving activities and would therefore be unwise. In fact, quite the opposite was found. *PI* students spending more time on concepts perform non-significantly (p=0.21) better than students in the control group in traditional problem solving. Thus, although less time is spent in algorithmic problem solving, providing the conceptual background allows students to be more effective in problem solving.

Lack of added effectiveness with clickers

One of the interesting and unexpected findings of this study is that the use of clickers does not provide any additional learning benefit to students. Previous users of clickers in university classrooms had reported benefits such as increased rates of attendance and decreased rates of attrition (Owens *et al., 2004;* Lopez-Herrejon & Schulman, 2004) since students may want to come in class to simply “play with the clickers”. However, no data was found in this study to support the claim that clickers increase conceptual learning. *PI* is an elaborate pedagogical approach that places a strong focus on basic concepts, requires students to commit to a conception and provides a setting for peer discussion to sort out correct concepts from misconceptions. Clearly, the technology is not the pedagogy. But if clickers don’t add to learning should they be abandoned?

In fact, clickers should be greatly encouraged. Although this conclusion seems to contradict the previous finding, there are three main reasons why clicker use should continue to be encouraged. First and foremost, clickers are responsible for much of the attention given to the *PI* approach. Indeed, much of the success of *PI* implicitly rest on the use of clickers (Burnstein & Lederman, 2003, 2001). Many instructors, including myself, have adopted the approach due to the appeal of using this technology in their classrooms. Using *PI* with clickers however forces instructors to reconsider their teaching, focus on concepts and thus fundamentally reshape their instruction. Since many instructors would not give proper attention to *PI* were it not for the clickers, one must continue to encourage their use.

Second, using clickers in the classroom allows instructors to archive ConcepTest data. Beyond data analyses and research questions that can be addressed, this data can be used instructionally to sort out useful ConcepTests from those that work poorly. Furthermore, ConcepTests of questionable effectiveness could be reformulated and the core set of questions can evolve from one semester to another. Using flashcards does not enable the instructor to collect any ConcepTest related data. Thus, reusing the same questions from semester to semester may differ in effectiveness from using questions that can be modified from one semester to the next. Since only one semester of implementation was compared no such differences were found although differences are expected to emerge over time.

The third reason for encouraging clicker use is to maximize the effect of peer discussion. Currently, 2-way clickers with a LCD display are available. These clickers allow students to send data but also receive data from the instructor’s computer (such as acknowledgment of vote reception). To maximize the effect of peer discussions, one may program the response displayed to students so that it pairs students of differing conceptions. The response could then relocate a student to another seat in the classroom where the adjacent student holds a different conception. Using the clicker display to pair students holding different conceptions would thus maximize the effectiveness of the approach.

Some instructors may be aware of *PI* methodology and willing to reshape their instruction to provide greater focus on basic concepts. Yet, the capital expense for the purchase of clickers and related hardware may not be available or passing the expense onto the students not possible or desirable. In this instance, *PI* should be implemented with flashcards as it is the *PI* pedagogy which is effective regardless of the modality used by students to report their answer.

Conclusion

PI is an effective pedagogical approach which must be widely disseminated and encouraged. It is simple enough to enable systemic change in relatively little time. This study confirmed the effectiveness of *PI* in Cegep contexts. Increased gains in conceptual learning were found and no difference in traditional problem solving skills were observed even though *PI* students had less class time devoted to problem solving activities. This study also found that clickers did not add significantly to students' learning. That is, although clickers have many advantages, their use does not increase the effectiveness of the *PI* approach. The conclusion is that the technology is distinct from the pedagogy.

Many science instructors teach today the way science was taught 100 years ago (Beichner *et al.*, 1999). Yet, the *PI* approach is slowly but surely changing the way instructors and students conceive instruction. Its methodology requires very little changes from traditional lecturing besides an extended focus on basic concepts. Its approach does not conflict with current institutional constraints as it is well received by administrators, teaching colleagues and students. By focusing on basic concepts it has taken away the perception that science (physics specifically) is about finding formulas. It has integrated Simon's (1996) notion that "*the meaning of* "knowing" *has shifted from being able to remember to being able to find and use*" by pushing students to find and use the basic concepts instead of remembering which formulas to use. Although its use of technology was not found to add to students' learning, it integrates the current culture looking for newer forms of technology applications in the classroom. From this perspective, *PI* is a sound pedagogical approach that must be warmly welcomed into Cegeps and universities.

References

Beichner, R., Bernold, L., Burniston, E., Dail, P., Felder, R., Gastineau, J., Gjersten, M. and Risley, J. (1999). “Case study of the physics component of an integrated curriculum,” Phys. Educ. Res., Am. J. Phys. Suppl. 67, 16-24

Burnstein, R.A. & L.M. Lederman. 2003. "Comparison of Different Commercial Wireless Keypad Systems," Phys. Teach. 41(5): 272-275; online at http://ojps.aip.org/dbt/dbt.jsp?KEY=PHTEAH&Volume=41&Issue=5

Burnstein, R.A. & L.M. Lederman. 2001. "Using wireless keypads in lecture classes." Phys. Teach. 39(1): 8-11; online at http://ojps.aip.org/dbt/dbt.jsp?KEY=PHTEAH&Volume=39&Issue=1

Crouch, C.H., and Mazur, E., (2001). *Peer Instruction : Ten years of experience and results.* American Journal of Physics, v. 69, p. 970-977.

DiSessa, A. (1982). Unlearning Aristotelian physics: A study of knowledge-base learning. *Cognitive Science* 6:37–75.

Fagen, A., Crouch, C.H., and Mazur, E. (2002). *Peer Instruction* : results from a range of classrooms. *The Physics Teacher,* vol. 40, p. 206-209.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66, 64-74.

Hake, R.R. (2001). Suggestions for Administering and Reporting Pre/Post Diagnostic Tests. Unpublished; online at http://www.physics.indiana.edu/~hake.

Hake, R.R. (2002). Lessons from the Physics Education Reform. *Conservation Ecology* 5(2): 28; online at http://www.consecol.org/vol5/iss2/art28

Halloun, I., Hake, R., Mosca, E. and Hestenes, D. (1995). Force Concept Inventory (Revised, 1995) in Mazur (1997) http://modeling.la.asu.edu/modeling.html

Halloun, I. & Hestenes, D. (1985). The Initial Knowledge State of College Physics Students. *American Journal of Physics*, 53, 1043-1055.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158 .

Kim, E and Pak, SJ (2002). *Students do not overcome conceptual difficulties after solving 1000 traditional problems*. American Journal of Physics **70** (7), 759.

Lopez-Herrejon, RE, Schulman, M.(2004). *Using Interactive Technology in a Short Java Course: An Experience Report*. Annual Joint Conference Integrating Technology into Computer Science Education: Proceedings of the 9th annual SIGCSE conference on Innovation and technology in computer science education.

Mazur, E. (1997). *Peer Instruction* : A User's Manual. Upper Saddle River, NJ: Prentice Hall

McDermott, L. C., & Redish, E. F. (1999). RL-PER1: Resource letter on physics education research. *American Journal of Physics, 67*(9), 755-767 .

Owens, K., McConnell, D.A., Steer, D., Van Horn, S., Knott, J., Borowski, W., Mcgrew, H., Dick, J., Greer, l., and Malone, M. (2004). C*hanging pedagogy to include conceptests and Peer Instruction in introductory geoscience courses: the impact on instructors and students* . The geological Society of America Annual meeting (Denver). Abstract available online at http://gsa.confex.com/gsa/2004AM/finalprogram/abstract_75839.htm

Simon, H.A. (1996). *Observations on the Sciences of Science Learning*. Paper prepared for the Committee on Developments in the Science of Learning for the Sciences of Science Learning: An Interdisciplinary Discussion. Department of Psychology, Carnegie Mellon University.