CHAPTER 5
OUTCOMES OF STUDY

This thesis is primarily concerned with the development of student problem-solving skill during a year of introductory college physics which had instruction in an explicit problem solving strategy. To help with the interpretation of the results from such a class, problem-solving skill data was collect from a class without this specialized instruction. Since different classes taught by different instructors were used, a case-study methodology was employed. Two cohorts of students were selected to act as the case units. The EPS cohort came from the explicitly taught class. The TRD cohort came from the other class. Both classes are far from traditional. Each instructor had their students solve context-rich problems using cooperative groups (Heller, Keith, and Anderson, 1991) in discussion sessions, in the laboratory, and on exams. The students in both cohorts were initially matched on several parameters in an attempt to create two equal cohorts of students. While the cohorts were generally equal to each other at the start of the study, it was observed that the TRD cohort performed better than their classmates did in the grades they received. In effect, the TRD cohort represents the better performing students in the TRD class.

Once the cohorts were selected, their solutions to twenty-six problems were coded for four problem-solving skills. These skill codes were then adjusted by difficulty ranks to get a better reflection of the students' actual skills. These scores will be plotted versus time to create development graphs. These problem-solving skills development graphs are the topic of this chapter. This chapter begins with a discussion of development graphs and then will move on to the development graphs of each problem-solving skill and their interpretations.

Skill Development

In this study, graphs will be used to show skill development with time. This section of this chapter will begin by explaining general shapes of growth curves. These curve shapes define development. This section will then discuss possible interpretations of curves on the development graphs used in this study. The above two discussions combine to justify the use of the problem difficulty ranks. Following this introductory section is the discussion of the four problem-solving skill graphs for each cohort.

Generic development graphs

In general there are only three possible shapes to skill development graphs when students actually get better at the skill. There can be linear growth, exponential growth or logarithmic growth. These three graphs are shown in Figure 5.1. A quick inspection shows that the linear and exponential graphs allow for continual growth of skills. There is no limit. This is not a realistic situation. Rather it is has been shown repeatedly that experts in all fields do not show limitless skill growth. Rather to observe a noticeable gain in skills, these experts must practice their craft. Conversely, when a competent person engages in a new skill, gain is easy to observe over short times. These observations lead to the only logical conclusion; skill growth is globally logarithmic over time.

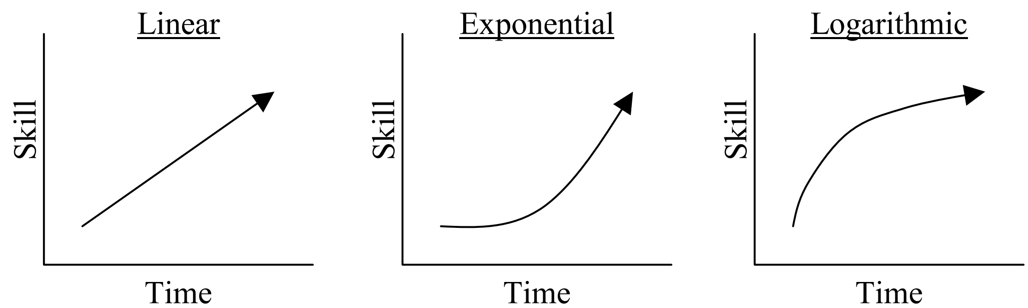


Figure 5.1

Generic skill development graphs.

However, even though skill development must be logarithmic over the lifetime of skill use, the development graph does not need to be strictly logarithmic over short time intervals. The combination graph shown in Figure 5.2 displays a reasonable development graph. The presence of plateaus on the curve allows for regions where linear and exponential growth can occur over short time periods. This kind of rejuvenated growth after a plateau in the sports world is called "reaching a new level." Therefore the existence of linear or exponential growth curves should not be viewed as wrong, but rather evidence of the learning moving between plateaus.

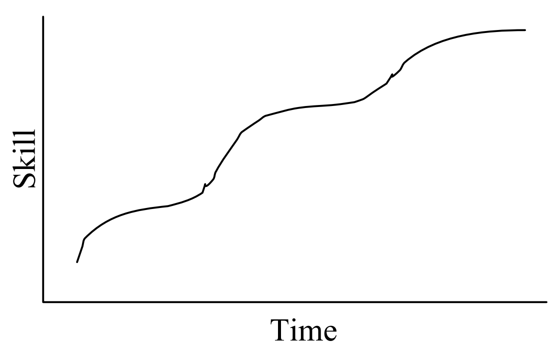


Figure 5.2

Combination development graphs showing plateaus.

Problem-solving Skill Development Graphs

As was introduced in Chapter Three, the students in both cohorts had their open-ended problem solutions analyzed in a two-fold process. First the problems were coded on four problem-solving skills. These skills were General Approach (GA), Specific Application of the Physics (SAP), Logical Progression (LP) and Appropriate Mathematics (AM). Then the difficulty rank of each problem was determined. This was a number between one and twenty-one with an average around 5 for each cohort. Combining these two analyses allows for the creation of problem-solving skill development graphs, where the students problem-solving skill score is plotted versus time with the quizzes grouped around the middle date. The development graphs also have a backdrop of difficulty-adjusted skill bands.

The skill bands displays how each problem-skill code was affected by the difficulty ranks. Recall that these are bands because there is a region surrounding each code-level line from the standard error of the measurement. However, drawing these uncertainties only served to clutter up the graphs. Generally speaking the error of each band slightly overlaps the error of the surrounding bands. Figure 5.3 shows a development graph with three hypothetical skill-development curves drawn on it. The General Approach skill for the EPS cohort provides the data for Figure 5.3, but it will be used for both cohorts and for each problem-solving skill.

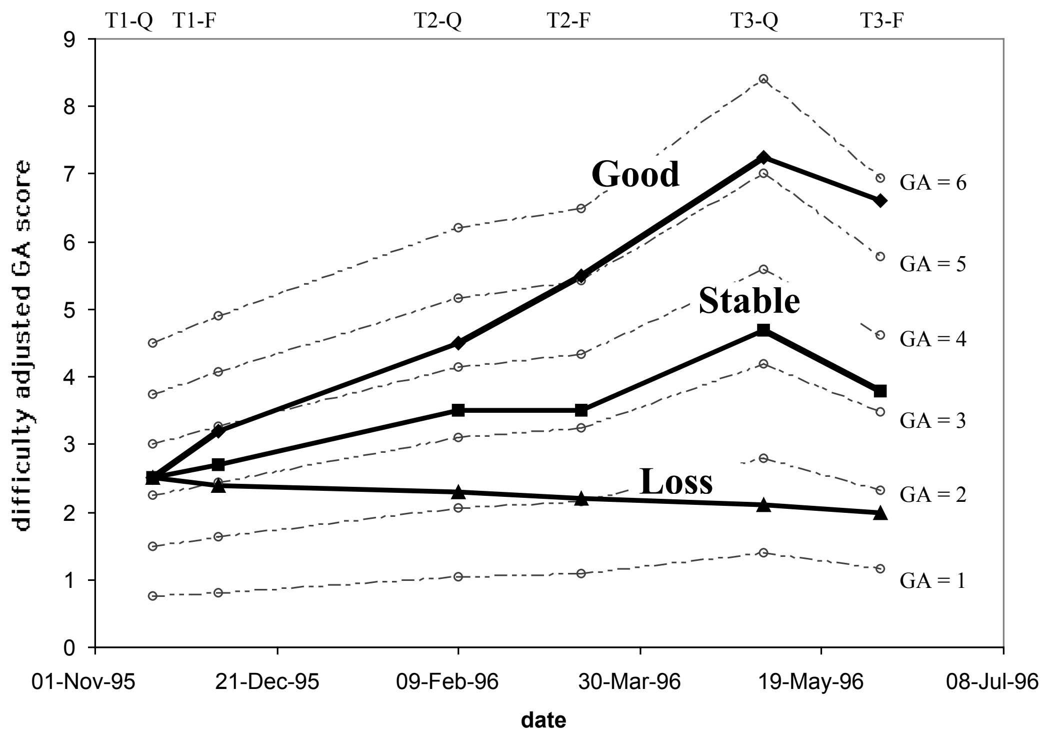


Figure 5.3

Problem-solving skill development graph with three hypothetical lines for the EPS cohort. The Good line shows clear improvement in problem-solving skill. The Stable line shows limited improvement over time, but the line stays within the same skill band. The Loss line shows a clear loss of performance.

The three lines on Figure 5.3 show three hypothetical development curves. The Loss curve shows what happens if the students did not improve over time. Without the skill bands, one might misinterpret the Loss curve as no change since the score values did not change significantly over time. However, the students clearly do not keep up with the material. Both the Good and Stable graphs show development. The course material gets harder for the students, but both groups improve. The Stable curve shows students who continue to make the same type of mistakes as the year progresses. The Good curve shows students who thrive. These students not only handle the more difficult material, but they actually get better with this skill over time. They make less severe mistakes as time progresses.

As can be seen from Figure 5.3, the problem difficulty plays a key role in the shape of the curve. The problem difficulty allows growth to be seen as such. Without it, one might misinterpret a flat line (such as the Stable line would be without difficulty adjustments) to show no development. Yet as the material gets harder, the students maintain consistency in their errors. This does not have to be the case. In fact it would be more probable based on the literature reviewed in Chapter Two that the students would get worse on these skills as the content gets harder. The difficulty ranks play an important role in giving students credit for the improvements they make.

The curves on the development graphs are the principle source for many of the interpretations of this study. Understanding both their shape and their position relative to the skill bands is important in interpreting the results that occur in the subsequent sections of this chapter. Development is defined not only as performance crossing skill bands, but includes staying even with skill bands as the year progresses.

The analyses in the following sections all have a common format. First the skill will be introduced. Then predictions will be made for each cohort's performance. Then the EPS cohort will be described followed by the TRD cohort. Each skill discussion will end with a set of interpretations.

General Approach

The first problem-solving skill assessed was General Approach (GA). The codes assigned to this skill measured the student's initial qualitative approach. Essentially this

Table 5.1

General Approach Codes

0	Nothing written
1	Physics approach is inappropriate. Successful solution is not possible
2	Physics approach is appropriate, but the manner of its application indicates a fundamental misunderstanding.
3	Physics approach is appropriate, but a wrong assertion is made as a serious misinterpretation of given information.
4	Physics approach is appropriate, but neglects one or more other principles necessary for the solution.
5	Physics approach is appropriate and all necessary principles included, but errors are evident.
6	Physics approach is appropriate and all necessary principles included without any conceptual errors.

skill determined if the students were in the right mental space and if this mental space was complete and correct for the given problem. It is in this skill that any conceptual error the student made will be identified. Table 5.1 shows the codes used to measure this skill.

Expectations – General Approach

In interpretative research, there are two overlapping traditions depending on how theories are generated and used. One tradition is called grounded theory and in this tradition theories are built from the data upward. The other tradition uses existing theories and looks for confirmation of these ideas. This study is definitely in the second tradition, with the occasional foray into hypothesis generation. An extensive review of the expert-novice problem-solving literature forms the basis of this study. Recognizing that a theory exists, predictions can be generated before examining the data. The expectation section of this and the next three problem solving skills will layout the predictions for each cohort's performance.

For the EPS cohort, the prediction for the General Approach problem-solving skill is good. Since the Minnesota Problem-solving Strategy emphases an initial qualitative analysis of the problem (Larkin, 1979), these students should be in the correct mental space and it should be reasonably complete. Also, given the high FCI and other multiple choice test scores by the EPS cohort, there is evidence that the students are not making excessive conceptual errors (Hestenes, Wells, & Swackhammer, 1992). Since perfection is never expected for an average, the prediction is that the EPS cohort should have a development graph in the GA = 5 skill bands for most of the study. This skill bands denotes almost perfect solutions. However, since the EPS cohort is a representative cross-section of the course, not all the students will do this well. The EPS cohort average might be lower, but there should be some students in the highest bands.

In the TRD cohort the students did not have the benefit of explicit instruction in problem solving. They did have the benefit of cooperative group problem solving in discussion sessions and in the laboratory. Therefore, one might expect the students to do well in their general problem solving, but be more prone to forget important principles. Both of these issues combined to produce curves around the GA = 4 skill band. However, since the TRD cohort over-represented their class, the average might be higher.

<u>EPS cohort – General Approach</u>

Figure 5.4 shows the EPS cohort skill development curve. The curve clearly shows growth. The students stay at the same level as the material gets more difficult.

However, the level of the curve was lower than would have been expected if all the students had been helped by the explicit problem-solving strategy. Since the students had to learn new physics and then perform nearly perfectly on these problems, the fact that the curve is in the GA = 4 skill band instead of the expected GA = 5 skill band was not too surprising in hindsight. This level suggests that the students in the cohort generally knew what they were doing, but still making errors of omission.

Figure 5.4 also shows some interesting fluctuations at a few points. The data point for the first two quizzes (T1-Q) may actually be in the GA=4 band. It certainly bridges the gap. The more interesting data point is the last one. On the third term final exam (T3-F) the curve drops into the lower part of the GA=4 skill band after being in the

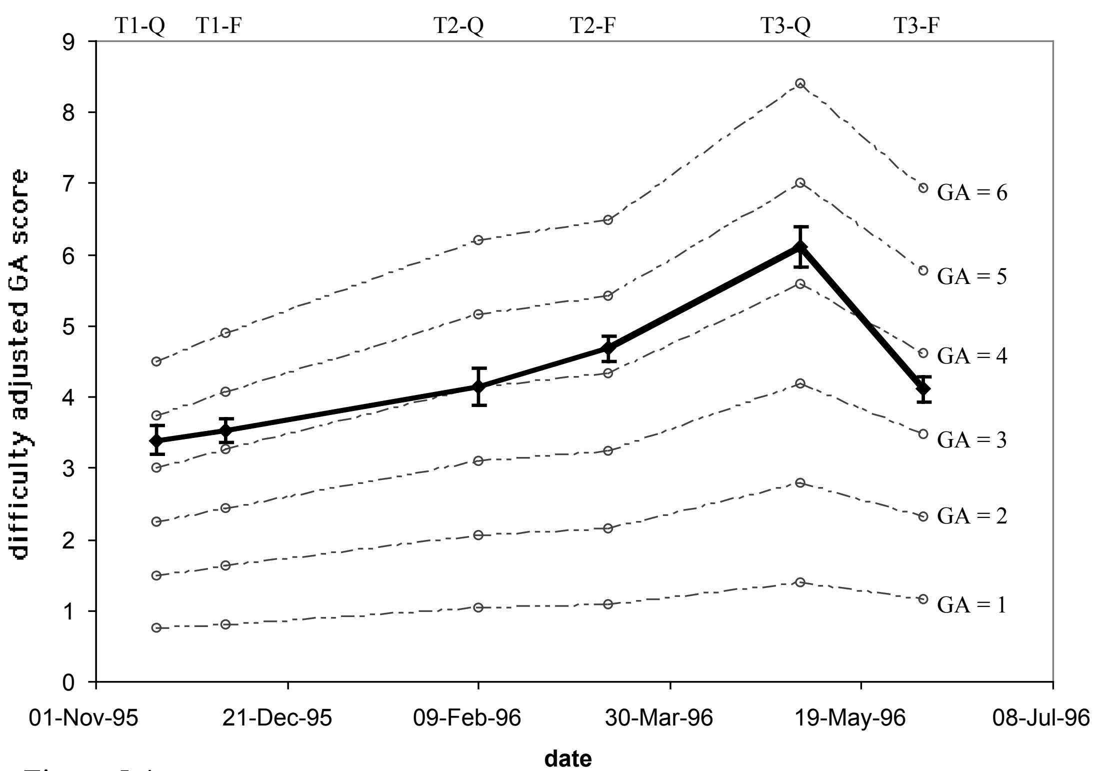


Figure 5.4

EPS Cohort development graph. N=24 students higher part. Clearly the students were making more fundamental errors on the final exam than they were making on the quizzes of that term. It is not obvious why this would be the case only for the third term final. This mini-drop off will be revisited later in this chapter. Additionally, more information is available about the EPS cohort by examining individual students within the cohort.

EPS Clusters – General Approach

The analysis presented so far has been a rough look with the cohort as the unit of analysis. This analysis has been the geological equivalent of considering two rocks to be equivalent without cracking them open to see if one is an agate. Consistent with the interpretative paradigm in which this study was based, the individual student development graphs (presented in Appendix B) were holistically sorted based on general shape and magnitude of values. These holistically sorted groups are referred to in this study as clusters. The cluster development curve was produced by averaging the scores by problem of each student in the cluster. The resulting clusters are reported in this section and the actual breakdown of which students are in each cluster can be found in Appendix C.

Figure 5.5 shows the EPS cohort clustered by their General Approach skill curves. There are three students in the "low" cluster and student EPS25 was the most characteristic of these students. There are twelve students in the "mid" cluster with student EPS22 as the most representative student. The four students in the "hi" cluster are represented best by EPS 23. There were five students who did not fit into any of these clusters.

The "hi" cluster is clearly hitting the ceiling through most of the year, except for the third term final exam. The "mid" cluster tracks the GA = 4 skill band pretty well and also shows a drop at the third term final. Both of these clusters show growth throughout the year. This is unlike the "low" cluster who waits until the second term quiz (T2-Q) to begin to seriously track the GA = 3 line. However any growth this group showed was lost on the third term final exam. The "low" students were not doing sensible things in their approach to the problems.

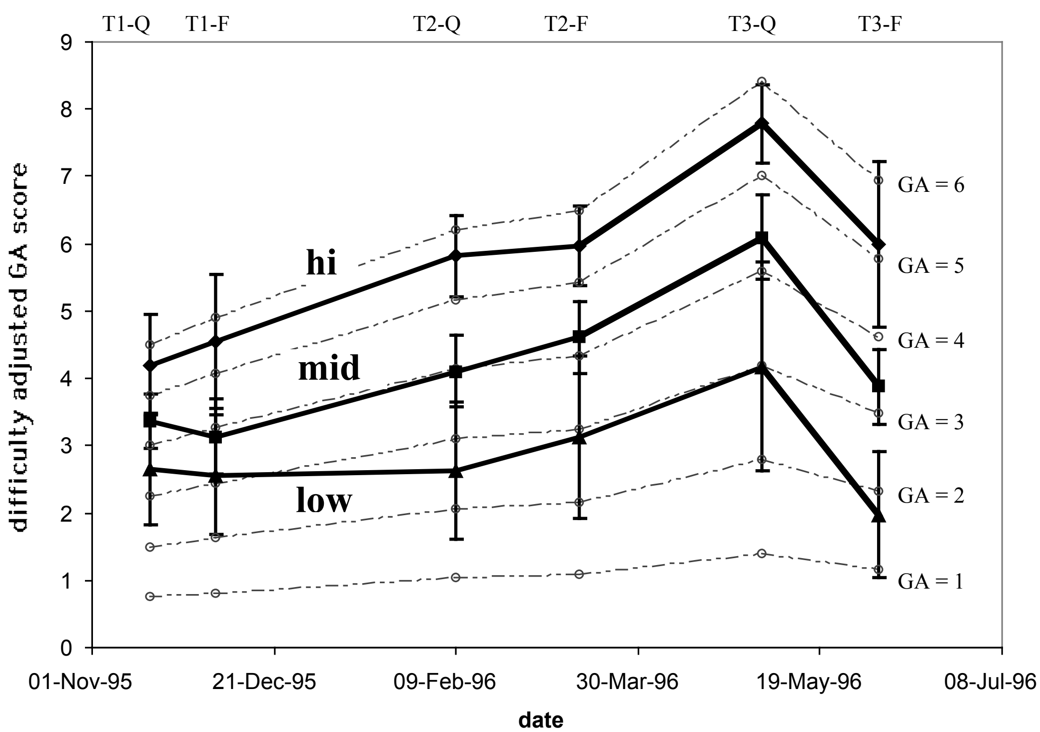


Figure 5.5

EPS Development graphs sorted by Clusters. N = (hi: 4, mid: 12, lo: 3)

TRD Cohort – General Approach

The TRD cohort was included in this study to provide reference for the EPS cohort. While it would have been possible to use this theory to provide the same reference, as we have already seen from the EPS expectations this is not always accurate. There are also some abnormalities in the EPS development curves that would remain mysterious until a replication study could be completed. A better technique would be to apply the same measures to another class of students. Unfortunately the TRD cohort was not a fair sample of the TRD class. It has more students from the top end of the class than the EPS cohort. This means that the TRD cohort represents a better than average pool of students. In spite of this shortcoming, the TRD cohort still provides the best context to judge how the EPS cohort did in developing problem-solving skills.

Figure 5.6 is the General Approach development graph for the TRD cohort. It was produced in the same manner as those for the EPS except that the TRD problem difficulty ranks and students are used. Since the TRD difficulty ranks are different than those used for the EPS cohort, slightly different skill bands exist. For this reason it is problematic to plot the EPS and the TRD cohorts on the same set of axes. Instead the skill bands will provide common ground between the cohort's development graphs.

The TRD cohort development curve on Figure 5.6 is similar to the EPS curve seen in Figure 5.4 in three important ways. First, it too shows growth during the year. The students keep pace with increasingly more difficult material. Second, the TRD stays in the GA = 4 band for most of the year. This was the same band the EPS cohort stayed near. Third, the TRD cohort also showed a drop on the third term final exam. There are also some differences. The TRD cohort began firmly at GA = 4, while the EPS cohort started almost in GA = 5. This is evidence that the General Approach problem-solving skills of the two cohorts did not start off the same. Another difference is that the TRD cohort appears to be in the GA = 5 skill band for the third term quizzes, while the EPS cohort doesn't appear to get this high.

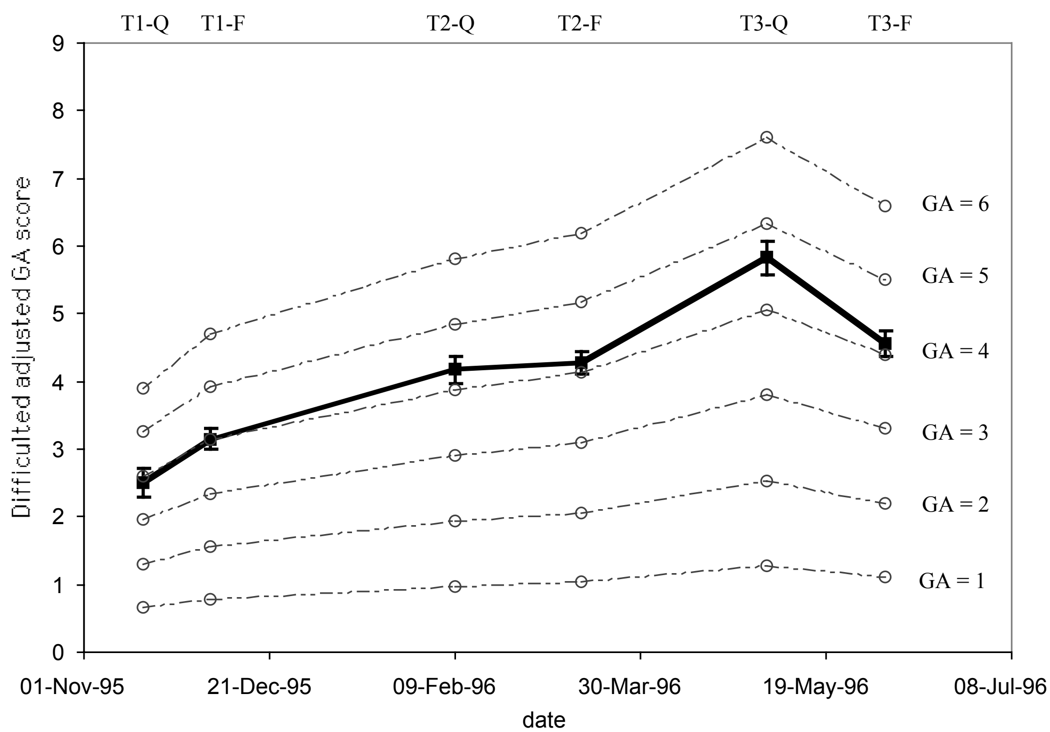


Figure 5.6

General Approach Development Graph for the TRD cohort. N = 24.

TRD Clusters – General Approach

As was done for the EPS cohort, the students in the TRD cohort were holistically sorted into clusters of similar growth and performance. The development curves of the clusters are shown in Figure 5.7. Student TRD11 best characterizes the low cluster of three students. The ten middle students are similar to student TRD19 and TRD13 was the student most similar to the five students of the high group. There were six students who did not fit into any of the groups. Appendix C details which cluster each student is in.

There are several important observations from these curves. All three clusters start at the same place in the GA = 4 band. The students in the "hi" cluster made

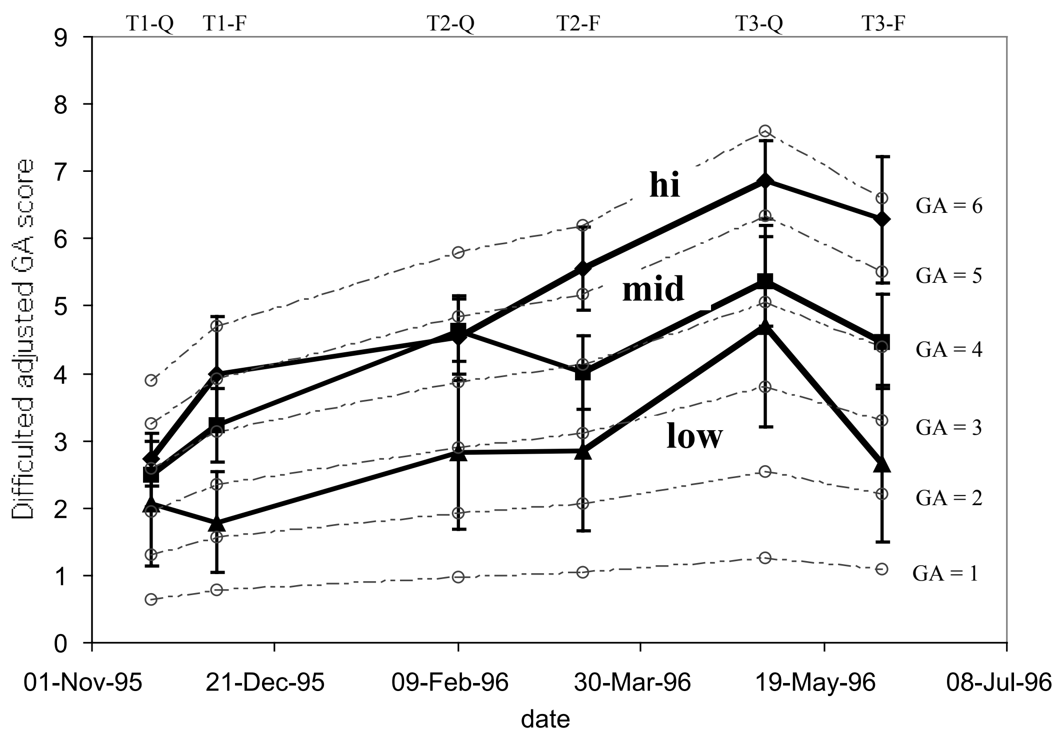


Figure 5.7

TRD Development graphs sorted by Clusters. N = (hi: 5, mid: 10, low: 3) remarkable progress getting into the highest bands. However, unlike the EPS "hi" cluster, the TRD "hi" cluster doesn't start at the ceiling, but rather the students work up to this height over time. It is at the ceiling by the second term final exam. The "mid" group made an odd improvement out of the GA = 4 band on the second term quizzes, but they could not sustain this improvement into the final exam. The TRD "low" cluster behaved similarly to the EPS "low" cluster. Again we see the "low" cluster hovering around the GA = 3 skill band after a fall from the higher skill bands on the first data point. Even the slip into the GA=2 skill band on the third term final was again repeated, although perhaps not as dramatically. The principle difference was the TRD "low" cluster peaked on the third term quizzes, which may have inflated the TRD cohort average on this datum.

General Approach Summary

The first problem-solving skill analyzed was General Approach. This was a measure of the appropriateness of the mental space the students were in when creating their problem solutions. The effect that teaching an explicit problem-solving strategy had on this skill beyond what cooperative group problem-solving gave the students was minimal at the cohort level. There was only one difference between the EPS and the TRD cohorts at the cohort level. This was the slighter higher average score on the first term quizzes. Otherwise, both cohorts were essentially at the GA = 4 level and showing development through the year. However, since the TRD cohort was an over representation of their class, this suggests that one effect of the problem-solving strategy was to equalize the cohorts. In effect, an average group of students looks like better-than-average students.

The cluster level analysis revealed another benefit to the explicit teaching of a problem-solving strategy. The best students in the EPS cohort get to the highest skill level faster and stay in this band throughout the year. The "hi" TRD cluster did eventually reach this same level, but the problem-solving strategy seemed to help the best students become better sooner. Again, the lower two cluster seemed about the same suggesting that the problem-solving strategy raises the level of average students to better than average performance.

One trend that is clearly evident in the EPS cohort's clusters yet missing in all but the low TRD cluster was a slight drop on the third term final exam. The student's lack of knowledge cannot be the cause. The students clearly have some knowledge of the third term content. They do well on the quizzes and they answered over half of the multiple-choice items correctly. A poor attitude about the course is also unlikely since the MPEX failed to show a dramatic drop. Since the students generally showed growth through exams, it cannot be a final exam effect.

It is a hypothesis of this study that the reason why there is a third term final exam slump is due to the organization of the course content of the third term. The standard presentation of electricity and magnetism concepts is very disjointed. Unlike through mechanics where energy and forces unite the various contexts, in electricity and magnetism the students are hit by one abstract and disconnected concept after another. They can prepare for quizzes by studying the relevant concepts, but by the final exam, the disjointed presentation takes its toll. The students lack an accessible framework for the topics and cannot put themselves in an appropriate mental space (Eylon & Reif, 1984). Their work suffers from the very start. Adding insult to injury it appeared that the students began to abandon the explicit problem-solving strategy at this same time.

Specific Application of Physics

The second skill measured was Specific Application of Physics (SAP). This skill assessed how well the students applied the physics they thought they needed to apply. Table 5.2 shows what is included in Specific Application of Physics. A student's Specific Application of the Physics score was dependent upon their General Approach, so even if the concepts applied are not wholly appropriate for a successful problem solution; the application of those concepts were still judged. However a quick inspection of the codes show that the lower codes of General Approach (GA = 0) and Specific Application of Physics (SAP = 0, 1 or 2) are actually dependant on each other. This statistical relationship between these skills was determined for the cohorts in Chapter Four. For the EPS cohort, General Approach and Specific Application of the Physics were correlated. The relationship is less clear for the TRD cohort were no such relationship was found.

Finally, if the students left their solution blank, then these codes were removed from the averaging for both the cohort and the cluster development graphs. A blank paper does not show development of this skill. Only when the students apply their knowledge is this skill assessed. However a blank solution is information. The number of data points removed are recorded in an accompanying table.

Table 5.2

Specific Application of Physics

0	Nothing written.
1	Difficult to assess (GA#2).
2	Solution does not proceed past basic statement of concepts.
3	Vector/scalar confusion, or specific equations are incomplete, or confusion resolving vectors into components.
4	Wrong variable substitution: Poor variable definition.
4.5	Wrong variable substitution: Difficulty in translating to a mathematical representation.
5	Careless use of coordinate system without a coordinate system defined.
5.5	Careless use of coordinate system with a coordinate system defined.
6	Careless substitution of given information.
7	Specific equations do not exhibit clear inconsistencies with the general approach, but hard to tell due to poor communication.
7.5	Specific equations do not exhibit clear inconsistencies with the general approach and the solution is clear.

Expectations – Specific Application of Physics

The explicit problem-solving strategy should have had a large impact on the students' SAP scores in the EPS cohort. The Minnesota Problem-solving Strategy (Heller & Heller, 1995) strongly encourages drawing diagrams, label axes, and defining variables. These specific traits are highlighted by the higher SAP codes. Therefore the EPS cohort ought to do well on this skill, perhaps as high as the SAP = 6 band and develop these skills quickly. Of course, this optimistic prediction assumes that the students internalize the strategy and not just use it as window-dressings. Also the lowest performing students might still have errors applying their skills to even create physics equations (SAP = 3). Conversely the TRD cohort does not have this scaffolding so these

Table 5.3

Number of blank solutions in the EPS cohort over time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Number of solutions	1	7	6	7	3	9
Percent of all solutions	2.1	4.9	8.3	4.9	4.2	5.4

Note: T1 denotes the first term. Q denotes quizzes and F is the final exams

errors might be prevalent. Typically scores might range around the SAP = 4 band, but given the TRD bias, perhaps SAP = 5 would be ultimately possible as the year progresses because there are few low performing students to drag the average down.

EPS Cohort – Specific Application of Physics

One important measure of the EPS cohort's performance was to examine the number of solutions students left blank. These solutions are removed from the analysis of the development curves since they do not represent skill development. Table 5.3 reports the number of blank solutions over time. From this table it can bee seen that except for the first term quizzes, about five percent of the solutions were returned blank. The second term quizzes were also atypically high, but given the small numbers this could be a fluctuation.

Of course the principle data is the development graphs. Figure 5.8 shows the EPS cohort's average scores for the development of their Specific Application of the Physics skill. The EPS cohort does not show growth on his skill. It starts at a very respectable SAP = 6 band. It was not the case that the cohort, on average, misread every problem. Rather the occasional error most likely involved not using the coordinate system (SAP

=5). However, as the year progressed, the EPS cohort left this skill band and hovered near the SAP = 5 skill band. The students were making errors involving coordinate systems and variable substitution errors.

This was a surprising result. As discussed in the expectation section, the explicit problem-solving strategy should have prevented the very types of errors seen in the cohort average. It is important to notice that the EPS cohort began at the next-to-highest skill band. This might suggest that the Minnesota Problem-solving Strategy, as taught, was most effective for the mechanics problems used in the first term. This conclusion would not be entirely surprising since the strategy was originally developed for this class

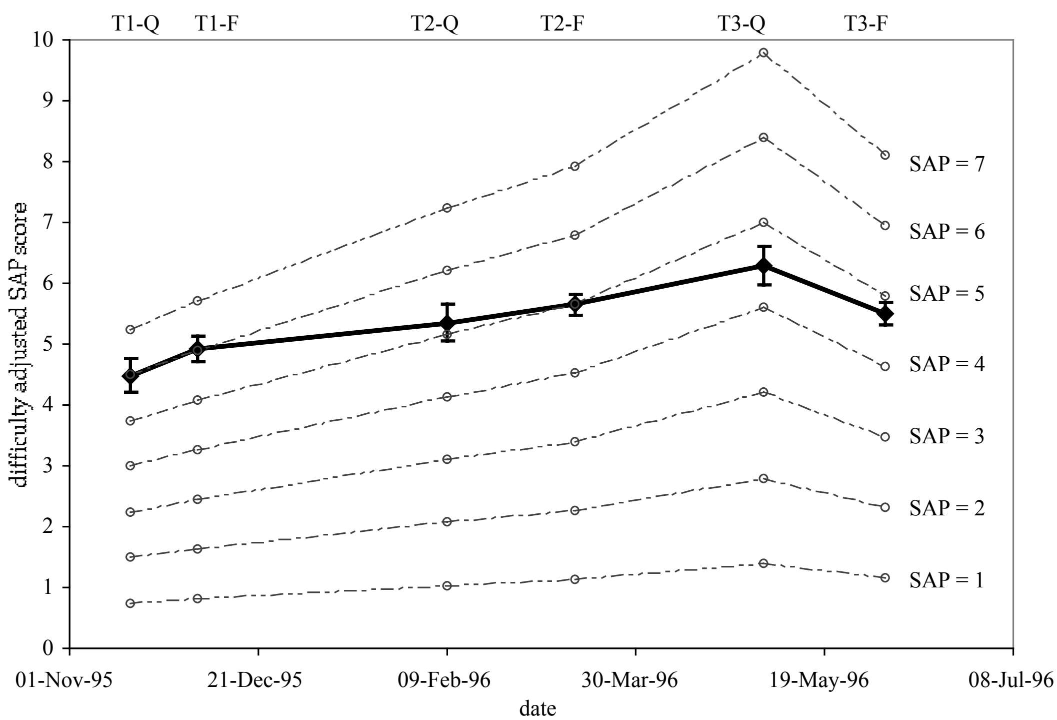


Figure 5.8

EPS Cohort Development Graphs for Specific Application of the Physics.

N = 24 students

of problems and content area.

EPS Clusters – Specific Application of Physics

The next step in the process is to examine the performance of the EPS clusters. Figure 5.9 shows the cluster curves. There were two students in the "low" cluster. The 13 middle students are similar to student EPS14. EPS5 was used to select the four high group students. One new clustering for the EPS cohort emerged for Specific Application of Physics. Three students fell into a group labeled "hhm." Appendix C detailed into which cluster each student fell.

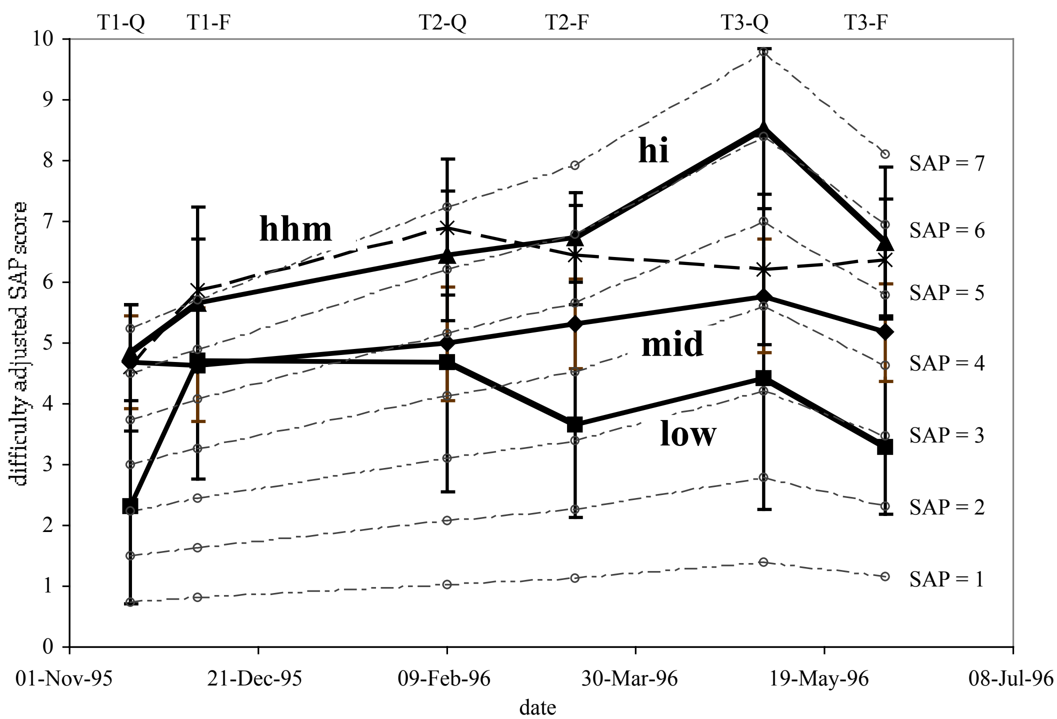


Figure 5.9

EPS Development graphs with Clusters. N = (hi: 5, mid: 13, low: 2, hhm: 3)

The students in the "hi" cluster had a surge of growth on the first term quizzes moving into the highest band and then settled in around the SAP = 6 band. This suggests that in general these students were doing well with occasional errors. It is unlikely that they were misreading every problem statement. The three "hhm" cluster students looked like the "hi" cluster students until the third term quizzes. This was the term where electro-magnetism was covered and it was not surprising to see some students, even good ones, have difficulty with this material. However, since they were good students, they rebounded to the "hi" level by the end. The "mid" and "low" cluster students showed little growth. The "mid" cluster students began around SAP = 6 and proceeded to hover under the SAP = 5 band for the second term until the third term where they slip to SAP = 4 without recovering. These students went from being able to apply their physics knowledge to making variable substitution errors and worse. The "low" cluster did worse. After a fabulous growth spurt on the first term final, these two students looked like the "mid" students for a while until they began to make very serious errors in applying their physics knowledge. All told, the students in the EPS clusters did not perform as well as predicted. These students were taught an explicit problem-solving strategy designed to prevent the class of errors they were making. Something is amiss, but before a stronger conclusion can be reached the TRD cohort should be inspected.

TRD Cohort – Specific Application of Physics

The TRD cohort also had students who did not complete the solutions. These numbers are shown in Table 5.4. Clearly the TRD cohort, which had better than average grades, had learned not to turn in blank quizzes. On the final exams, which had six

Table 5.4

Number of blank solutions in the TRD cohort over time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Number of solutions	2	4	0	4	0	8
Percent of all solutions	4.2	2.8	0	2.8	0	4.8

problems, some students apparently decided that a few blank problems wouldn't hurt their grade too much. Or the added burden of a comprehensive final exam left a few students unable to do all the problems.

Figure 5.10 shows the development curve for the TRD cohort. The TRD cohort

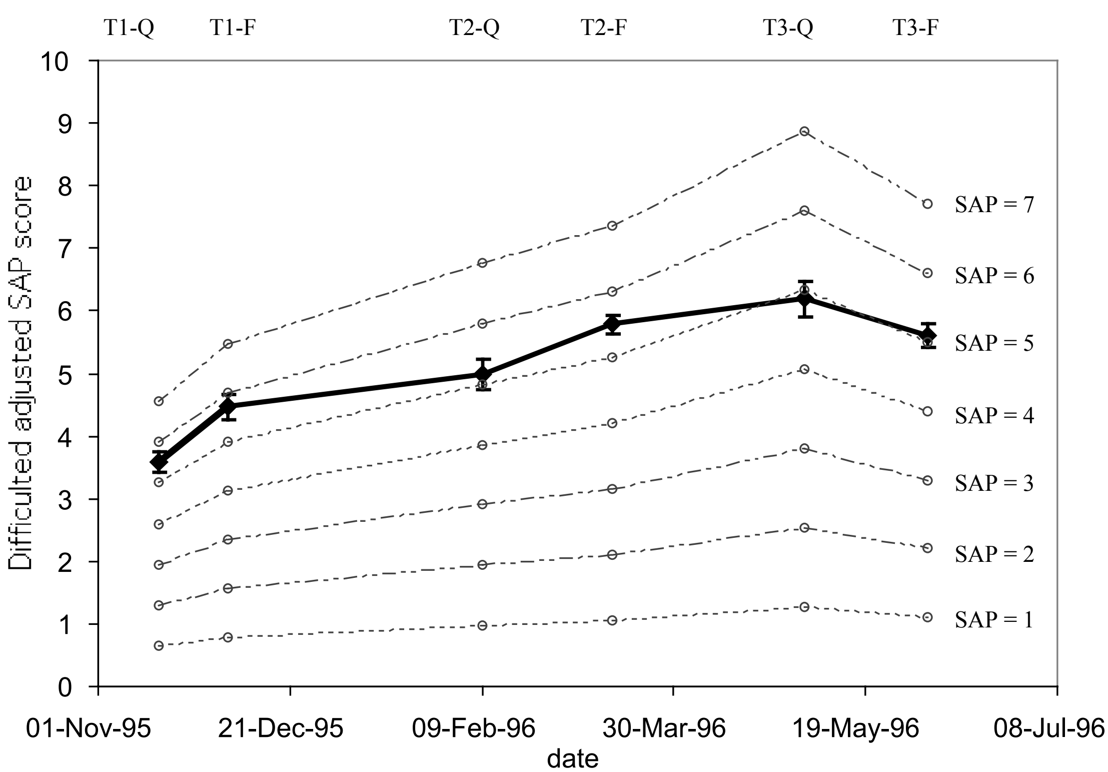
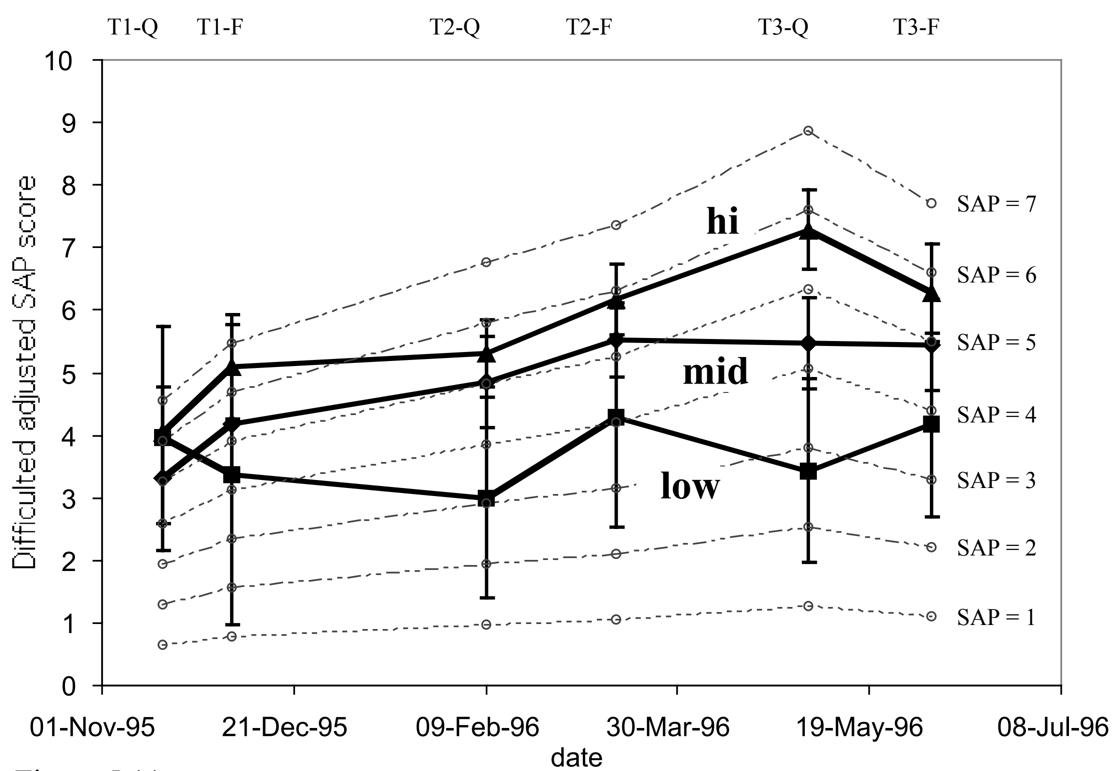


Figure 5.10

TRD Development graphs for SAP skill. N = 24 students looks very similar to the EPS cohort. They both started strong and then hovered around the SAP = 5 line. Therefore we see that the EPS cohort, on average, looks like the better-than average TRD cohort. The explicit problem-solving strategy did have an impact on this SAP skill because it equalized the cohorts, but this performance is not as high as expected. Minimally, the problem-solving strategy did no harm to the students.

<u>TRD Cluster – Specific Application of Physics</u>

Figure 5.11 shows the development curves for the TRD cohort. There were two students in the "low" cluster. The nine middle students are similar to student TRD10 and TRD9 was the student most similar to the nine students of the "hi" cluster. There were



<u>Figure 5.11</u>

TRD Development graphs with Clusters. N = (hi: 9, mid: 9, low: 2) four students who did not fit into any of the groups. Appendix C detailed into which group each student fell.

The TRD "hi" cluster and the EPS "hi" cluster are similar. Both clusters were around the SAP = 6 band, only the EPS "hi" cluster tended to be above it while the TRD cluster tends to be beneath it. The TRD "mid" cluster in contrast seemed to stay with the SAP= 5 much better than the EPS "mid" cluster did. The TRD "mid" cluster showed growth up until the third term final where they slipped into the SAP =4 band. Again it was not surprising to see students slip when entering the third term. The two "low" cluster students were erratic. They began the study very well, but quickly fell to the SAP = 4 band where they stayed for every final. During quizzes these two students slipped into the SAP = 3 band. Clearly on quizzes these students had a hard time applying their physics, but on the final exams they did slightly better than on the quizzes. This trend was not seen in the EPS low grouping.

In addition to the interpretations seen from the development graphs, other interpretations are available. For example, the higher number of students in the TRD Specific Application of the Physics high cluster as compared with the TRD General Approach high cluster was partial confirmation of the lack of correlation between General Approach and Specific Application of physics for the TRD cohort discussed in Chapter Four. Also there is rhetoric that teaching an explicit problem-solving hurts the best students. This is evidence from comparing the EPS and TRD cluster development graphs that teaching a problem-solving strategy did not hurt the best EPS students.

<u>Specific Application of Physics Summary</u>

The second problem-solving skill measured was Specific Application of Physics. This was a measure of how well the students applied the physics they thought they needed to apply. Therefore it did not matter if the students had made a conceptual error, only what they did with it. This measure includes items such as variable substitution errors and direction errors.

This skill also measured if the students produced blank solutions. It was seen that students in both cohorts only rarely turned in blank solutions for the problems. The EPS cohort had a moderately constant level of blank solutions while the TRD cohort only turned in blank solutions on exams.

At the cohort level, the EPS cohort did not perform as well as expected. The cohort hovered around the SAP = 5 band suggesting that on average the students were not making good use of coordinate systems. This in spite of the explicit strategy designed to help the students avoid these classes of errors. When compared to the TRD cohort, both cohorts performed at about the same level, including falling to the SAP = 5 skill band after starting near the SAP = 6 skill band. Clearly, as the term progressed, both cohorts began to make application errors. The TRD cohort again demonstrated that the explicit strategy did have an effect since the EPS cohort was as good as the better grade-performing students of the TRD cohort.

At the cluster level, both cohort's "hi" clusters were in the SAP = 6 band, only with the EPS cluster near the top of the band and the TRD nearer the bottom of the band. Neither "mid" cluster performed well on this skill with both of them leveling out when the third term started. This effect was also seen in a new cluster ("hhm") in the EPS cohort. The two "low" clusters were sparsely populated and did not give consistent feedback except to say that these students make serious application errors.

The net effect of the EPS performance on the Specific Application of Physics skill suggests a problem. This skill should have been a strength for the EPS cohort given the structure of the Minnesota Problem-solving Strategy. However, this was not the case. Perhaps the Minnesota Problem-solving Strategy does not adequately address the errors (either by design or execution) the students are making. Or perhaps the students have not internalized the strategy. For example, some students made direction errors in spite of having drawn a coordinate system. Whether the problem is with the instruction or the students' use of the strategy, a problem seems to exist.

Logical Progression

The third problem-solving skill assessed was the students' Logical Progression. Logical Progression is a measure of the planfullness of the student's solution. The codes measure the cohesiveness of the solution. It also measured whether a student worked forward or backwards. Table 5.5 lists the codes used to assess the student's Logical Progression. Also, if the students left their solution blank, then these codes were removed from the averaging for both the cohorts and the clusters. The number of data points removed will be recorded in an accompanying table.

Table 5.5

Logical Progression

0	Nothing written.
1	Not applicable - one step problem.
2	The use of equations appears haphazard and the solution is unsuccessful. Student may not know how to combine equations.
3	Solution is somewhat logical, but frequent unnecessary steps are made. Student may abandon earlier physics claims to reach answer.
4	Solution is logical, but unfinished. Student may stop to avoid abandoning earlier physics claims.
5	Solution meanders successfully toward answer.
6	Solution progresses from goal to answer.
7	Solution progresses from general principles to answer.

Expectations – Logical Progression

The EPS cohort should have high values for Logical Progression for two reasons. Not only did the explicit problem-solving strategy provide a framework that the students used to proceed through a problem in a logical fashion, but the EPS students were graded for their logical progression. This should be the real strength of teaching an explicit problem-solving strategy. The best students in the EPS cohort should be at the LP = 7 band. The EPS cohort might also possess more LP = 4 instead of LP = 3 codes since it was the process and not the final answer which was graded. The very lowest codes might vanish all together somewhere in the middle of the course. The average for the EPS cohort will probably be around LP = 6 given the above.

The TRD cohort should have an average Logical Progression score around the LP = 5 band at least initially. The TRD cohort may reach the LP = 6 band sometime during the year because these were better than average students and they could be expected to

Table 5.5

Number of blank or one-step solutions in the EPS cohort over time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Number of LP = 0 solutions	1	6	8	10	13	11
Number of LP = 1 solutions	0	0	0	5	0	18
Total	1	6	8	15	13	29
Percent of all solutions	2.1	4.2	11.1	10.4	18.6	17.3

progress toward expert like behavior on their own. However, since the final answer it typically important in grading, the TRD students may have more LP = 3 than LP = 4 since stopping a solution would not get them much partial credit.

EPS Cohort – Logical Progression

To describe the Logical Progression of the EPS cohort, more than just the development graph needed to be examined. There are two additional data sources. First, Table 5.5 shows the number of students who received codes LP = 0 and LP = 1. Neither of these codes measures the student's Logical Progression. For example, a student who started the problem, but did not write anything else (i.e. SAP = 2, LP =0), did not display any progression to their solution. Or if a student turned an otherwise multi-step problem into a one-step solution through an error, then discussing progression for their solution in a misrepresentation of this skill. These codes are not included in the development graphs.

The second additional source of data is shown in Table 5.6. This table displays the number of students who were coded with the terminated solution codes. This codes reflect students who either could not finish what they started, or abandoned earlier

Table 5.6

Percent of students with the three lowest LP codes in the EPS cohort by time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Percent of LP = 2 solutions	2.1	1.4	1.4	0.7	0.0	0.0
Percent of LP = 3 solutions	6.2	4.2	8.3	6.2	6.9	7.1
Percent of LP = 4 solutions	12.5	9.7	18.1	11.8	23.6	14.3

physics, or stopped. Students who stopped typically stopped because they ran out of time or they caught an error and chose to stop.

Tables 5.5 and 5.6 reflect different aspects of the EPS cohort. From Table 5.5 it is evident that as the year progressed the students in the EPS cohort were increasingly unable to progress very far on their solutions. When Table 5.6 was also considered a sizable portion of students did not finish their solutions. Table 5.6 shows that the number of students who completed haphazard solutions dwindled to zero as the year progressed. Also there were more students who choose to stop than violate physics. Furthermore, the percent of students who scored an LP =3 were generally about seven percent.

The data for Table 5.6 is included in the development graphs for the EPS cohort shown in Figure 5.12. The most striking feature of this curve is its magnitude. The average is quite high for the students who could completed their solution. There also appeared to be at least stable growth through the new and more difficult material. The cohort began the year firmly in the LP = 6 band. There was a drop on the second-term quizzes, which was repeated on the third-term quizzes. It is likely that the time pressure on quizzes forced some students to rush and be less systematic or stop (Table 5.6).

However, by the respective final exams, the EPS cohort was closer to the LP = 6 band again. This result fit what was expected for a representative selection of students from a class where an explicit problem solving strategy was taught. The next step is to examine the performance of the clusters.

EPS Clusters – Logical Progression

The cluster curves shown in Figure 5.13 are not as diverse as the other problem-solving development graphs. This is due to the removal of the lowest two codes. This caused the low grouping (Appendix C) not to be plotted. For example, on the third term final exam only about one-third of these students' solutions could be included. The three

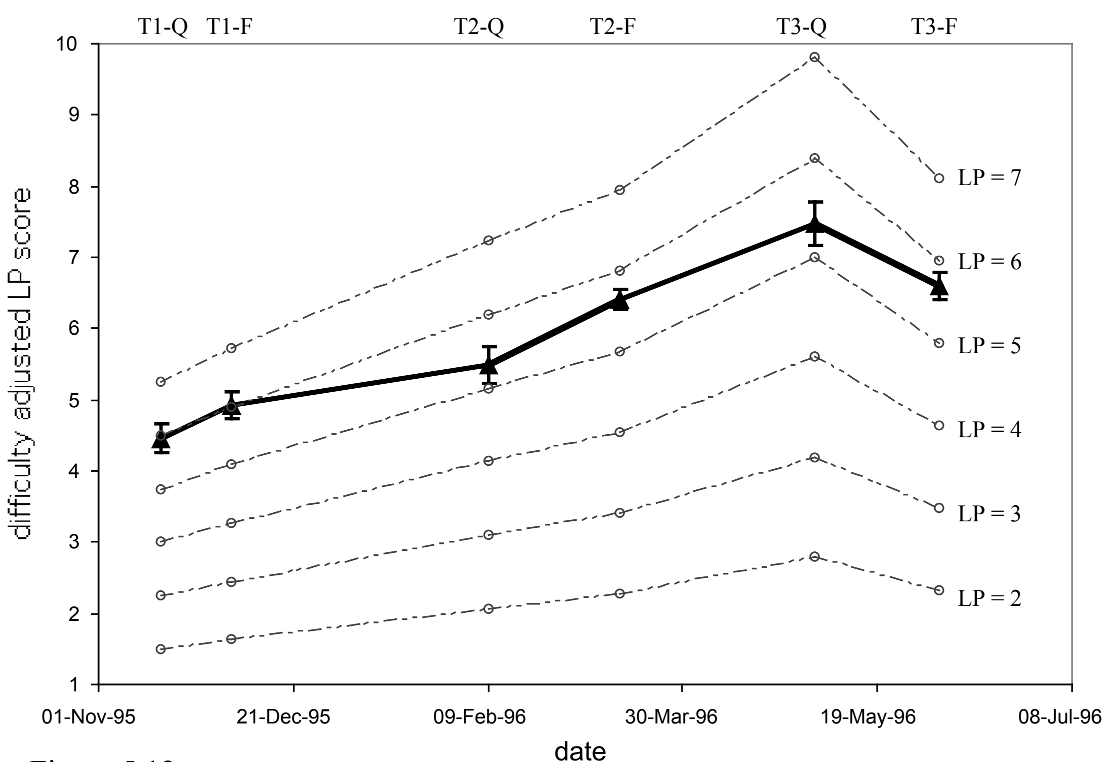


Figure 5.12

EPS Cohort LP Development Graph. N = 24.

low students, most similar to student EPS10, accounted for most of the lowest codes. In contrast, the "lo-mid" cluster benefited just slightly from this effect with over five-sixths of the problems included in the curve. This lower mid-range cluster was labeled "lo-mid," had four students, and EPS2 was the representative student. The higher mid-range cluster also had four students and was labeled "hi-mid." EPS3 was the most characteristic student. EPS20 was used to represent the five high group students. One new cluster for the EPS cohort emerged for Logical Progression. This cluster is labeled "mms," where the "s" denotes the spike on the third term quizzes. Five students defied grouping.

All four clusters drawn show a slight loss of logical progression during the year.

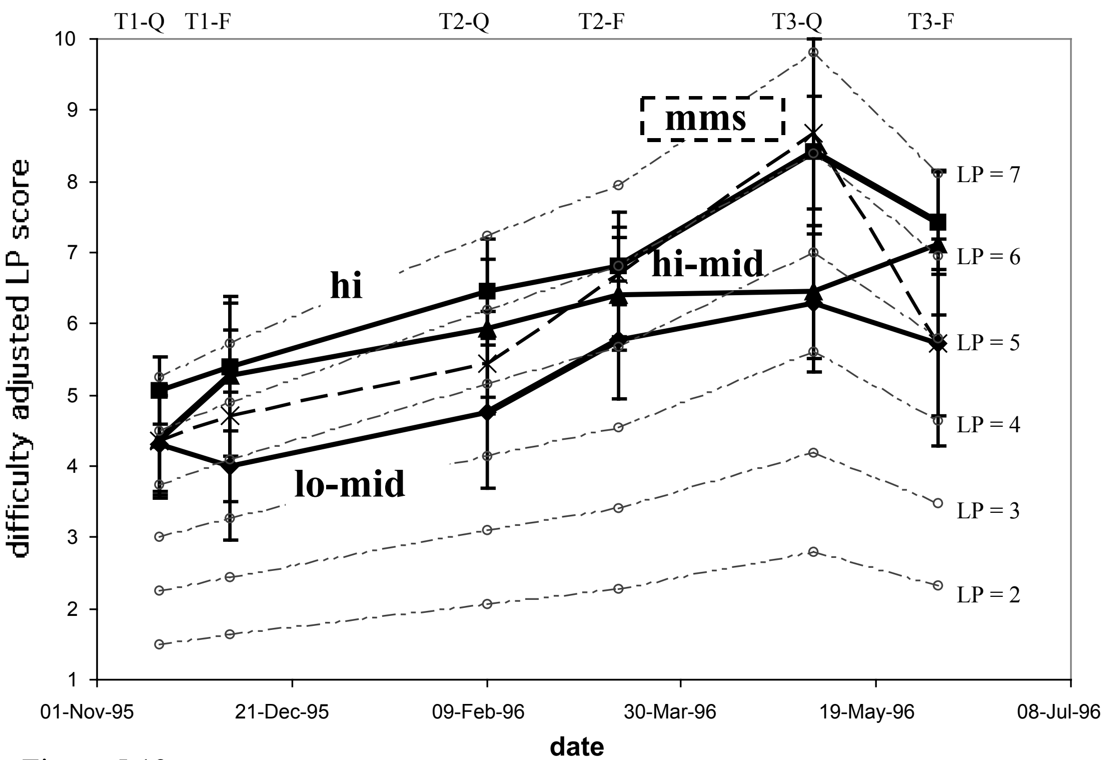


Figure 5.13

EPS Development Graph by Cluster. N = (hi=5; hi-mid=4; lo-mid=4; mms=3)

They also, in general, show higher scores on finals than quizzes. The "hi" cluster starts the year off at the ceiling, but by the second term they are firmly at the LP = 6 band. The "hi-mid" cluster was generally around the LP = 6 band, except for the quizzes of the third term where they slipped down a band.

The "lo-mid" cluster was similar except for the initial slip on the first term final. From this point onward, the "lo-mid" cluster looked like the "hi-mid" cluster, except it was around the LP=5 band. The third-term quiz slip was not surprising given the difficulty of the third term material. The "mms" cluster was a surprise. Here were three students who appeared to be "mid" cluster students until the third term where they did very well on the quizzes, yet lost any year-long gain on the third-term final. Why they did so well on creating a logical solution for the last set of quizzes, yet fail to do so for the subsequent, untimed exam remains a mystery, even after considering the TRD cohort.

TRD Cohort – Logical Progression

The discussion of the TRD cohort begins by examining the number of students who started problems (i.e. had a non-zero SAP score), but either wrote nothing or misconstrued the problem and made it a one-step solution. This information is in Table 5.7.

The numbers in Table 5.7 for the TRD cohort look very different than those seen for the EPS cohort in the third term of the course. The EPS cohort had about twenty percent of the students' solutions for each of these last two columns. At first glance the explanation that the TRD cohort was a better-than-average sample would seem to suffice.

Table 5.7

Number of blank or one-step solutions in the TRD cohort over time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Number of LP = 0 solutions	0	3	7	1	5	9
Number of LP = 1 solutions	0	1	2	11	2	7
Total	0	4	9	12	7	16
Percent of all solutions	0	2.8	12.5	8.3	9.7	9.5

However, on closer inspection it was seen that while these numbers occurred uniformly for the TRD cohort, they were concentrated on particular problems for the EPS cohort. Thirteen percent of the third-term quizzes LP = 0's occurred on one problem (T3-Q3). In contrast only 2 students in the TRD cohort left this problem with only an approach. The EPS cohort students did not proceed past a statement of the problem for this very difficult problem which was a direct result of most of the students not having a complete understanding of the problem (GA=4). It appeared that they wrote nothing because they knew they didn't know enough. These maybe proto – LP = 4 students, but they did not write enough to code them as LP =4. The other abnormal problem was T3-F1. This problem accounted for seven percent of the EPS third-term final exam LP = 1's. Here the error was a conceptual one. These students used the concept of electric potential energy instead of electric potential to solve the problem. Such an approach leads to a one-step solution. With these two problem removed, the EPS cohort did not appear to be much different than the TRD cohort.

Table 5.8

Percent of students with the three lowest LP codes in the TRD cohort by time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Percent of LP = 2 solutions	8.3	1.4	4.2	1.4	2.8	1.2
Percent of LP = 3 solutions	20.8	4.2	4.2	5.6	12.5	10.1
Percent of LP = 4 solutions	10.4	9.7	9.7	9.7	15.3	7.1

One of the predictions for the TRD cohort was that there would be students who would abandon their earlier physics claims to get to any answer. This data is in Table 5.8. When compared to the EPS cohort in Table 5.6, two trends are evident. The LP =2 codes do not taper off for the TRD cohort. There are still students who flounder (perhaps looking for some partial credit) when attempting problems instead of writing nothing. The second trend is that there are proportionally more LP=3 solutions than in the EPS cohort in the third term. These codes did not have a pattern for the TRD cohort. They occurred through out the third term problems in roughly equal numbers and not always with the same students. These better-than-average students had somehow learned that getting an answer was more important than consistently using physics. This is a troubling result for the TRD cohort, but may show that teaching an explicit strategy had a positive effect on the EPS cohort.

There is more to the Logical Progression skill presenatation than just these two tables. The TRD cohort development graph is shown in Figure 5.14. This curve looks vaguely similar to the EPS curve.

The TRD development curve in Figure 5.14 begins very low reflecting the high percent of low scores seen in Table 5.8. After this point, the students clearly learned the game, much quicker than predicted. They grew quickly to LP = 6 band and stayed there showing the same slight decrease on quizzes as seen in the EPS cohort. Both cohorts apparently felt the time pressure on quizzes. It is clear that the two cohorts became nearly identical on the Logical Progression measure as the year progressed. Again this shows that the explicit problem-solving strategy seems to raise the level of an average sample of students to a better-than-average sample of students.

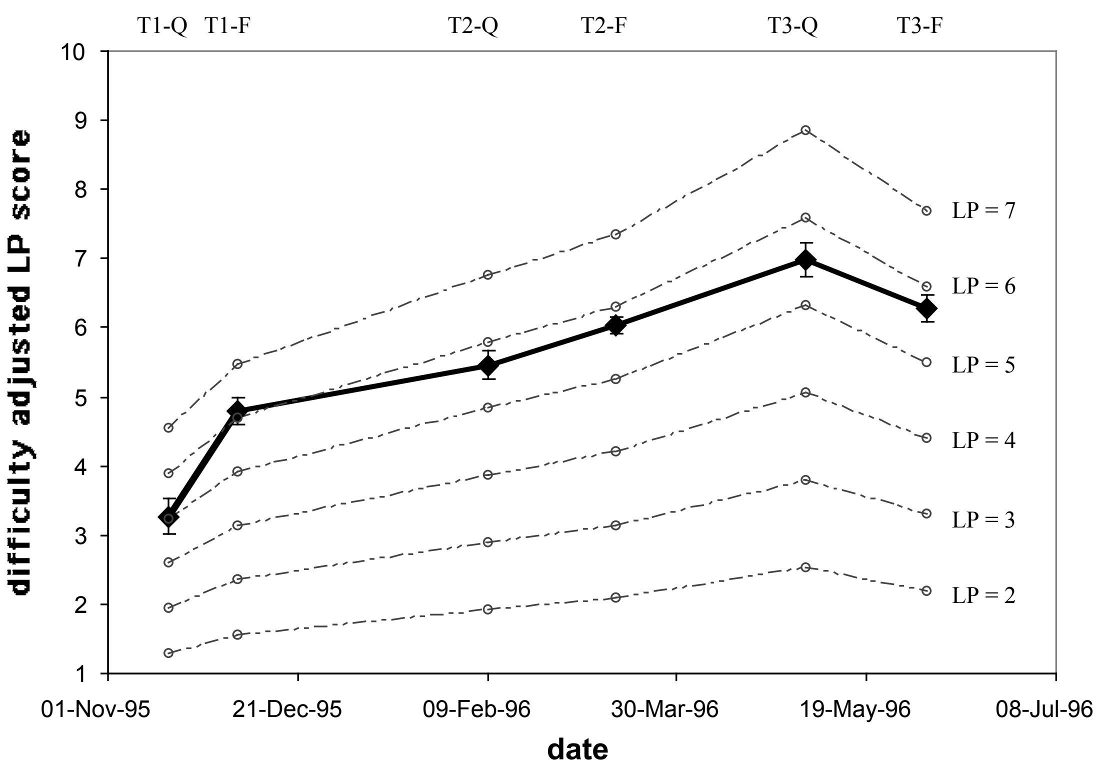


Figure 5.14

TRD Cohort LP Development Graph. N = 24.

TRD Clusters – Logical Progression

Figure 5.15 show the principle clusters of the TRD cohort. Most of the students are shown in this graph. TRD11 represented the nine middle students and TRD9 characterized the ten high students. A new grouping emerged for the TRD cohort. Three originally high students slipped on this skill at the start of the third term. They are labeled "hhm." With the exception of this last group, the other groupings showed improvement. In addition, these curves, except for the low starting point, do not differ from the EPS cohort. However, it is clear that the TRD cohort has more students doing better than the EPS cohort.

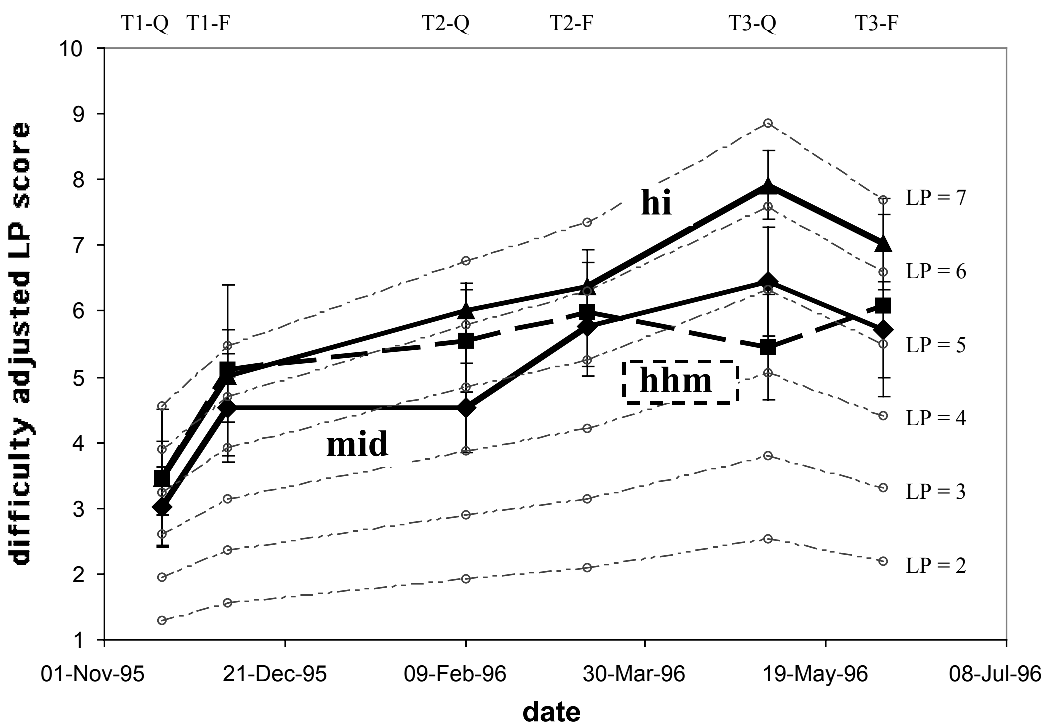


Figure 5.15

TRD Development Graph by Cluster. N = (hi=10; mid=9; hhm=3)

A hypothesis for why the TRD cohort had better Logical Progression (at least in number of students scoring in the high cluster) may have had to do with the consequences of the grading used in the two courses. In the TRD class, the students had to get to the answer to receive most of the credit. Recall from Chapter Four that there was a large exodus of "C-grade" students from the TRD course. This implies that students who could not reach an answer would not be in the cohort. Perhaps the students who made it into the TRD cohort had to be smart enough to realize that they needed to learn to solve problems. If this was the case, the TRD cohort students possess impressive metacognitive skills which would further elevate all of their problem-solving skill scores.

In the EPS cohort, the students received most of the points for a problem by setting it up. Even though this ratio changed during the quarter, never did the points favor the final answer over the physics approach. The students knew they did not need to get to an answer to receive the lion share of the partial credit. The consequence of this grading scheme meant that being able to solve the problems at any cost was not critical to success in the EPS course. This may have given more diversity in the cluster graphs. It must be re-emphasized that the EPS cohort represented a fair cross-section of the EPS class while the TRD cohort was a better-than-average sample of students.

Logical Progression Summary

The third problem-solving skill assessed was the students Logical Progression. This was a measure of the students' planfullness when solving problems. The codes measured the cohesiveness of the solution. It also measured whether a student worked forward or backwards and if the students left solutions incomplete. Teaching the students to be planful is one of the goals of any explicit problem-solving instruction. Therefore it was expected that the EPS cohort would excel in this measure.

There were preliminary indications that the EPS cohort was doing well. The number of students who floundered before ending their solution went to zero in the EPS cohort. There were also proportionally fewer students in the EPS cohort who violated previous physics claims than in the TRD cohort. The EPS students apparently preferred to just stop than violate physics. And with the exception of two problems, the EPS cohort left as many solutions unattempted as the TRD cohort. The two outlier problems in the EPS cohort suggest an instructional effect since on one of these problems (T3-F1) an incorrect approach was actually graded as correct.

However, when the development curves (both averages and clusters) were plotted, there was no difference between the two cohorts, except that the TRD cohort had more students in the highest cluster. It was hypothesized that a survival effect was the reason becuase there were essentially no C-students who stayed with the TRD instructor after the first term. Therefore the TRD cohort students had to learn how to solve problem on their own or they would have dropped out of the course. Also since the EPS and TRD averages looked the same while the clusters had different numbers of students in them, this suggests that several of the unclustered students in the EPS cohort must have had high scores on some of the problems. Finally there is the realization that even if the EPS and TRD cohorts appear similar, the fact that the TRD is an over-representation of the class suggests that the explicit problem-solving strategy may be raising the planfulness of the EPS students to the level of the better portion of the students in the other class.

Appropriate Mathematics

The last problem-solving skill measured was Appropriate Mathematics. This was an assessment of the students' ability to correctly apply what they know about mathematics in the context of physics. The codes used to measure Appropriate Mathematics are in Table 5.9. These codes include looking for students who made convenient math errors, such as dropping a minus sign under a radical operator, allowing them to continue the solution. This trait is dubbed "math magic." This skill also checked to see when students put numbers into their solutions. Most physics instructors would prefer students to wait until the last practical step before substituting values in for variables. This measurement is clearly dependant on the students actually producing a solution. Therefore the lowest codes account for these incomplete solutions and will not be used in the development graphs.

Expectations – Appropriate Mathematics

The explicit problem-solving strategy doesn't teach applying mathematics. Its only explicit impact might be that it encourages the students to hold-off on inserting numbers until the last step. There has been evidence from earlier studies (i.e., Huffman 1994) that students who were taught an explicit problem-solving strategy have better math performance. This effect might raise the EPS scores. This suggest a prediction for the EPS cohort average would be around the AM = 6 band. This accounts for the students holding off inserting numbers while still making mistakes. Also there should be a drop in the EPS curve in the third term to account for the increased reliance on integral calculus.

Table 5.9

Appropriate Mathematics

0	Nothing written
1	Solution terminates for no apparent reason
2	When an obstacle happens, “math magic” or other unjustified relationships occurs
3	When an obstacle happens, solution stops.
4	Solution violates rules of algebra, arithmetic, or calculus
5	Serious math errors
6	Mathematics is correct, but numbers substituted at each step
7	Mathematics is correct, but numbers substituted at last step.

A representative TRD cohort would be have an average curve lower than the EPS curve and may take longer to reach its maximum height. The TRD students may also not reach the highest score because they were not discouraged from entering numbers at every line of the solution. But since the TRD cohort was better than average, its curve may grow as quickly as does the EPS cohort's curve.

EPS Cohort – Appropriate Mathematics

As with Logical Progression, the lowest two codes were not included in the development curves. The code of AM = 0 was redundant with the LP = 0 code and is not reported. However the code AM = 1 is informative. This code is used when a solution unexpectently stops. This would also occur if the student had run out of time. If the student encounters an obstacle, these are separate and interesting codes. If AM = 2, then math magic occurred to pass through the obstacle and the solution continued. If AM = 3, then the solution stopped at the obstacle. The frequency of all three of these codes are listed in Table 5.10

Table 5.10

Percent of students with the three lowest AM codes in the EPS cohort by time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Percent of AM = 1 solutions	8.3	5.6	16.7	11.1	22.2	10.1
Percent of AM = 2 solutions	4.2	7.0	0	5.6	7.0	2.4
Percent of AM = 3 solutions	8.3	2.8	2.8	1.4	4.2	1.8

The percent of students getting AM = 1 is comparable to the data seen in Table 5.6. This implied that the students could not finish their solutions. There is also little evidence of students encountering obstacles (codes AM =2 and AM = 3) and there is no

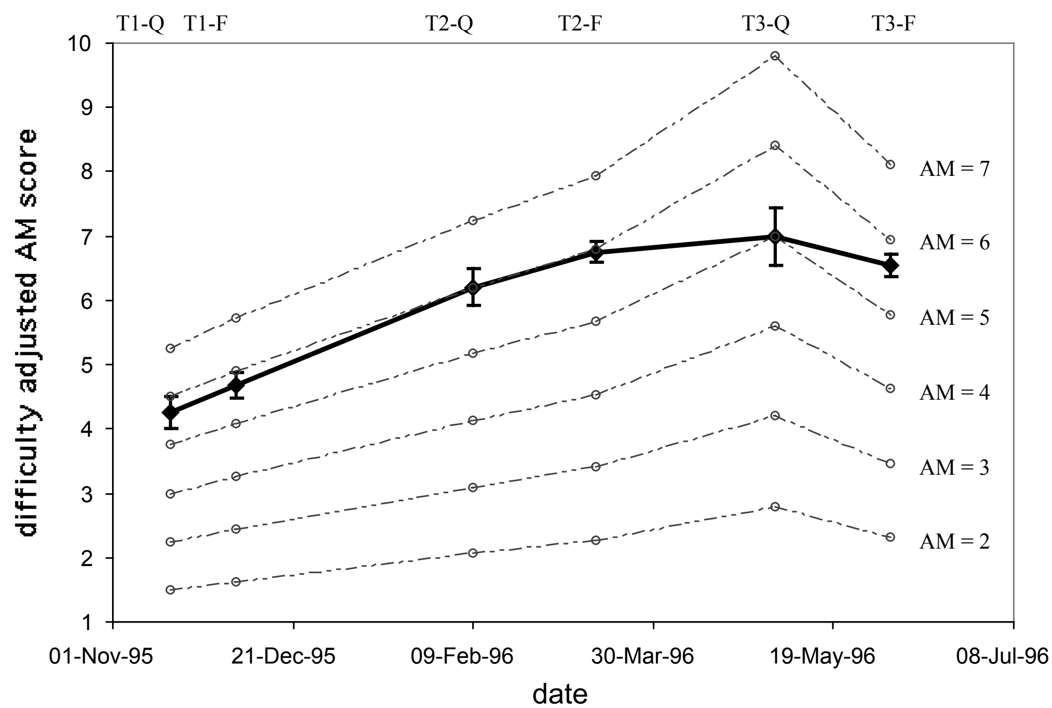


Figure 5.16

EPS Cohort AM Development Graph. N = 24.

trend to their occurrence. They seem to happen with about the same low frequency and among many of the students.

Figure 5.16 shows the development curve of the EPS cohort on the Appropriate Mathematics skill. The curve clearly shows growth of this skill over time. It also has the expected decline during the third term where integral calculus is frequently used. The level is also as predicted. The EPS cohort begins on the low side of the AM = 6 band and stayed there through the second term. The third term drop is very evident on the third term quizzes, but the cohort nearly recovered by the final exam. The EPS cohort was clearly comfortable with applying mathematics in the context of physics.

EPS Clusters – Appropriate Mathematics

There were many clusters found for the EPS cohort Appropriate Mathematics skill measurement. As has been typical, only clusters with three or more students are shown. However, a full breakdown is available in Appendix C. As happened with the Logical Progression clusters, the low cluster is not shown because it has too few scores once the codes of AM = 0 and AM = 1 are removed. There were five students in this low cluster. Fortunately not all clusters had this problem. Those that did not are shown on Figure 5.17.

The two clusters shown are rather distinct but both show growth. The three "hi" cluster students had EPS9 used to select these students. They clearly are at the ceiling until the third term. There they fall and remain for the term. The use of calculus had a pronounced impact on them. The wide error bars in the third term suggests that the students are making the occasional math error (AM = 4 or AM = 5 codes). The 12 "mid" cluster students also show growth until the third term, only hovering around the AM = 6 band. The "mid" cluster was selected using EPS23. The "mid" cluster was different from the "hi" cluster since it nearly rebounds back to the AM = 6 band on the third term final exam. Both clusters confirm what was seen with the average curve; in general the students did not have problems applying mathematics to the problems when they got that far in the solution. The only exception was the third term when calculus became central to the solutions.

TRD Cohort – Appropriate Mathematics

Meeting theoretical expectations is one way of quantifying development, but in this thesis a cohort of students who were not taught an explicit problem-solving strategy

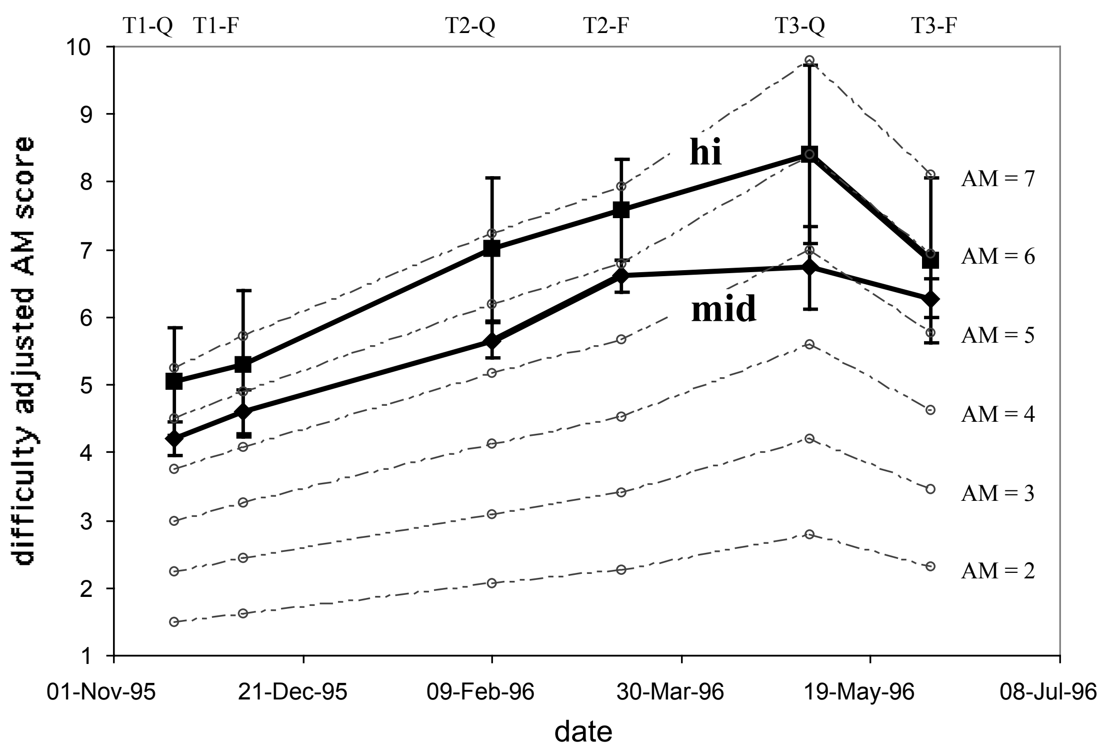


Figure 5.17

EPS Development Graph by Cluster for AM. N = (hi = 3; mid = 12)

Table 5.11

Percent of students with the three lowest AM codes in the TRD cohort by time

	T1-Q	T1-F	T2-Q	T2-F	T3-Q	T3-F
Percent of AM = 1 solutions	2.1	6.2	7.0	11.1	18.1	4.7
Percent of AM = 2 solutions	16.7	7.6	2.8	4.9	4.2	5.4
Percent of AM = 3 solutions	8.3	0.7	1.4	0	1.4	0

(but who still benefited from cooperative group problem solving) was used to help understand the results. The frequencies of low scores seen in Table 5.10 for the EPS cohort were certainly an example of data needing a context. The TRD cohort provides this context. Table 5.11 shows the number of students in the TRD cohort who got the lowest codes. Again AM = 0 is excluded since it is the same as LP =0 reported earlier.

Table 5.11 definitely helped interpret Table 5.10. The EPS quizzes are seen to be a bit higher than the TRD quizzes for AM =1. This higher rate of unfinished solutions again reinforces the extra time pressure felt on the quizzes for the EPS cohort since they wrote longer solutions (an average of half a page more). There was also an interesting comparison to be made for the AM = 3 codes. Unlike for the EPS cohort, which stayed around three percent for most of the course, the TRD cohort clearly learned that stopping before making a mistake would not be rewarded. After the first set of quizzes, this code only occurs three times for the TRD cohort. In a similar vein, the TRD cohort may have done slightly more "math magic" on average than the EPS cohort. This again supports the notion that the TRD cohort was rewarded for getting to an answer.

Figure 5.18 shows the TRD cohort development graph. The TRD cohort began in the AM = 5 band, grew into the AM = 6 band by the first term final. The TRD cohort fell into the AM = 5 band for the third term quizzes and started a climb back into the AM =6 band on the final.

This description of the TRD cohort matches the description of the EPS cohort. Both cohorts grew through the year and experienced a drop at the third term when calculus becomes important. The TRD cohort again started around a lower skill band and climb quickly into a higher band. This may be typically of better-than-average students. In general both cohorts can apply mathematics in the context of physics.

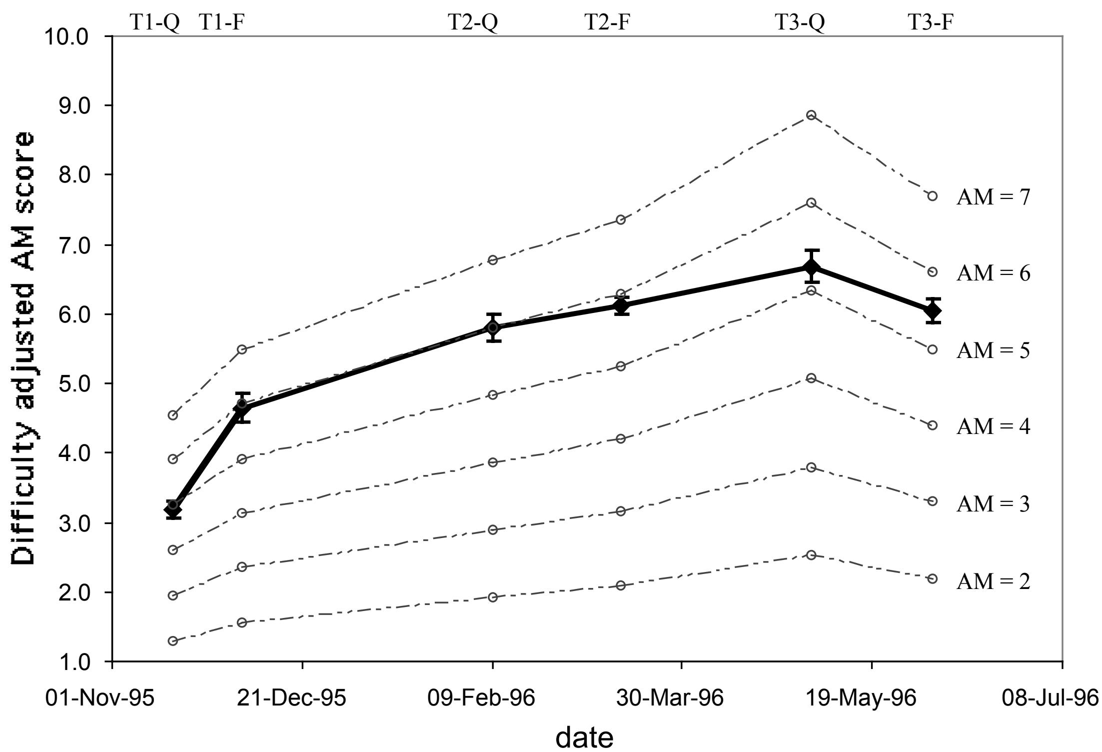


Figure 5.18

TRD Cohort AM Development Graph. N = 24.

TRD Clusters – Appropriate Mathematics

There was very little diversity in the TRD cohort's Appropriate Mathematics curves. Essentially all the students fell into two clusters and even those two clusters are barely distinguishable from each other. Only their starting points and third term bands are different. The clusters with three or more students are shown in Figure 5.19.

The "mid" cluster curve represents nine students who were selected by comparing them with TRD23. The 12 "hi" cluster students were similar to TRD12. As was seen in the earlier Logical Progression discussion, this is most of the cohort and may be a result of adapting to the TRD grading emphasis. Also, the similarities in population sizes

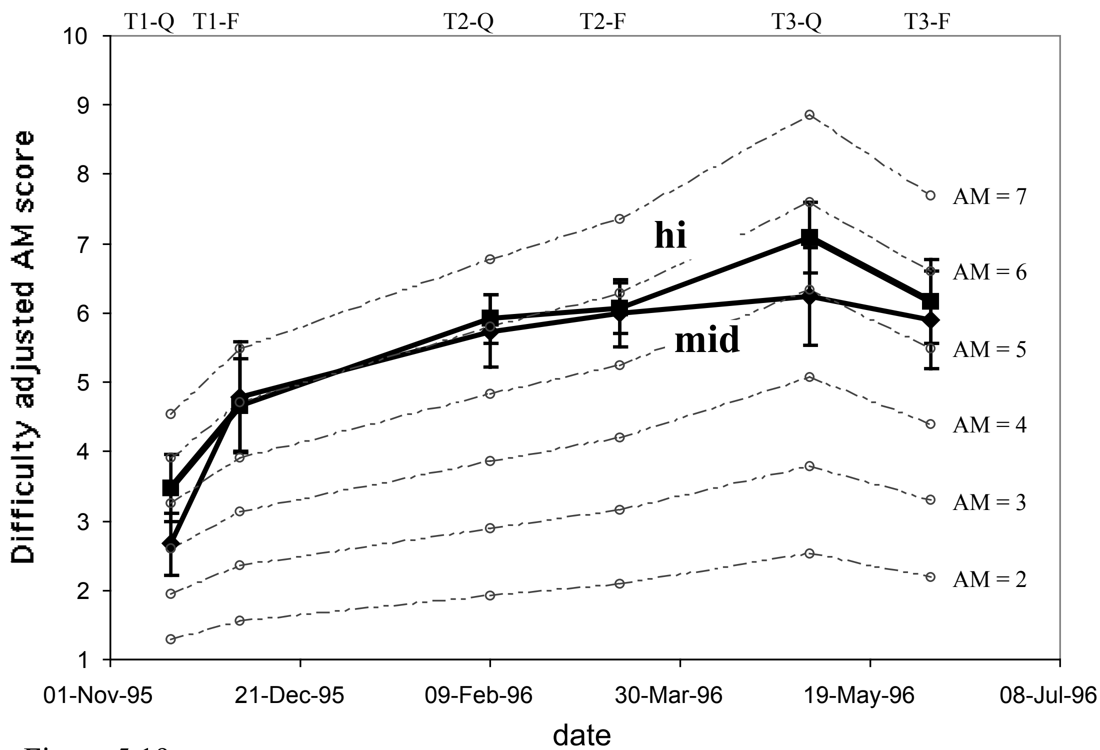


Figure 5.19

TRD Development Graph by Cluster for AM. N = (hi = 12; mid = 9) between the Logical Progression and Appropriate Mathematics clusters reinforces the correlation between these two skills reported in Chapter 4.

The are two differences between the TRD and the EPS cohort on their Appropriate Mathematics cluster curves. The first difference is the proportion of the cohort in each cluster. There are far more TRD students in the "hi" cluster. However, since the TRD "hi" cluster is in the same skill band as the EPS "mid" cluster, this is really a difference in the skill bands occupied by each cohort. Not surprisingly, only the EPS students occupied the AM = 7 band. There was nothing in the TRD course that encouraged students to wait until the last step before entering numbers. The explicit problem-solving strategy helped some of the EPS students wait until it was useful to enter values. It was also evident that the problem-solving strategy raised the level the EPS cohort to that of the TRD cohort.

<u>Appropriate Mathematics Summary</u>

The Appropriate Mathematics skill was the last of the problem-solving skills measured for this study. It is a measure of how well the students can apply their mathematics knowledge in the context of physics. The codes used for this skill cover everything from blank solution to "math magic" to working the solutions algebraically.

The frequency of some of the lower codes showed that the TRD students seemed more likely to use "math magic" than to stop when faced with an obstacle. The EPS cohort clearly had more blank and incomplete solutions. This was probably due to the longer solutions the EPS students tended to produce. Writing clear and logical solutions takes time.

The most salient conclusion drawn from the cohort development graphs is that the students can apply their math skills to physics. All the curves showed stable development through the course until the third term when calculus became important. And even in the third term the cohorts essentially rebounded by the final exam, which had even more calculus than the quizzes.

Another important conclusion drawn from the curves is that the explicit problem solving strategy helped students apply their math skills. Recall that the grade for the EPS cohort was negatively related to math background (Chapter Four). This fact, coupled with the similarity in the cohort development graphs, suggests that teaching the problem-solving strategy provided the necessary structure for many students: enough students so that the average EPS cohort appeared similar to the better-than average TRD cohort.

Summary of Results

The most obvious conclusion drawn from the inspection of all the problem-solving ability graphs for both cohorts was, in general, that the problem-solving skills of the students in both cohorts improved during the physics course. This was an important conclusion since one of the principle goals assigned to this course by the engineering faculty who require it was problem-solving (Foster, Heller, & Heller, in press). However, the development was generally stable, meaning that neither cohort generally moved between the skill bands on the development graphs. A more detailed discussion is had by taking each problem-solving skill in turn.

General Approach

The first skill was General Approach which was a measure of the adequacy of the student's representation of the problem statement. At the cohort level, both cohorts did about the same and stayed around the GA = 4 skill band. This suggested that the students in both cohorts generally understood the physics, but it was not a solid understanding. This conclusion is reinforced by the student's performance on the multiple-choice tests (Chapter Four). Both cohorts could answer these questions, although the EPS cohort did consistently better on the multiple-choice questions. Students made errors in their approach, but they generally knew the concepts. However, since the TRD cohort was composed of better than average students, while the EPS cohort was more representative, it appeared that the EPS instruction was effective in raising the level of the EPS students.

There were also an important detail seen in the General Approach development graphs. Both cohorts began to show growth out of the GA = 4 skill band on the third term quizzes, yet both fell back into the skill band on the final exam, with the EPS cohort being the most dramatic. Notice that this final exam effect was not observed on the earlier two final exams. Also the MPEX did not show a dramatic change in student attitude at this time. Rather there is something unique about the third term. It is hypothesized that the third term did not present the students with a useful knowledge structure. In the first two terms, the theme of mechanics provided a useful structure. In the third term the topics were presented as a cornucopia of concepts and perhaps the students could not inter-relate the topics very effectively.

The development graphs of the student clusters also provided insights. In the TRD cohort, the principle student-clusters began the term at the same skill band. The clusters did not fully stratify until the second term final exam. Once they did stratify, the clusters generally stayed within a skill band. The highest cluster did not encounter ceiling effects until the second term final exam. The benefits of the explicit problem-solving strategy become evident by contrasting the EPS cohort clusters to the TRD cohort clusters. The EPS clusters began this study stratified. This suggests that the strategy helped students start the course more successfully and reach their skill band sooner. This result was also interesting in light of the TRD cohort's better-than-average status.

Specific Application of Physics

The next problem solving skill examined was the Specific Application of the Physics. Specific Application of the Physics was a measure of how well the students did what they thought they needed to do. This skill (as well as the next three) all accounted for students who wrote nothing for a solution. Since this code could mask the development of students who did development, it was removed. This removal of blank solutions impacted both cohorts, although the EPS cohort produced more blank solutions. This was not surprising given the above average composition of the TRD cohort.

Both cohort development graphs differed significantly from the General Approach graphs. Both cohorts started in the SAP = 6 skill band but fell into the SAP = 5 skill band at about the same time during the year. They students were showing more direction and variable definition errors. The TRD cohort showed a non-significant gain on each final exam compared to the preceding quiz. Clearly as topics were presented that the students did not encounter in high school were introduced, they had increasing more trouble with applying these concepts. In the EPS cohort, the average Specific

Application of Physics score was related the General Approach score suggesting that as their knowledge wavered, so did their approach. This relationship was not seen in the TRD cohort.

The student skill clusters showed more detail, but very few new conclusions. The highest cluster in the EPS began the year with a definite ceiling effect (which none of the TRD clusters had), but dropped firmly into the next lowest skill band by the end. The lowest clusters in both cohorts showed very little (if not actually negative) improvement during the year. While the overall results are promising for the EPS cohort since they were no different than the better-than-average TRD cohort, their performance was still disheartening. The explicit problem-solving strategy should have helped the students with this skill since many of them drew diagrams and discussed variable definitions, yet they still made mistakes. For these errors to have occurred in spite of the work they have done suggests that the students had not internalized the explicit problem-solving strategy either by using the page as their extended memory or by really utilizing what they wrote.

Logical Progression

The third skill was Logical Progression. This was a measure of how planfull the students were in their solutions. This skill is at the heart of problem solving since it measures if, and how well, the students can construct a reasoned argument. Since the issue is about constructing a solution, only solutions that proceeded past a basic statement of concepts will be analyzed. There were a few more solutions excluded for the EPS cohort than the TRD cohort. This was probably due to the difference between the two cohorts in terms of student grades.

There were also codes for students who stopped their solutions in mid-stride. Again there were more of these students in the EPS cohort. This wasn't surprising given the difference in the cohorts and since the EPS students wrote longer solutions than the TRD students did. What was encouraging was that the EPS students encountered fewer obstacles (LP =2 and 3) while a few students kept this habit in the TRD cohort. Apparently the EPS students preferred to stop than write logically incorrect solutions.

There were some interesting differences seen in the Logical Progression development graphs. The curves for both cohort were similar in values, with the both cohorts in the lower LP = 6 band and with the EPS cohort in the higher LP =5 band for the quizzes. The students' pension to stop in mid-solution may have affected the EPS cohort development graphs, especially on the quizzes when there was less time to write long solutions. Also, the TRD cohort began in the LP = 5 band and quickly grew into the LP = 6 band. The Minnesota strategy may have helped the EPS cohort learn this skill sooner. The clusters provided little new evidence except that the TRD cohort had more students in the high cluster. These high scores imply that both cohorts could generally construct logical arguments and the growth came early in the class. However, it was hypothesized that the TRD students maybe metacognitively superior to the ES cohort because the TRD cohort had to invent good problem-solving on their own.

On the surface of the cohort results both cohorts looked very good. However, there was an interesting result. In a cohesive solution one might expect the four skills to be correlated. A cohort of experts, the research literature implies, would do all four skills together. They could not score well on one skill and not the others for each problem.

This was seen in the EPS cohort but not in the TRD cohort. For the TRD cohort only Logical Progression and Appropriate Mathematics were correlated. This implies that in spite of the lack of individual variation in the Appropriate Mathematics graphs, the TRD cohort solutions were only planful insofar as they were mathematically correct. This lack of cohesiveness was an interesting part of the overall description of the TRD cohort. They improved in general, but were not consistent within solutions. The Minnesota Problem-solving Strategy presumably helped the EPS students write more cohesive solutions than they would have without the strategy.

Appropriate Mathematics

The fourth skill was Appropriate Mathematics. This was a measure of the students' mathematical skill while doing physics. Both cohorts generally had no difficulty (AM = 6), except on the third-term quizzes, in using their math skills in physics. The TRD cohort started one band lower, but quickly climbed and mirrored the EPS cohort. This suggests that the EPS instruction helped the student's apply their math sooner in the course than they might have otherwise. Another benefit of the problem-solving strategy was to reduce the incidence of "math magic" in the EPS cohort.

The cluster curves were also similar between the two cohorts, except that the high cluster was hitting the ceiling, something not seen in the TRD cohort. Hitting the ceiling was possible only if students generally waited until the end of the solution to substitute numbers. Conversely several of the TRD cohort students frequently did their math on their hand-held calculators, which made those solutions difficult to code. The similarity between the two cohorts is encouraging since the representative EPS cohort appeared similar to the better-than-average TRD cohort. The explicit problem solving strategy made average students better.