The Use of Tracking for Student Placement in Calculus

MAJ Randal Hickman
Department of Math

MTP 2007 Graduate

Introduction:

The purpose of this research is to offer a fresh perspective on the tracking debate (whether or not students benefit from sectioning them by ability.) Two specific aspects of my research separate it from the majority of other papers on this subject. First, my research focuses on a university setting, considering the relative merits of tracking in a required calculus course at the United States Military Academy. Second, I base my conclusions on a broad analysis of student surveys and faculty surveys, in addition to a statistical analysis of test scores with and without tracking. I conclude that tracking was enormously successful in the calculus curriculum at West Point, as evidenced by a statistically significant improvement in test scores for students on all ability levels and overwhelmingly supportive feedback from students and faculty alike.

Background:

Few topics in math education have been as thoroughly studied as the controversial topic of tracking. A quick search on-line will yield a plethora of papers passionately arguing either for or against this sensitive issue. The majority of these papers focus on educational theory or academic principles regarding tracking in junior high or high school programs.

Many of the papers that consider tracking argue from the perspective of theoretical pedagogy or faculty experience. My research considers the input of faculty members, but it also includes two of the most important sources of information that many studies ignore: input from the students and the statistical analysis of test scores with and without tracking. While the theoretical debate may be stimulating, a more careful analysis of consumer (student) opinion and test results offers stronger evidence of whether or not tracking benefits the student population.

I conducted my research in the Math Department at the United States Military Academy. All cadets must complete (or validate) a challenging math curriculum that includes math modeling, differential and integral calculus (in single and multiple variables), probability, and statistics. My research addressed the relative benefit of a tracking policy within the required sophomore course in integral calculus. This class was a prime candidate for ability-group sectioning because of the broad spectrum of calculus experience observed in our student body. Many of the cadets arrive at West Point with a strong calculus background (pre-calculus courses or AP Calculus experience from High School). However, many of our students also start lesson one of this course having never seen an integral symbol in their previous math classes. Furthermore, the large enrollment (approximately 800 students) relative to the small class size (18 students per section), allows for a well-defined, homogeneous level of academic ability in each section if

tracking is employed. Finally, historical analysis of cadet grades has shown that this particular course has proven to be the most difficult of the core courses in the math department. By tracking sections into ability groups, our intent was to focus the educational experience of each section towards the needs of the students in that particular section in an attempt to improve the overall course experience for all cadets.

Although the calculus course has been in existence for many years, the course underwent content changes in 2005. During the first year of the revised calculus course, a small test group of 188 students studied the revised curriculum. This group of 188 cadets was randomly selected from the top two-thirds of students' scores on the math placement exam. During the second year, (2006), the new curriculum was used for all cadets, and cadets were assigned sections purely randomly. During the third year (2007), the new curriculum was again used for all cadets, but cadets were assigned their sections based on their cumulative academic GPA (their class rank) at the start of the semester.

Statistical Results:
My initial goal in this research was to determine whether or not there were any statistically significant differences in academic performance between the three years. Since the course-end grades include group projects and homework assignments that change from year to year and are graded by a variety of different faculty members, I was concerned that it would be extremely difficult to isolate the effect of tracking versus a variety of confounding factors if we consider final grades. In response, I only considered the students' performances on the final exam. This exam was virtually unchanged during the three year period, and it is always graded by a faculty team to ensure equity of grades across the course. It is also a very challenging, comprehensive exam that serves as a good measurement of overall student understanding of course material. Table one shows the descriptive statistics for the three years of final exams.

	Without Tracking (top 2/3 only)	Without Tracking	With Tracking
Year	2005	2006	2007
Mean	81.22%	68.88%	78.27%
Median	83.08%	68.94%	79.00%
Standard Deviation	10.85%	12.46%	11.58%
Minimum	9.16%	12.24%	40.50%
Maximum	98.16%	97.88%	99.67%
Range	89.00%	85.64%	59.17%
Number of Students	188	836	847

Table 1: Descriptive Statistics of Final Exam Performance

Table 2 shows the percentage of students who earned each of the possible letter grades on the final exam during the same three years.

	Without Tracking (top 2/3 only)	Without Tracking	With Tracking
Year	2005	2006	2007
Percentage A Grades	16.49%	4.19%	17.59%
Percentage B Grades	49.47%	16.15%	29.40%
Percentage C Grades	19.68%	25.72%	27.63%
Percentage D Grades	9.04%	15.43%	11.57%
Percentage F Grades	5.32%	38.52%	13.81%

Table 2: Final Exam Performance by Letter Grade

We will first consider only the years 2006 and 2007 when the entire cadet population studied the revised curriculum. Both the mean and the median for 2007 (with tracking) are approximately ten percentage points larger than in 2006 (without tracking). While the percentage of C grades remains relatively constant in both years, table two shows us that the percentage of A and B grades is much higher (and the percentage of D and F grades is much lower) when students were sectioned by ability. This seems to contradict the most common criticism of tracking, the notion that tracking benefits the strong students at the expense of the weak students. Rather, my data shows that the number of D or F grades on the final exam decreased by more than twenty-eight percent when students were sectioned by ability. This remarkable result occurred when everything else in the course (other than the use of tracking) was held constant. The same curriculum, the same final exam, and the same grading processes were used both years. Furthermore, the student body was made up of similar, high-quality cadets both years. In fact, the entrance exam scores on the math sections of both the SAT and the ACT were essentially identical for both year-groups. The only difference was the use of tracking.

An even more remarkable result becomes apparent when we compare the final exam results from 2007 (with tracking) to the final exam results of the test group in 2005. Recall that the test group was a random selection of 188 cadets who scored in the top two-thirds on West Point's math placement exam, while the 2007 group included the full spectrum of cadets. However, the performance of these two groups is statistically very similar. In fairness, almost all of the key statistics (mean, median, and percentage of each letter grade) are better for the test group in 2005 than for the full student population with tracking (in 2007). However, the students in 2007 (with tracking) were statistically much more similar to test group in 2005 (students from the top two-third of their class) than they were to the full student body in 2006. This suggests that one of the statistical effects of tracking is to "pull-up" the bottom one-third of a student body such that the overall performance of the entire collection of students is statistically similar to the "upper two-thirds" of a student group that does not experience tracking.

I used Minitab to conduct an ANOVA on the final exam scores for the three year-groups to reinforce the statistical significance of these results.

One-way ANOVA: 2007, 2006, 2005

```
Source    DF       SS      MS       F      P
Factor     2   4.6899  2.3449  165.17  0.000
Error   1868  26.5204  0.0142
Total   1870  31.2103

S = 0.1192   R-Sq = 15.03%   R-Sq(adj) = 14.94%


                           Individual 99% CIs For Mean Based on
                           Pooled StDev
Level    N    Mean   StDev    +---------+---------+---------+---------
2007   847  0.7827  0.1158                          (--*-)
2006   836  0.6888  0.1246    (-*--)
2005   188  0.8122  0.1085                               (-----*-----)
                             +---------+---------+---------+---------
                           0.680     0.720     0.760     0.800
```

This output reinforces the dramatic statistical difference between the three years. The low p-value (0.000) shows that there is essentially 0 % chance that such a difference in performance can be attributed simply to random processes.

A separate comparison of 2007 scores to 2006 scores is below. This removes the effect of the test group in 2005, and isolates the comparison between 2006 test scores (without tracking) to 2007 test scores (with tracking).

One-way ANOVA: 2007, 2006

```
Source    DF       SS      MS       F      P
Factor     1   3.7090  3.7090  256.37  0.000
Error   1681  24.3196  0.0145
Total   1682  28.0285

S = 0.1203   R-Sq = 13.23%   R-Sq(adj) = 13.18%


                           Individual 99% CIs For Mean Based on
                           Pooled StDev
Level    N    Mean   StDev  ----+---------+---------+---------+-----
2007   847  0.7827  0.1158                                (---*--)
2006   836  0.6888  0.1246  (---*--)
                           ----+---------+---------+---------+-----
                             0.690     0.720     0.750     0.780
```

As expected, we see a powerful statistically-significant difference between the two years. Since all other factors were held constant, we must attribute this change to the use of tracking in 2007.

A separate comparison of 2005 (test group from top 2/3^{rds} of students) to 2007 (full student body with tracking) is below. This removes the effect of the 2006 data, and isolates the comparison between 2005 and 2007.

One-way ANOVA: 2007, 2005

```
Source    DF       SS      MS      F      P
Factor     1   0.1341  0.1341  10.22  0.001
Error   1033  13.5469  0.0131
Total   1034  13.6810

S = 0.1145   R-Sq = 0.98%   R-Sq(adj) = 0.88%


                           Individual 99% CIs For Mean Based on
                           Pooled StDev
Level    N    Mean   StDev  -------+---------+---------+---------+--
2007   847  0.7827  0.1158  (-----*------)
2005   188  0.8122  0.1085            (-------------*------------)
                           -------+---------+---------+---------+--
                                0.784     0.800     0.816     0.832
```

The p-value is still very low (0.001) indicating a strong statistical difference in the data sets. However, we see an overlap in the 99% confidence intervals suggesting that the data sets are much more similar than the previous comparison between the 2006 data and the 2007 data. While the test group in 2005 outperformed the full student body in 2007, the difference in performance is much smaller than the difference between 2006 and 2007. This suggests that the full student body that experienced tracking (in 2007) had final exam scores that more closely resembled the test group from the top 2/3^{rds} of the class in 2005 than the full student body in 2006.

Up to this point, my intent was to compare the overall performance of each of the three groups. However, since the 2005 group was made up of students from the top 2/3^{rds} of their peer group, this has been a somewhat unfair comparison. In the next section, I compare only the top 2/3^{rds} of the students from 2006 and 2007 with their peer group – the test group of students in 2005 (all of whom represented the top 2/3^{rds} of the student population in 2005.)

One-way ANOVA – Top 2/3^{rds} of Year-Groups: 2007, 2006, 2005

```
Source    DF        SS       MS       F      P
Factor     2   2.37202  1.18601  178.25  0.000
Error   1307   8.69624  0.00665
Total   1309  11.06826

S = 0.08157   R-Sq = 21.43%   R-Sq(adj) = 21.31%


                           Individual 99% CIs For Mean Based on
                           Pooled StDev
Level    N     Mean    StDev    +---------+---------+---------+---------
2007   565  0.84931  0.06819                                  (--*--)
2006   557  0.75757  0.08346    (--*-)
2005   188  0.81218  0.10849                    (----*----)
                             +---------+---------+---------+---------
                           0.750     0.780     0.810     0.840
```

As before, we see a very low p-value (0.000) indicating a statistically significant difference between groups. However, we now see that the students with tracking (in

2007) significantly outperformed **both** of the other groups. I attribute this difference to the levels of tracking experienced in each of the three groups. The top 2/3^{rds} of the students in 2006 did not receive any benefits of tracking. Not surprisingly, their scores tended to be the lowest of the three groups. The students in the experimental group in 2005 received some benefits of tracking. All of their peers came from the top 2/3^{rds} of their year-group. However, there was no additional refinement of the tracking within that group. In 2007, cadets were sectioned according to their class rank. All cadets in every class section (18 students per section) were within a very close range of academic abilities with each other. Given this tracking strategy, final exams scores of the top 2/3^{rds} of the students were significantly higher than the top 2/3^{rds} of either of the other two groups.

If we consider only the top 1/3rd of each population, we see a similar result. The analysis below compares the top 1/3rd of 2006 and 2007 with the top half of the hand-selected group from 2005. (i.e. It compares the top half of the "top 2/3^{rds} data for each of the three years.) The Minitab output is below.

One-way ANOVA — Top 1/3^{rds} of Year Groups: 2007, 2006, 2005

```
Source   DF       SS       MS       F      P
Factor    2  0.93537  0.46768  205.38  0.000
Error   653  1.48695  0.00228
Total   655  2.42232

S = 0.04772   R-Sq = 38.61%   R-Sq(adj) = 38.43%


                              Individual 99% CIs For Mean Based on
                              Pooled StDev
Level    N     Mean    StDev  --+---------+---------+---------+-------
2007   283  0.90694  0.04067                                    (--*--)
2006   279  0.82692  0.05661  (--*--)
2005    94  0.88830  0.03734                       (----*----)
                              --+---------+---------+---------+-------
                              0.825     0.850     0.875     0.900
```

As before, we see that the most aggressive tracking policy (in 2007) produced the highest test scores for the students in the top 1/3rd of their peer group. The moderate tracking policy in 2005 produced the next highest test scores. The top 1/3rd of the randomly sectioned students (in 2006) had the lowest scores.

The comparisons above seem to offer conclusive evidence that aggressive tracking helps students in the top 2/3^{rds} and top 1/3rd of their peer groups. However, this is not surprising. The overwhelming bulk of the tracking literature supports this notion. The greater concern is typically with whether or not the bottom 1/3rd of students suffer at the expense of the top 1/3rd of students.

The analysis below compares the bottom 1/3rd of the final exam test scores for 2006 (without tracking) to the bottom 1/3rd of final exam test scores for 2007 (with tracking.)

One-way ANOVA: 2007, 2006

```
Source   DF       SS       MS       F      P
Factor    1  1.33900  1.33900  313.89  0.000
Error   559  2.38456  0.00427
Total   560  3.72355

S = 0.06531   R-Sq = 35.96%   R-Sq(adj) = 35.85%


                              Individual 99% CIs For Mean Based on Pooled
                              StDev
Level    N     Mean    StDev    +---------+---------+---------+---------
2007   282  0.64912  0.06493                                    (--*---)
2006   279  0.55141  0.06570    (---*--)
                                +---------+---------+---------+---------
                              0.540     0.570     0.600     0.630
```

This Minitab output demonstrates one of the most important outcomes of this research. Cadets in the bottom 1/3rd of the student population had a dramatic improvement in their test scores when they learned calculus in a tracking-structured environment. More precisely, the average test score of the bottom 1/3rd of students increased by almost 10% when students learned calculus in a tracking-structured setting. This is approximately the same increase in the mean test scores that we had previously observed in the entire population when exposed to tracking.

The statistical evidence provides overwhelming support towards the benefits of sectioning calculus students by math ability. In the second section of this paper, I will address the most commonly overlooked aspect of the tracking debate – the students' opinions.

Student Survey Results:
Most of the research with tracking focuses on theory and expert opinion. Many studies also include statistical evidence from test scores. However, most research completely avoids any attempt to gather and understand student opinion on the subject. While the literature is rampant with *expert* opinions about how tracking may affect student psyche, there is very little effort to survey students and record their candid opinions. This was precisely my goal in this section of my research.

Throughout the semester, the faculty did not tell the students that they had been sectioned according to their class rank. Students in very "high" or very "low" sections may have noticed an unusual distribution of similar performers in their section, but most students later expressed no previous knowledge about their manner of sectioning. On the last day of class, the cadets were told about the tracking strategy that had been used during the semester, and all students were given an anonymous survey with three questions on it.

1. Describe any benefits/advantages (if any) that you observed this semester due to the tracking policy.
2. Describe any difficulties/problems (if any) that you observed this semester due to the tracking policy.

3. Given your personal experience with tracking this semester in your calculus class, circle one of the choices below to describe your overall opinion of the tracking policy: Terrific OK Don't Care Bad Idea Horrible

Of the 847 students who took calculus that semester in 2007, 798 students attended class during the last class period and completed the survey. Given the truly anonymous nature of the survey, the cadets were remarkably candid in their replies. I read each of the 798 surveys and compiled the results. Given a population as large as 798 students, at least one or two of the surveys included almost every conceivable viewpoint and opinion. However, the trends associated with cadet opinion became very clear: the students overwhelmingly supported sectioning their calculus classes by ability groups.

Considering the third question first, the histogram below shows the number of students who answered the question with each of the possible choices.

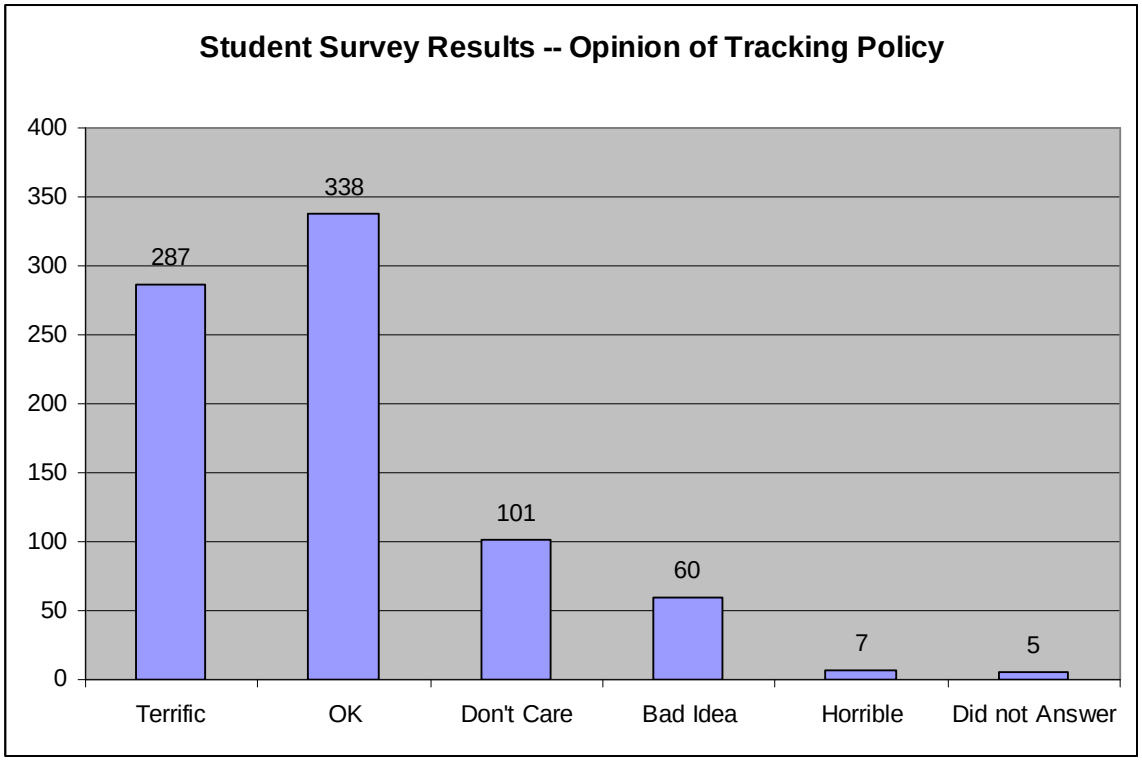

Figure 1: Student Survey Results – Opinion of Tracking Policy

I had expected a bell-shaped curve to the histogram, anticipating that most students would reply that they did not care either way and that the less extreme statements of "OK" or "Bad Idea" would have been more frequent than the more extreme replies of "Terrific" or "Horrible." Instead, almost 36% of the students thought that sectioning was "Terrific" and another 42% said that it was "OK." Even if we combine the "Bad Idea" and "Horrible" replies, they only account for about 8% of the total responses. The

cadets' answers to the first two questions on the survey reflect the overall trend seen in the histogram.

The vast majority of the surveys were very favorable of the tracking policy. The comments below are taken from cadet surveys, and they reflect the overall trends of supportive comments. Comments such as these were typical of most cadet responses.

1. "Cadets experience similar problems, and the instructor is concerned about everybody at the same level."
2. "The benefits are enormous because the cadets are on the same level, and they can work at the same pace rather than some cadets being held up by others."
3. No Disadvantage. The speed my class went was perfect. I am currently receiving the best math grade I have ever received."
4. "One person won't be left behind and another cadet won't be held back by slower people."
5. "The pace of learning was just right."
6. "It keeps people from being bored or uncomfortable.
7. "We all got along better, and it made class more enjoyable."
8. "We had better discussions in class."
9. "I was not scared to ask 'stupid' questions."
10. "The ability to comprehend course material was surprisingly consistent in the class. The class was more efficient because the instructor rarely had to give 'remedial' explanations in class."

Comments of this type were the norm, regardless of whether the student survey came from an "upper," "middle," or "lower" section.

The principle concerns with the tracking policy fell into one of three areas:

1. Students in the lower sections had to work harder on projects and homework assignments because there were not any "superstars" in the class who would do the lion-share of the group work.
2. Cadets tend to form study groups with other cadets in their sections. Cadets in lower sections needed to look beyond their class section to form study groups.
3. A small group (less than 2%) of surveys expressed concern about students in "low" sections feeling discouraged either by a "low section" label or because their section (as a group) was struggling with the course content.

Interestingly, the students addressed each of these concerns in their own comments on the surveys. I will respond to each of these concerns using actual replies from the student survey.

Many of the students who expressed the first concern in their survey immediately followed that statement by a realization that they had benefited from the tracking policy. One student in a "lower" section commented: "There were no people to do our work for us [on the projects]. We had to figure it out on our own. I think we do benefit."

Another cadet mentioned that the sectioning policy "kept one person from doing all the work." Cadets recognized that this short-term challenge had long term benefits since it forced them to learn the material. This better prepared them for success on course-wide exams and the final exam. As one student from a "lower" section commented, "it made me more responsible for my learning."

While many cadets expressed some *frustration* with the second concern (difficulty forming study groups), many of these same students recognized that this was easily overcome by studying with peers in their cadet companies, sports teams, or clubs. Furthermore, several cadets in lower sections commented that they had found themselves in the position of teaching their section-mates (a wonderful new experience for many of these cadets who had not typically been in this position.) As one such cadet wrote in the survey: "the best way to learn it to teach."

The third concern is one of the primary objections in scholarly research that opposes the notion of tracking. In fairness, these concerns did manifest themselves in a small set of the student surveys. However, a fascinating realization occurred when I cross referenced these comments with the sections from which the surveys came. The vast majority of these concerns came from "higher" level sections who were expressing concern for their fellow cadets who were in "lower" sections. In some respects, these comments were oddly reminiscent of the scholarly articles written by *experts* who speak of how the lower-tracked students *probably* feel. While it is admirable to be concerned for these students, the reality is that the overwhelming majority of the surveys from "lower" sections were exceptionally supportive of tracking. Instead of discouraging students in "lower" sections, being tracked with academic peers actually encouraged most of these students. Actual comments from students in "lower" sections are below:

1. "Cadets are more comfortable with asking questions in a class where everyone learns at about the same pace."
2. "I was able to communicate better – there was no identification as the 'stupid kid' in a class of stronger performers."
3. "The attitude for the students was improved. When you see someone you know of equal ability next to you doing well, you have a better confidence in doing well."
4. "Yes we benefited – there are no overly-arrogant know-it-alls who answer all the questions, draw all the teacher's questions, and intimidate the rest of the class into silence."
5. "I felt I belonged to class. Nobody was seen as too smart or too dumb."
6. "I felt more comfortable in this class."
7. "We were able to take the necessary time to ask pertinent, relative questions without pressure from more advanced students."
8. "I didn't feel as dumb as last year when there were a few people that were dominating the class."

One survey from a "lower" section described his/her tracking experience particularly eloquently: "I felt much more confident in class, and I noticed that cadets were more

willing to help each other succeed. Classes were more fun, I got more out of instruction, [I] left feeling like I understood the concepts, and [I] didn't feel stupid for asking questions. Most of the time, when I asked a question, somebody else was wondering the same thing. I did not observe ANY disadvantages."

In summary, the overwhelming majority of "lower" sectioned students were extremely supportive of tracking. In contrast to the fears of damaged psyche and discouraged performance, these students said that tracking boosted their confidence, increased their classroom participation, and ultimately helped them succeed. Combined with the strong statistical evidence for their improved performance on the final exam, it becomes clear that the students in "lower" sections experienced a tremendous benefit from tracking.

Faculty Survey Results:
A third information source used in this research was faculty surveys. Because of the large number of students enrolled (847) and the small class size (18 or fewer students per section) there were 19 faculty members teaching the calculus course in 2007 during the semester with the tracking experiment. Each faculty member completed an anonymous survey at the end of the semester that asked a variety of questions about sectioning.

I asked faculty members to rate the tracking policy on a scale of 1 (horrible) to 7 (terrific). The results are in the histogram below.

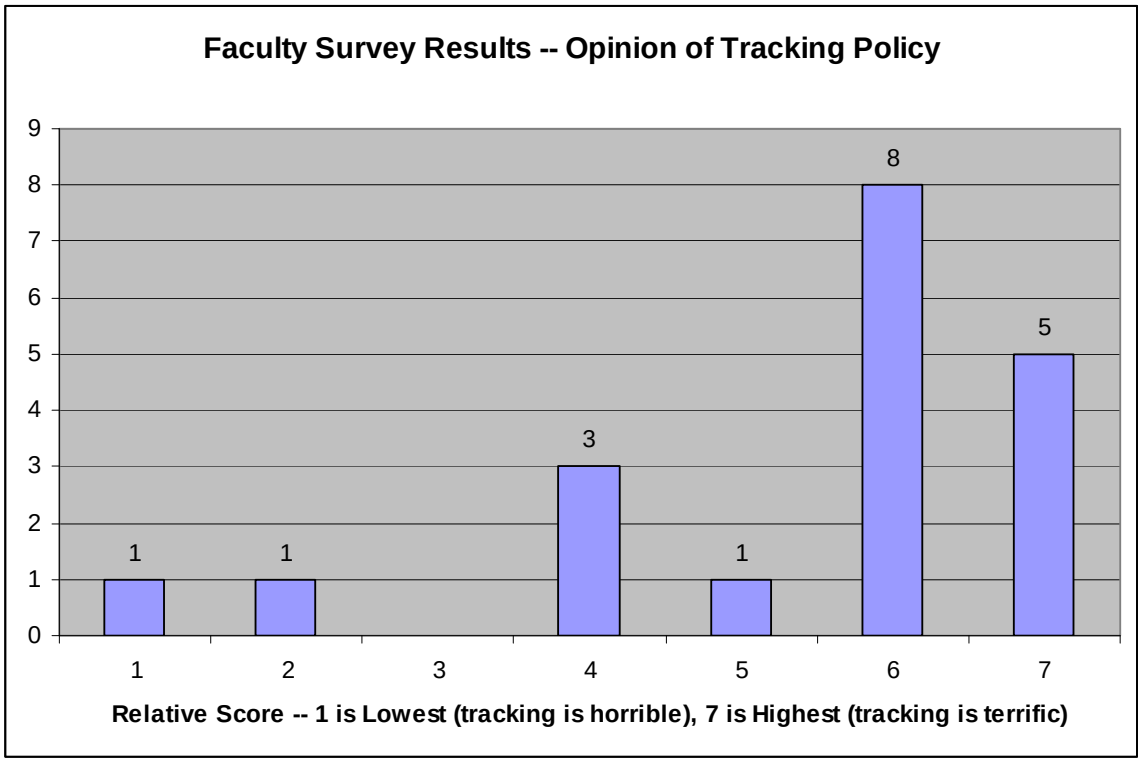

Figure 2 – Faculty Survey Results – Opinion of Tracking Policy

The comments of the 14 faculty members who rated sectioning as a 5, 6, or 7 essentially mirrored the students' comments. These faculty members taught a mixture of "upper," "middle," and "lower" sections, and their comments included the same favorable language that dominated the cadets' responses. The 3 faculty members who rated the tracking policy with the "neutral" score of 4 either stated that they were "unsure" or saw both benefits and disadvantages that also mirrored the comments made by cadets. I was most interested in the comments made by the two faculty members who gave the policy the low scores of 1 or 2.

The faculty member who scored sectioning as a "1" taught upper third students. This faculty member had the following comment: "It has been very beneficial to my (upper third) students. I was able to cover all the suggested materials and even some advanced materials. On the other had, I am concerned with the cadets in the lower third, who apparently have a hard time since there are few peers to learn from." The faculty member who scored tracking as a "2" taught a mixture of upper and lower sections. His main concern was a lack of motivation for success in "lower" sections. This faculty member commented: "I think it is not a good idea for "F" students. However, I think advanced students could benefit greatly from sectioning."

If we consider these comments, both faculty members stated that they thought tracking was beneficial to advanced students. Their sincere concern was for students in "lower" sections. While this concern is admirable, we must remember that one of these faculty members was speaking in theoretical terms only (having only taught in the upper sections), and that the students who were actually in the "lower" sections tended to be tremendously supportive of tracking in spite of the concerns of these two faculty members. I appreciate the sincere concern expressed by these two faculty members, but the "lower" sectioned students replied to these concerns through their own overtly-supportive comments in the student surveys.

As a final comment, there may have been some level of teaching experience associated with the trends in the faculty responses. Although our faculty enjoys a very congenial relationship in which all are seen as colleagues, it was an interesting observation that all five of the faculty members who rated tracking as a 4 or below on the faculty scale were members of the junior faculty. Of the six members of the senior faculty who taught calculus that semester, 100 percent rated the sectioning policy as a 5, 6, or 7 (average rating of 6.167.)

Many faculty members who read this paper may wonder about the difference in faculty workload associated with using a tracking policy. I also addressed this concern on the faculty survey. Most faculty members said that their workload was largely unchanged or actually less because they could more easily target their student audience. Some faculty said that their work load had increased, but this was most typically associated with a faculty member who was teaching numerous sections of different ability levels. One faculty member commented: "I basically have to adjust my lesson plan for each level. It's almost like teaching three courses at once." This is a valid concern that may be addressed by assigning faculty members sections which are on a similar ability level.

While some may speculate that it would be difficult to find volunteers to teach the lower sections, this was not the case in my observation. Many of our faculty members particularly enjoy the opportunity to work with the struggling students in a teaching and mentoring capacity. The more challenging task appeared to be teaching several sections of largely dissimilar abilities simultaneously.

Relevance Beyond USMA:

This paper has offered both a qualitative and a quantitative assessment of the use of tracking in a college calculus curriculum. I readily acknowledge that many colleges and universities are not inclined to address this issue because it may be seen as irrelevant in a college-setting where lecture halls full of undergraduate calculus students have the opportunity to sink or swim according to their capabilities and study habits. However, many colleges and universities are re-emphasizing the importance of "teaching" rather than "research," and many of theses schools are using small classroom settings to encourage student/teacher interaction. A tracking policy is ideal in this setting, as demonstrated by the success of the tracking policy at West Point. I would also suggest that many High Schools may recognize similarities with the West Point system of academics, and that this study may provide evidence for the benefits of tracking in High School mathematics departments.

Conclusion:

This research has demonstrated the immense success of tracking in a calculus curriculum at the United States Military Academy. This was evidenced by a striking, statistically-significant increase in final exam scores. This increase in test scores applied to students of all ability levels, including the students in the "lower" sections. My research also included the (typically overlooked) student perspective on tracking, and found that an overwhelming majority of students favored tracking, regardless of their ability group level. Finally, this research considered faculty input and found that more than 70% of the total faculty (and 100 percent of the senior faculty) supported the use of tracking in the calculus curriculum. Most importantly, this research challenges the commonly-held opinion that tracking benefits the top few at the expense of the many at the bottom. In contrast, the students in the "lower" sections experienced dramatic improvements in final exam scores as compared to previous years that did not employ a tracking policy, and student comments from "lower" sections were immensely supportive of the tracking system. I will readily acknowledge that the West Point academic experience is somewhat unique when compared to other colleges and universities. However, the exceptional success of tracking in the calculus curriculum at West Point provides strong evidence to continue the policy at USMA and to consider (or reconsider) a similar policy at many other academic institutions.

Notes:

http://www.usma.edu/Class/2008/profile.asp
http://www.usma.edu/Class/2009/profile.asp

Acknowledgement:

This work was conducted in collaboration with the United States Military Academy (USMA) Master Teacher Program, offered through the USMA Center for Teaching Excellence (CTE), West Point, New York.