Why Teach Kinematics? An Examination of the Teaching of Kinematics and Force—II†

Dewey I. Dykstra, Jr.(a

Department of Physics and Physical Science, Boise State University, Boise, ID 83725-1570

(Received ?? October 2002; accepted ?? ???? 200?)

The new Motion and Force Units of the Powerful Ideas in Physical Science Project (PIPS) sponsored by the American Association of Physics Teachers (AAPT) and funded by the National Science Foundation (NSF) are described in this pair of papers. In part I the two units are described along with the typical new understandings of motion and force developed by the students. It is shown that classes engaged using these materials can routinely make changes nearly *2 standard deviations* in size on established diagnostic evaluations of their ideas concerning force and under instructional circumstances which are far from optimum. The performance of groups within the class ranges widely from almost *no change* to the order of *5 standard deviations*. In part II factors are identified which appear to differentiate the student performances. These results suggest yet another round of modifications, departing even further from a *content-driven* orientation toward an even more *radical, student understanding-driven* approach. The resulting instruction induces routinely *2.5 standard deviations change* in the class average on the pre to the post instruction diagnostic scores with *normalized gains of 0.60*. In contrast, effect sizes of *0.60 standard deviations* and *normalized gains of 0.15* are typically generated by traditional, *content-driven* instruction. In this *radical, student understanding-driven* approach, the percentage of students in a large university class who appear to make working sense of acceleration-time graphs increases from about 25% to about 45%. Groups of students who meet this last criterion routinely make *5 to 7 standard deviations* change in their scores and *normalized gains of 0.90* on the force diagnostic. In the traditional, *content-driven* approach, 60% or more typically *do not* make useful sense of the meaning of acceleration used by scientists. In such courses the class as a whole makes *no significant progress* at developing new understanding either of motion or force. Such results raise ethical issues in our decisions as to how we teach.

I. Setting the Stage from Part I

In part I it was established that, using items from the Force and Motion Conceptual Evaluation (FMCE),[1] one can compute two different scores for each student's performance on the diagnostic. One of these scores gives the degree of correspondence between the student's choices and those choices consistent with the typical person-on-the-street view of force or "old" view as described in part I. The other score gives the degree of correspondence between the student's choices and those choices consistent with the "new" view worked out by the students in class, an essentially Newtonian view of force. With these pairs of scores it is possible to plot scatter graphs, "new" view of force versus "old" view of force, representing the student's scores at the beginning of the semester and then again after instruction. The distance and direction between the pre-instruction and post-instruction positions on the scatter graph indicate the magnitude and nature of change in the pattern of choices made by the student. The magnitude and nature of change in the pattern of choices is taken as indicative of change in the student's understanding of the views involved. Such scatter graphs for the whole class are illustrated in Figure 2 from part I.

One might be satisfied at the results generated, but having achieved *2 standard deviations* with a class and not knowing what limits there may be if any, one wonders how much more change can be induced. A glance at the plots in Figure 2 of part I of this pair of papers naturally leads any instructor to speculate about the students distributed along the diagonal toward the lower-right. Most instructors would prefer that at the end of the semester all of the students be at the upper-left. How can we get everybody up there? It should help to identify some of the factors differentiating the students along the diagonal. A factor that springs to mind is the students' grasp of the view of motion represented in the MBL graphs. There may be other factors including other content issues and psychological and cultural issues, but we have data on hand concerning their grasp of the motion issues.

II. What if anything distinguishes students distributed along the diagonal on the scatter graph at the end of the semester?

One might take from Figure 1 of part I that coming to a view of velocity and acceleration consistent with the graphs from MBL in a sense makes the shift in view of force possible. Figure 1a of part II contains two scatter plots taken from the post-instruction scores of students in the two PIPS semesters being considered so far in Figures 2 & 3 and Table 6, all of part I. The plots in Figure 1a of part II are of new view of force score vs. velocity score, both scores are post-instruction. The velocity score was the number of velocity questions answered consistently with the behavior of the graphs in the MBL system. There is another sub-set of 8 kinematics questions about acceleration in the FMCE used to generate an acceleration score. Figure 1b of part II contains scatter graphs for the same two semesters. These scatter plots show the distribution of post-instruction scores of acceleration vs. post-instruction scores of velocity. Finally Figure 1c of part II reveals scatter plots of post-instruction new view of force scores vs. post-instruction acceleration scores. With these graphs we can begin to tease out relationships between motion scores and force scores.

At first glance all of the scatter plots suggest that at the post-instruction measure, if one is not doing well at velocity one is not likely to do well at acceleration and if one is not doing well at acceleration one is not likely to have a high new view of force score. This is probably no big surprise, but let us look more closely at the plots. The plots in Figure 1a suggest that if one does well on velocity one is equally likely to do poorly as well on the new view of force, but that almost everybody does well on velocity-time graphs in the end. Figure 1b while it is not exclusive shows a preponderance of those who did well on velocity also doing well on acceleration. There is some scatter and there are people scoring the maximum on velocity who score all the way down to zero on acceleration, but these are clearly less in number than those doing well on the acceleration questions. Figure 1c shows a difference. The "bubbles" indicating more people and indeed "bubbles" in general seem to be clustered along the diagonal. It appears there is a correlation, that if one does well on acceleration, one is not equally likely to do poorly as well on the new view of force.

Figure 1c is consistent with the possibility that performance on (understanding of) acceleration may be what distinguishes students distributed along the diagonal in Figure 2 in part I on the post-instruction measure. It appears that once one is scoring at the 7 or 8 out of 8 level (greater than 6) on acceleration at post-instruction then the likelihood that one will do poorly on the new view of force is reduced.

What if we look at plots similar to Figure 2 of part I, but include only students who at the post-instruction measure scored 7 or 8 out of 8 on the acceleration questions? At some risk of a slightly misleading label we shall call these the high acceleration score students. Figure 2 of part II shows such plots. These look like what we wish for the whole class. Something about these students at least in part indicated by their success on the acceleration questions makes possible this performance.

Does this convince us there is a connection between understanding acceleration as it is displayed in the MBL graphs and developing the new view of force? While on the face of it we find support for this conjecture in Figure 2 of part II, the very question gives rise to the need to look for more support. What comes to mind is the students who do poorly on acceleration at the post-instruction measure. How did they fare on the force questions? Let us look at the same type scatter plot as in Figure 2 of part I, but this time confine ourselves to students from the bottom end of the performance scale on acceleration. In order to have comparable groups the cut-off in acceleration was chosen to give approximately equal numbers in the low acceleration range as we had in the high acceleration score group. This turned out to be acceleration scores of 0 to 3 (less than 4). We will refer to those who scored in this range as low acceleration score students. The result is Figure 3 of part II. It appears that the students in these scatter plots made little or no change in either their new or old view of force from beginning to end of instruction. This gives us much more confidence that there is some correlation between having figured out this version of acceleration and coming to the new view concerning force.

A. Understanding of Acceleration is a Major Determining Factor

It appears there is empirical evidence here supporting the conjecture that understanding acceleration is important for understanding the new view of force. We do not know why these latter students performed differently on acceleration, but there remains the impression that if we could get more students to a level of understanding acceleration as it appears in the graphs, their performance on the new view of force could be expected to go up as would the class averages in these measures. In fact an experiment to increase the performance on acceleration to generate an increased performance on the new view of force, if successful, would lend support to the conjectured relationship between understanding of acceleration as it appears in the graphs and the understanding of the new view of force.

We can come to the same conclusion that there is empirical support for the importance of understanding acceleration looking at the other analyses of the data. The results of Thornton's conceptual dynamics analysis in Table 1 of part II show major changes on the part of the high acceleration students and very little or no changes for the low acceleration students. The scatter plot in Figure 4 of <g> vs. <L> for the two sets of students reveals that the high acceleration students are doing exactly what we would like the whole class to do and the low acceleration students are making almost no change on average. In Table 2 we see effect sizes of the order of *5 standard deviations* on the part of the high acceleration students and effect sizes of slightly less than one standard deviation on the part of the low acceleration students. Imagine: if we could get most of the students performing like the high acceleration students we could have whole class performance effect sizes of *5 standard deviations* and *normalized gains of 0.90* or so!

III. An Experiment to Test the Conjecture

No experiment can prove more than that conjecture fits experience, but it is fit to predicted experience that is a mark of the useful conjecture. On the basis of both theoretical speculations such as those represented in Figure 1 of part I and empirical findings such a those presented in Tables 1 & 2 and Figures 1 – 4, all of part II, it is conjectured that understanding the new view of force depends on understanding acceleration as it occurs for example in the MBL graphs. If this were the case, if we could conduct a course in such a way as to increase the numbers of students who apparently understood acceleration in this form, then we would expect that the numbers of students who understand the new view of force would also increase. If this happened then we could argue by abduction,[2] that it is as if understanding the new view of force depends on understanding acceleration is a reasonable description of the situation.

A. How to improve performance on acceleration?

In the two semesters of the course described so far it is a matter of record that student performance on velocity graphs started out moderately good at the beginning of instruction (about 3 out of 5) and ended very good at the end of instruction (greater than 4.5 out of 5). Since spending more time on issues of acceleration might be expected to result in higher acceleration scores and they already seem to find velocity not all that hard, why not reduce the time formally spent on velocity and make possible more time on acceleration? The same question might be asked about position, freeing up even more time for acceleration.

It was suggested above that there may be psychological and cultural/social factors which explain the poorer performance on acceleration by some of the students. The instructor should attempt to address issues of motivation and self-image. Information can be a factor in change. It makes sense that if students have a reason why something is important, namely that there is evidence that doing well on topic A is correlated somehow with doing well on topic B and that doing well on both topics is the path to the best possible grade, then there are some students who will respond by taking topic A more seriously.

Students come to the course many of them apologizing for “not being good in science.” The atrocity from which they derive this belief was perpetrated on them by their previous schooling as an agency of society. It is not solely due to their schooling, but sadly it is institutionalized in the selection, training and indoctrination program that passes for science “education” today. They have learned to immediately assume they are wrong in an absolute sense because they are not good at science when what they first predict about acceleration-time graphs does not match the behavior of the graphs.

If the conclusion is that they are not good at science, then there is no basis for hope to actually understand acceleration. The best one can do is get by and certainly not to attempt to make sense of the phenomenon. The particular self-image they have learned is a serious problem. While this might effectively select and train a scientific elite, it is an ineffective strategy if *understanding* acceleration by all students is the goal. What is needed is some alternative that might enable the students to believe that they have the requisite capabilities and that the task is tractable. Again, by abduction, if we do find ways that noticeably improve their apparent understanding of acceleration with the expected concomitant change in understanding of force, then the notion that they are “not good at science” cannot be considered a rationally valid basis for pedagogical practice. In fact such a notion would be unethical to tolerate in pedagogical practice.

Clearly, for whatever reasons, in the instruction previously described in part I, in spite of the impressive overall class performance, there was something the low acceleration scoring students needed that went missing. One of the problems associated with the course format (large class enrollment, separate “lecture” and lab meetings) is that the course moves from full-class meetings to lab sessions whether issues raised in the full-class meetings are resolved or not and whether these issues need to be resolved or not before the next lab activity. The same is true after lab and we move back to the full-class meeting. These shifts give those who believe they cannot really make sense of things a convenient "out," an excuse to stop worrying about the present issue, resolved or not, usually not. The course format works against those who believe they cannot make sense of the phenomena and who may be less practiced at the kinds of skills that are best suited to working on making sense of the phenomena. In any fixed class-sized period of time it is simply not the case that everyone reaches the same state of understanding or even the same understanding.

B. Alterations in the Instruction

1. Reallocation of time

In order to allow more time for the class to try and make sense of the acceleration-time graphs, eight out of the 10 activities on position-time and velocity-time graphs were eliminated from the experimental course. It is always necessary to give students some sort of experience with the apparatus that enables them to trust and believe the graphs and to just get used to making the motions and operating the computers. In order to accomplish this familiarization, one activity (Activity M1.4) from Motion Investigation 1 and one activity (Activity M2.4) from Motion Investigation 2 were kept as an introduction to lab and the apparatus.

Both of these activities are "matching" graph tasks. In each case an idealized graph is given to the students and they are asked first to describe what motion could produce the graph and then to try actually making the predicted motion. If the attempted motion does not match the graph then the students are challenged to work out what they have to do to actually match the graph.

Because these activities served to introduce the students to the apparatus and because they were the only ones specifically on position-time and velocity-time graphs, somewhat more time was spent on them than had they been immersed in the full set of 10 activities, but no more than about twice the time.

In Activity M1.4 a position-time graph is drawn and the students are asked first to predict how they would have to move to make the graph and then to make the motion to produce the graph. The students were pressed to have each member of each lab group do an acceptable job. This is possible but not immediately easy. They had to recognize that points on the graph correspond to specific locations and times. To be accurate they had to figure out just what locations and times and to mark them. They had to become aware of delays in the timing process and how these relate to the actual motion and what is seen on the screen. They also come to grips with the significance of different slopes on the graph.

In Activity M2.4 a velocity-time graph is given to the students. Again, they are asked to predict what motion would produce such a graph and then to try their predictions. They were pressed to have each member of each lab group do a satisfactory job of moving to make the match. In both cases, Activities M1.4 & M2.4, there were a number of details the students discovered they had to pay attention to. One of the primary factors is that a velocity-time graph does not "behave" like a position-time graph. This is usually quite a surprise to most students, especially when they try to move to produce a pre-determined graph. There is evidence here to support the interpretation that there is a distinction between our cognitive/declarative knowledge and a kind of knowledge associated with kinesthetic experience. Often a student can say what is required having worked with a group of three others, but when it comes to the first attempt or so it appears the body still wants to move according to the old declarative knowledge instead of the new. As a consequence, they have to work out what one has to do to remain at some constant non-zero value on the velocity-time graph. The significance of the difference in meaning between the slope on a position-time graph and the slope on a velocity-time graph is another factor that is tough to take into account. They grapple with the difference between one meter and one meter per second.

With 80% of the activities in the first half of the unit omitted, there was a significant amount of time available to focus on figuring out what acceleration meant in the acceleration-time graphs. With the time available it was possible to allow the class to attend to as many issues as people were willing to bring up as often as they wanted to bring them up. Some of the time gained was spent dealing with the issues described in the next two sections of the paper, but it was also spent attending to issues raised by students and requests by the instructor to consider additional details.

Examples of student raised issues concern comparisons between acceleration-time graphs and the other two types of graph. In one case in Motion Activity M3.2 when the students have for the first time seen all three graphs of the same motion (a single step away from the motion detector), the question of why all three graphs do not rise from the zero axis or the original position at exactly the same time comes up. Figure 5 shows the three graphs from Motion Activity M3.2. This is asked after a first tentative definition of acceleration has been generated and a definition for velocity is pretty firm. A common initial response is something like: "You have to have some acceleration before you can have a velocity." and "You have to have a velocity before you can change your position." A lengthy discussion ensues when the instructor asks: "Given what we have decided about acceleration so far, can the velocity really be remaining at zero and the acceleration be non-zero?" The students end up resolving the issue when it is suggested by someone that maybe the graphs in effect have different sensitivities. This leads to an exploration involving changing the scales on the graphs and resulting in general agreement with the suggested resolution, but not without more discussion about the possible meaning for the terms, velocity and acceleration, and additional close inspection of the set of graphs.

Another example of a student-raised issue comes up in the same activity. Some students notice that when the acceleration-time graph has peaked in the positive region and is returning to zero, decreasing, the velocity-time graph is still increasing. This can be seen in the region between the two vertical dashed lines in Figure 5. They wonder how that could be. It is clear that if their notion of acceleration was still tied to strictly increasing velocity, then as long as the velocity increased they might expect the acceleration to increase at the same time. It is quite a subtle issue but a profoundly different idea of acceleration in the graphs than the acceleration of everyday language that the acceleration is tied to the slope of the velocity-time graph.

The instructor raised issues which were intended to get the students to look closely at various details or to operationalize in terms of features on the graphs ideas students had articulated. When they first encounter the acceleration-time graph (Figure 5) the instructor raised the issue of what it appears one has to do to remain on zero on the velocity-time graph and the acceleration-time graph and what you have to do to stay off of zero on the two graphs. He also asked if it would suffice to give exactly the same answer for the two types of graphs. In another example, when the students suggested that acceleration apparently had something to do with “how fast” the velocity was changing, the instructor would ask: “What would two different constant accelerations look like on a velocity-time graph?”

2. Blurring the distinction between lab and class

The goal was to generate and settle as many unresolved issues and queries as possible in the class discussions. Making time in class and lab makes this possible, but in and of itself this does not resolve the related issue of interruptions and mismatches at the transition between full-class meetings and lab.

In order to reduce the problem of “hanging” issues when the class switched from full-class meeting to lab and back, the distinction between lab and class was blurred. What’s the problem with “hanging issues?” They leave students to believe they are not good at science since they do not “get it” before the class moves on. When an issue remained that seemed necessary to achieve a greater class consensus and clarity and which appeared to be a kind of conceptual precursor to what they were going to see in lab, but which was as yet unsettled, then each of the lab sections began as a continuation of the full-class discussion. If there things yet to be seen in lab that would support the on-going discussion in the full-class meeting or if there were “experiments” deemed necessary by the class during a full-class discussion, then we did them in Interactive Lecture Demonstration (ILD) mode in the big hall where full-class meetings were held.[1] There are issues of technique to facilitate these alternative uses of lab and the full-class setting.

In the case of lab we face the problem of constructively continuing the discussion in lab when from 10 minutes to 28 hours and 10 minutes have elapsed since the discussion in full-class ended, depending on the lab section. The method used was to engage the students in the first two steps of the PIPS cycle: “What do you think?” and “What does your group think?” within their lab groups with the assignment of preparing a “poster” to report to the lab class their position on the issue at hand. The “poster” could be prepared on a large sheet of paper, on a whiteboard, or a portion of the blackboard in the room. This appears to refresh many students to the mental state they were in during the full-class meeting. The ensuing lab class discussion often makes such progress toward resolution that convinces one the course really *should* be run in strictly lab-sized classes. The discussion having reached some sort of resolution for more students it is expected that the lab activity then is more relevant to the development of the ideas.

In the case of moving back from lab to the full-class meeting with some of the phenomenon as yet to be observed there is a two-fold problem. One is enabling everyone to experience the phenomenon. We can do any of the force or motion activities in the lecture hall and everyone can see the computer output as we have dual, large screen projection of the computer output. Students are engaged in helping with the demonstrations as most require more than one set of hands, but we cannot have every student release and catch a cart on each demonstration as is possible in lab. On the other hand we can entertain and test suggested alterations to any of the things we try.

The other problem of lab in the full-class meeting is keeping everybody aware of what we are trying and why. This is crucial. It is important that everybody be able to see and experience what happens, but all is lost if the experience is not tied into the *elicit, compare, resolve* cycle involving the students’ own conceptions. This is where the ILD strategy[1] comes into play. While not exactly the same, in spirit it follows closely the strategy of the cycle we use in PIPS to focus the students’ attentions on their own ideas with respect to the details of the demonstration of the phenomenon at hand.

3. Self-image issues

In this work a particular position with respect to the nature of learning, coming to understand, was taken. At the heart of the process of an individual constructing new ideas is the recognition by that individual that some aspect of their existing ideas leaves something to be desired. The natural response to this state of affairs is a response to resolve the discrepancy. This notion is the foundation of the PIPS cycle, the *elicit, compare, resolve* and *apply* cycle, and the steps of the ILD approach described above. Some readers will note that this, in different terms, is Piaget's description of the formation of new mental structures involving equilibration, accommodation and assimilation.[3]

Students who have been taught that they are likely to be wrong (in an absolute sense) because they are not "good" at science are at a severe disadvantage in that their natural learning processes are blocked at the first step. If one is just naturally inclined to be wrong, then this is a natural state to be accepted. To attempt to "get it right" in addition to being ill fated is unnatural and unnecessary because one can always find out from a really smart person. This is yet another reason for the claim that the most prevalent outcome of science "education" is a rude violence perpetrated on the majority of students in spite of the best intentions of teachers.[4] It is a monumental task to counter a lifetime of indoctrination in less than one semester, but to recognize the situation is to be responsible to take some action.

In this experimental course a thread already in PIPS Motion was amplified. The thread is an attempt to resituate the implications of predictions that do not come true. At *every* opportunity were the standard initial view led the students to make a prediction that did not pan out, the instructor took the opportunity to help the students see that their predictions were perfectly consistent with their meanings for the words they were using. Given the chance to interact with each other as they are in the PIPS cycle, the students generally do come to a logical conclusion. An explanation for the discrepancy between prediction and outcome in a number of cases in the activities is a difference in meaning for the words used by the students and the meaning intended by the authors of the MBL software.

For example, in the discussion over Motion Activity M2.4 to predict what will produce the given velocity-time graph, many students ask: "How can we get a negative velocity? How could a velocity be less than zero?" We stopped to analyze the reasoning involved. In the discussion students were able to articulate that the questions implied the assumption that the signs had something to do with the size (magnitude) of the velocity; a notion that seems to have its origins in elementary school.[5] Based on this realization they decided that either there could be no negative velocities or the signs must mean something else. Very quickly after they were allowed to move in front of the motion detector with the velocity-time graph running, they came to a resolution of the issue that they had first made clear for themselves. The signs refer to directions. Once resolved the issue could be described as a difference in the meaning being used for the signs.

Another example in the PIPS Motion materials comes up in their first encounter with acceleration-time graphs. The students' initial meaning for the word acceleration is profoundly different than the one used to make the acceleration-time graphs produced by the MBL system. Again, to shunt the tendency of the students merely concluding here again they are wrong and shutting off their sense-making processes to wait for some explanation from the instructor, exercising the dependency learned in previous science classes, an example is used to suggest that a different meaning for acceleration is being used and that their predictions are perfectly consistent with their existing notion of acceleration.

In analogy to acceleration, it is pointed out that the same word, biscuit, is used in both the United States and in British Commonwealth countries such as Canada, but that it does not mean the same thing in both settings. What you get for a biscuit in Canada is what we call a cookie. Who is wrong about the word biscuit? Clearly, no one, it is not a matter of right and wrong.[6] There are two different cultures that happen to have two different meanings for a particular word.

The notion of cognates from two different languages is illustrated with examples such as "ciencia" in Spanish and "science" in English. Then, notion of false congnates is explained with the examples of "embarasada" in Spanish and "embarrassed" in English.[7] It is suggested to the students that "acceleration" in everyday language and "acceleration" in the acceleration-time graphs are another example of false cognates as are "biscuit" at high tea and "biscuit" in the morning in the U.S. With the two versions of acceleration, the difference is not in two languages but in the way language is used differently by two sub-cultures. It is not a matter of right and wrong, but a matter of different meanings for words. We can all figure out the meaning of new words, given enough context. We have all been doing this since birth. It is how we first learned language. Just as we generally do not reject the "cookie" offered to us as a "biscuit", instead we take a bite of the cookie to see what kind of cookie it is. We should strive to figure out what this new and different meaning for "acceleration" might be.[8]

An attempt was made at letting the students know they are important enough to learn their names. This is not an easy task for most of us and there are limits to one's capacity, but for this number of students there are ways of accomplishing this in reasonable time. A piece of writing is collected in each class meeting. This is used to generate evidence of each student's presence in class and to gain feedback on what they were thinking each class period. These items are handed back in lab sections individually to each student. In addition the instructor personally records attendance in each lab section. This means that each lab period the instructor interacts with each student, by name, at least three times. In a few weeks time the instructor generally can refer to students by name in the large class meetings. This makes it much harder, although sadly not impossible, for students to feel anonymous and resist engagement.

In their previous "education" the students are trained that there are particular "right" answers to every question. When a teacher asks a question, student responses are considered until the "right" one comes out, the one the teacher wanted, then the process moves on to another question. This fosters a guess-the-answer approach which takes the students' attentions off of their own conceptions and directs it toward "clues" the teacher strategically arranges before them. Obviously, if the teacher "goes on" with the same question after hearing the response of a student then that student's response *must* have been wrong. This is extremely incompatible with real progress at making sense of the world. Standard procedure in the course is to attempt to generate a list of all the reasonable possibilities when an issue or question comes up. With this list a discussion can ensue in which the pros and cons of each possibility are considered. Because the initial generation of the list is often misconstrued as "not hearing the right answer yet," the issue of the difference between the typical instructor behavior and that of the instructor in the course is made the object of discussion several times at the beginning of the semester and reminders are given periodically during the rest of the semester.

4. Sharing the rationale with the students

In each of the semesters of the modified PIPS instruction at the first full class meeting after the acceleration-time graphs are encountered for the first time, the only formal presentation was delivered by the instructor. In this presentation, it is pointed out at this early point in the unit that as a community of learners in this course we have as yet to figure out what this acceleration means in these graphs, but it is clearly not the same as the meaning for the word we all brought to the course. Figures similar to Figure 2 in part I and Figures 2 – 4 in part II in these papers were shared with the students along with brief suggestions as to how they were generated. The presentation made use of the clear possibility of two different views of acceleration *already* encountered by the class at that point in the semester and suggested there also may be different views about the nature of force without getting into any specifics either with respect to the views of acceleration or of force.

The intent of the presentation was to point out that there is evidence to believe that figuring out what is meant by acceleration in the MBL graphs is key to figuring out a new view of force. Instead of just suggesting that acceleration would be useful later, the effort was made to share with the students the basis for the belief while acknowledging that the details would be clear later with hindsight. It was expected that a number of the students would respond with greater diligence if they were treated with respect by having this explanation shared with them. At the end of the semester some students voluntarily reported doing so.

IV. Results

The modified version of the instruction was repeated two semesters in the same course as before (Fall, 2000 and Spring, 2001). The same diagnostic, the FMCE, was used before and after instruction.

A. Thornton's Conceptual Dynamics

Table 3 shows the results of Thornton's conceptual dynamics analysis on not only all four semesters described so far, but also two others in addition. In one of those semesters, Spring, 1996, the PIPS materials did not exist. In the Spring of 1996 the students used the TST Project kinematics materials and the Minstrell-inspired force materials from which the PIPS materials were developed. The Fall, 1998 results are from a semester in which the first draft PIPS Motion materials were used, but the topic of force was not one of the topics taught. Tables 3a,b, and c have the same design as Table 6 in part I. The first four semesters listed in Table 3a are duplicates of the data in Table 6 in part I for comparison.

From Table 3a it is fairly clear that significant numbers of students moved through the conceptual positions described by Thornton's conceptual dynamics when the topic of force was taken up. When force was not treated as in the Fall of 1998, very little change in conceptual positions about force happened, as might be expected. The pre-PIPS semester, Spring of 1996 appears to have similar results to that of the normal PIPS instruction semesters, Spring and Fall of 1999. Comparing these latter two semesters with the two semesters of the modified PIPS instruction, Fall of 2000 and Spring of 2001, reveals that in each motion category, the modified PIPS instruction appears to have more thoroughly cleared the earlier positions in favor of the later positions compared to the normal PIPS instruction.

Comparing the normal PIPS semesters, Spring and Fall of 1999, with the modified PIPS semesters, Fall, 2000 and Spring, 2001, in Table 3b looking at the performance of the students with high acceleration scores at the end of the semester ($a > 6$ out of 8) reveals a slightly more complete clearing out of the earlier conceptual positions by students in the modified PIPS instruction. This is especially the case in the decreasing velocity category. It is also important to note that in the modified PIPS instruction, the number of those doing well on acceleration was noticeably higher. The results in Table 3c reveal a similar trend between types of instruction among the low acceleration scorers (a < 4 out of 8 at the end of the semester). It is important to note that the numbers doing poorly at acceleration remain roughly the same to ever so slightly less. A clear distinction between those who do poorly at acceleration and those who do well is the number who remain in the earlier conceptual stages of Thornton's conceptual dynamics.

B. Scatter Plots of New vs. Old View of Force

Figure 6 shows the scatter plots of the high and low acceleration performers, respectively, before and after instruction for both the Fall, 2000 and Spring, 2001 semesters during which the experimental instruction was used. Display of such data is difficult as three dimensions are really needed. Here as in previous figures the size of the circles ("bubbles") is a rough indication within the graph of the relative number of students falling on the same point in the graph. The sizes of the bubbles are not to the same scale as one moves from graph to graph. To aid the reader the number of students at a point represented by the "bubble" is printed to the right of the bubble.

It is also instructive to consider the average positions for the distributions in these figures with respect to each other and with respect the different semesters represented. The scatter plots for the Spring, 1996 semester, the pre-PIPS semester of data are not included, but the averages positions of distributions from that semester and the other four in the study are included as well in Table 4. Each column labeled "new" contains average new view scores out of a maximum of 15. Each column labeled "old" contains average old view scores out of a maximum of 15. The column labeled "N" contains the total number of student with matched pairs of data each semester. The columns labeled "N%" give the percentages of students in the relevant acceleration score category.

One slight possibility is suggested by the data in Table 4. A comparison between the average starting positions (pre-instruction) reveals slight differences. The low acceleration scorers' averages are always slightly weaker in old view score and generally in new view score than the class average and the high acceleration scorers' average. The high acceleration scorers' averages are generally although not always slightly higher than the class average and always higher than the low acceleration scorers' average. Hence, part of the difference in performance of the two groups by the end might be explained by the notion that the low acceleration scorers' conceptions have to organize into the old view sufficiently before moving on to the new view. The differences in average positions are very small, but it might suggest a difference in developmental starting point for the various groups, but the argument must be considered weak in the light of the changes made in numbers of students in the categories.

Over the period of the studies the high acceleration scorers went from 26% of the class to 44% of the class and the low acceleration scorers went from 41% of the class to as low as 20% – 24% of the class. It appears that for these students if such a developmental difference as described in the previous paragraph is present, it is compensated at least for some via appropriate changes in instructional approach. We have demonstrated that adjustment in instructional approach has compensated for the difference in the case of the order of 50% of these particular students. Whether further adjustments can result in the remaining students changing their conceptions remains to be seen. Given that these are the sort of students previously written off as not capable, because "after all physics is hard," it should not be surprising if further instructional approach adjustments are found to even further reduce the numbers of low acceleration scorers. There may be factors outside the influence of the instructor which account for the relatively small change in some of the remaining students.

Figure 6 reveals the same sort of bunching of the high acceleration scorers at the end of the semester up in the upper left of the plot as we might like all students to be doing. If anything the bunching is to a greater degree in these plots from the modified PIPS instruction than it was for the normal PIPS instruction.

One of the data points in the Fall, 2000 scatter plot in Figure 7 demonstrates again that a student must actually engage in the instructional process on force in order before we can expect that student might change views about force. There is a post-instruction point on the graph at the position: (15 old, 0 new), yet this was a high acceleration scorer (8 out of 8). As it turns out in the last five full-class meetings that semester we videotaped the discussion. There is visual evidence that this particular individual slept through all five of the taped class sessions. Given this fact and the conceptual dynamics data (Table 6 in part I) from the Fall, 1998 semester in which force was not taught, this particular student's performance on the diagnostic comes as no surprise.

Compare in particular Figure 7 with Figure 3, both from part II, the low acceleration scoring students from Fall, 2000 and Spring, 2001 as opposed to the low acceleration scoring students from the Spring, 1999 and Fall, 1999 semesters. It appears that the students in the modified PIPS instruction in Figure 7 who still did poorly on the acceleration questions, made more headway at trying to make sense of force. There is not as much overlap between the pre and post instruction "bubbles" in Figure 7 as there are in Figure 3. The post instruction distributions have shifted. This is also illustrated by a greater change in average positions of the distributions discernable in Table 4.

C. Normalized gain in new view vs. Normalized loss of old view of force scatter plots

Figure 8 shows the scatter plots of <g> vs. <L> for each of the semesters of modified instruction. Again, the reader might want to compare Figure 8 with Figure 4, looking in particular at the low acceleration groups. This different view of the data again shows both the high acceleration scorers and the low acceleration scorers making greater changes in the modified PIPS instruction.

The large positive values of <L> represent students *gaining* in old view. While such students clearly have gained very little in new view, coming to think more clearly and explicitly in terms of the old view is a kind of improvement which might be necessary before these students are in a position to make the next "move."

In Table 5 in part II we have a listing of the average normalized gains and losses for the whole class, the high acceleration scorers ($a > 6$) and the low acceleration scorers ($a < 4$). Again, caution should be urged in comparing gains on the FMCE such as we have here and gains on the FCI reported elsewhere. There is a clear increase in all categories of the magnitudes of the gains and losses when we compare the standard PIPS and the modified PIPS instruction. Many of the high acceleration scorers are close to "topping out" on this diagnostic.

ALL of the instruction reported on parts I & II, including the pre-PIPS Spring 1996 class, was conducted in a far more *student understanding-driven* manner than is typically available. Yet, each move to make the instruction more *radically student understanding-driven* has resulted in more of the class giving evidence of conceptual change concerning acceleration and force. The number of high acceleration scoring students has nearly doubled and the low acceleration-scoring group has nearly halved. The low acceleration scoring group of non-science majors has gone from essentially no change as indicated by the normalized gains and losses to normalized gains and losses which as we shall see surpasses traditionally taught science and engineering student class averages. One might speculate that this group is now trying, but trying to take the course as a traditional one whereas before their predecessors were convinced they could not do well and hence were not trying very hard.

D. Effect Size

Table 6 in part II shows effect sizes in change in view of force by full class, high acceleration scorers, and low acceleration scorers for the pre-PIPS semester (Spring, 1996), the normal PIPS instruction semesters (Spring & Fall, 1999) and the modified PIPS instruction semesters (Fall, 2000 & Spring, 2001). In this representation of the results, we can again see that the performance of the whole class, the high acceleration scorers and the low acceleration scorers is roughly the same for the pre-PIPS instruction and the normal PIPS instruction. Yet, the pre-PIPS instruction at least for the semester of data available appears not quite as effective at moving as many students on the issue of acceleration and consequently force. It certainly is intended that the normal PIPS instruction be well suited to this particular type of student in this type of course. The data suggests a great deal of success at this goal. Hence, even though the TST activities were being used somewhat out of context, they were able to aid the students to a whole class effect size of nearly 2 standard deviations. Yet, the normal PIPS instruction enabled more students to be successful at acceleration and thus at force.

A close look at Table 6 reveals that while the whole class effect size in the pre-PIPS semester is in the same ballpark (see Table 7 in part I) as that of the normal PIPS semesters and the class average position at the end of instruction on the scatter plot is also in the same ballpark as the normal PIPS semesters, the number of students not yet successful at the acceleration questions is noticeably higher in this pre-PIPS semester. This might be explained by the fact that the TST materials were not really developed for the specific purpose for which they were used in the course under study. The TST activities were originally intended to be a “non-invasive” intervention in lab setting that could be introduced into the laboratory portion of typical courses for science majors without requiring the instructor of the lecture section to change behavior. Nonetheless, the use of the TST materials “as they are” still enable a whole class effect size of nearly 2 standard deviations in this course for non-science majors centered around the activities themselves. Given the results one might argue that the normal PIPS materials are slightly better suited to courses such as the one under study with non-science majors, but not necessarily strongly.

Table 5 also reveals an effect apparently due to the modified PIPS instruction. This modification appears to enable the class to extend the effect size by another half standard deviation. This is apparently accomplished by a somewhat stronger performance on the part of the high acceleration scorers, but more importantly because 40 – 45% of the class becomes high acceleration scorers. It also appears that the low acceleration scorers though reduced in numbers perform somewhat better on force under the modified PIPS instruction.

V. Discussion

A. Maximizing educative aspects of schooling.

If we believe that teachers go out and teach as they were taught, then it makes sense for us to consider just exactly what we might want them to do for the children in schools and then try to do the same for the teacher candidates before they leave us.[9] We should decide that we wish the teachers to *educate* the children.

What might the goals of an *education* be? What distinguishes an education from training is that an education should leave children with the ability to genuinely develop for themselves new understandings of the physical world around them. Not only should it leave them with the ability, but it should also leave them with the confidence that they know how to do this and can when they need to. It is utterly reprehensible that if science "education" does anything to any significant degree for students, *as it does today,* it appears to do exactly the opposite for a vast majority of the students.

Of what should *education* consist instead? The teacher candidates cannot have developed the requisite skills and beliefs about themselves without having participated in the process on numerous occasions, working at times individually and at other times collaboratively. In so doing they should have developed many new notions of the physical world and be able to show evidence of this. There is no way that teacher candidates can have the necessary experiences when the courses they take in the content area are watered-down survey course versions of the selection and training, *content-driven* courses for scientists and engineers that currently pass for science "education" today.

Such an *education* cannot be based on a rationale meant to justify training. It will not do to start with the notion that automatically some might not be able to make it.[10] It does not instill the positive self-image we desire if we take the attitude that the students will never know something unless we tell them. Such an *education* cannot be more about making sure students know some designated subset of some official professional canon of knowledge as given than it is about cultivating the beliefs and habits-of-mind of an *educated* person.

What evidence do we have concerning the outcomes science education now? On every significant topic that has been studied by physics education researchers, there is clear evidence that in general most students leave school with essentially the same notions about those phenomena they had when they came to school.[11] Most students leave their school experience with the clear notion that physics is not something they really understand and the notion that only really smart people can understand physics.[12] Their school experience cannot be classified as an *education* in the sense described above.

On the other hand we do produce scientists and at the graduate level of schooling, the world comes to us. We must be doing something right. Our record of Nobel prizes continues to support this thesis, as does the presence of many foreign students in our graduate schools majoring in the sciences. The issue is what are we *really* doing?

Recently, Goodstein[13] questioned how a science education system could produce such dichotomies. It is fairly clear how this could be. We do not have an *education* system. What we have is a *vocational filtering and training* system, a production system for elites. Since we are really talking about selecting and training a very small percentage of the population, it is easy to see how we ignore or dismiss the effect on everyone else. It is hard to see how we could have designed a better system for the vocational selection and training of physicists. Sadly, we have perfected a system which short changes the rest of society. In so doing, it is possible that this discounts a significant fraction of the good physics as a profession might otherwise be doing. As long as we can identify the problem and imagine a solution we have the potential of doing better and the obligation to try.

In preparing elementary education teacher candidates we should not be giving them a one-semester version of a watered-down vocational filtering and indoctrination program for physicists and engineers. Such an offering can only reinforce what the teacher candidates have already learned from their science classroom experiences. This leaves the teacher candidates only prepared and indoctrinated to "teach" in a fashion consistent with this experience, which is tantamount to drafting them into the filtering and indoctrination enterprise without giving them a choice. We have no right to do this to the elementary education teacher candidates and it is certainly wrong in its potential effect on their students. It is hard to imagine elementary school teachers, given an informed choice, choosing the option to which they are restricted at present.

B. The problem with the typical approach and attitudes about it

What stands out most in the typical approach is that instead of being engaged in making sense of the world, students are being told someone else's view. The implication, more importantly the lesson learned, being that they would never come up with such a view on their own and that this official view is so important that engaging students in making their own sense can be dispensed with.[14]

Each step in the presentation is the next piece of the physicists' present official picture of the world, regardless of whether the physicists sorted it out in this order or not. It is presented this way because it is considered to be the logical and efficient way to explain the physicists' view. This is the *goal-driven* approach at its worst. It has become assumed teaching lore that students could never come up with something resembling the physicists' view on their own, that one cannot "know" acceleration until one "knows" velocity first, and anyone who cannot "get it" when it is presented logically probably is not capable of "getting it." Hence, the sentiment: "How can they know if we don't tell them?" often expressed as a display of genuine concern for the students and the subject, turns out to be a call to maintain the training and selection status quo. "How can they know if we don't tell them?" *is not* a call to *educate* however sincerely it might be intended to be so. The new views developed by students, examples of which are Tables 1 and 2 in Part I, demonstrate the utter groundlessness of this elitist assumption.

C. Why and how to teach kinematics?

We can imagine there are multiple possible reasons for teaching kinematics in different situations calling potentially for different methods of instruction. The traditional reasons and methods seem to have worked in their way successfully for the few people who have successfully completed the training to be physicists and engineers. Sadly, since these traditional methods were also perpetrated on everyone else, many for whom the methods were not successful not only failed to "learn" the kinematics taught, but they were cognitively and psychologically damaged in the process.

If we *must* use the traditional reasons and methods for the training, we should reserve their use for whom they appear to work, those who have already decided they wish to be physicists and engineers. The PIPS materials and approach offers an effective alternative. We should be using PIPS-like materials and methods at the pre-college level and with non-majors at the university.

In the example of the present work, three factors served as the foundation for pedagogical decisions about the teaching of kinematics: 1) the intent to produce an *educative* experience (the essence of a core course) as opposed to a training experience, 2) theory explaining how and why people formulate new ideas and data indicating their initial ideas, and 3) the intent to engage the students over their ideas concerning physical force and its relationship to motion. Not only was the teaching of kinematics called for, but also what is conceptually important was indicated. The method of instruction was consistent both with the learning theory and the desire to produce an educative experience. Pushing these factors to the extent they were pushed in the modified PIPS instruction can routinely achieve an additional *half standard deviation* larger effect size beyond the impressive *two standard deviations* already possible with the standard PIPS instruction.

It is important to consider the effect on the students beyond what they "learn for the test."[12,15] It is clear that the traditional methods are not satisfactory for the majority of students.[11,16] *Is it really the best for physics and engineering majors to learn this elitist view of the world?* Does this really put them in the best position to enable the professions of physics and engineering to serve society? Regardless of one's answers to these two questions with respect to future physicists and engineers, it should be clear that the traditional methods damage the majority of the students. The majority do not experience any lasting change concerning their ideas about the phenomena, but they do experience lasting negative change in their views of themselves and others. They learn that they cannot make sense of the phenomena themselves but are dependent on so-called, special people, supposedly smarter than them to do so.

This belief about themselves gives them and their teachers the excuse not to try. It is easy to argue that this is a major reason why conceptual change research of the last 25 years that started about a decade and a half after the big U. S. national effort to improve science education in the late 50's has consistently shown that science instruction typically results in no significant change in student understanding of the phenomena studied. This not necessarily to suggest that the curricula developed such as PSSC Physics by the Physical Science Study Committee and others were not good, but *how could any curriculum be effective given the prevailing view that only certain students can really learn physics?*

In the present work it is clear that kinematics can be taught with a focus on its intellectual challenges: acceleration as any change in velocity and both velocity and acceleration as vector quantities. Taught in a *student understanding-driven* approach such instruction in kinematics can constitute an *educational* experience. Data presented in this work suggest that if force is to be a topic of consideration then such an educational experience in kinematics involving a focus on acceleration is clearly in order. It is apparently not so necessary for all of the details of position and velocity to be taught directly in order that students be able to figure out what is meant conceptually by acceleration and to develop a new view of force. Furthermore, in this work it is demonstrated that the very students typically "written off" are the ones who have risen to these challenges.

D. Is the approach transportable?

There are other examples, but one may suffice. In the Fall of 2001 a high school physics teacher with six years experience used the standard PIPS materials in his classroom for the *very first time*. This teacher characterizes his previous teaching as based on the transmissionist model. In the terms of this work one can describe his previous teaching as *content-driven*. He was dissatisfied with the results and was searching for an alternative. The PIPS materials include extensive instructor notes about the cognitive and affective strategies intended. The teacher also had access to other publications by the present author describing features of the approach.[17,18] The author and this teacher have never actually met face-to-face, but corresponded occasionally via e-mail during the time the teacher used the PIPS materials and since. He had two classes of high school physics students numbering approximately 25 students each. The students were about equally divided female and male, juniors and seniors. The two classes are distinguished by the time of day they were met by the instructor. This work by the high school teacher was in part to satisfy requirements for a master of science in education degree at a university in his state.

First, it is important to note that the pre-instruction diagnostics again speak volumes. Here we have some of the top students in the school as seniors. They all had physical science as junior high students, yet the diagnostic given at the beginning of their senior year had the following averages: person-on-the-street view of acceleration (an old view) = 6.8 out of a possible 8, scientist-like view of acceleration (a new view): 0.20 out of 8; person-on-the-street view of force (old view) = 11.0 out of 15, Newton-like view of force (new view) = 0.74 out of 15. These results are not particularly different from those of the non-science majors of the main study. They indicate that a view not available to the students "on the street," a view which should have been part of the content that drove their previous instruction, was not present to any significant degree in the students' thinking as they answered the pre-instruction diagnostic. For all of the good will and sincere intent that may have been on the part of their junior high teachers, the result of the instruction these students actually received is as if it had *never happened* at least with respect to these understandings. In fact the teacher reports surprise at the students' pre-instruction performance in the high school class because he was the physical science teacher of some of them and thought he knew what they did in his class.[19]

Figure 9 shows the scatter graphs of the views of force for the pre and post diagnostics. Because there may be more than one student at any point the arrows show the average positions of the distributions. Table 7 shows effect sizes and normalized gains and losses for the two classes. The graphs in Figure 9 and data in Table 7 look like the graphs and data for the high acceleration scoring students from the main study. This is not surprising because as one can see from Table 7 on the order of 75 – 80% of the students in each class were high acceleration scorers and only 2 (about 8%) of the students in each class were in the low acceleration scorer category. Notice that the conjecture made in section II.A. is supported here. A large majority of these students scored very high on the acceleration questions and we see effect sizes greater than 5 and normalized gains of nearly 0.9

Considering the results in Figure 7 and Table 9, this performance is one of which to be very proud. The class size was conducive to the process and the classes had the option to move between discussion and lab activities at will. There is evidence that small class size and the ability to move at will between discussion and lab activities makes a difference in the performance based grades of the college level course which is the object of the main study. Three times the course has been offered in a format of blocked time all in the lab room with the enrollment limited to 24 for that particular section of the course. In the large class format the grade distribution usually centers on "C" with slightly fewer "D's" than "B's" and slightly fewer "F's" than "A's." The small section met for blocks of time all in the lab room. In three different sessions of the small class version there were a total of 2 "D's," zero "F's" and the distribution of grades centered on "B." Apparently the greater the individual interaction between students and between students and instructor the harder it is to avoid engagement. Once students are engaged the chances of conceptual change are great.

These results suggest that the method is indeed transportable. The success of the transfer is attributable to several factors. First and foremost, is the desire of the teacher for more satisfying results, his willingness to try another approach and his understanding of the principles behind the method. Another factor is the favorable teaching conditions: small class size, no interruption between full class discussions and lab. The third factor is possibly the student sample. Generally one expects the high school physics students to be more motivated, have a positive self-image, especially with respect to science, and have better study skills than one might find on the average in a large university class of non-science majors taking a science course because they have to in order to satisfy a distribution requirement for graduation.

E. *What are typical results for a traditionally taught class?*

1. Algebra-Trigonometry Level Introductory Physics Courses

This level of introductory physics is generally required of science majors (not physics, chemistry, geophysics or engineering), pre-health professionals, *including pre-medical*, kinesiology (physics education) majors, and sometimes biology. Among the ranks of these students there are some who have had high school physics, but it is not considered a prerequisite for the course. The traditional instruction generally includes three lecture hours a week and a two or three hour lab conducted to teach lab skills and verify the descriptions of the phenomena from text and lecture. Sometimes, for certain of the students, lab is optional. This course is generally a two-semester course. Some of the majors are only expected to take the first semester.

The course is oriented toward problem solving using applications of mathematics to physical situations, but the mathematics in text, lecture and problems generally does not go beyond algebraic methods, up to the quadratic formula and simultaneous linear equations, and right-angle trigonometry. Homework consists of solving such problems and exams generally consist largely of the same activity. It is generally asserted that one learns via problem-solving and one demonstrates what one has learned via problem solving. In an ideal world this might be the case, but in the real world it is not as we shall see when we look at the data below.[20]

The data comes from two different institutions and a dozen years apart. The more recent data is from a large state university in one of the “prairie” states designated “Prairie State Public University.” The older data is from a large west coast state university. Because there is data from the calculus level introductory physics from another west coast state we will designate this data as being from “West Coast Public University A.”

The version of the FMCE used at Prairie State University was the version published by Thornton and Sokoloff[1] in 1998. The version of the FMCE used at West Coast Public University A did not include the same acceleration questions as the current version. As a result it was not possible to make the classification of high acceleration ($a > 6$) scoring and low acceleration scoring ($a < 4$) students.

There was another difference between the version of the FMCE used at West Coast Public University back in 1990. From the first 21 items, items 1 – 7 & 14 – 21 exactly matched the published version.[1] Items 8 – 13 involving gravitation matched except the set of choices available to the students was actually less challenging as it was less detailed.[21] Choices clearly consistent with the new and old views are easily identifiable and the actual choices students made were consistent with these. Calculating scores to determine effect size and normalized gains and losses gives results that can be directly compared with the results of the later version of the FMCE and a scatter plot of the pre and post instruction distribution of scores can be made for comparison.

2. Calculus Level Introductory Physics

This course is typically required of physics, engineering, geophysics, and chemistry majors. Sometimes pre-med majors take it. Generally the percentages of students in this course who have taken high school physics is somewhat higher, but it is still the case that at many institutions high school physics is not a prerequisite to enter this course. Traditional instruction consists of 3 to 5 lecture hours per week and a two or three-hour laboratory session per week. Sometimes one of the lecture hours is set aside as recitation or problem solving question and answer session in smaller sections. Lab is generally not optional and is intended to teach laboratory skills and verify the assertions made about the phenomena in text and lecture. At different institutions this course is from two to four semesters long.

The course is oriented toward mathematical problem solving using applications to physical situations. The mathematics in text, lecture and problems generally makes use of elementary differential and integral calculus and advances in mathematical sophistication as students are generally expected to be taking advancing levels of the calculus as they take successive semesters of this level introductory physics. In practice many of the problems assigned can be solved using the same level of algebra and trigonometry expected in the other level course. Homework consists of solving such problems and exams generally consist largely of the same activity. The same myths concerning the role of problem solving are even more certainly applied to this level course.

In both of these types of courses kinematics and force are studied in the first semester of the course. All of the data from these traditionally taught classes is from the first semester of the respective courses.

The data for the traditionally taught, calculus level introductory physics comes from three institutions: a large northeastern state university designated "North East State Public University," a large west coast state university designated "West Coast Public University B," and a smaller private university in a west coast state designated "West Coast Private University."

3. Evidence of Previous Instruction on Physics Topics

At the North East State Public University, it is significant in this context to note that this student who happened to score an 8 out of 8 on the acceleration questions at the post-test also scored a 7 out of 8 on the pre test. This person's new view of force score was 12 on the pre-test and 15 on the post-test. No one else in this group scored higher than 4 out of 8 on the pre-test. This student's understanding did not come from the course, but from some previous experience.

At West Coast Public University B there is also some evidence of students who show some evidence of previous physics instruction on the pre-diagnostic. In each semester there were 5 to 8 students who scored higher than 10 out of 15 on the new view of force on their pre-instruction diagnostics. It is interesting to note that for most of these students like the one from North East State Public University, many also scored at least 7 out of 8 on the acceleration questions at the beginning of the semester. These students apparently did "learn" some physics in their previous instruction. Even more intriguing are 3 or 4 of these students who scored several points lower on their post-new view score. All of these latter students had a post-acceleration score less than 6 out of 8.

One might estimate that since about 25% of the high school graduates take at least one year of high school physics, maybe as much as 25% of the students in the algebra-trigonometry level of introductory physics and possibly more than 25% of those in the calculus level introductory physics might have had high school physics. The numbers of students who demonstrate in the pre-instruction diagnostics here that they pick choices consistent with a new view (essentially Newtonian) of force falls far short of 25%. It is also important to remind ourselves that essentially *all* of the students had force and motion in their junior high physical science course. Clearly that course had no lasting effect on student understanding of force or motion.

4. Results

Table 8 shows the results of the analysis. One observation immediately stands out. The class average results for the 7 traditionally taught courses from 5 different institutions across the country over a period of a dozen years are essentially all the same. In fact they differ generally by less than 0.25 standard deviation on all measures. For all intents and purposes they are the same.

For comparison figures for the PIPS instruction, both in the conceptual physics course and the high school, given at the bottom of Table 8 are the averages of 6 sections worth of data collected under that approach to instruction. First of all looking at the scatter plot average positions, we can see that none of the groups of students start at a significantly different state of understanding of force, even though all of them have had at least one round of instruction on motion and force in junior high and of the order of 25% have had high school physics.

We see that the effect size for the traditional instruction classes is nearly *2 standard deviations less* than the effect size for the modified PIPS instruction. Hence, the *radical, student understanding-driven* instruction meets Bloom's 2-sigma challenge even when taken as direct comparison to traditional, *content-driven* instruction. Furthermore, it appears that the effect sizes for *high* acceleration scoring students who experience *content-driven* instruction are only about the same as that for the *low* acceleration scoring students under *radical, student understanding-driven* instruction.

The normalized gain and loss for the traditional instruction is essentially insignificant. This might be okay if most students were already scoring high on the new view of force, but one can see that the typical average starting new view of force score is somewhere between 0 and 3 out of 15. In other words for all intents and purposes, in terms of this diagnostic, most students start with an essentially old (person-on-the-street) view of force and in the traditional instruction the science majors (including the physics and engineering majors) experience no real change in understanding of the nature of force. This is entirely consistent with the relatively small effect size. When we can see the students separated out by high and low acceleration scores, we see that more than 40% of the traditionally taught students (sometimes pushing 90%) make *no* normalized gain or loss in their understanding of force and apparently fail to understand acceleration as it appears on the graphs. Those with high acceleration scores make normalized gains and losses only about *half* that of their peers in the modified PIPS instruction. Furthermore, they do not even rise in this respect to the level of the *average* of the modified PIPS instruction. A comparison of the scatter plot average positions (the average pre and post, new and old view scores) yields the result that the whole class average shifts, pre to post, for the traditionally taught students are most similar to the low acceleration scoring students in the *student understanding-driven* instruction. These high acceleration scoring, traditionally-taught students, who appear to have learned an new view of acceleration have made disappointing progress learning a new view of force from their university physics professor.

We do not have enough information on the West Coast Public University A students' responses to the acceleration questions to make a direct comparison but it is revealing that their pre instruction average on acceleration is 0.8 out of 4 questions and the post instruction average was 1.4 out of 4 questions. It appears a fairly safe bet that few of these traditionally taught science majors would have done well at all 8 of the acceleration questions and many would be found doing poorly. We see for the other group of traditionally taught students at the algebra-trig level that 65% scored poorly on the 8 acceleration questions and this same sub-group scored an average normalized gain and loss that are essentially indistinguishable from zero; no conceptual change at all. In the other places in Table 8 where there were too few students it is significant to note that the traditionally taught, *content-driven* classes would fail to populate the high acceleration scoring category and the *student understanding-driven* instruction would essentially depopulate the low acceleration scoring category.

The results from the traditionally taught science majors support the assertion that conceptual change with respect to force is dependent on understanding a new view of acceleration. The striking aspect is how poorly traditional, *content-driven* instruction compares to the modified PIPS, *radical, student understanding-driven* instruction at these different large and small institutions and over a dozen years. These results support the both the assertion that we have been getting the same results for many decades and that there is resistance to change in teaching physics despite the problems being pointed out by the research in physics education community for now more than two decades.

If this traditional instruction is to be the time-honored standard by which some would have all other instruction judged, *it is a very low standard at which to aim.* We could easily attribute the poor initial showing to the results of people who might not have undergone significant conceptual change themselves (such as possibly pre-high school teachers) about force attempting to teach what they understand. *How do we explain university professors of physics, who we believe do understand this new (Newtonian) view of force, teaching so as to make essentially no difference in their students either?* The results of this study suggest strongly that the problem is the instructional practice, based on views of the nature of knowledge and the nature of learners adopted largely without critical examination. *Content-driven* instruction, *the presentation of the canon in the approved fashion for the deserving*, does not result in conceptual change about the phenomena nor does it leave most students believing they too can make sense of the phenomena. *Radical, student understanding-driven* instruction does both. There may be other practices that are effective,[22] but clearly traditional, *content-driven* instruction is not one of them.

We have had some who have looked into the PIPS instruction briefly and rejected it because there is not enough mathematics and no standard problem solving, suggesting that maybe this is okay for lower level, younger students, but certainly not for the serious physics students. In the context of such rendered opinions the data in Table 8 begs the question: What should we call a student who can manipulate high level formal mathematics but has no idea what the numbers mean? The same data suggests the answer should be that such a person is at best, as yet a *dangerous* technician. It is worth noticing again that the *status quo* is a mighty low standard at which to aim.

5. Is this just a case of difference in time on task?

There are several reasons for believing that the differences in performance are not exclusively due to a difference in time on task. In the examples of *student understanding-driven* instruction most of the semester is taken up by the units on motion and force. In typical *content-driven*, two semester introductory physics courses, these topics take up from slightly more than a third up to about half of the first 15 week semester. What is the evidence that this particular variable is not the primary factor in the difference in performance observed?

To test the issue for completeness sake one would need to look at two types of comparisons: (1) shorten the time spent on the *student understanding-driven* approach to match that of the traditional *content-driven* approach and (2) extend the time spent on the topics in the traditional *content-driven* approach to match that spent in the present examples of *student understanding-driven* instruction. As for the first type of comparison work has already been done that enables us to see and outcome. The second type of comparison poses at least some ethical problems. Under what circumstances would it be acceptable to someone, who believes the traditional *content-driven* approach is appropriate, to spend a whole semester with students in a real course on "just" kinematics and force in one-dimension? Without doing this second type of experiment, it may be reasonable to extrapolate from the first type of experiment. Some work very close in design to the first type of experiment has already been published.

There are at least two examples of using an approach consistent with the student understanding-driven approach for "brief" interventions. The original microcomputer-based laboratory activities from the Tools for Scientific Thinking (TST) Project, referred to in Part I, were the inspiration for the PIPS activities. The TST activities were devised for insertion in place of traditional laboratory activities in a traditional *content-driven* course without having to ask the lecturer to change behavior at all. In the first example of the use of these materials in brief intervention, we have these alternative activities, albeit pre-PIPS in design, being used as the lab portion of a standard course moving at the standard pace. In this work, Thornton and Sokoloff found "strong evidence for significantly improved learning and retention by students" who used the TST materials "compared to those taught in lecture."[23] The lab portion of such courses is about 35% of the weekly class time. In the second example of the use of these materials in brief intervention, the original TST activities were incorporated into a method called Interactive Laboratory Demonstration (ILD). The ILD was devised for situations in which students taking introductory physics courses are not required to take the laboratory portion of the course. The ILDs were devised to bring the essential features of the activities into the lecture portion of the class. The approach used in ILDs is very consistent with the underlying ideas behind the *student understanding-driven* approach. Again, Sokoloff and Thornton have tested the ILDs and found "There is strong evidence for significantly improved learning and retention of fundamental concepts by students who participate in ILDs as compared to those taught in traditional lectures."[24] These ILDs are used in a course that is proceeding at a normal pace.

In the two examples just cited we have evidence that when the time spent in activities consistent with the student understanding-driven approach is brief, comparable to time spent on task to the traditional content-driven courses, the result is "significantly improved learning and retention of fundamental concepts."[24] We could on this basis make the claim that approaches like the student understanding-driven approach can make a difference even when used briefly. The experiment to test whether the traditional *content-driven* approach could "catch up to" the results of the *student understanding-driven* approach given enough time has not been conducted and as has been pointed out, there are ethical questions surrounding such an experiment.

In very real ways, the answer to this last question is academic. The results of teaching in the traditional *content-driven* approach are loath to "go slower" as they put it, yet one can see the results in Table 8 are most of the students do not understand. It is in these courses, the algebra-trig and calculus levels of introductory physics, that nearly all teachers teaching middle/jr high school physical science and high school physics receive their terminal experience in physics content. An extremely small percentage of these teachers have any more physics in college than this. So, they come to the university having been taught about motion and force in 8th or 9th grade and possibly in high school physics, but these experiences leave their understanding essentially unchanged. Then they take a couple of semesters of physics at the college or university level and the experience again leaves their understandings unchanged. All of this instruction is in the traditional *content-driven* approach, that not only fails to induce any change in their understanding, but teaches them an elitist scheme as an excuse for why people fail to learn. They take a teaching methods course, usually given in the traditional *content-driven* approach in which they may *hear* about alternative teaching and learning strategies. Is it any great surprise then that the teacher candidates that graduate from our institutions go out to teach in the middle/jr high schools and high schools to use exclusively the traditional *content-driven* approach, teaching so as to leave *their* students' understandings essentially unchanged? Do we have any responsibility to their students for what we teach and the way we teach teachers as they are prepared by us to teach?

The plain fact of the matter is that there is plenty of instruction time between grades 1 and 12 to accomplish foundational conceptual understanding goals in all of the sciences. We fail to take advantage of all this instruction time, because we fail to prepare the teachers we train by example to make adequate use of it. We have clear evidence that the traditional *content-driven* instruction fails and we have examples of how to take advantage of this time to actually induce change in understanding in the students. Yet, there are those who insist on continuing this process that results in psychological violence against most who experience physics instruction and fails to induce any real change in understanding of the physical phenomena. This makes sense if the enterprise of *content-driven* instruction is *the approved presentation of the canon for the benefit of the deserving*, based in a *realist* view of knowledge and backed by an *elitist* view of people. Is this really any way to discharge our responsibilities to society as instructors or as physicists?

F. Can the results be expected to last?

It was not a part of this particular study to investigate this issue, but there is some evidence beyond anecdotal to suggest that such results can be expected to last. Francis, Adams, and Noonan conducted a three-year study of the modified instruction in the algebra-based introductory course at their institution.[25] They used a different diagnostic (the Force Concept Inventory[26], than the one used in this study. In their study the typical post instruction score was about 68%. Follow-up administrations of the diagnostic, up to three years after instruction gave typical average scores of 60%. To the extent that one can claim both diagnostics are similare measures of student conceptions, although not exactly the same, one can claim Francis and colleagues could expect to see similar results with either diagnostic.

VI. Conclusion

A. The value of a student understanding-driven *approach in PIPS*

Using the PIPS materials in the manner they were intended with a *student understanding-driven* approach, even under adverse conditions, can lead to routinely to nearly a *two standard deviation* effect size in the difference between pre and post instruction performance by non-science majors on the Newton's first and second laws portion of the FMCE. If we even further depart from the normal dictates of content coverage and let the students' apparent conceptual issues lead us to focus on the meaning of acceleration, we blur the distinctions between "lecture" and "lab", we attend to self-image problems brought on by previous instruction, and we explain why we have made these changes to the students, then under the same adverse teaching and learning conditions students can routinely make a *2.5 standard deviation* shift in how they respond to the diagnostic. This is all possible because principled efforts to understand the nature of learning in physics—physics education research, have been applied to the study of students learning topics in physics to result in *radical, student understanding-driven* instruction. It clearly outstrips the results of traditional, *conten--driven* instruction and not just by an insignificant amount. Given the common starting point all students in the study had on this diagnostic, the students who experienced *radical, student understanding-driven* instruction finished instruction with an average that ranks roughly at the 98th percentile of the post distribution in the *content-driven* instruction.

B. New ways of displaying data

A well-designed diagnostic made possible several new ways of looking at the diagnostic data collected in this study. These ways of looking at the data helped settle the conjecture that motivated the experiment involving the modification of the PIPS materials. The diagnostic used, the FMCE, concentrates on a small number of fundamental ideas. It asks several questions coming from "different directions" on each aspect or facet of the phenomenon as the students parse it. The choices provided enable students to pick answers that make sense to them even though several points of view are present in typical student populations. There are not many validated diagnostics that do this. More are needed.

The FMCE enabled the calculation of more than one score for a given student performance: one score for the degree to which the choices were consistent with an "old" view of force[27] and another score for the degree to which the choices were consistent with a "new" view of force. The "old" view of force is what might be called an everyday or person-on-the-street view. The "new" view of force can be recognized as a "Newtonian" or Newton-like view of force. The fact that the new view of acceleration is scientist-like is a consequence of the fact that the students were engaged in making sense of acceleration-time graphs as scientists would make them. The fact that the new view of force is Newtonian in nature is a consequence of the aspects of the phenomena to which the students' attentions were directed and that they were 'primed' by first making sense of the acceleration-time graphs. Two scores for each student's performance enables the plotting of scatter graphs which can be produced to view pre and post instruction distributions of the scores.

The existence of an "old" view score that we might expect to decrease as a result of the instructional experience makes possible a "loss" analog of sorts to the normalized gain <g> score calculated by Hake to display class average changes in FCI scores. Since it is intended that the "old" view decrease during instruction and it does decrease under the influence of instruction that actually has an effect on student understanding, then one can calculate a normalized loss <L> in "old" view. Again, these occur in pairs for each student and we have another type of scatter graph (a gain vs. loss or <g> vs. <L> scatter graph) to display the distributions of students.

C. In defense of the status quo*?*

Since its beginning, research in physics learning has been about the study of students understanding of topics in physics in order to improve the learning. If all that was required was an "arm chair" approach then the research in physics learning would not have grown to become a research area in the profession of physics.[28] It is clear from Hake's meta-analysis[29] that change in physics instruction is called for if the goal truly is more students learning physics. McDermott's Oersted Medal acceptance speech[16] contains powerful ammunition in favor of changes in physics instruction-driven by research. The research in physics learning community is diverse. As such, a number of themes can be seen in the work.

1. Against constructivist influences?

The present work is certainly consistent with a particular constructivist theme of the nature and origin of knowledge and of the process of coming to know. Particular views of the nature and origin of knowledge can be used to develop particular pedagogical practices and be used to guide decision making during the practice of such pedagogies. The pedagogies are not necessarily completely unique in that other beliefs about the nature and origin of knowledge might lead in some cases to similar pedagogical practices. Nonetheless the pedagogy practiced in the *radical, student understanding-driven* course studied here is consistent with the central tenets of what is called radical constructivism.[30]

Now it might be that the pedagogy studied here is an example of "constructivist pedagogy" referred to by Gross when he wrote:

> "In any event the, literature of science teaching has seen, during the two decades past, a flood of constructivism. Whatever interest these productions may or may not have as philosophy, the scores of papers, reviews, and dedicated journal provide little evidence that constructivist pedagogy has made or will make a difference in student performance."[31]

It might also be the sort of pedagogy referred to by Aldridge when he wrote:

> "Radical constructivism as used in science education is a peculiar relativism of scientific knowledge and a perversion of the common wisdom that when you teach someone something, you must start with what they already know. … If purveyors of constructivism knew the distinctions between empirical science and the theories or models we create to account for that empirical knowledge, their focus could legitimately be on helping students modify their preconceptions to better match those scientific models and theories scientists have created and tested."[32]

It is clear from the present work that such pedagogy can indeed make a difference in student performance. Furthermore, while we cannot tell what Gross believes about constructivism from his quote, we can see that Aldridge has a very different view of constructivism than is represented in the present work. As for traditional methods, we see from the examples above that the results of traditional instruction of science majors are not worth defending in comparison to the demonstrated results of the *radical, student understanding-driven* instruction studied in this work. Ehrlich waxes eloquent in his skepticism of physics education research-driven alternative instruction methods,[33] yet he conveniently ignores the state of student understanding after already experiencing traditional *content-driven* instruction at the pre-college level and worse the lack of increase in understanding demonstrated after their first college level *content-driven* instruction in the phenomena. Thornton and Sokoloff,[1] who developed the diagnostic used in this study, give us convincing evidence that traditional approaches defended by Gross and Aldridge compare only very poorly with the results reported here. The same conclusions can be drawn from the examples Hake[26] provides and those of Wittmann[34]. The inadequacy of the traditional approaches with respect to content is rendered *insignificant* in comparison to the damage done to students otherwise.[12]

We certainly cannot justify this treatment as exemplary teaching for our teacher candidates who will be teaching all students. Can we really continue to afford subjecting the physics and engineering majors to this *status quo* of instruction given that so few even "get" the physics even after what for a noticeable percentage of this is the *third spiral* in the curriculum? It is reasonable to assume that at least the same percentage (if not more) of the physics and engineering majors in our charge in the calculus level introductory classes have had high school physics as the percentage of high school graduates who have had physics, namely about 25%. The reader will note that four out of the six examples in Table 8 for which it could be calculated had *less than 25%* achieving high acceleration scores at the end of their semester of instruction. This amounts to a *third pass* in the traditional, *content-driven* "spiral curriculum," junior high physical science, high school, and now introductory college level. There is only *one pass* left at formal instruction before some of these students become practicing engineers. There are only *two passes*, upper division undergraduate courses and graduate courses, left before we unleash others of these students as physics professors on the next round of unsuspecting students. If history is any indication many of these physics professors will admit they never understood until they had to teach the material. Too bad their students have to wait until some future experience to "get" it too. It is no coincidence that in *one pass*, 44% of the non-science majors in a very large class setting and 75% in a small class setting achieve high acceleration scores and average normalized gains in the 90% range in the *radical, student understanding-driven* instructional practice. There is evidence that a noticeable percentage of students could have constructed a good understanding for themselves as of the first pass they all get in junior high.

2. Not the way we learned it…

There is no doubt that some who read this will be saying: "but the author forgets that he must have learned 'that way.'" The author is the first to express his gratitude to his teachers and the fact that he was lucky to have so many who personally had his and his peers' well being in mind. One of them was on the Board of Directors for the National Science Teachers Association and the author of nationally published textbooks, another was state Biology Teacher of the Year, and at least one received the Presidential Award for State Science Teacher of the Year. They all made a career of doing the best by their students they knew how at the time. From the author's point of view there is not enough gratitude that can be shared with these teachers.

Not to attribute any mal-intent on his teachers' parts, because this is clearly not the case, but the author was a student who was among the few identified as "good" in science. He was the beneficiary, if such can be, of the system and very much in the minority. If we continue to teach teachers in the fashion we have, inducting them into the filtering and indoctrination system, then the ultimately destructive cycle will continue for the majority of students. The *student understanding-driven* approach is a cycle breaking model for teacher candidates now in our charge.

As for other readers who will be pointing out that most physicists learned "that way," given the data on traditional instruction presented, it is clear that we really *did not* learn that way. There are at least two responses that can be made in addition. One is that most of us in fact forged our own path-driven by our own desire to understand, wrestling with and resolving dilemmas of our own making because we were the ones the process "taught" that we *could* do such things as make sense of this material. The other is that we physicists constitute on the order of 1/2% of the total population that experiences physics instruction during their lives in the US at least. For many of *us* in physics we had to experience 5 formal cycles of the material from junior high through grad school and then we had to try to teach it before we figured out many of the basics.[35]

When the results are pretty overwhelming that only some tiny percentage of the students experience significant change in their understanding[11] in normal instruction and at that only after multiple exposures,[36] *researchers* in physics education are strongly inclined to reject the premises underlying the instruction that produces such results. It is our experience that a physical theory with such results would suffer a similar fate. In the face of alternatives such as that so clearly demonstrated in the present work, the act of determined defense of the *status quo* instruction by those such as Aldridge, Cromer[37], Ehrlich, Geilker, and Gross becomes profoundly unscientist-like, yet they claim to be *scientists* defending a mode of instruction as the well-spring of science. Sadly instead it appears to be the well-spring of a very unscientist-like ideology which truly produces neither good scientists nor healthy members of society.[38] Unfortunately data does not have the effect on an ideology-bound mind that it does on an open mind.

D. Those who can and those who cannot

The examples provided in Tables 1 & 2 in part I of this pair of papers were actual documents developed by students in various semesters of the course under study. There is no text for the course. There was no lecture for the course. Anyone present during the semesters of the study[39] can vouch for the fact that each semester the students developed a consensus document the hard way. They dug it out of themselves, testing new ideas against experience and keeping what seemed to fit. They did not have to be told although many expressed the desire to be told on a regular basis. If anything could be improved about the modified version of the course it would be the inclusion of methods facilitating consensus development and additional attention to self-image issues.

What is important here is that students who are not science majors, who have "learned" from their previous science classes that they are not any good at science, nonetheless developed powerful new meanings for the terms, acceleration and force. Furthermore, these new meanings are recognizable as consistent with what scientists have worked out in the past and which many would hold that these students would have to be told because they could not be expected to work out such ideas or understandings for themselves. There is no evidence that these students could not have made these conceptual changes at an even earlier age.

It is abundantly clear that the non-science majors taught in the *radical, student understanding-driven* approach in this study did far better at developing a new view of both acceleration and force than did the science majors taught in the traditional, *content-driven* fashion. It is also significant that the numbers of students who figure these things out can be changed by changing features and conditions of the instruction. These findings utterly destroy the validity of the construct of student ability as a significant explanation for student performance.[10] Once this construct is destroyed the *elitist* ideology which justified physics teachers abdicating responsibility for everyone learning in their classes can no longer honestly or ethically be practiced. Clearly, Ehrlich's assertion that "…that some students have what it takes to develop a mastery of a difficult subject such as physics, while others simply did not, even after expending considerable effort"[31] is not supported by anything but ideology. The notion is clearly incompatible with the data. This is not surprising given an ideology in which there is no distinction between *educating* with physics and making sure the undeserving are adequately excluded from a profession. The notion of traditional content-driven physics instruction as *properly presenting the approved canon for the benefit of the deserving* could not be more clear.

The evidence comes at us from several directions. By 1995 there were several studies of college students who "did not persist" in their studies of science, math or engineering.[40,41] The work is summarized in some detail in the *American Journal of Physics*,[42] yet teaching physicists *who win awards for their physics teaching from national organization*s still function as if the myths this work dispels are still true.[31] Sadly, investigators "did not find switchers and non-switchers to be two different kinds of people: they did not differ by attributes of measured ability or moral character. … What distinguished survivors from those who left was their development of particular attitudes or coping strategies—both legitimate and illegitimate. Serendipity also played a part in persistence." Even at the level of attempting the major, it is not a matter of some can and some cannot. It is a matter of some do and some do not. These ethnographic studies show quite clearly it matters how we treat the students and the knowledge.

Instead of the world being made up of those who can and those who cannot, the present work suggests that some instructional methods are good for many, while others, in particular the traditional, *content-driven* one, apparently only "benefit" the few. Although it is comfortable to blame the victims, we cannot continue to rationalize the failures of instruction on the basis of student ability. One reason should be sufficient: it is profoundly unethical for us as teachers to lead students to believe that they are inadequate when it can be demonstrated that they can in fact perform perfectly well given a different instructional approach. We owe it to our students to be able to distinguish between this and selecting and training physics majors. We owe an *education* to the students we teach. We should reserve the selection and training for those who come to us specifically asking to be physics majors. To fail to make this distinction and act accordingly is to fail our obligations to society as teachers and as physicists.[43]

E. "There is no neutral pedagogy."[44]

Many physics instructors probably prefer to believe that their work is neutral and as much as they can make it, objective. They believe as they have been led to believe that physics teaching is ideology free. Sadly, this is not true. As Fereirro points out, every pedagogy is based on a notion of the nature of knowledge and possibly more than anything else students learn to think of knowledge in the way *implied* by the pedagogy. While Fereirro's field is early literacy, her insight into the nature of education is profound. It applies to the teaching of every subject, but how does it apply to the present work?

The violence previously described has the opportunity to occur, and does occur to some extent, *whenever* someone is told something as "true" and is not given the chance to construct their own understanding as to why someone might believe such a thing. *Content-driven* pedagogies, whose goal is to present approved versions of the canon in an acceptable fashion for the deserving, are prone to this situation. What are the underpinnings of these pedagogies? One is a belief in the *objectivity* of the knowledge being dealt with which entails *realism*: the notion that we can, by our methods, know the real nature of physical reality and judge what knowledge is closer to this "truth."[45] The other is an *elitist* view of the nature of human capacities for understanding.

We can see from the results presented in this work that there is at least one approach which actually leaves most of the students with new understandings they did not have when they entered the instruction and which leaves them in the position to recognize that they accomplished this feat themselves. This pedagogy, a *radical, student understanding-driven* approach is based on neither realism nor elitism. Instead it is based on a much more skeptical and humble view of the nature of knowledge, namely radical constructivism,[27] and on an egalitarian view of the nature of human capacities to make sense of their experiences.

While some might actually arrive at a similar pedagogy via other avenues and wish not to claim the position of radical constructivism for themselves, it is extremely hard if not impossible to imagine how one could really arrive at a sufficiently similar pedagogy and not reject realism for some more skeptical position. As such the results presented here constitute evidence for an argument to reject pedagogies based on realism and to support among others pedagogies based on the radical constructivist view of the nature of knowledge and learning. Realism and radical constructivism are both constructs adopted by choice. Neither is uniquely dictated by our experiences. As such we have the responsibility for the consequences of our choices. The *status quo* is neither defendable nor is it a worthy goal, either on the basis of actual learning results or on the basis of some supposed superior philosophical foundation.

F. A moral dilemma entailed by these results

Many readers may be responding that *they do not intend this harm* to the students and that *they really do intend that students "get" the content* they try to transmit. It is clear that intent is not sufficient. As Ferreiro points out: "Every single [pedagogy] is based on a given conception of the learning process and of the object of such a process. Most probably, those practices much more than the methods themselves are exerting the greatest lasting effects in the domain of literacy, as in any field of knowledge."[46] In the traditional *content-driven* approaches while students do not learn new understandings of the phenomena, they do seem to learn a *realist* view of scientific knowledge and an *elitist* view of human capacities with respect to that knowledge. Regardless of an instructor's intent the pedagogy has this effect. One cannot hide behind a neutral pedagogy because a neutral pedagogy does not exist. We have in this work demonstrated the superior results of one alternative: *radical, student understanding-driven* instruction. There may be others, but one has to be very careful. Realism and elitism are insidious players in this game. Even just a little of either of these has the potential to reduce the effect of the instructional practice to the unacceptable level of traditional, *content-driven* instruction.

Many readers might also be responding that teachers "teach best in the ways they are most comfortable" in defense of continuing to teach in the *content-driven*, traditional way they have taught for many years. When it is clear that teaching in the traditional way does not result in people learning any real physics, instead they learn what must be useless manipulations because the meaning of these manipulations cannot be that for which they are intended and they learn to perpetuate an *elitist* class system which discriminates against most of society, there is no defense for continuing to teach in this fashion, certainly not the excuse of convenience. Our friend from the high school teaching example mentioned above was able to induce a profound change in his students on his first try. Given the data presented herein, if you think you cannot change how you teach physics, then maybe the best course would be to terminate both teaching and advising others how to teach, even if it is just by example.

Knowing the results of the traditional methods and the existence of an alternative and its results sets up an ethical issue for everyone who teaches topics in physics, from elementary school teachers to professors in physics departments. One must explicitly choose between pedagogies with their observed results in mind. None are neutral. This is an obligation automatically entailed by being willing to take responsibility for teaching. It is an obligation to society, to one's students, and one's discipline. Furthermore, the argument does not rest on uniqueness. It applies to any sort of teaching including physics teaching. Can we continue to maintain explicitly or implicitly a *realist* stance with respect to the nature of knowledge in physics in our teaching? Can we continue to maintain explicitly or implicitly an *elitist* view of the nature of human capacities with respect to knowledge in physics in our teaching? Each of us who teaches must answer these questions personally and in our practices.

VII. Acknowledgements

With a few exceptions the author has had the privilege of working directly with those whose works are cited in the references to these two papers. The author is deeply indebted to these colleagues, mentors and predecessors. In addition to those specifically mentioned above in the references, several others had an important impact on this work. While only the author can take the blame for errors, the following have also had additional significant impact in some way on this particular work: John J. Clement, Robert Cohen, Fernando Flores, Robert Fuller, Igal Galili, Fred Goldberg, Richard Gunstone, Richard Hake, Sandra Harpole, Peter Hewson, Andy Johnson, Don Kirwin, Gayle Kirwin, Pam Krause, James Krueger, John Layman, Darryl Macomb, James Minstrell, Valerie Otero, Robert Poel, Edward Redish, William Smith, David Sokoloff, Dale Sweet, Ronald Thornton, Ernst von Glasersfeld, and Michael Wittmann. One should not take these acknowledgements to imply that anyone named necessarily agrees with the positions taken above.

The author is also indebted to the Scientific Reasoning Research Institute in the Physics Department at the University of Massachusetts at Amherst, the Center for Science and Mathematics Teaching at Tufts University, the Physics Learning Research Laboratory in the Physics Department at the University of Maine at Orono, and the Physics Education Research Group in the Physics Department at the University of Washington in Seattle for hosting my visits during a recent sabbatical while this manuscript was being developed.

Figure Captions

Fig. 1a. These "bubble" plots illustrate the possible relationships between students' new view of force post-instruction versus their score on velocity questions post-instruction. Within each graphs the size of the bubbles indicate how many students are represented at each point. The scale of the bubble sizes is not the same from graph to graph. The number of students represented at each point is shown to the right of each bubble. It is clear that most of the students post-instruction score the maximum possible on the questions concerning velocity graphs, but they seem to range pretty equally from very low new view of force scores to the maximum new view of force score (15).

Fig. 1b. These "bubble" plots illustrate the possible relationships between students' post-instruction score on the acceleration questions versus their post-instruction score on velocity questions. As before, within each graphs the size of the bubbles indicate how many students are represented at each point. The scale of the bubble sizes is not the same from graph to graph. The number of students represented at each point is shown to the right of each bubble. It appears that for these questions mostly about velocity-time graphs and acceleration-time graphs, the students who do best on velocity questions are more likely to do better on the acceleration questions.

Fig. 1c. The "bubble" plots illustrate possible relationships between students' post-instruction new view of force scores versus their post-instruction score on acceleration questions. As before, within each graphs the size of the bubbles indicate how many students are represented at each point. The scale of the bubble sizes is not the same from graph to graph. The number of students represented at each point is shown to the right of each bubble. It appears that there is a loading of points on the lower-left to upper-right diagonal axis. This suggests that students who do well on the acceleration questions are more likely to do well on their new view of force scores. It also suggests that students who do poorly on the acceleration questions are less likely to do well on their new view of force scores.

Fig. 2. In these two "bubble" plots from Spring and Fall semesters in 1999, only the students who successfully answered more than 6 out of the 8 acceleration questions were included from each class. Again the bubbles are not quite to the same scale in each plot so the numbers of students represented at each point is to the left of the corresponding bubble. Notice that the pre-instruction distributions (the un-shaded bubbles) are in the same region as the corresponding pre-instruction distributions for the whole class. The post-instruction distributions (the shaded bubbles) here are centered where we would desire the whole class to be at the end of instruction. These students appear to have made significant change in the view of force they are using to respond to the diagnostic questions. Students in these semesters experienced the standard PIPS instruction.

Fig. 3. In these two "bubble" plots from Spring and Fall semesters in 1999, only the students who successfully answered less than 4 out of the 8 acceleration questions were included from each class. Again the bubbles are not quite to the same scale in each plot so the numbers of students represented at each point is to the left of the corresponding bubble. Notice that in these plots there is a significant amount of overlap between the pre- and post-instruction distributions.

This suggests these students did not make much change in their views of force and are still largely using the old view of force to respond to the diagnostic questions. Students in these semesters experienced the standard PIPS instruction.

Fig. 4. These normalized gain versus normalized loss plots for the two semesters contrast the performance of the students who scored well on the acceleration questions at the end of the semester (the black diamonds) and students who scored poorly on the acceleration questions at the end of the semester (the open squares). The arrows point to the centers of the respective distributions. No change in view of force would be represented by an arrow of zero length centered on the origin. The high acceleration scoring students appear to make most of the change possible to be indicated in the diagnostic. The low acceleration scoring students make such little change that it can hardly be seen to be significant. Students in these semesters experienced the standard PIPS instruction.

Fig. 5. These position-time, velocity-time, and acceleration-time graphs illustrate a single step away from the detector. The two vertical, dashed-lines identify a period of time during which the velocity is increasing while the acceleration is decreasing. When noticed by students, this is often a cause of disequilibration. Students question how the acceleration can be decreasing, if the velocity is still increasing.

Fig. 6. These "bubble" plots illustrate the performance of students during two semesters of the modified PIPS instruction (Fall, 2000 & Spring, 2001). In comparison with the plots in Fig. 2. one can see that there are more students who scored well on the post-instruction acceleration questions. It is also clear the there is an even greater separation between the pre- and post-instruction distributions.

Fig. 7. The "bubble" plots in this figure illustrate the performance of students during the same two semesters of modified PIPS instruction (Fall, 2000 & Spring, 2001) as Fig. 6. These are the students who scored poorly on the post-instruction acceleration questions. In comparison with the performance of the students in Fig. 3., these students appear to have made slightly greater change from pre- to post-instruction, but the final performance is still not what we would wish for our students.

Fig. 8. These normalized gain versus normalized loss plots illustrate the performance of both sets of students, high acceleration scoring (the black diamonds) and low acceleration scoring (the open squares), as a result of the modified PIPS instruction (Fall, 2000 & Spring, 2001). A comparison with the corresponding plots from the standard PIPS instruction in Fig. 4 reveals that not only have the high acceleration scoring students actually have an improved performance, but so have the low acceleration students, somewhat.

Fig. 9. The new view of force versus old view of force plots for the two high school classes that experienced the standard PIPS instruction look essentially like the high acceleration scoring students from the conceptual physics courses. This is most likely due to the fact that very few of these students scored less than 4 out of 8 on the post-instruction acceleration questions and a very high percentage scored greater than 6 out of 8 on these same acceleration questions.

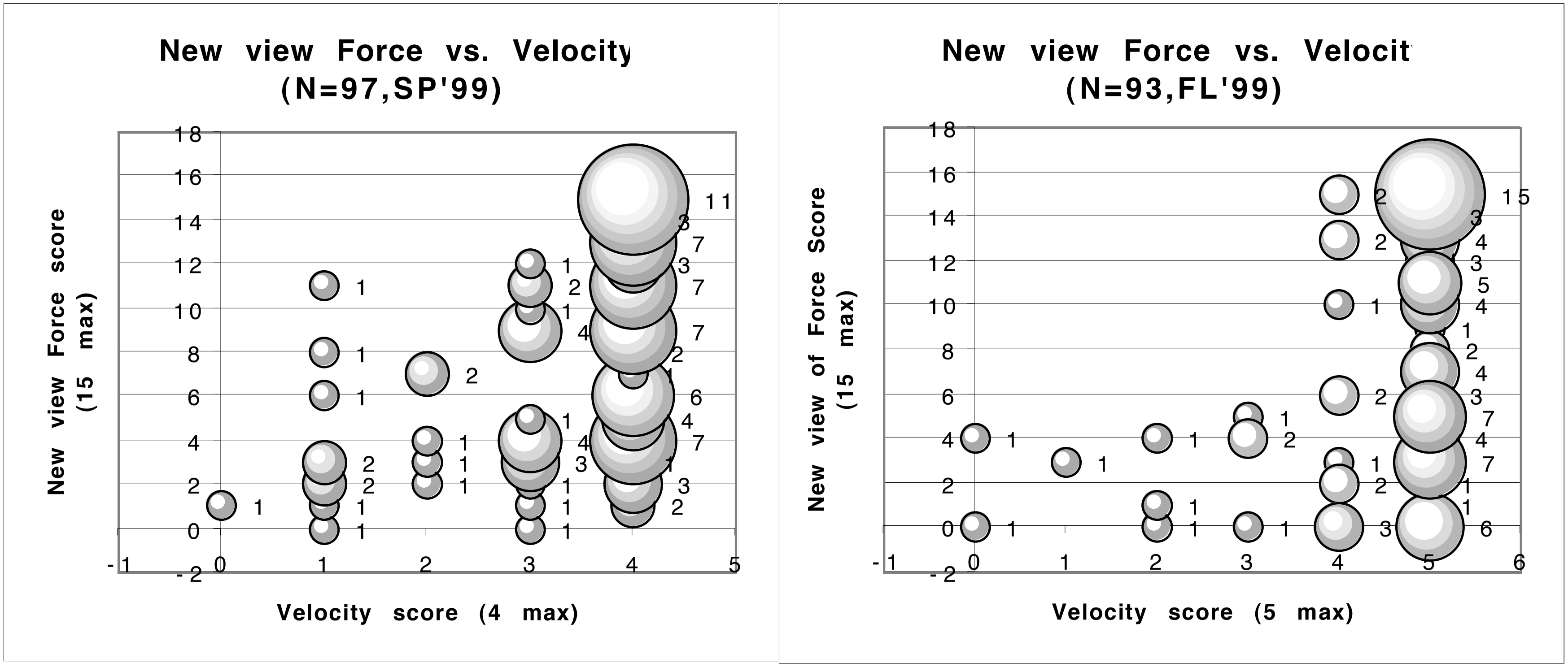


Figure 1a Part II

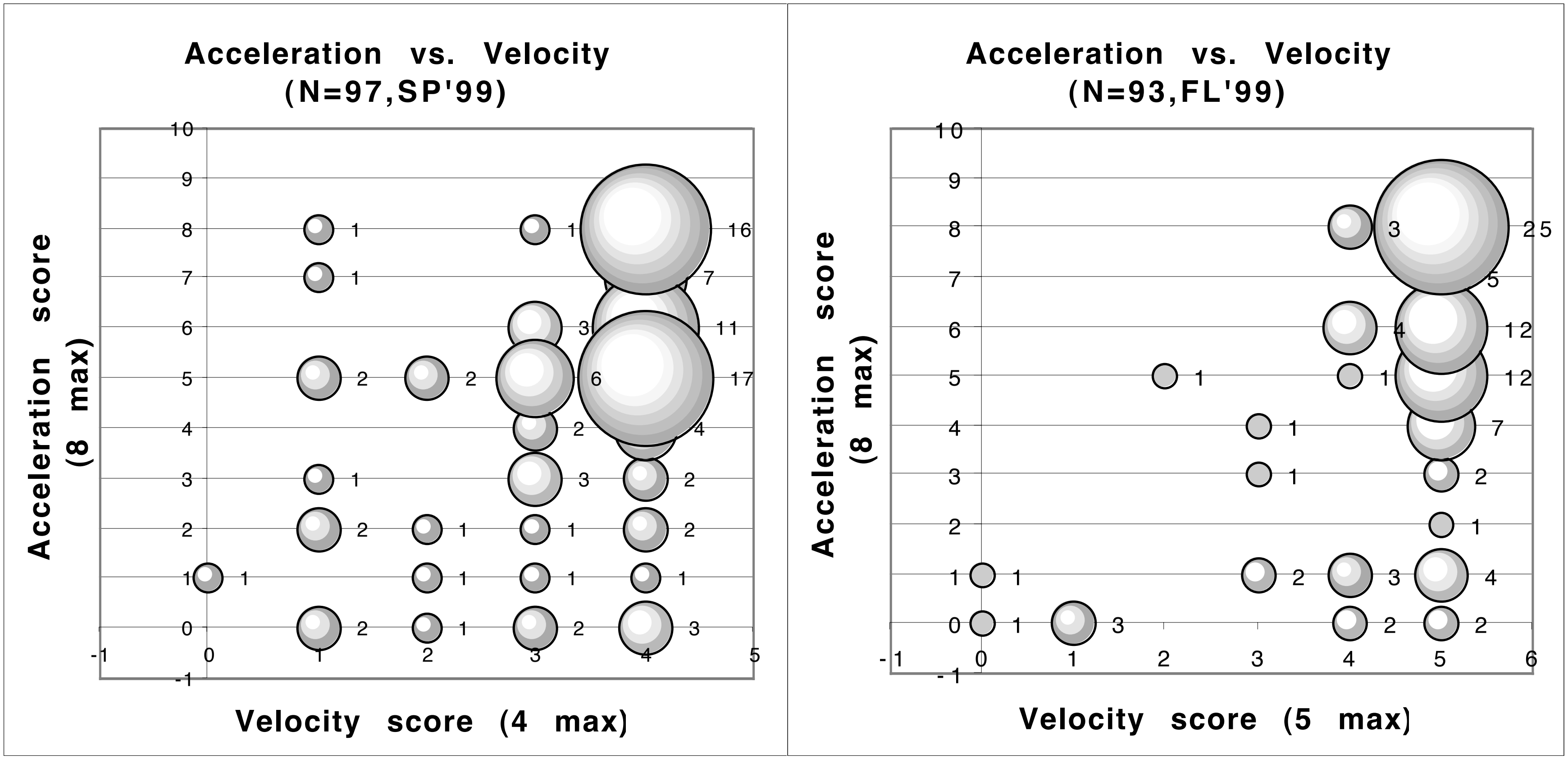


Figure 1b Part II

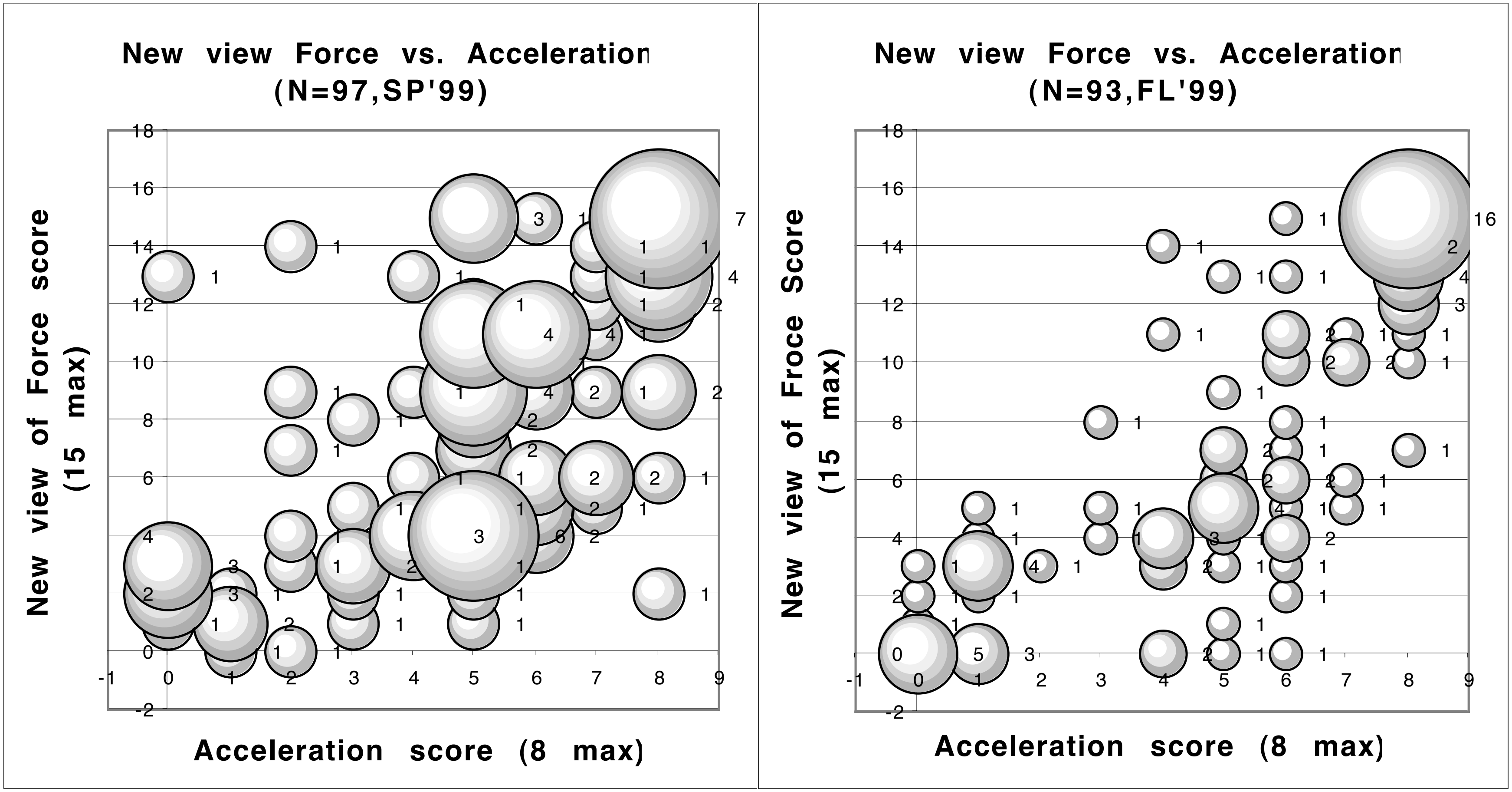
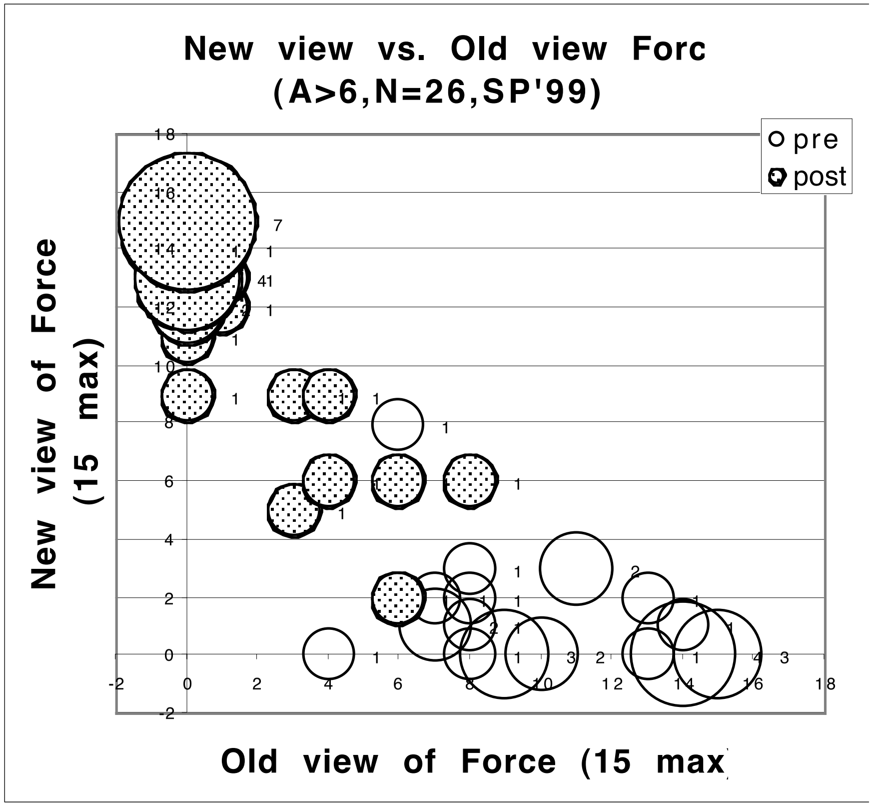


Figure 1c Part II



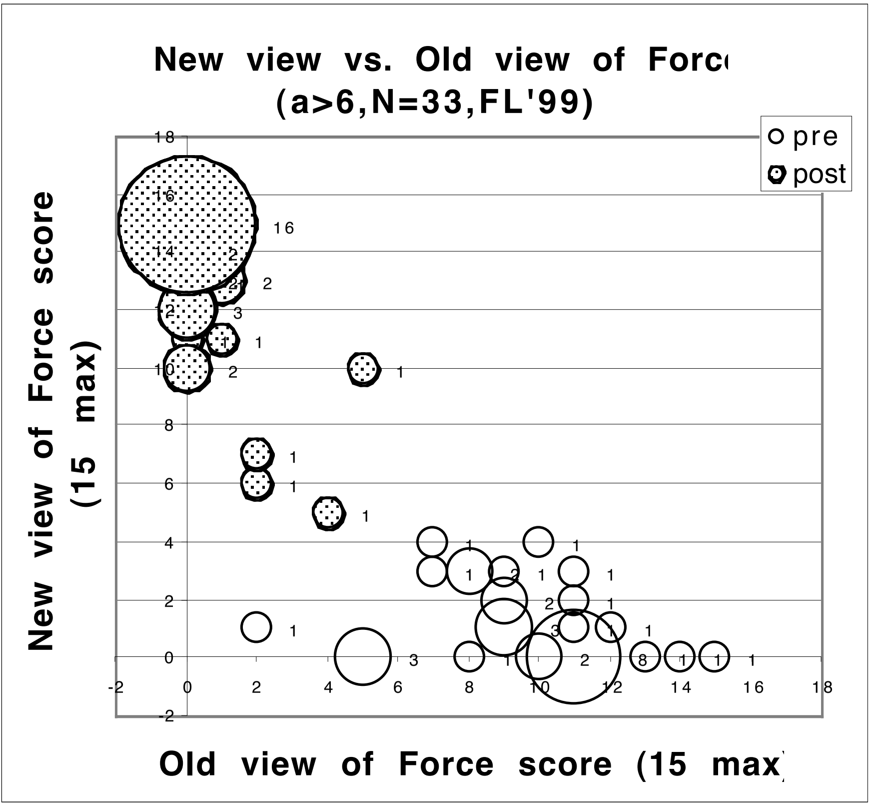
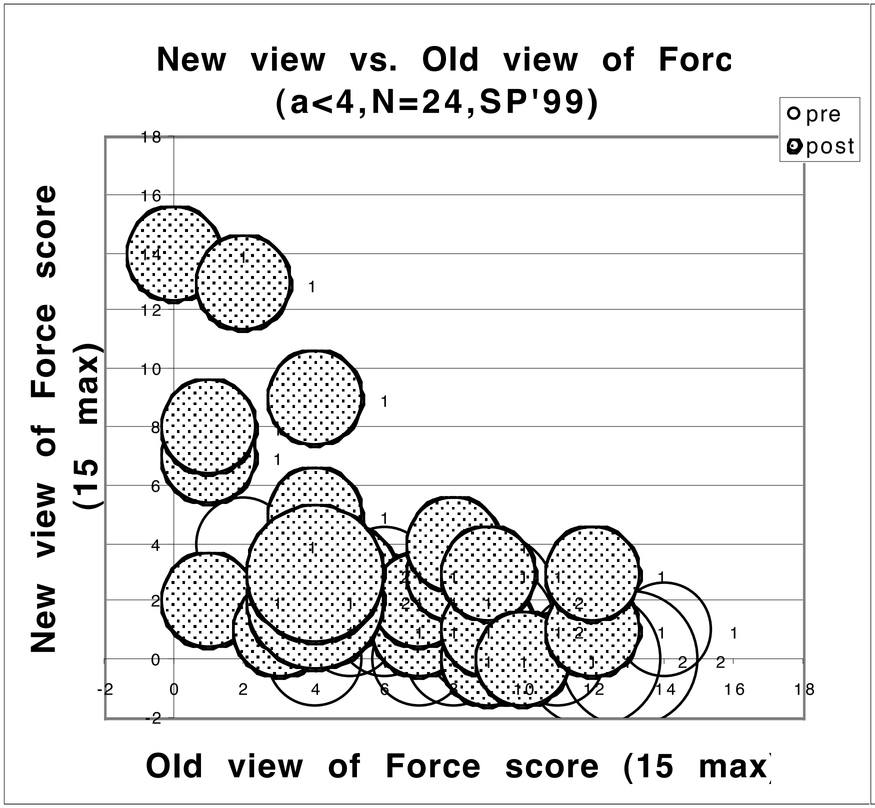


Figure 2 Part II



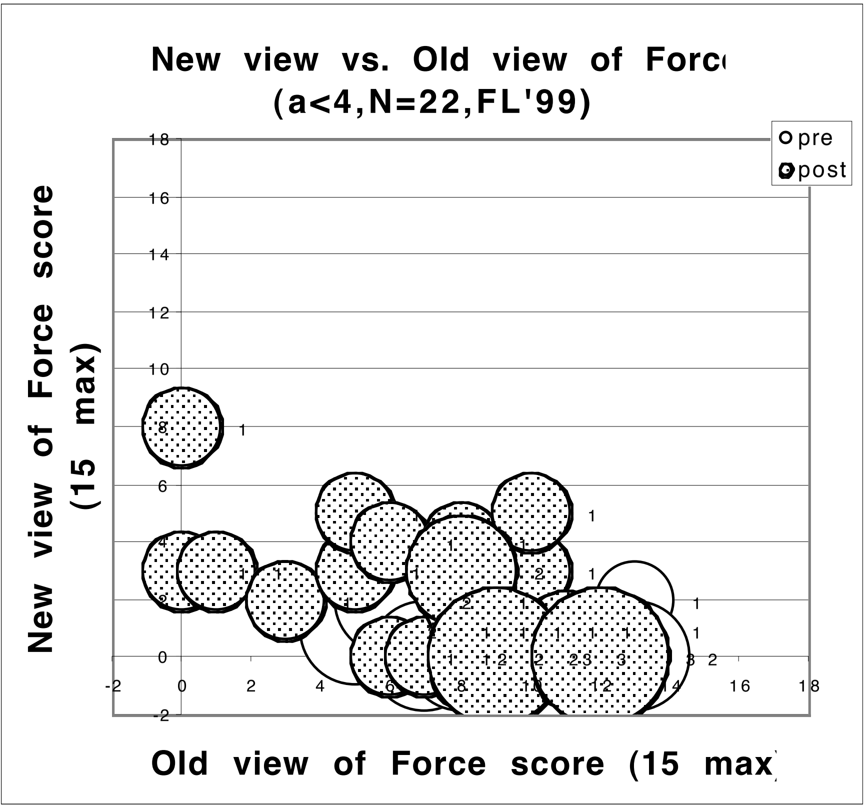


Figure 3 Part II

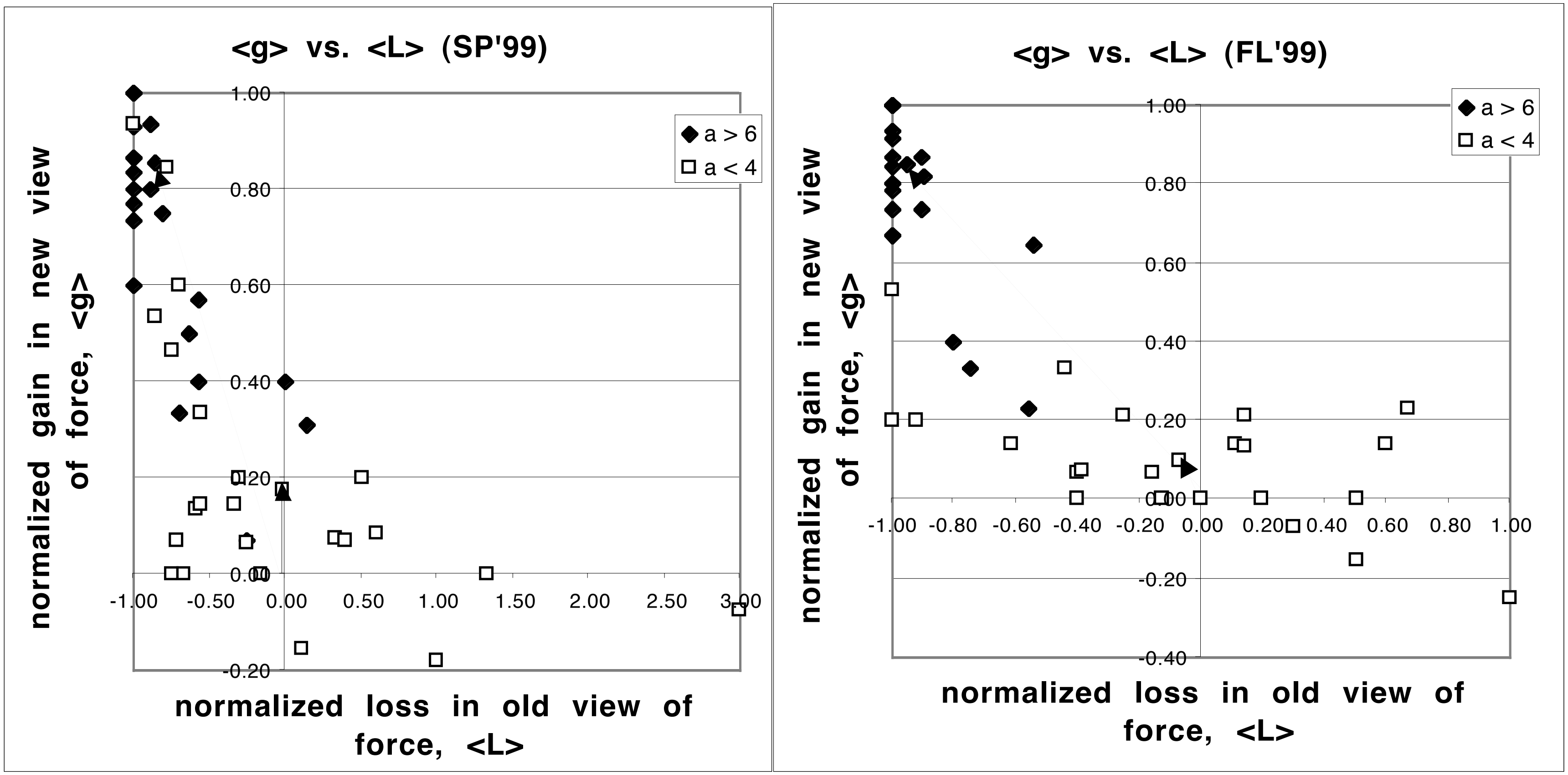


Figure 4 Part II

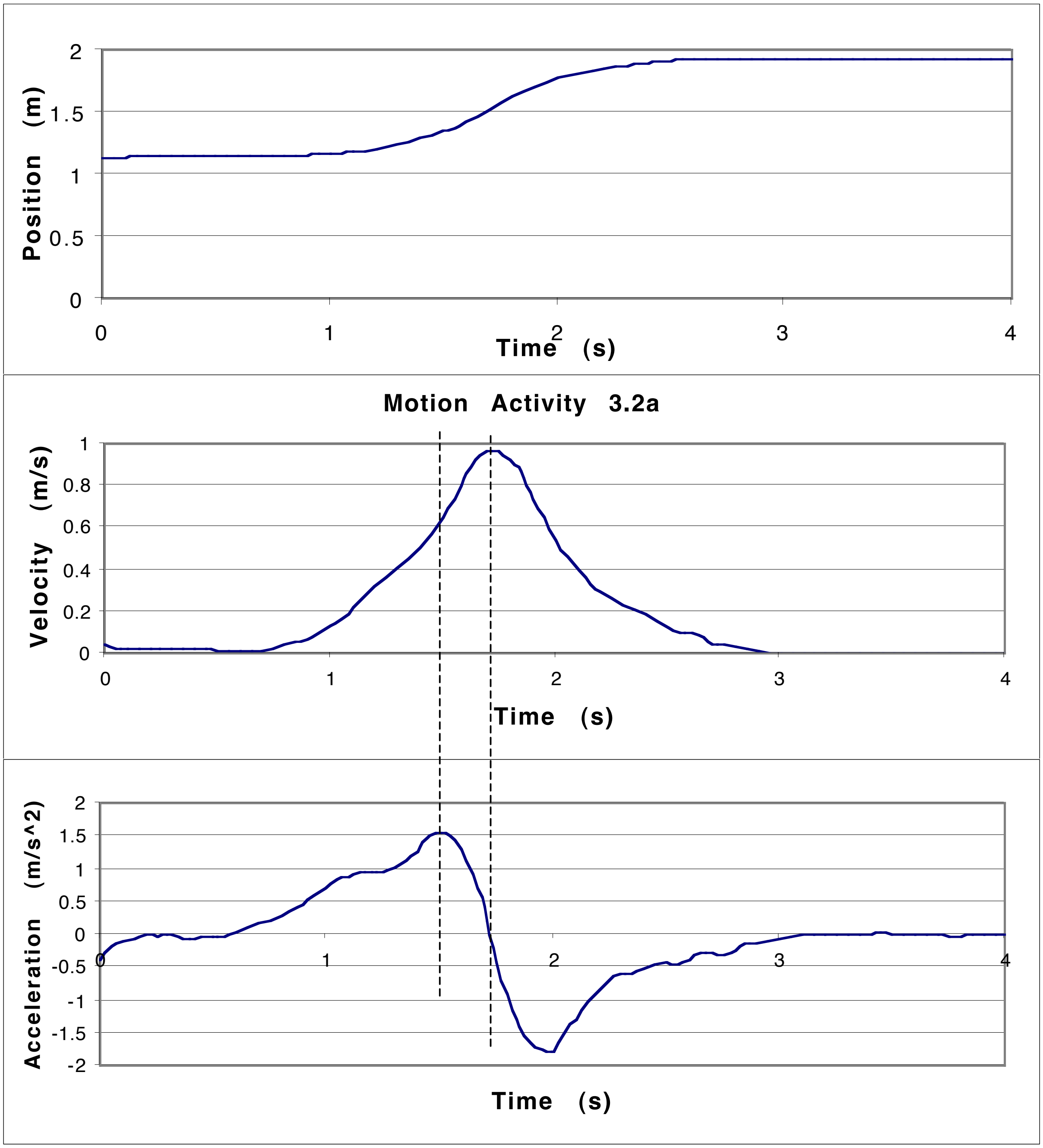


Figure 5 Part II

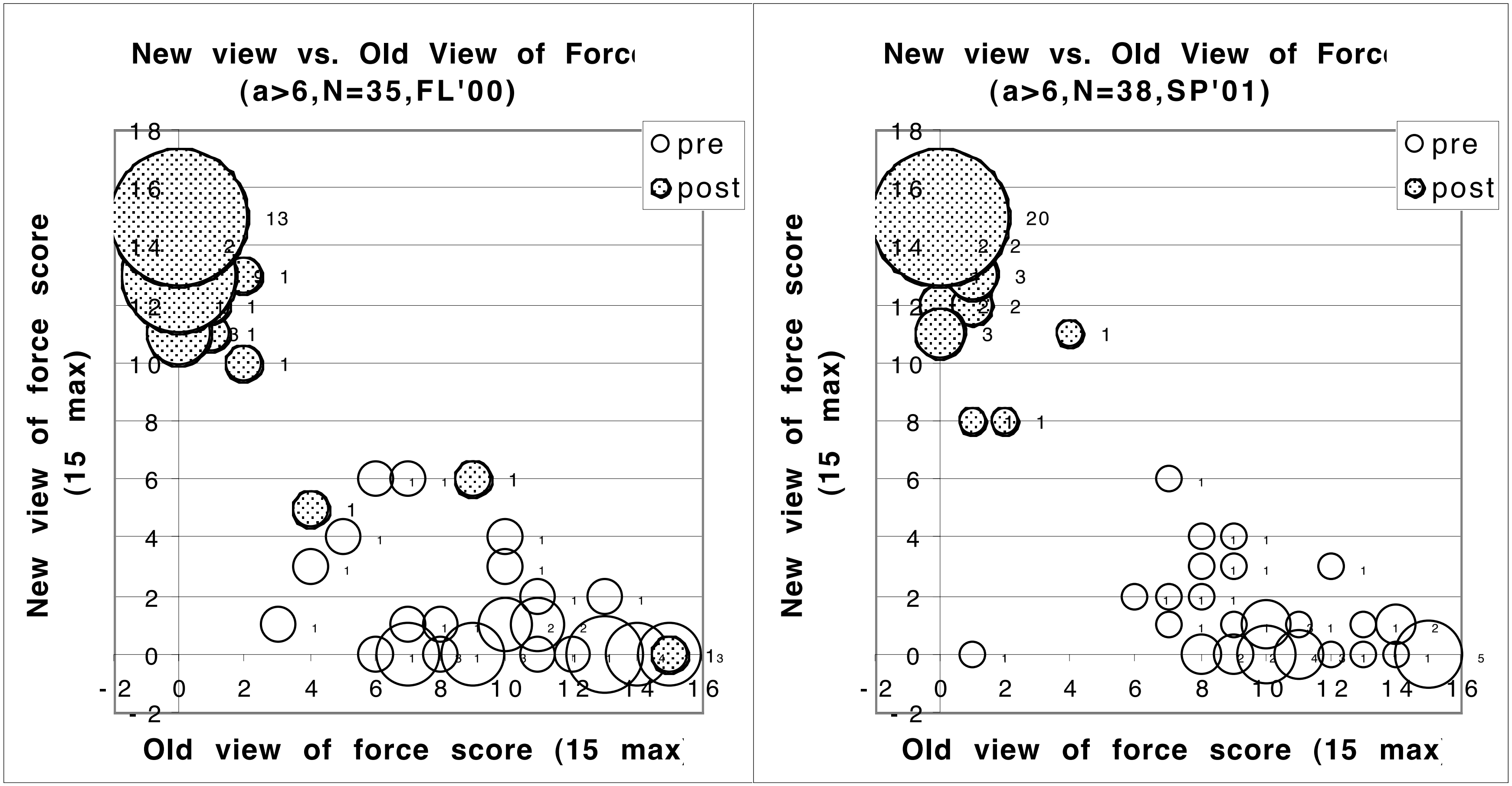


Figure 6 Part II

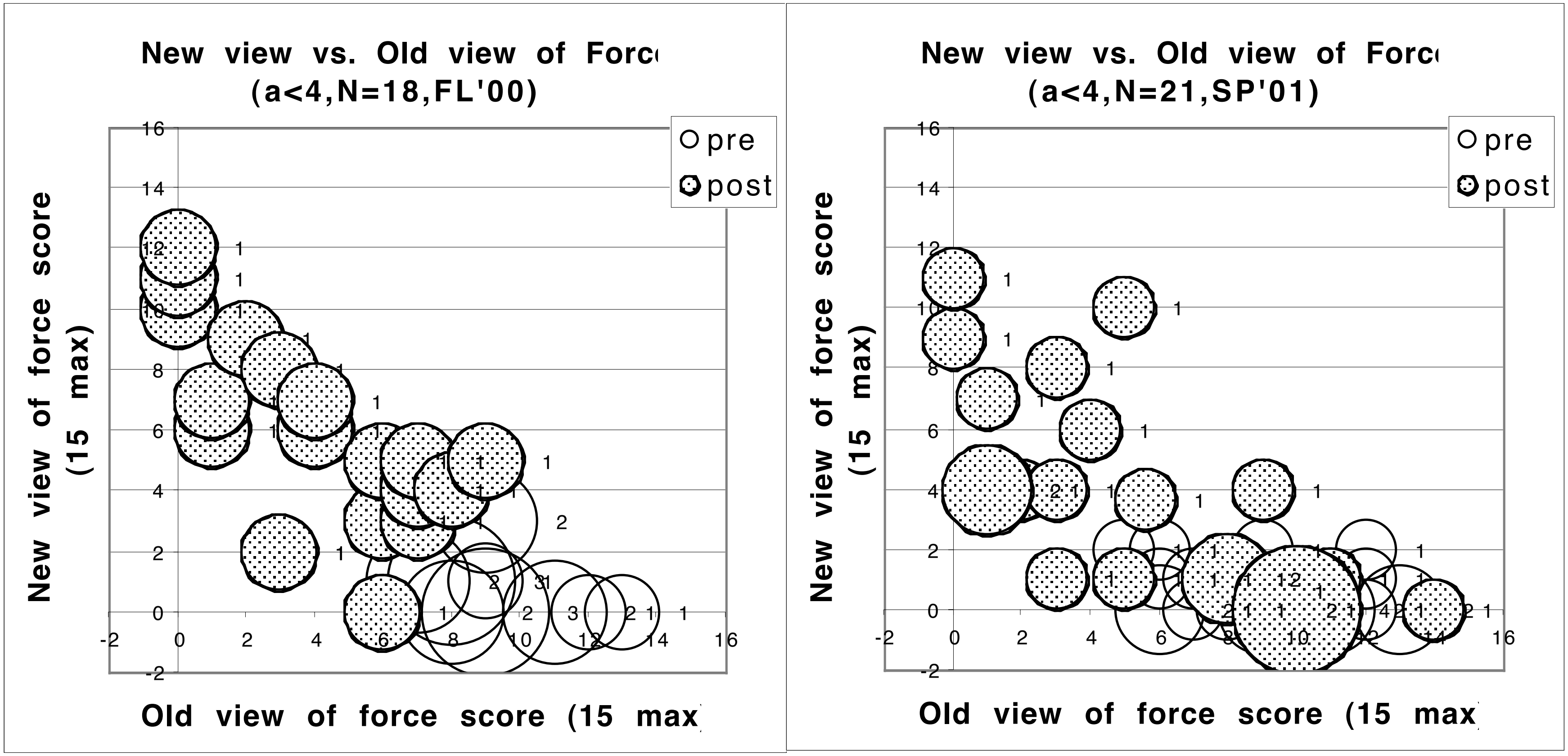


Figure 7 Part II

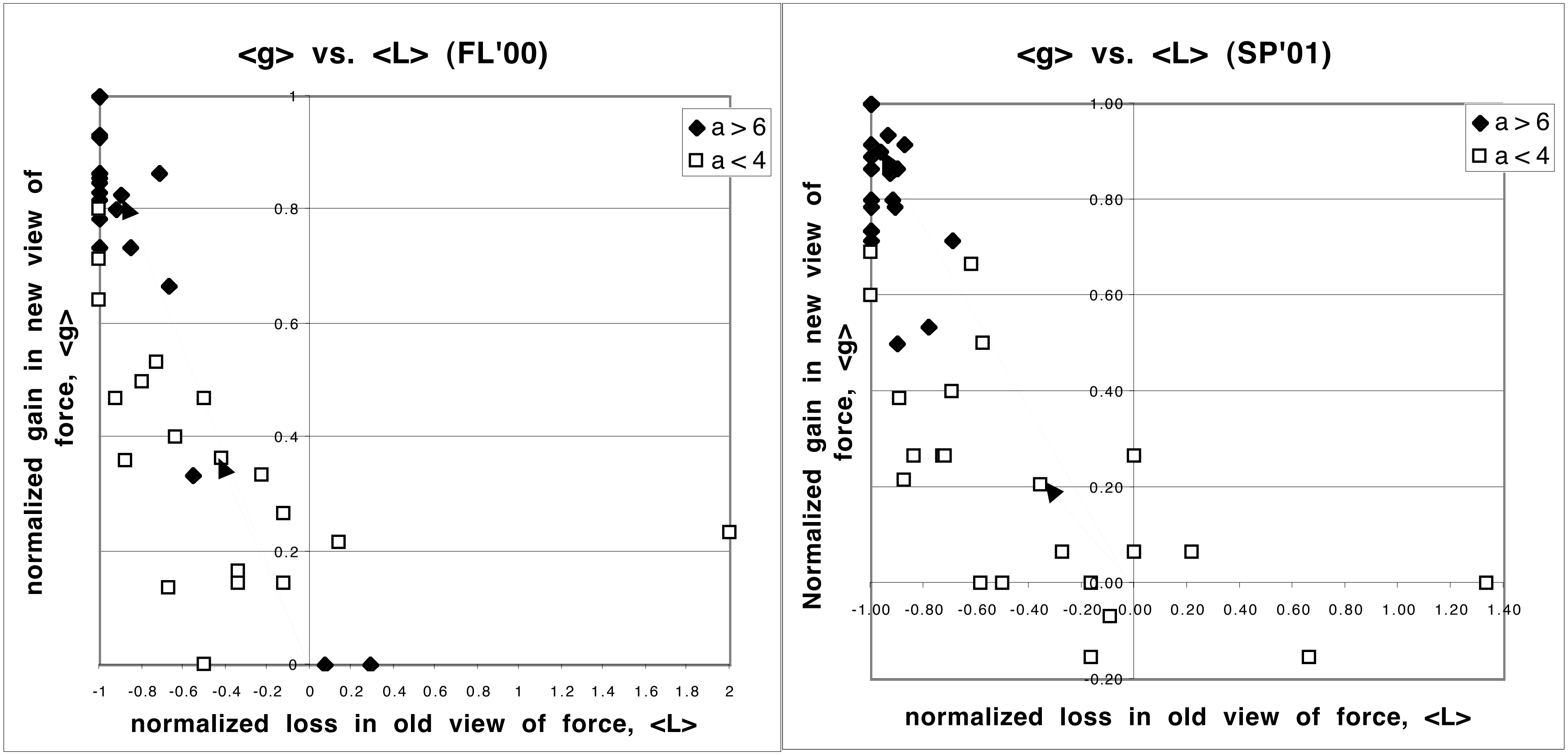


Figure 8 Part II

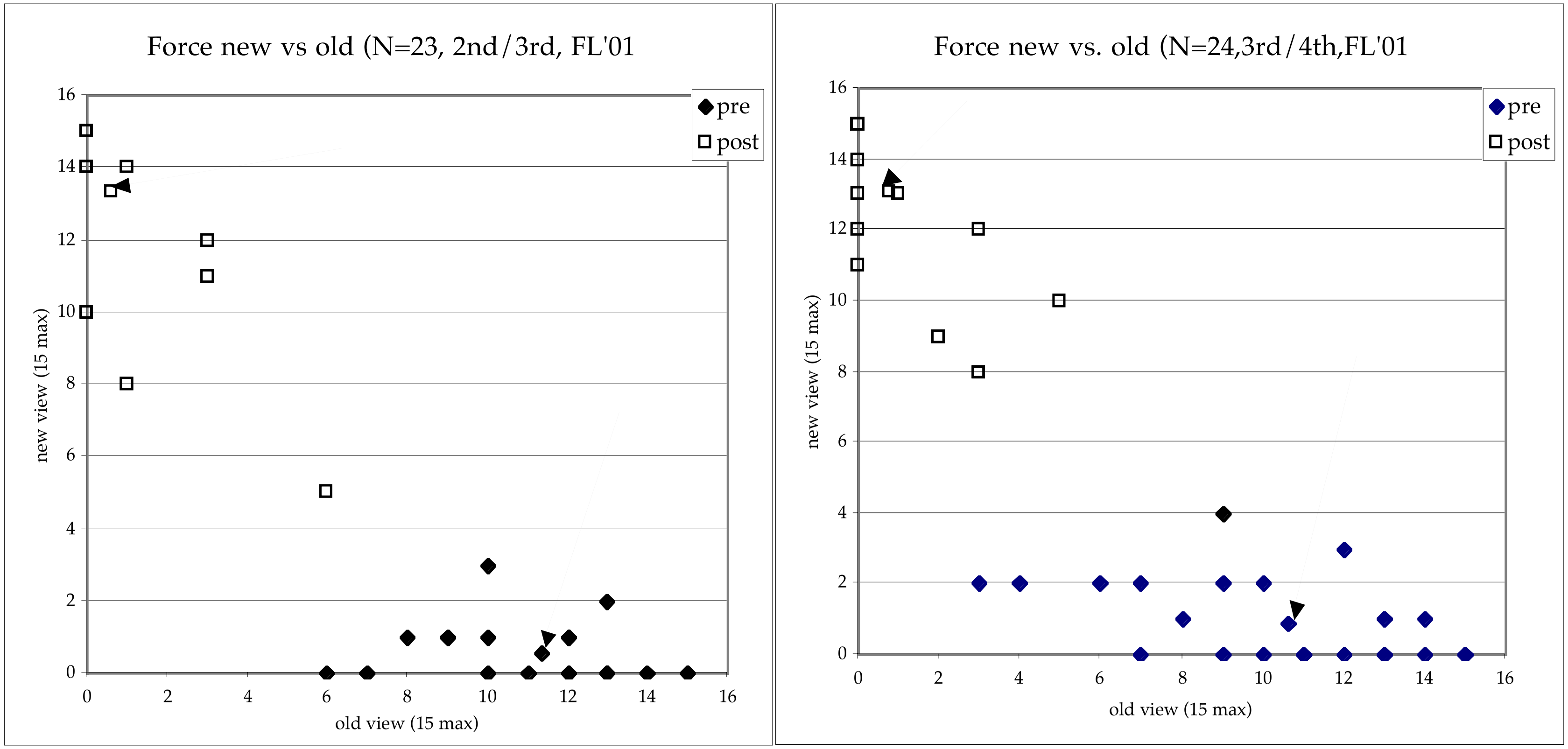


Figure 9 Part II

Table 1a: Conceptual Dynamics, a > 6 Students

		V incr (percentage				V const (percentage)							V decr (percentage)								mixed		
	N	I	OR1,2	2	NI	1	OR1,2	2	OR2,3	OR1,3	3	NC	1	OR1,2	2	OR2,3	3	OR3,4	4	OR4,5	OR	5	ND
SP'99																							
pre	26	81	12	0	92	88	0	0	0	4	4	96	12	8	42	0	4	0	0	4	8	19	96
post	26	0	23	77	100	4	0	0	4	15	77	100	4	0	8	0	0	0	19	4	0	62	96
FL'99																							
pre	33	85	6	0	91	82	3	0	0	15	0	100	3	0	39	0	6	3	15	0	6	15	88
post	33	0	9	88	97	0	0	0	6	9	82	97	0	0	0	0	3	0	6	3	6	76	94

Table 1b: Conceptual Dynamics, a < 4 Students

		V incr (percentage				V const (percentage)							V decr (percentage)								mixed		
	N	I	OR1,2	2	NI	1	OR1,2	2	OR2,3	OR1,3	3	NC	1	OR1,2	2	OR2,3	3	OR3,4	4	OR4,5	OR	5	ND
SP'99																							
pre	24	63	29	4	96	63	4	0	0	21	0	88	4	13	46	0	0	4	4	0	4	8	83
post	24	25	46	17	88	33	13	0	4	17	8	75	0	13	29	13	4	0	4	4	8	21	96
FL'99																							
pre	22	59	36	0	95	77	9	0	0	0	0	86	5	9	68	0	0	0	0	0	5	5	91
post	22	41	32	9	82	59	0	9	9	9	0	86	0	9	36	0	0	5	5	5	0	27	86

Table 2: Effect Size High and Low Accleration Scoring Students

		%	%	whole	class	a>6		a<4	
	N	(a>6)	(a<4)	new	old	new	old	new	old
SP'99	97	27	25	2.02	-1.92	3.54	-3.20	0.94	-0.59
FL'99	93	35	24	1.78	-1.80	5.38	-4.58	0.90	-0.42

Table 3a: Conceptual Dynamics Analysis Whole Class

		V incr (percentage				V const (percentage)							V decr (percentage)								mixed		
	N	I	OR1,2	2	NI	1	OR1,2	2	OR2,3	OR1,3	3	NC	1	OR1,2	2	OR2,3	3	OR3,4	4	OR4,5	OR	5	ND
SP'96																							
pre	82	88	12	0	100	78	4	1	1	10	2	96	10	5	41	2	10	0	7	1	2	17	96
post	82	13	27	60	100	17	4	2	2	20	44	89	4	1	20	0	1	1	15	5	2	51	100
FL'98																							
pre	74	81	11	1	93	81	3	1	0	8	1	95	14	7	42	0	4	1	12	0	4	12	96
post	74	76	16	3	95	78	1	0	1	11	1	93	9	3	39	4	7	1	3	1	11	15	93
SP'99																							
pre	97	75	18	2	95	78	3	0	0	10	2	94	10	6	44	1	4	2	4	2	3	9	87
post	97	8	40	48	97	12	5	0	2	22	47	89	1	5	20	3	1	2	12	9	4	36	94
FL'99																							
pre	93	81	15	0	96	83	3	0	0	8	0	94	4	2	55	0	5	1	9	1	9	8	94
post	93	16	31	46	94	27	0	2	8	23	41	100	0	2	19	0	2	1	10	3	8	49	95
FL'00																							
pre	90	90	6	1	97	73	3	0	1	13	1	92	8	8	47	3	4	4	3	2	11	4	96
post	90	7	28	61	96	7	2	0	0	29	58	96	0	0	8	0	0	2	9	3	6	70	98
SP'01																							
pre	87	77	14	0	91	77	7	0	0	14	0	98	5	7	39	7	6	1	2	2	17	7	93
post	87	13	23	59	94	14	0	0	0	15	66	94	2	0	7	2	1	1	9	2	5	68	98

Table 3b: Conceptual Dynamics Analysis, a > 6 Students

		V incr (percentage				V const (percentage)							V decr (percentage)								mixed		
	N	I	OR1,2	2	NI	1	OR1,2	2	OR2,3	OR1,3	3	NC	1	OR1,2	2	OR2,3	3	OR3,4	4	OR4,5	OR	5	ND
SP'99																							
pre	26	81	12	0	92	88	0	0	0	4	4	96	12	8	42	0	4	0	0	4	8	19	96
post	26	0	23	77	100	4	0	0	4	15	77	100	4	0	8	0	0	0	19	4	0	62	96
FL'99																							
pre	33	85	6	0	91	82	3	0	0	15	0	100	3	0	39	0	6	3	15	0	6	15	88
post	33	0	9	88	97	0	0	0	6	9	82	97	0	0	0	0	3	0	6	3	6	76	94
FL'00																							
pre	35	83	14	0	97	77	3	0	3	14	3	100	14	3	26	6	6	9	3	3	17	6	91
post	35	3	9	89	100	3	0	0	0	6	89	97	0	0	6	0	0	0	0	0	3	91	100
SP'01																							
pre	38	87	11	0	97	84	8	0	0	8	0	100	5	3	34	3	8	0	5	0	26	11	95
post	38	0	8	87	95	3	0	0	0	5	92	100	0	0	0	0	0	0	0	0	0	95	95

Table 3c: Conceptual Dynamics Analysis, a < 4 Students

		V incr (percentage				V const (percentage)							V decr (percentage)								mixed		
	N	I	OR1,2	2	NI	1	OR1,2	2	OR2,3	OR1,3	3	NC	1	OR1,2	2	OR2,3	3	OR3,4	4	OR4,5	OR	5	ND
SP'99																							
pre	24	63	29	4	96	63	4	0	0	21	0	88	4	13	46	0	0	4	4	0	4	8	83
post	24	25	46	17	88	33	13	0	4	17	8	75	0	13	29	13	4	0	4	4	8	21	96
FL'99																							
pre	22	59	36	0	95	77	9	0	0	0	0	86	5	9	68	0	0	0	0	0	5	5	91
post	22	41	32	9	82	59	0	9	9	9	0	86	0	9	36	0	0	5	5	5	0	27	86
FL'00																							
pre	18	89	0	6	94	72	6	0	0	17	0	94	0	6	72	0	6	0	0	0	0	6	89
post	18	17	50	22	89	11	6	0	0	50	33	100	0	0	11	0	0	0	11	6	11	61	100
SP'01																							
pre	21	81	5	0	86	71	5	0	0	14	0	90	5	10	43	14	0	5	0	0	10	10	95
post	21	38	33	19	90	43	0	0	0	10	29	81	5	0	19	10	0	5	19	0	10	33	100

Table 4: Average Positions for Distributions in Scatter Plots

	whole class			a > 6			a < 4		
semester	N	old	new	N%	old	new	N%	old	new
SP'96									
pre	82	9.2	1.1	26	8.8	2.3	41	8.5	0.6
post	82	3.4	6.8	26	0.6	11.9	41	6.1	3.5
SP'99									
pre	97	9.4	0.9	27	10.5	1.0	25	7.6	1.2
post	97	3.0	7.8	27	1.4	11.4	25	5.5	3.8
FL'99									
pre	93	9.9	0.8	35	9.6	1.1	24	8.7	0.7
post	93	3.6	7.6	35	0.5	12.9	24	7.3	2.2
FL'00									
pre	90	9.3	0.8	39	10.0	1.1	20	8.8	0.8
post	90	2.5	9.2	39	1.0	12.6	20	4.1	5.9
SP'01									
pre	87	9.8	0.8	44	10.4	1.0	24	9.1	0.6
post	87	2.2	9.6	44	0.4	13.6	24	5.6	3.6

Table 5: Normalized Gains and Losses

		%	%	Whole class		$a > 6$		$a < 4$		Type of Instruction
	N	$a > 6$	$a < 4$	gain	loss	gain	loss	gain	loss	
SP'99	97	27	25	0.49	-0.68	0.75	-0.82	0.18	-0.02	standard PIPS
FL'99	93	35	24	0.48	-0.63	0.85	-0.95	0.10	-0.07	
FL'00	90	39	20	0.59	-0.66	0.83	-0.90	0.36	-0.42	modified PIPS
SP'01	87	44	24	0.62	-0.74	0.90	-0.97	0.21	-0.36	

Table 6: Effect Size - Whole Class, a > 6 & a < 4 Students

		%	%	whole class		a > 6		a < 4		Type
	N	a > 6	a < 4	new	old	new	old	new	old	Instruction
SP'96	82	26	41	1.7	-1.7	3.4	-3.2	1.3	-0.8	pre-PIPS
SP'99	97	27	25	2.0	-1.9	3.5	-3.2	0.9	-0.6	Standard PIPS
FL'99	93	35	24	1.8	-1.8	5.4	-4.6	0.9	-0.4	
FL'00	90	39	20	2.5	-2.2	4.5	-2.9	2.2	-1.8	Modified PIPS
SP'01	87	44	24	2.4	-2.4	7.3	-4.5	1.1	-1.0	

Table 7: High School Class Results

	Effect Size		<g>	<L>	class	a > 6	Effect Size		a < 4
class	new	old			N	n	new	old	n
2nd/3rd	6.34	-5.42	0.89	-0.95	23	17	17.8	-6.5	2
3rd/4th	6.12	-3.66	0.86	-0.93	24	19	7.91	-3.47	2

Table 8: A comparison of the results of two types of teaching practice
Traditional, *Content Driven* Instruction

Algebra-Trig Level Intro Physics

Whole class			Effect Size		Normalized		Scatter Plot Averages				**High Accel a > 6**			Normalized		Scatter Plot Averages				**Low Accel a < 4**			Normalized		Scatter Plot Averages			
			(st dev)		Gain	Loss	Pre (0 - 15)		Post(0 -15)			Effect size		Gain	Loss	Pre (0 - 15)		Post(0 -15)			Effect size		Gain	Loss	Pre (0 - 15)		Post(0 -15)	
Year	Term	N	New	Old	<g>	<L>	Old	New	Old	New	N%	New	Old	<g>	<L>	Old	New	Old	New	N%	New	Old	<g>	<L>	Old	New	Old	New
West Coast Public Univ. A																												
1990		99	0.59	-0.47	0.14	-0.09	10.1	1.5	8.5	3.3	Data not available. (1)									Data not available. (1)								
"Prairie State" Public Univ.																												
2002	SP	112	0.66	-0.40	0.13	-0.06	10.3	0.9	9.0	2.7	17%	1.53	-1.51	0.42	-0.52	10.9	1.6	5.4	7.2	65%	0.51	-0.04	0.05	0.08*	10.2	0.6	10.0	1.3
Calculus Level Intro Physics																												
North East State Public Univ.																												
1998		72	0.47	-0.30	0.15	-0.06	9.6	1.7	8.5	3.5	Data not available. (2) Too few students.									89%	0.56	-0.27	0.13	0.01*	10.0	1.2	9.0	2.8
West Coast Public Univ. B																												
1999	Wint.	87	0.60	-0.62	0.22	-0.30	9.3	2.6	6.5	5.4	37%	0.76	-0.65	0.34	-0.39	6.8	5.4	3.8	9.4	34%	0.66	-0.42	0.05	-0.06	11.0	0.4	9.9	1.2
1999	SP	73	0.38	-0.36	0.12	-0.01	9.1	2.3	7.6	4.0	23%	0.94	-0.86	0.51	-0.41	8.2	5.1	3.6	10.6	47%	0.25	-0.06	0.02	0.2*	9.4	1.1	9.2	1.6
2000	SP	115	0.59	-0.50	0.15	-0.10	9.2	2.4	7.2	4.8	32%	1.10	-0.82	0.44	-0.32	7.8	4.1	4.1	9.4	42%	0.28	-0.09	0.02	0.12*	9.8	1.2	9.5	1.6
West Coast Private Univ.																												
2000	SP	38	0.54	-0.08	0.09	0.05	9.8	0.6	9.6	1.9	Data not available. (2) Too few students.									74%	0.08	0.29	0.00	0.18*	9.5	0.6	10.3	0.7

Student Understanding Driven Instruction

Conceptual Physics College Level--Standard PIPS Instruction

Whole class			Effect Size		Normalized		Scatter Plot Averages				**High Accel a > 6**			Normalized		Scatter Plot Averages				**Low Accel a < 4**			Normalized		Scatter Plot Averages			
			(st dev)		Gain	Loss	Pre (0 - 15)		Post(0 -15)			Effect size		Gain	Loss	Pre (0 - 15)		Post(0 -15)			Effect size		Gain	Loss	Pre (0 - 15)		Post(0 -15)	
Year	Term	N	New	Old	<g>	<L>	Old	New	Old	New	N%	New	Old	<g>	<L>	Old	New	Old	New	N%	New	Old	<g>	<L>	Old	New	Old	New
Boise State University																												
1999	SP	97	2.00	-1.90	0.49	-0.68	9.4	0.9	3.0	7.8	27%	3.50	-3.20	0.75	-0.82	10.5	1.0	1.4	11.4	25%	0.94	-0.59	0.18	-0.02	7.6	1.2	5.5	3.8
1999	FL	93	1.78	-1.80	0.48	-0.63	9.9	0.8	3.6	7.6	36%	5.40	-4.60	0.85	-0.95	9.6	1.1	0.5	12.9	24%	0.90	-0.42	0.10	-0.07	8.7	0.7	7.3	2.2
Conceptual Physics College Level--modified PIPS Instruction																												
Boise State University																												
2000	FL	90	2.50	-2.20	0.59	-0.66	9.3	0.8	2.5	9.2	39%	4.50	-2.90	0.83	-0.90	10.0	1.1	1.0	12.6	20%	2.20	-1.80	0.36	-0.42	8.8	0.8	4.1	5.9
2001	SP	87	2.40	-2.40	0.62	-0.74	9.8	0.8	2.2	9.6	44%	7.30	-4.50	0.90	-0.97	10.4	1.0	0.4	13.6	24%	1.10	-0.99	0.21	-0.36	9.1	0.6	5.6	3.6
High School Level--Standard PIPS Instruction																												
2001	FLa	23	6.30	-5.40	0.89	-0.95	11.3	0.6	0.6	13.3	74%	17.80	-6.50	0.98	-0.99	11.4	0.6	0.2	14.6	Data not available. (2) Too few students.								
2001	FLb	24	6.10	-3.70	0.86	-0.93	10.6	0.9	0.8	13.1	79%	7.90	-3.50	0.92	-0.94	10.4	1.1	0.5	13.9	Data not available. (2) Too few students.								

(1) The version of the FMCE used with this class did not make use of all 8 of the acceleration questions making it impossible to categorize the student performances.

(2) There were less than 3 students in this category rendering calculations meaningless for this category.

* A positive Loss in old view is a gain in old view.

[†] Partial support from National Science Foundation Grants Nos. MDR-8950313, MDR-9153989, MDR-9154015 and DUE-9554625. Experiences during the work under these grants led to the opinions expressed herein. Opinions expressed herein are not necessarily those of the National Science Foundation or the American Association of Physics Teachers.

(a email: ddykstra@boisestate.edu

[1] R. K. Thornton & D. R. Sokoloff, "Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula," *Am. J. Phys.* **66**(4) 338 – 352 (1998).

[2] C. S. Peirce, "Abduction and Induction," in *Philosophical Writings*, Justus Buchler (ed), (Dover Publications, New York, 1955), 150 – 156.

[3] J. Piaget, " Equilibration Processes in the Psychobiological Development of the Child," in *The Essential Piaget*, H. Gruber and J. Vonèche (eds) (Basic Books, New York, 1977), 837 and R. Vuyk, "5.2.4 The Final Question: 'Do Structures Exist?'" in *Overview and Critique of Piaget's Genetic Epistemology 1965 – 1980, Volume 1*, (Academic Press, New York, 1981), 60.

[4] B. Shapiro, *What Children Bring to Light: A Constructivist Perspective on Children's Learning in Science*, (Teachers College Press, New York, 1994) 163 – 164 "My conversations with Donnie yielded particularly good examples of how a student's development of a view of herself as science learner creates anticipations and expectations about participation in science and science learning. … Donnie was disturbed when she could not answer questions. She did not realize that some of the questions were in many ways similar to those asked by physicists themselves in the search for fundamental understanding. She described a recurring sense in science that "I never quite get all of it." But in fact, Donnie did demonstrate a partial grasp of may difficult concepts. … Often children make incomplete connections and have no opportunity for continuing discussion or extended experiences with materials. For many children like Donnie, the struggle to understand can lead to a feeling of incompleteness. Donnie's own efforts to put ideas together were not recognized and encouraged. She did not see her own struggle as beneficial nor did she find satisfaction in overcoming the challenge, for she never had a sense of understanding the problem. Her struggle led to a lowered sense of the possibility of being a successful science learner and on many occasions she spoke of science learning with discouragement and despair."

[5] In elementary school students encounter signed numbers usually in conjunction with the symbols, > & <, which they are told are the "greater than" and "less than" symbols, implying that the signs, + & –, are about the magnitudes of the numbers. Some students are even taught to draw little "teeth" inside the symbols to "represent an alligator eating the bigger number." All is fine until one encounters a pair of numbers such as –9 and –5 and one is told that –5 > –9, negative 5 is greater than negative 9. Most of us are told and accept or just decide that in the case of negative numbers, one "does it backwards" or "opposite to the way that makes sense." As it turns out the signs on numbers on a number line probably should be interpreted as directions on a number line; the negative meaning toward the negative infinity from zero and the positive meaning toward positive infinity from zero. In this case the symbols, > & <, should not mean "greater than" and "less than;" instead they should mean "is toward positive infinity from" and "is toward negative infinity from." Hence we have negative 5 is toward positive infinity from negative 9, –5 > –9.

---

[6] A reaction from a student essay written between the motion and force units during modified PIPS instruction, Spring, 2001: "Well, as you know, the graph did not turn out like we expected. We were all a bit shocked and confused. My idea of acceleration was completely wrong! I was especially confused because this is how everyone seemed to think of acceleration. Were we all wrong? Well, in class after this lab, …, we were told that this is the cultural meaning for the word acceleration. We were not exactly wrong, we just used the word differently than the way scientists use the word. I felt a little better that I was not being stupid by using the word the way I had been, but now I had to figure out how to use the word in a different way! Here came the challenge."

[7] To be "embarasada" in Spanish means to be pregnant, not the same as to be embarrassed in English.

[8] It would be a fundamental error to conclude that the position advocated here means that all answers or all possible meanings are right or okay. Clearly to use the word "biscuit" at tea to refer to the breakfast pastry would be inappropriate and confusing to some. Each of the terms described has an established, taken-as-shared meaning in its own cultural setting. In a situation in which the velocity of an object traveling in a straight line is increasing constantly, it may be appropriate in the everyday meaning of acceleration to say the acceleration of the object is increasing, but it is not appropriate to say the acceleration in the sense of what is portrayed in the MBL graphs is increasing because that acceleration is not changing. Still, regardless of which version of acceleration is being used, the object is still moving in a straight line with its velocity constantly increasing.

[9] L. C. McDermott, P. S. Shaffer, and C. P. Constantinou, "Preparing teachers to teach physics and physical science by inquiry," *Phys. Educ.* **35**(6), (2000); L. C. McDermott and L. S. DeWater, "The Need for Special Science Courses for Teachers: Two Perspectives," in *Inquiring into Inquiry Learning and Teaching in Science*, J. Minstrell and E. H. van Zee (eds) (AAAS, Washington, DC, 2000); and L. C. McDermott, "A perspective on teacher preparation in physics and other sciences: The need for special science courses for teachers," *Am. J. Phys.* **58**(8), 734 – 742, (1990)

[10] R. Millar, K. Klaassen and H. Eijkelhof, "Teaching about radioactivity and ionising radiation: an alternative approach" *Phys. Educ.* **25**, 338 – 342 (1990).

[11] H. Pfundt & R. Duit, *Bibliography: Student Alternative Frameworks and Science Education*, (Institute for Science Education (IPN) University of Kiel, Kiel, Germany, 2002). Available on the web at: http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html

[12] P. Trabal *La Violence de L'Enseignement des Mathematiques et des Sciences: Une autre approche de la sociologie des sciences.* (L'Harmattan: Montreal, Canada. 1997) and D. W. Blades, *Procedures of Power & Curriculum Change: Foucault and the Quest for Possibilities in Science Education* (Peter Lang, New York, 1997), 186 – 187.

[13] D. Goodstein, "Science Education Paradox: How can the same system produce scientific elites and illiterates?" Technology Review, September, 2001.

[14] We can see an example of this view of instruction in C. D. Geilker, "Guest Comment: In defense of the lecture-demonstration method of teaching physics," *Am. J. Phys.* 65(2), 107 (1997). In it he says: "We have a higher obligation to let our students glimpse the whole domain of physics, from sea to shining sea. Awareness, with partial comprehension, is still better than omission, with certain ignorance." He also states: "Physics is a *spiral* subject and *significant* misconceptions tend to get corrected with further exposure and experience. … Until

---

the day when lecture disappears completely, note-taking will continue to be the survival skill of college—particularly so, if *I* think something is important enough to write it on the board. … Highlighting your open text will not suffice—you'll have nothing to show when I do an note inspection during the next hour-long class exam." (*emphasis* in the original) With the best of intentions he repeats the mantra without examining it to see that it is circular and non-unique.

[15] B. Shapiro, *What Children Bring to Light: A Constructivist Perspective on Children's Learning in Science* (Teachers College Press, New York, 1994)

[16] L. C. McDermott, "Oersted Medal Lecture 2001: 'Physics Education Research—The Key to Student Learning,'" *Am. J. Phys.* 69(11) 1127 – 1137 (2001)

[17] D. I. Dykstra, Jr., C. F. Boyle, and I. Monarch, "Studying Conceptual Change in Learning Physics," *Science Education* **76**(6), 615 - 652 (1992)

[18] D. I. Dykstra, "Teaching Introductory Physics to College Students" Chapter 12 in *Constructivism: Theory, Perspectives and Practice*, C. T. Fosnot (ed) (Teachers College Press, New York, 1996)

[19] J. Krueger, "PLEASE!!!! …Just tell me the answer!!!" M.S. Ed. Project Report, University of Wisconsin-White Water, to be submitted July, 2002, 18 - 19: "Even after my review of the literature I was still surprised at the pervasiveness of non-Newtonian thinking in my **own** students. All of the students studied Newton's Laws at some point in their previous years of education. In fact, many of the students who took part in this study are former physical science students of mine. We spend an entire quarter their freshman year studying motion and force. Most of them could recite Newton's Laws perfectly and did well on the assessment I gave them. However, it now appears that the ninth grade physical science curriculum did little to change their basic conceptions on motion and force." (**emphasis** in the original)

[20] E. Kim & S-J. Pak, "Students do not overcome conceptual difficulties after solving 1000 traditional problems," *Am. J. Physics* **70**(7), 759 – 765 (2002).

[21] In the early version of the FMCE used with the traditionally taught science majors, the choices consisted of just the possible directions of the acceleration and zero, for example: A. up the ramp, B. zero, C. down the ramp. In the published version of the FMCE used to collect all the rest of the data in this study, on these two groups of three questions the students had to indicate the direction of the acceleration and whether it was increasing, decreasing or constant. This results in a total of seven choices (including the acceleration being zero) compared to the three choices of the older version.

[22] For example see the work of J. Minstrell. Because of his early influence on the author and certain similarities between their points of view it should not be surprising that their approaches would have similar results. Both can easily generate a list of differences between their practices and approaches, but both will agree that what is important is the underlying aspects of honoring students' existing understandings and the students' abilities to make reasonable sense of the phenomena. Information available on the web at http://www.talariainc.com/train.html

[23] R. K. Thornton & D. R. Sokoloff, "Learning Motion Concepts Using Real-Time Microcomputer-Based Laboratory Tools." *Am J Phys* **58**(9) 858 – 867 (1990)

[24] D. R. Sokoloff & R. K. Thornton, "Using Interactive Lecture Demonstrations to Create an Active Learning Environment." *The Physics Teacher* **35** 340 – 346 (1997)

[25] G. E. Francis, J. P. Adams, and E. J. Noonan, "Do They Stay Fixed?" *The Physics Teacher* **36** 488 – 490 (1998).

---

[26] D. Hestenes, M. Wells, & G. Swackhammer, “Force concept inventory,” *Phys. Teach.* **30**(3), 141 – 158 (1992). Updated version: Halloun, I., R.R. Hake, E.P Mosca, D. Hestenes. Force Concept Inventory (Revised, 1995; password protected at http://modeling.la.asu.edu/R&E/Research.html

[27] It is important to remember that this new view/old view analysis lumps a number of possible epistemological states or positions with respect to force and motion together. Thornton’s conceptual dynamics analysis separates out some of these positions phenomenologically. (See R. K. Thornton, “Conceptual Dynamics: Following changing Student Views of Force and Motion,” In *The Changing role of Physics Departments in Modern Universities: Proceedings of ICUPE*, E. F. Redish & J. S. Rigden (eds) (American Institute of Physics, College Park, MD, 1997) Abstract online at http://ase.tufts.edu/csmt/html/abstracts/icupe_cd.html.) Regardless the arguments made on the basis of the new and old views of force herein still stand.

[28] 99.2 RESEARCH IN PHYSICS EDUCATION
(Adopted by the Council, 21 May 1999) http://www.aps.org/statements/99.2.html

In recent years, physics education research has emerged as a topic of research within physics departments. This type of research is pursued in physics departments at several leading graduate and research institutions, it has attracted funding from major governmental agencies, it is both objective and experimental, it is developing and has developed publication and dissemination mechanisms, and Ph.D. students trained in the area are recruited to establish new programs. Physics education research can and should be subject to the same criteria for evaluation (papers published, grants, etc.) as research in other fields of physics. The outcome of this research will improve the methodology of teaching and teaching evaluation.

The APS applauds and supports the acceptance in physics departments of research in physics education. Much of the work done in this field is very specific to the teaching of physics and deals with the unique needs and demands of particular physics courses and the appropriate use of technology in those courses. The successful adaptation of physics education research to improve the state of teaching in any physics department requires close contact between the physics education researchers and the more traditional researchers who are also teachers. The APS recognizes that the success and usefulness of physics education research is greatly enhanced by its presence in the physics department.

[29] R. Hake, “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses,” *Am. J. Phys.* **66**(1) 64 – 74 (1998)

[30] Two sources should be considered as possible starting points for understanding what is meant by radical constructivism. The first is an article by E. von Glasersfeld, available on the web: http://www.douglashospital.qc.ca/fdg/kjf/17-TAGLA.htm. It was originally published in *Research and Reflexivity (Inquiries into Social Construction)*, F. Steier (ed), (Sage Publications, London, 1991). A more lengthy introduction is in the book by von Glasersfeld *Radical Constructivism: A way of knowing and learning*, (Falmer Press, Washington, DC, 1995)

[31] P. R. Gross in “Politicizing Science Education” at http://www.edexcellence.net/library/gross.html sponsored by the Thomas B. Fordham Foundation of Dayton, Ohio.

[32] B. G. Aldridge, “Invited Paper” *The Science Teacher*, 8 (October, 1995)

[33] R. Ehrlich, “How do we know if we are doing a good job in physics teaching?” *Am. J. Phys.* **70**(1), 24 – 28 (2002)

---

[34] M. Wittmann, "On the dissemination of proven curriculum materials: Real-Time Physics and Interactive Lecture Demonstrations" submitted to the *American Journal of Physics* (2002) Available on the net: http://perlnet.umephy.maine.edu/research/02RTPpaper.pdf

[35] The old "saw" repeated by some in our physics teaching community that you never really understand it until you have to teach it, appears to be a truism, but one dependent on the way we teach and the way we assess our students.

[36] Maybe the fabled necessity of a "spiral curriculum" is more a necessity brought on by an inappropriate approach than the "true" nature of learning.

[37] A. Cromer, *Connected Knowledge: Science, Philosophy and Education* (Oxford University Press, New York, 1997).

[38] "There is no reading problem. There are problem teachers and problem schools…If talking and walking were taught in most schools, we might end up with as many mutes and cripples as we now have non-readers." H. Kohl, *Reading: How to*, (Clarke, Irwin & Company, Ltd., Toronto, Canada) xi, 1973

[39] Observers are welcome to attend and observe any semester of the course as the author teaches it. Interested readers are invited to make contact with the author.

[40] E. Seymour & N. M. Hewitt, "Talking About Leaving: Factors Contributing to High Attrition Rates Among Science, Mathematics, and Engineering Undergraduate Majors," Final Report to the Alfred P. Sloan Foundation on an Ethnographic Inquiry at Seven Institutions, April, 1994. (Westview Press, Boulder, CO).

[41] S. Tobias, *They're Not Dumb, Just Different*. (Research Corporation, Tucson, AZ, 1990)

[42] E. Seymour, "Guest Comment: Why undergraduates leave the sciences" *Am J Phys* **63**(3)199 – 202 (1995).

[43] For those who might be unmoved by the moral or ethical issues, maybe some numbers will make an impression. Using data from the American Institute of Physics and the National Center for Educational Statistics, one can find that there are very close to one thousand times more elementary school teachers than college or university physics professors. The ratio of the number of student-hours of instruction on physics topics taught per week taught by those who are not university or college professors of physics to the number of student-hours of instruction on physics topics per week taught by those who are university or college professors of physics is about 4.5. If we consider the number of student-hours of instruction on physics topics to all who are presently non-physics majors to all who are declared physics majors the ratio is 8.8. These last two figures are under estimated by the fact that the physics major numbers are swelled by the engineering majors who take calculus level introductory physics along with the physics majors and who were not separately accounted for. In other words the ratios are greater than 4.5 and 8.8 respectively. The *real* enterprise of teaching physics is largely out of the hands of professors of physics. This state of affairs is exacerbated by the traditional stance of physics departments who essentially wash their hands of any possibility of influence on the process of preparing or assisting teachers except a token effort for high school teachers and the expenditure of the minimum resources necessary to teach non-majors at the colleges and universities. One of the few ways in which this makes sense is if physics "education" is really a selection and training program and not really for the *education* of any who come to us.

[44] E. Ferreiro in "Literacy Acquisition and the Representation of Language." a chapter in *Early Literacy: a constructivist foundation for whole language* (NEA Professional Library, Washington, DC) 45 – 46 (1991). (*emphasis* in the original)

At the heart of the controversy in traditional approaches to the problems involved in the teaching of reading and writing lies the question of the method to be used: the analytic *versus* global approaches, and so on. The classical controversy does not take into account what we now know about the conceptualizations that children have regarding the writing system. For this reason, it is imperative that we examine teaching practices from a new perspective. If we are willing to accept that the child is not a *tabula rasa* upon which letters and words are going to be inscribed in the order determined by the method employed, that what is "easy" and what is "difficult" to learn must be defined from the perspective of the learner and not in terms of the adult, and that whatever information received must be assimilated (and therefore transformed) before the child may operate with it, then we must also accept that teaching methods (understood as a sequence of steps ordered in such a way as to attain a goal) can at best offer suggestions and hints (when they are not just reduced to the imposition of ritual practices or to a set of restrictions). The method cannot produce knowledge.

It is fundamental that we understand the problems as children pose them and the sequence of solutions they find acceptable (that give rise to new problems) before we can even imagine the kind of pedagogical intervention that should be designed to meet the real needs of the learning process. To reduce these interventions to what is traditionally designated as "the method employed" would put too great a restraint on our inquiry.

Instead of asking about the method employed, it is more useful to look at the *practices* used to introduce the child to written language, and how this object is presented in the classroom. There are practices that lead children to think that knowledge is something that *others* possess and that they must turn therefore to *others* to obtain it without ever participating in the construction of such knowledge. There are also practices that make them think that "what has to be known" is given once and for all, as if it were a closed, sacred, and immutable set of elements that are to be transmitted but not modified. Yet other practices place the children "outside" the knowledge, making them passive spectators or mechanical receivers who can never find the answers to the whys and wherefores that they don't even dare to formulate aloud.

There is no neutral pedagogical practice. Every single one is based on a given conception of the learning process and of the object of such a process. Most probably, those practices much more than the methods themselves are exerting the greatest lasting effects in the domain of literacy, as in any field of knowledge. Certain practices may appear "normal" and others "aberrant" depending upon how the relation between the subject and the object of knowledge is understood and how both terms of this relation are characterized. It is at this point that psychopedagogical considerations must be supported by epistemological reflections."

[45] "…we postulate the objective existence of physical reality that can be known to our minds…with an ever growing precision by the subtle play of theory and experiment." A. C. de la Torre & R. Zamorano, "Answer to Question #31. Does any piece of mathematics exist for which there is *no application whatsoever in physics*?" *Am. J. Phys.* **69**(2) 103 (2001).

[46] E. Ferreiro, ibid., 46