PART IV. CONCLUSION

Chapter 11. Conclusion

In this dissertation, I have examined the methods used to evaluate instruction in terms of the hidden curriculum and used these methods to study classes at ten schools taught with one of three research-based curricula or traditional lecture instruction. In this chapter, I will summarize what I have learned in this study, discuss the implications for introductory physics instruction and physics education research, and suggest directions for future studies in this area.

I. SUMMARY

A. Why are Research-Based Curricula Necessary?

Physics Education Research (PER) has demonstrated that traditional lecture, recitation, and lab instruction is not helping many students in introductory physics classes develop a functional understanding of physics concepts. In chapter 2 (Mazur[1] and Hammer[2]), chapter 4 (Halloun and Hestenes,[3] Hestenes et al.,[4] and Hake[5]), and chapter 6 (example of a students' solution to a quantitative end-of-chapter problem and a qualitative problem), we have seen that many students have difficulty with basic physics concepts on simple, qualitative questions even when they can successfully solve mathematically complex end-of-chapter problems. The studies by Mazur,[6] Hammer,[7] and Tobias[8] suggest that the emphasis on typical end-of-chapter problems and the structure of the traditional lecture method encourage students to see learning physics as memorizing and applying the facts and equations without understanding the underlying concepts.

To the students in these traditional lecture courses, the main learning goal appears to be to demonstrate mastery of the material by solving typical end-of-chapter problems on exams and homework. However, most physics instructors want their students to achieve much more than that. Among other things, they want a majority of their students to achieve the following:

- to understand the main physics concepts including knowing when and where specific concepts apply;
- to be able to express what they have learned consistently in multiple representations including graphs, equations, and words;
- to see physics knowledge as a connected, coherent framework of ideas where a few key principles can be used to understand many physical situations; and
- to be able to apply what they know to new physical situations in and out of the classroom.

Learning goals like these, which are often neither stated explicitly to the students in class nor encouraged through grading and testing, are part of what we call the "hidden curriculum."[9]

In the last twenty years, PER has made significant progress in identifying and understanding student difficulties with introductory physics. One of the main findings is that students come to the introductory physics sequence with beliefs and attitudes based on years of experience with school and the world around them. In particular, they have their own ideas about how to solve physics problems, common sense beliefs about how things work, and cognitive beliefs on learning, physics, and mathematics. Many of these views are incompatible with what instructors want the students to learn, hinder the students' learning, and outlast traditional lecture instruction. PER has also demonstrated that if students' initial views are taken into account, it is possible to design active-learning activities that induce most of the students to develop a good better understanding of many of the basic concepts or develop a more-expert problem-solving approach.

This last result has led to the development of new curricula to improve student learning by focusing on what is happening in the student rather than on what the teacher is doing. Most of these PER-based curricula use a strategy similar to Posner *et al*.'s four conditions for conceptual change (discussed briefly in chapter 2)[10] including a component where the students are actively involved in debating and discussing the course material in peer groups of 2-4 students. The three PER-based curricula examined in this study, Tutorials,[11] Group Problem Solving,[12] and Workshop Physics,[13] all make use of cooperative student group activities. Tutorials and Workshop Physics are designed to improve students' conceptual understanding of physics while Group Problem Solving emphasizes developing expert problem solving skills.

The developers of these three curricula have each presented evidence that their methods significantly improve student performance on problems and/or multiple choice tests in the areas the curricula were designed to address. In addition, students taught and engaged with one of these three curricula have been shown to do as well, if not better on conventional end-of-chapter problems as students taught with traditional lecture. However, to be effective, each of these PER-based curricula require additional resources and major changes in teaching style. Because of the effort and cost involved in implementing one of the research-based curricula as well as the difficulties of adopting a curriculum developed elsewhere, it is important to learn how to evaluate these PER-based curricula to see what students are learning, particularly with regard to the hidden curriculum.

B. How Do We Evaluate Research-Based Curricula?

In this study, I evaluated two aspects of the hidden curriculum, conceptual understanding and expectations, by collecting four types of assessment data to evaluate four curricula used in calculus-based introductory physics courses. The types were multiple-choice concept tests, the Maryland Physics Expectation (MPEX) survey, student interviews, and specially designed qualitative exam problems. Concept test results and surveys were collected from at least one class at each of the ten schools. Student interviews were conducted at five of the ten schools and qualitative exam problem results were available from classes at the University of Maryland..

From a research standpoint, the four methods are most useful for learning about what happens when students go astray and don't learn what was intended. By studying what these students have learned from the class and why, we can begin to understand the nature of the students' difficulties in learning physics and how to modify instruction to address these difficulties. Each of the four methods tells us different things about what students in introductory classes learn during the sequence.

In chapter 4, we saw that concept tests like the Force Concept Inventory (FCI) give an indication of how well students know basic concepts. In the case of the FCI, Halloun and Hestenes demonstrated this by comparing FCI results with results from interviews and open-ended questions from individual students.[14] Hake demonstrated that the fractional gain between the pre- and post-course FCI results is a more consistent measure of how much students' conceptual understanding has improved than the absolute gain or raw post score.[15] However, Steinberg and Sabella found that a large minority of students did not respond consistently to similar FCI and open-ended qualitative exam problems.[16] This suggests that while the FCI is a good indicator of whether students know Newtonian force concepts, it may not be a good indicator of how students use their conceptual understanding in problem solving.

An analysis of exam problem solutions can be a good indicator of students' ability to apply what they know. A variety of carefully constructed problems like those in chapter 6 can reveal much about what students have learned to apply. However, as we saw in the case of the waves-math pretest in chapter 7, interviews may also be needed to understand what the students are thinking and why they answer a particular way. Interviews are the most effective way to determine how students are thinking about physics.

As we saw in chapter 2 (Hammer and Tobias) and chapter 10, interviews are also a good way to learn about student expectations. While interviews are the most effective research tool for determining how students think about physics and what they are learning, the time required to transcribe and analyze interviews makes them impractical for evaluating classwide effects in all but the smallest classes. Since one of the goals of this investigation is to study the distribution and evolution of student expectations in introductory classes, we developed and used the MPEX survey, as discussed in chapter 5, refined through the use of interviews.

The MPEX survey was constructed to probe student expectations with a focus on six structures: independence, coherence, concepts, the link between physics and the real world, understanding of the role of math in physics, and the kind of effort students expect to make. The survey was calibrated using five groups. The calibration group expected to be most sophisticated was in strong agreement (better than ~80% on almost all the items) as to the desired responses on the items of the survey. Their preferred response was defined as favorable. The other calibration groups showed decreasing agreement with the expert group as predicted. Over 100 hours of interviews were conducted with student volunteers to validate the students' interpretations of and responses to the 34 survey items. In addition, reliability of the survey was established by a Cronbach alpha of 0.81 and by demonstrating that the overall survey results and the results for the six dimensions were reproducible for several similar main sequence classes. The calculated standard deviations of each of the survey results for similar classes were comparable or less that the estimated distribution widths.

While the survey can provide a measure of student expectations are and how they change, interviews are also needed to see why they change and to better understand the process of change.

C. Evaluation of PER-Based Curricula

In this dissertation, the evaluation procedures described in the previous section were used to compare three PER-based curricula (Tutorials, Group Problem Solving, and Workshop Physics) to traditional instruction. We find explicit answers to the three research questions posed in chapter 1.

Q1. What are the characteristics of different student populations coming into the calculus-based introductory physics class?

a. Conceptual understanding of physics

To measure conceptual understanding, FCI or Force and Motion Conceptual Evaluation (FMCE)[17] data were obtained from first-term classes at all ten schools. FCI data were collected from eight of the ten schools. The average overall pre-course FCI results at the three large state universities (UMD, MIN, and OSU) and one of the small liberal-arts colleges (DRY) were all very similar at approximately 50%. This was significantly larger than the average overall pre-course FCI score at the other three liberal arts schools (DC, NWU, and SKD) and the community college, where the average score was about 40%. The pre-course FMCE results show a similar gap. The students at MSU start with an average FMCE of 37.5% while the students at DC and CAR both start with an average of about 25%. For the FCI results it is worth noting that the average pre-course FCI scores at all ten schools are well below the Halloun and Hestenes' 60% entry threshold for thinking about motion in terms of Newton's laws.[18] Halloun and Hestenes suggest that students who have not reached this threshold are not yet ready to solve physics problems with an understanding of the underlying concepts.

Of the 8 schools that supplied FCI data, 6 of them (UMD, MIN, OSU, DC, NWU, and SKD) provided itemized data, allowing the Newton's third law FCI cluster to be evaluated. The students at UMD had the highest initial score (40%) on the Newton 3 cluster and the students at SKD had the lowest (26%). The classes at OSU, MIN, and DC averaged about 33% on the Newton 3 cluster while classes at NWU initially averaged 29%.

The overall concept test results indicate that students in introductory physics classes at the large public universities (UMD, MIN, OSU, and MSU) start off with a small but significant advantage in their understanding of Newtonian force over their counterparts at the smaller liberal arts colleges (with the exception of DRY). The differences on both overall concept test averages and on the Newton 3 cluster do not seem to correlate with the overall selectivity of the school.[19]

b. Student expectations

MPEX survey data were also collected from classes at the ten schools. The initial state of students at the ten schools deviated significantly from that of the expert calibration group with overall responses of the students ranging from 50-65% favorable in compared to 87% for the expert group. The expert responses on the clusters varied from 85-93% favorable. The student responses deviated most strongly on the independence, coherence, and concepts clusters. These varied from 40-60% favorable with the community college near the bottom for all three clusters. The most favorable student response was on the reality cluster where responses ranged from 60-80% favorable.

Many of the student populations showed some differences when compared with Maryland students. Starting with our three large public universities (UMD, OSU, & MIN), the Minnesota students had significantly more favorable expectations overall and in three of the clusters than the Maryland students. The Ohio State students had similar expectations to the Maryland students overall and in all the clusters except the reality link cluster where the OSU students were slightly but significantly better. Except for MSU, which is really a medium sized public university, the students at the small liberal arts colleges all responded more favorably initially than the Maryland students on two to five expectation dimensions of the survey. Note that none of the student populations at any of the nine other schools consistently responded more favorably than the Maryland students for all six clusters.[20] The NWU students responded more favorably than the Maryland students overall and for the reality and effort clusters. The DC students responded more favorably than the Maryland students to the independence, reality, and effort clusters. The students at CAR and SKD both responded more favorably than the UMD students to the reality and effort clusters. It is worth noting that the students at all five of the liberal arts schools and MIN responded more favorably to those two clusters than the University of Maryland students.

Q2. How do we determine if students are improving their knowledge of physics concepts and expectations?

Q3. Are the research-based curricula more effective for helping students to improve their conceptual understanding and their expectations of physics?

a. Conceptual understanding of physics

Multiple-choice concept tests, specially designed open-ended exam problems, and interviews were used to evaluate students' conceptual understanding of physics. The change in the mechanics concept test scores from the beginning to the end of the first term of introductory physics are used to determine if students are improving in their knowledge of physics concepts. In this study, we used Hake's fraction of the possible gain h to measure the change.[21] The larger the fractional gain, the greater the improvement in understanding of the basic concepts.

Classes that used one of the three research-based curricula (RBC) had significantly better overall fractional gains on the multiple-choice concept tests on Newtonian Force than the traditional lecture classes at UMD, CAR, and PGCC. This was even true for classes at schools that were in their first or second year of implementation of research-based curricula. In fact, all the classes using RBC had overall fractional gains on the FCI as good or better than the best Traditional lecture class in this study. The best results came from classes using Workshop Physics (DC & DRY) or Group Problem Solving (MIN). At Maryland, two instructors taught classes both with Tutorials and without in different semesters. In both cases, the classes taught with tutorials had significantly better overall fractional gains on the FCI.

Similar results were obtained with the Newton's third law cluster on the FCI. The classes taught using one of the three RBC had significantly better average fractional gains on the Newton 3 FCI cluster than the classes taught with traditional lecture instruction. With two exceptions, the average Newton 3 FCI fractional gain for every RBC class was larger than the gains for the traditional classes. One exception was a small (≈ 40 students) traditional class taught by an award winning instructor at Maryland that achieved a larger fractional gain than any of the GPS classes and one of the Tutorial classes. This class also had the best overall fractional gain on the FCI of any traditional class. The other exception was a Workshop Physics class at Dickinson that had a fractional gain typical of traditional lecture classes. This class had unusually severe attendance problems. The Tutorial and Workshop Physics' classes tended to have higher fractional gains than the Group Problem Solving classes. This may be due to the fact that the Group Problem Solving curriculum emphasizes problem solving, while Tutorials and Workshop Physics emphasize conceptual understanding.

As we discussed earlier, concept tests are an indication of how well the students know the concepts, but not necessarily how well they can apply the concepts in problem solving. To address this issue, four specially designed problems were given on exams in Traditional and Tutorial classes at Maryland. The students are said to have improved in their ability to use their conceptual understanding if a significantly greater fraction of the students in the Tutorial class use the concept correctly in the context of a problem. Since the total and fraction of time spent on active-engagement activities in Tutorials (one hour per week) is the least of the three RBC, we would expect these results to be suggestive of the results that would be obtained from classes at other schools.

The results from the exam problems at Maryland were consistent with the concept test results. Tutorial students displayed a better understanding of velocity graphs, Newton's third law, graphs and equations describing the motion of harmonic oscillators, and 2 slit interference. The ratio of correct responses to the Newton's third law problem from the Tutorial class and the Traditional class were very similar to the ratio of correct responses to the Newton 3 FCI cluster, although the number of correct responses on the exam problem was significantly less.

Hammer's two-rock problem[22] was used in interviews with student volunteers at Maryland, Dickinson, and Ohio State. Here the issue was not so much to see how students had improved, but how well they were able to use their physics knowledge. Very few students were able to solve the problem in the interview, but this was not surprising because of the nature of the problem. What was surprising was the inability of many of the students to derive the kinematic equations or to recognize the condition that acceleration is constant. In addition, many students became stuck when the kinematic equations produced an expression that was hard to evaluate. In contrast, the few students who did solve the problem were able to come up with alternative approaches that indicated a better, more flexible understanding of physics concepts.

The concept tests and the exam problems clearly indicate that RBC are more effective for helping students learn and apply physics concepts than traditional instruction. However, all three types of evaluation indicate there is still need for further improvement. Only a few of the classes participating in this study achieved a fractional gain of at least 50% and none of the classes achieved an overall average score of 85% on the FCI, the threshold suggested by Hestenes and Halloun for confirmed Newtonian thinkers.[23] Moreover, the results from the exam problems and the interviews suggest that while more of the RBC students than the traditional students are able to apply concepts correctly, many of the RBC students were still not able to use the concepts correctly.

b. Student expectations

The students participating in this study took the MPEX survey at the beginning (pre) of the introductory physics sequence, at the end of the first semester or quarter (mid), and again at the end of the first year (post). The mid and post responses are compared with the pre responses. If the students in a particular group gave significantly more favorable responses on the overall survey or in one of the six clusters, we say that the students have improved in those expectations. A change Δ is calculated by subtracting the percentage of pre favorable responses from the percentage of post favorable responses. The change is considered to be statistically significant if $\Delta > 2\sigma$ where our estimation of the uncertainty σ is discussed in chapter 5. However, if students' expectations deteriorate significantly as a result of traditional instruction while the expectations of students taught with one of the RBC do not change significantly, then we can say the RBC has improved the students' expectations relatively.

At every school we studied, the overall MPEX survey results deteriorated as a result of instruction, although only five of the classes decreased significantly (CAR-TRD, MSU-WP, MIN-GPS, MIN-TRD, & OSU-GPS). Note that this group includes all the traditional classes for which data was available.[24] A major part of this deterioration was the significant decrease in favorable responses (deterioration) to the effort cluster at every school tested. In their judgments at the end of a semester or the end of the year, students felt that they did not put in as much effort as they had expected to put in at the beginning of the sequence. This part of the result is well known and neither surprising nor particularly disturbing. What is more troublesome is the result that half of the schools showed significant deterioration on the math link and the reality link dimensions. There was no significant increase in any of cognitive dimensions after one year of instruction except for the concepts cluster where UMD Tutorials and DC Workshop Physics both improved significantly.

It is interesting to note that the survey responses of both GPS classes deteriorate significantly overall and in three of the cognitive clusters including coherence. The students who had traditional instruction at Minnesota deteriorated significantly overall and in all clusters except the coherence or concepts clusters but did better than the Minnesota GPS students who deteriorated both overall and in every cluster.

These results suggest that students who have had a year of Workshop Physics and Tutorials have maintained better expectations than students taught with Group Problem Solving or Traditional lecture instruction. Only two schools showed any improvement in expectations, both in the concepts cluster.

Interviews with students at some of the schools were used to try to understand the interaction between the curriculum, student learning, and student expectations as well as validating the student interpretations of the MPEX survey items. Two of the four students interviewed from the top third of the Workshop Physics class at NWU had expectations that prevented them from developing a flexible, functional understanding of physics but did not prevent them from succeeding in the class. In addition, some of the interviewed students had favorable expectations but ran into difficulties with other members of their group who did not. The students with unfavorable expectations had different learning objectives in class, namely to get through the material, not to understand it and think about it. Since the groups work together as a team, this caused some of the interviewed students with favorable expectations to rush through the activity as well. The interviews also suggest that expectations, mathematical ability, and success in the course were independent in the NWU class.

II. IMPLICATIONS

In this dissertation, we compare three RBC with traditional instruction to determine how well these curricula improve student learning in terms of conceptual understanding and expectations. Compared with the three RBC, traditional instruction is clearly not working for many of the students. The students taught with traditional instruction had lower fractional gains on the FCI and FMCE and did not do as well as the RBC students on the qualitative exam problems. In addition, the MPEX results show that after a year of instruction the fraction of favorable responses on the reality link, math link, and effort clusters as well as the overall survey result decreased significantly. The results for the remaining clusters also showed some deterioration as well. This confirms the need for change discussed earlier in the chapter.

First, let us consider the effect of the three RBC on conceptual understanding. The average fractional gain on the overall FCI and FMCE was significantly better for the three RBC than the traditional classes at CAR, MIN, and UMD. In addition, each class' average fractional gain was as good or better than the best fractional gain for a traditional class. These results are consistent with the earlier findings of Hake.[25] These results strongly suggest that that the three RBC were more effective than traditional instruction for teaching the students the concepts of Newtonian force and motion. On the Newton's third law cluster on the FCI, the average fractional gain was higher for the students taught with any of the three RBC than for those taught with traditional instruction. Although the GPS sequences had overall fractional gains on the FCI as good as the best classes using one of the other RBC, the Tutorial and Workshop Physics classes had higher fractional gains on the Newton 3 cluster. This implies that the Tutorials and Workshop physics may be more effective for teaching difficult concepts.

The evaluation of the Maryland students with both concept tests and specially designed exam problems showed that students taught with Tutorials had not only improved their knowledge of physics concepts, but also their ability to apply their conceptual understanding in problems. The Tutorial students demonstrated a better functional understanding of several concepts than the Traditional students. The scores on the Newton's third law problem were reflective of the differences on the Newton 3 cluster on the FCI. The students performed significantly better on the concept test than on the comparable problem. This suggests that while the concept tests indicate how well students know the concepts, they are not necessarily an indication of how well students can use what they know, an important component of functional understanding and the hidden curriculum.

The expectation results were less encouraging. Only 40 to 60% of the student responses at the ten schools were favorable on each of the three cognitive dimensions described in Hammer's study.[26] This suggests that many of the students enter the introductory physics course as "binary" learners who believe that knowledge is composed of facts (in this case, equations) that are transferred from the authorities (the instructor and the textbook) to the students. This in turn implies that the type B attitudes and beliefs that Hammer observed in his small sample are prevalent in a large fraction of students in calculus-based introductory courses at community colleges, liberal arts colleges, and large state universities.

With regard to the effects of the RBC on expectations as measured by the MPEX results, only the Tutorials and one of the Workshop Physics sequences showed any significant improvement, both in the concepts cluster. However, the Tutorial classes and most of the Workshop Physics classes did better than the other classes by not deteriorating as much. The MPEX results for the GPS classes were not as good as the TUT and WP classes. The GPS classes had MPEX results that were no better and in some cases worse than traditional instruction.

While this analysis of the MPEX results is complete, there is still more to learn from the data we collected in this study; for example, how the MPEX survey results correlate with other evaluations such as grades, FCI performance, or majors? Laws has commented that junior and senior biology majors seem the most resistant to RBC in physics classes.[27] There are preliminary results from a Flemish study using the MPEX survey that biology majors have substantially different expectations than other majors.[28] Also, the results presented in this dissertation are only part of what has been collected to date. We have also started to collect data from traditional and innovative algebra-based courses.

This study has several implications for developers and adapters of RBC. The best results (fractional gains) on the FCI and FMCE were achieved by classes taught by people involved in developing the curriculum (Redish-UMD, Heller-MIN, Laws and Pfister-DC) or people who had effectively incorporated the active-learning activities with cooperative student groups (Riley-DRY). Part of this may be due to the skill of the instructor, but I believe this result suggests students learn more effectively when the active learning activities are well-integrated into the course and the student groups are functioning well discussing their understanding of the material. However, even the FCI gains of the best classes in this study and Hake's show there is ample room for improvement of PER-based teaching methods in helping students learn the concepts of physics.

Our understanding of expectations is considerably less developed than our understanding of students' difficulties with conceptual understanding. Therefore, it is not surprising that efforts to address expectation issues are also less developed than the

PER-based curricula to address conceptual understanding and problem solving. The MPEX results of the RBC classes and particularly the GPS classes indicate that even while students' conceptual understanding is improving, their expectations about what physics is and how to learn it are not. The MPEX interview at NWU showed that even when students' expectations did not prevent them from succeeding in the class, it did affect their perception of what they were learning and what they took away from the course. More research is needed on the role of student expectations in learning physics, particularly on the interaction of student expectations and the curriculum, to help curriculum developers and adapters address this issue of the hidden curriculum more effectively.

This study represents a first step is exploring the issue of assessment of the hidden curriculum and in expanding our understanding of what is really going on in our classrooms.

---

[1] E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, New Jersey, 1997).

[2] D. Hammer, "Two approaches to learning physics," *Phys. Teach*. **27** (9), 664-670 (1989).

[3] I.A. Halloun and D. Hestenes, "The initial knowledge state of students," *Am. J. Phys*. **53** (11), 1043-1055 (1985).

[4] D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," *Phys. Teach*. **30** (3), 141-158 (1992).

[5] R.R. Hake, "Interactive-engagement vs. traditional methods: A six thousand student study of mechanics test data for introductory physics courses," *Am. J. Phys*. **66** (1), 64-74 (1998).

[6] See Ref. 1.

[7] See Ref. 2.

[8] S. Tobias, *They're Not Dumb, They're Different: Stalking the Second Tier* (Research Corporation, Tucson AZ, 1990).

[9] E. Silver, "Foundations of cognitive theory and research for mathematics problem solving instruction," in *Cognitive Science in Mathematics*, edited by A. Schoenfeld, (Erlbaum, Hillsdale NJ, 1987), 33-60.

[10] G.J. Posner, K.A. Strike, P.W. Hewson, and W.A. Gertzog, "Accomodation of a scientific conception: Towards a theory of conceptual change," *Sci. Educ*. **66** (2), 211-227 (1982).

[11] L.C. McDermott and P.S. Shaffer, *Tutorials in Introductory Physics* (Preliminary edition) (Prentice Hall, New York NY, 1997).

[12] P. Heller, T. Foster, and K. Heller, "Cooperative group problem solving laboratories for introductory classes," in *AIP Conference Proceedings No. 399 The Changing role of the Physics Departments in Modern Universities: Proceedings of the International Conference on Undergraduate Physics Education*, edited by E.F. Redish and J.S. Rigden (AIP Press, Woodbury NY, 1997), 913-934; P. Heller, R. Keith, and S. Anderson, "Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving," *Am. J. Phys*. **60** (7), 627-636 (1992); P. Heller and M. Holloabaugh, "Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, *Am. J. Phys*. **60** (7), 637-644 (1992).

[13] P.W. Laws, *Workshop Physics Activity Guide* (Wiley, New York NY, 1997).

[14] I.A. Halloun and D. Hestenes, "Interpreting the Force Concept Inventory: A response to the March 1995 critique by Huffman and Heller," *Phys. Teach*. **33** (8), 502-506 (1995).

[15] See Ref. 5.

[16] R.N. Steinberg and M.S. Sabella, "Performance on multiple-choice diagnostics and complementary exam problems," *Phys. Teach*. **35** (3), 150-155 (1997).

[17] R.K. Thornton and D.K. Sokoloff, "Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and curricula," *Am. J. Phys*. **66** (4), 338-351 (1998).

[18] See Ref. 14.

[19] Selectivity as determined by the Princeton Review. See E.T. Custard, *The Princeton Review Student Advantage Guide: The Complete Book of Colleges* (Princeton Review Publishing, New York, 1996).

[20] Although, the seven DRY students came close. The DRY students responded more favorably than UMD students both overall and for every cluster except coherence.

[21] See Ref. 5

[22] D. Hammer, *Defying Common Sense: Epistemological Beliefs in an Introductory Physics Course*, Ph.D. Dissertation, University of California at Berkeley, (1991, unpublished).

[23] See Ref. 14.

[24] The third was not measurable since no data was collected from consecutive traditional classes at Maryland.

[25] See Ref. 5.

[26] See Ref. 21.

[27] P.W. Laws, P.J. Rossborough, and F.J. Poodry, “Women’s responses to an activity-based introductory physics program,” in *Fostering Student Success in Quantitative Gateway Courses*, edited by J. Gainen and E.W. Willemsen (Josey-Bass Publishing, San Francisco CA, 1995), 77-88.

[28] Private communication with E. Van Zele at Ghent University, Belgium (Sept. 1997).