Predicting Academic Success at WPI

An Interactive Qualifying Project Report

Submitted to the faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree Bachelor of Science

On April 24, 2003

by

Kevin Rogerson

Approved By:

Professor John Wilkes, Advisor

Abstract

This project is an attempt to determine if there is a way to predict a student's level of academic success at WPI. By analyzing data for the class of 2001 this project looked at factors such as the student's MBTI type, SAT scores, and high school academic performance to see if they offered any insight into how the student would perform at WPI. The level of achieved success was measured using the students' freshman year WPI grades.

Table of Contents

Project Overview

It is not hard to believe that certain students experience a greater level of academic achievement than others. Past studies of WPI students from the class of 2002 and 2003 have even been able to determine which groups and types of students are more likely to have a difficult freshman year. These studies have been based on analyzing freshman year grades while looking at the students' psychological type as determined through the Meyers-Briggs Type Indicator (MBTI).

One result of the original class of 2002 study, performed by Greg Doeschler, was the revelation of a phenomenon given the name B-Term Drop. When looking at the average grades for each term the students had a significantly lower GPA in B-term but recovered by the end of the year. It was found that students with the partial MBTI type SP not only showed the weakest overall performance of the groups studied but also had the most dramatic drop in their GPA for B-term.

A second study (*1st Year Experience for Class 2003*) was performed by Tara Murphy who conducted an identical analysis on the class of 2003 and was able to replicate the B-term Drop for the SP's but not for the rest of the students who instead suffered their worst performance during C-term.

This project originally intended to investigate the sophomore year of those two classes to see if similar patterns emerged or if the B-term plunge could be considered strictly a freshman year incident. Due to delays in receiving the data required, that plan was scrapped and a replication study was designed to examine the class of 2004 and 2005 to verify the B-term drop. Neither data set is received so a

third option is presented; attempt to revive the class of 2001 dataset that was plagued with problems and use it to execute the replication study.

Motivating this project is its impact on academic advising practices at WPI. If it can be found that certain groups of students are at a greater risk of failure (or of struggling with their classes and workload) then it would be possible to flag them before problems begin, and ensure that programs such as study groups or tutoring are readily available for those who need it most. In identifying students this way resources will not be directed to students who will recover on their own, or those who are not cut out for the college environment and do not benefit from assistance.

For this study, several tests and indicators were used to create a measurement of success. The widely used standard for determining personality type, the Myers-Briggs test, was essential to this study and has been an integral part of educational studies and human resource initiatives since its development. In addition, the Scholastic Aptitude Test (SAT) was used as a predictor of a student's potential success in a college setting.

The Myers-Briggs test was developed by Isabel Briggs-Myers and her mother, Katherine Briggs. Based on the work of Carl Jung, it sorted each person into a defined personality type based on their responses to carefully crafted questions in the survey. Carl Jung developed his theories after extensive study of literature, history, mythology, psychopathology and other sources of observations on human nature and personality. He believed that all personalities could be divided into categories that worked in conjunction with a second set of traits. The main group

contains personality attitudes; namely, introversion and extroversion. The second contains "modes of orientation" that operate in either an introverted or extroverted way. The modes of orientation are classified as thinking, sensation, intuition and feeling. Since these can operate in conjunction with introversion or extroversion, there are eight classifications under this system. According to Jung, each person adapts to the world in a specific way, and that way characterizes how they relate to the world and in which direction their focus is projected. An introverted person tends to keep to himself or herself and have a quiet, reflective, shy nature that hesitates to interact too much with the outside world. Extraverts, on the other hand, are outgoing, confident and adapt quickly to new people and situations. For this person, external forces tend to be the motivator behind their thoughts, feelings, judgments and perceptions. Conversely, introverts tend to use internal motivations and their own thoughts as their rationale for behavior.

The Myers-Briggs type indicator takes Jung's ideas further. In their model, there are sixteen types. Introversion and extroversion still characterize each type, but the four functions are given a type based on their relative dominance within each personality. For example, a person may have an extraverted orientation with a greater reliance on sensation as a basis for their perceptions and judgments, and a lesser reliance on intuition. The Myers-Briggs indicator breaks down Jung's work and makes it more specific, so that personalities can be more narrowly categorized. For this reason, this psychological indicator can be a useful tool for anyone attempting to sort people by personality type. Human resources personnel find it

useful in determining a person's potential work habits so as to place them in the environment or position in which they will be most productive, or to place them on work teams with others who have similar personalities instead of with personalities with whom they will clash. The Myers-Briggs indicator is also useful in determining learning styles, which ultimately assists teachers in tailoring their teaching strategies to their individual students. For example, the current educational system tends to favor introverted and intuitive personalities. They tend to learn best when theories and concepts are presented first and practically applied later on. However, the majority of people are Extraverts and have an ideal learning style different from what are considered the traditional teaching methods. Extraverts learn better in group settings, where they can express their ideas out loud and share their ideas with others. Introverts learn best if they can think about their work on their own and process information at their own pace. If students were sorted into personality groups, classes could be structured around personality groups that learn best with a certain style so that each student's learning potential could be maximized.

In addition, personality type can indicate how a person responds to conflicts, crises and change. It is in this function that the test most applies to my study. Each personality type has a different way of approaching challenges and stresses, and each will handle a certain situation in a different manner. In their career at WPI, students face the same hurdles at about the same time due to the nature of the curriculum. As such, their reactions to that same situation can be observed and categorized. If conclusions can be drawn from this study, recommendations can be

made about how the curriculum can be tailored and academic advisement resources allocated to those students who will benefit most from them and to identify students who are at the most risk for academic failure as a response to the stress and conflict the curriculum presents.

Also used in this study as a measure of academic success is the standard predictor of college performance, the Scholastic Aptitude Test, or SAT. The SAT was developed by the College Board to standardize the applications process at the nation's colleges and universities. Each college had its own set of admissions procedures by which it judged candidates. In addition, each high school had its own courses and measures by which they assessed their students, making it difficult for college admissions personnel to judge between their applicants' varying qualifications. Each college often had their own entrance examination, which required a student to travel to the school and sit for the exam. If they applied to several colleges, they had to take several different exams, which quickly became a hardship for many students.

The College Board was developed to create a standard entrance exam that could be taken at any location and would be accepted by all colleges. In the early 1900s, as college admissions quickly grew, it became necessary to make more standardized determinations of a student's academic merit. At the same time, a fad for intelligence testing swept the nation. In this environment, standardized college entrance exams gained headway. The SAT was designed not only to measure what a student had learned, but also to determine whether a student could overcome

factors such as a poor high school or poor study habits and still succeed in a college environment. The first test was administered in 1926, but was slow to gain popularity as colleges were not sure of its effectiveness and were reluctant to use it widely in their admissions procedures. However, with the flood of college applicants after World War II, a newly implemented, more updated and more comprehensive test was widely administered and became the standard for college admissions.

In 1947, Educational Testing Services took over the development and administration of the test. They continue to administer the SAT, and published the Myers-Briggs indicator for some time before it was sold to Consulting Psychologists Press (Palo Alto, CA).

Method and Analysis

The first step in this project was to examine the class of 2001 dataset to see what additional information would be needed. The original dataset was constructed by Keith McCormick and a group of students looking to study the predictive power of the SAT immediately after the class of 2001 completed its freshman year. At that time only 45% of students granted permission to view their high school transcripts. To complete my analysis it would be necessary to have access to the remainder of the students' records.

Dealing with this type of information raised some potential conflicts. The largest issue I had to confront involves confidentiality of the information contained within student records. In order to organize all the data effectively there would have to be some way to tell "who" the data belonged to without using names. The obvious solution to this was to use the students' WPI ID number as a key on any records collected and use that key to combine the data. The biggest issue in dealing with confidentiality, however, would be allowing me, as a student, to have access to any records that contain this ID number as it is also that student's Social Security number. Before any data collection could begin I had to agree to the strictest confidentiality practices to ensure that no personal information would be released in an unauthorized manner.

Once access to the student records was granted, the process of data collection was expedited by the work of Jonathan Oexner, who as part of an earlier study, *The Potential in Data Mining: WPI Class of 2001 Case Study*, was able to create a very

detailed dataset for the class of 2001 that had been stored in Microsoft Access. Oexner's efforts were able to produce a collection of datasets that included the students' name, ID number, gender, freshman grades from WPI, high school grades, MBTI scores, major of study, and a variety of other information. The dataset, however, was still missing high school grade data for many of the cases. My next goal was to fill in those gaps.

The first step was to determine which students were "missing" from the dataset. This was done by comparing the master list of students to the data set containing high school grades; a new list of "missing" students was compiled. Armed with that information I contacted the registrar's office to gain access to their archives of student records. With permission granted, the long process of collecting the required data could begin. Over the course of a few weeks I was able to find many of the missing students and obtain a copy of their high school transcripts. It is important to note that the transcripts contained within these files were not always complete. It was decided that any incomplete transcripts would be used. The reasoning behind this decision was that if the Admissions office was able to make a judgment based on the partial transcript then there must be enough information available to determine that the student will meet the challenges of WPI, which means that there is enough information present for that high school performance to be used as a predictor of potential.

After all of the new transcripts were collected, the information had to be entered into a new dataset. This was yet another time consuming process that

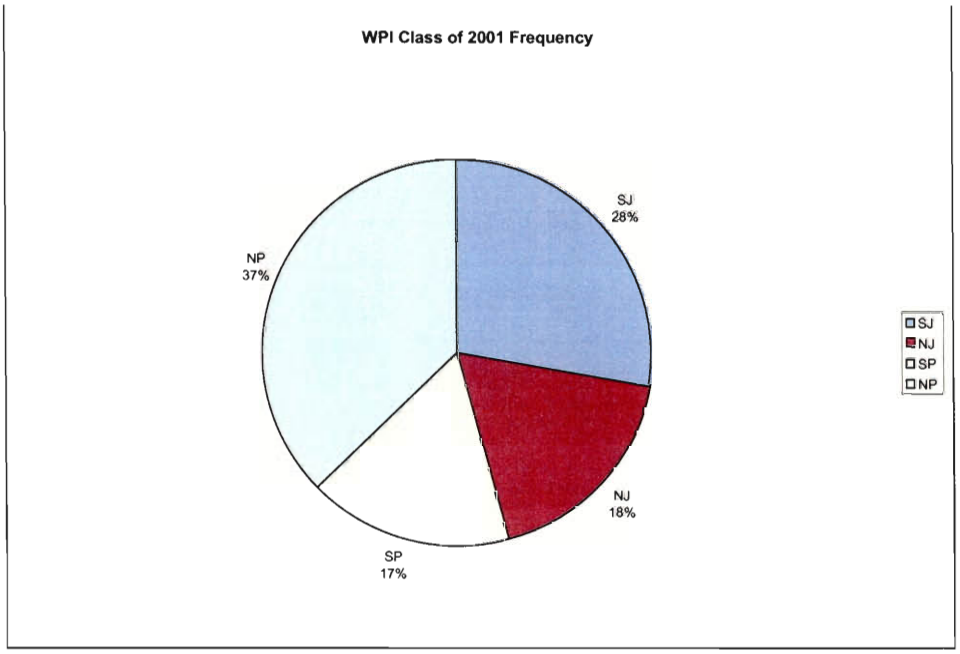
involved the manual entry of the student's name, which was later changed to ID number, and each student's grades broken down by year, (freshman sophomore etc) and subject of study (math, science, history, language and English). This data was entered into Microsoft Excel as it allowed for easy conversion to a dataset in SPSS (Statistical Package for the Social Sciences), a powerful software package which was used to perform the analyses to be described later. Because grading policies can vary from school to school, and to maintain consistency between this and other datasets, the grades were converted to a uniform numeric equivalent as summarized in the table below. If a course was not taken in a particular subject during that year, a grade of 0 was entered.

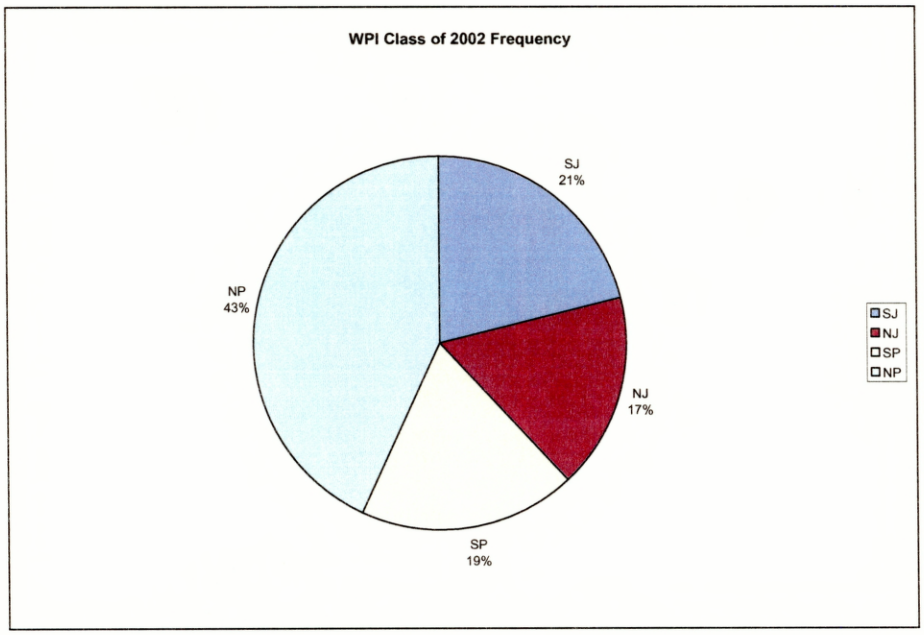
Grade	Numeric Equivalent
A+	98
A	95
A-	95
B+	88
B	85
B-	82
C+	78
C	75
C-	72
D	65
No Course Taken	0

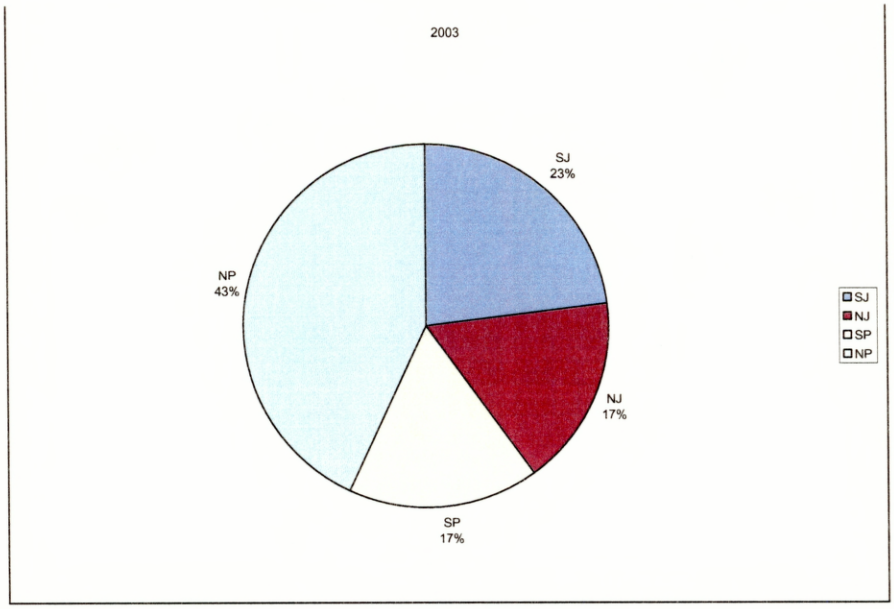
As part of this data collection effort it was also recorded whether or not a student had any exposure to calculus courses while in high school. The level of exposure was rated on a scale of 1-3, with 1 representing that the student had no exposure to calculus courses, 2 signifying that the student participated in a pre-calculus course, and 3 meaning that the student had studied Calculus in high school.

This information was then converted to an SPSS data file and merged with the data from Oexner's collection. Once in SPSS the data had to be manipulated to create more useful variables. High school grades for each subject were averaged to create the student's overall score in that subject. For the purpose of this averaging only completed courses were included. A grade of zero, meaning a course not taken, was not averaged in.

With the high school dataset completed the analysis could be begin. My first study revolved around comparing the class of 2001 to the other classes studied to see if it "looked" similar in terms of the SN&JP variable combinations. Information for all three classes is presented in the following charts.

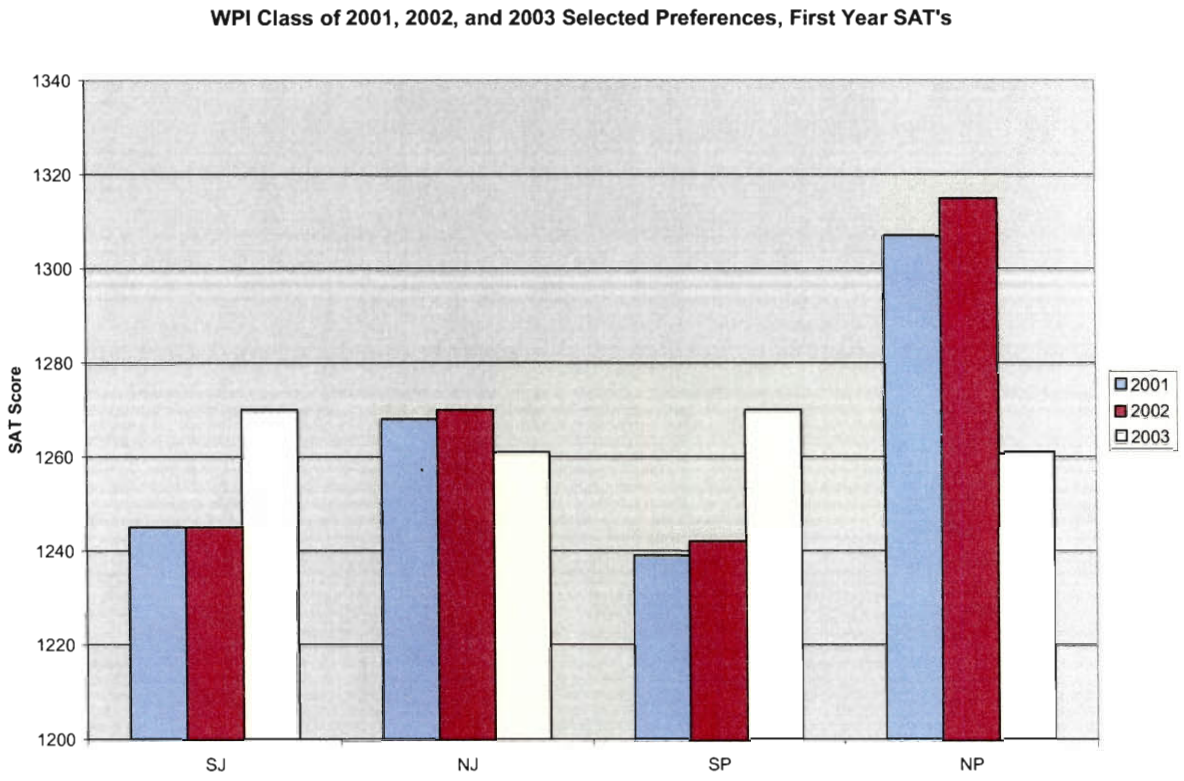






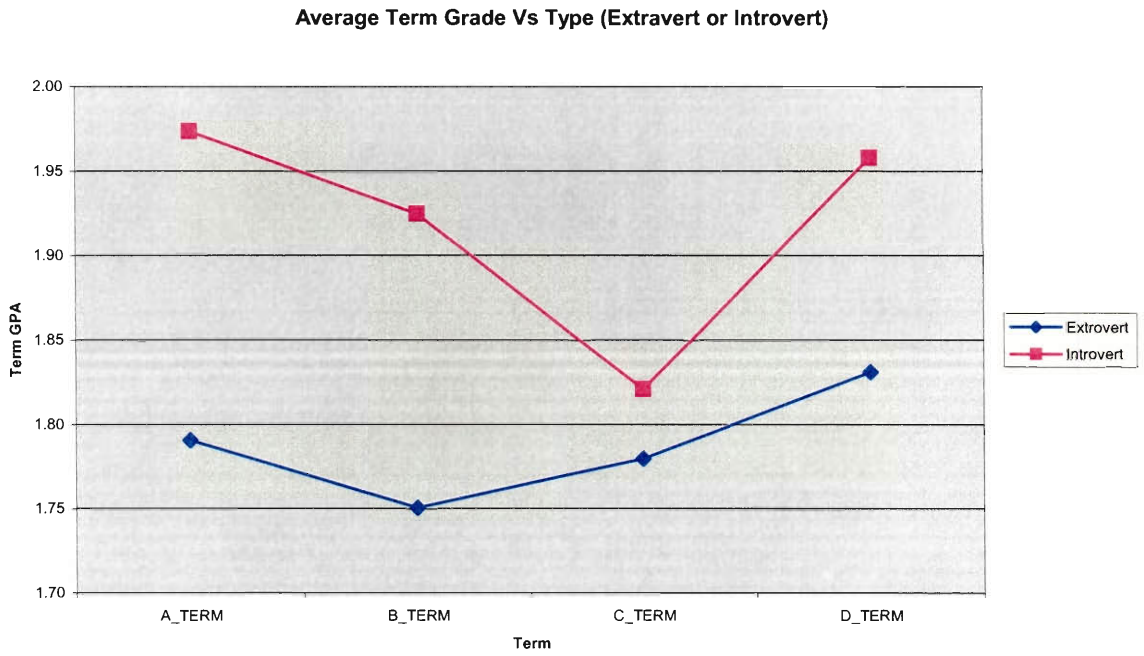
The make-up of the classes is nearly identical. Another reason for this comparison is due to the fact that prior to this study a myth surrounded this class that they had higher average SAT scores than any other. The result of that comparison is shown

below.



It appears that any mystery regarding the class of 2001 has proven false since it is nearly identical to the two others.

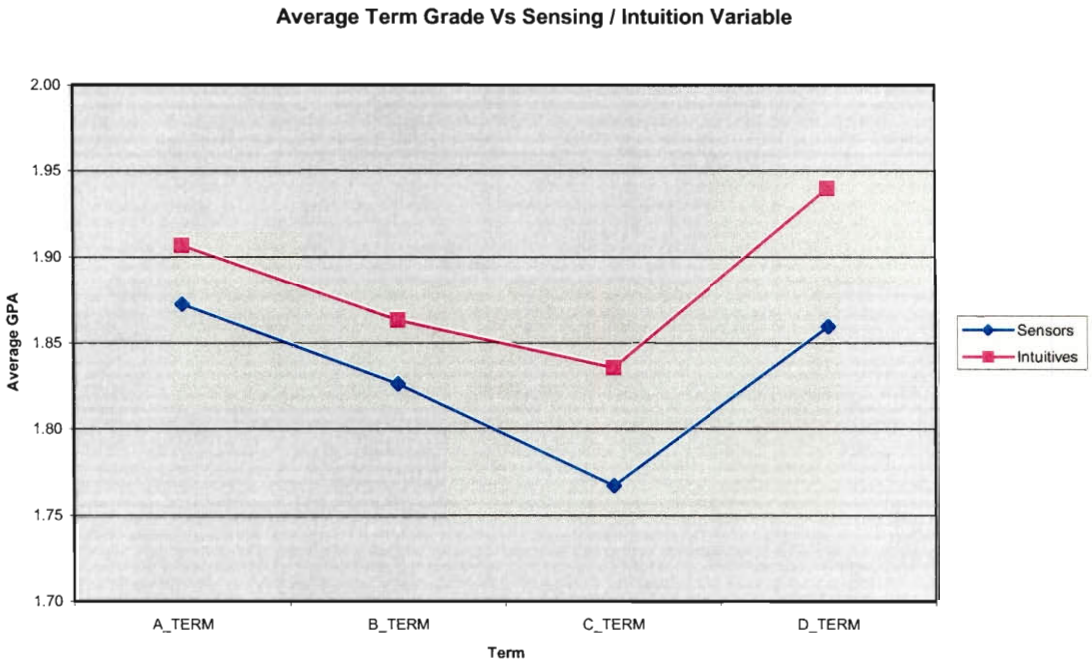
My next attempt was to examine a student's academic performance based on their other MBTI preferences starting with introversion and extroversion. As was previously stated, the current trends in teaching emphasize an introverted learned style so my theory was that the Introverts would demonstrate a greater level of academic success. The graph below shows the average term grades of the Introverted and Extraverted students. As a procedural note, when calculating the averages, grades of zero were not calculated in.



From this representation it is fairly clear that Introverts experience noticeably greater academic achievement.

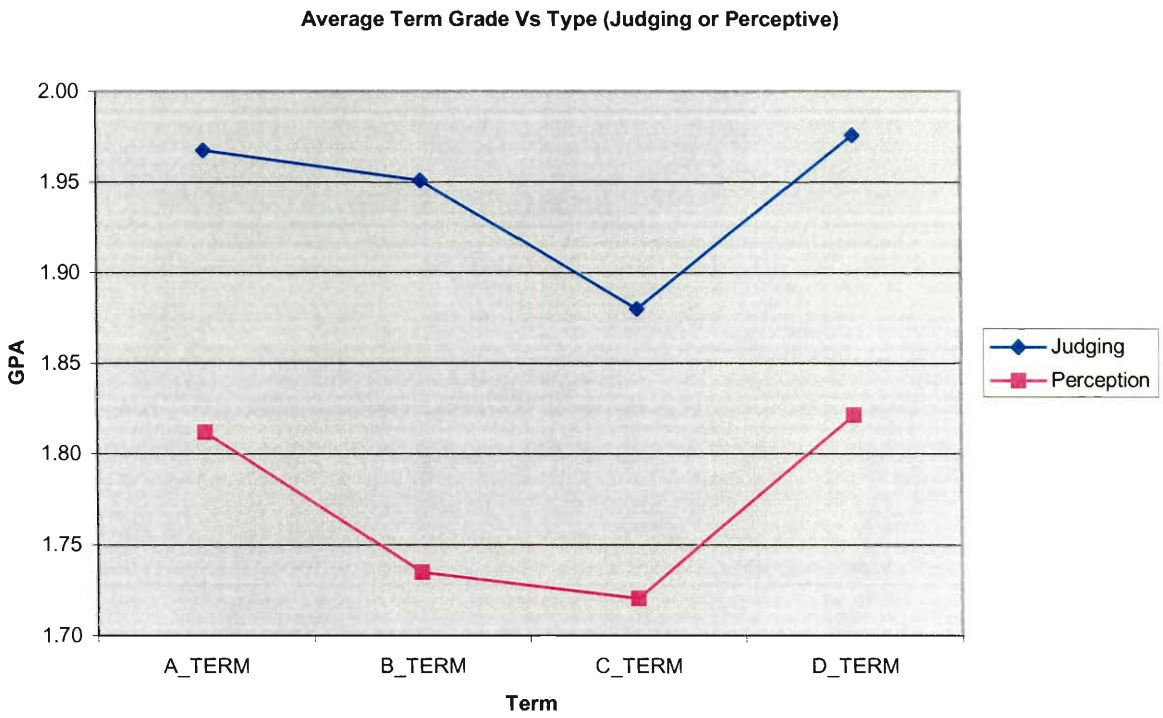
A similar analysis was performed by breaking the dataset into two groups representing the Sensing and Intuitives. Sensing students, as the name implies, learn and understand through the use of their five senses. An ideal learning environment would involve hands on experience with concrete answers. Intuitive students prefer to see a concept through the theory of its operation and understand the possibilities of its application. With the exception of lab sessions, most classes are taught by presenting theory first and then working through an example. For instance, a professor may derive a complicated mathematical equation first and then demonstrate the result using real numbers. This style of teaching would appear to favor the Intuitive students who could mentally process the theory of the equation

while the Sensing students would struggle to follow along until the usefulness of the equation could be applied. While the difference between the two types is only a few points it is clear from the graph below that during each of the four terms the Intuitive students performed better.



A third analysis was carried out broken down by the students' preference for Judging or Perception. Judging students ideally live a life that is structured and orderly. Problems that arise are faced decisively and resolved quickly. Perceiving students prefer to gather as much information as possible before deciding on a course of action. Perceptives will also flourish in a more flexible and uncontrolled environment. Prior to performing my analysis it was unclear which, if either, of the two groups would demonstrate an advantage in the academic environment. Since most courses at WPI require daily attendance, this would appear to favor the

Judging students who could easily fall into a routine of class time and study. Looking at the breakdown of the grades received it is indeed that case that Judging students outperform their Perceptive counterparts.

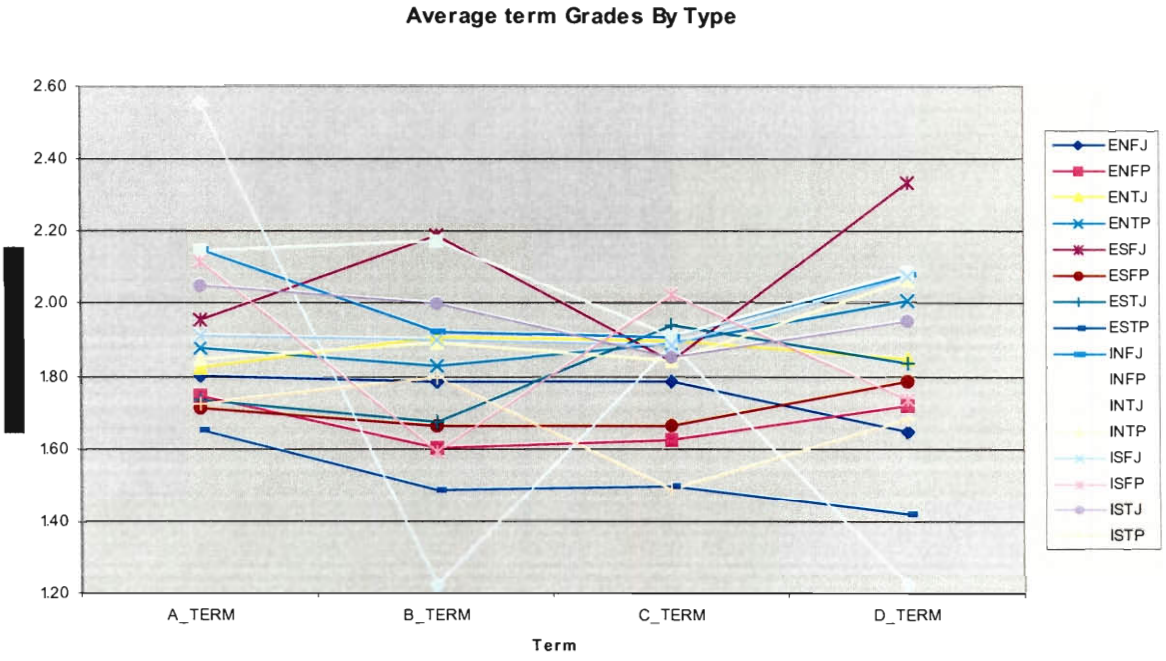


Another explanation for this discrepancy may be that Perceptive students are reluctant to ask for help when they experience difficulty. While the Judging students would recognize a dip in performance and make a quick decision whether or not to seek help, Perceptive students may take time to analyze the reason for the slump and debate a future course of action to correct it.

It is interesting to note that all the types studied so far, with the exception of the Extraverts, have experienced the greatest difficulty during C-term, with a clear improvement to end the year. While I cannot offer an explanation for this, the cause

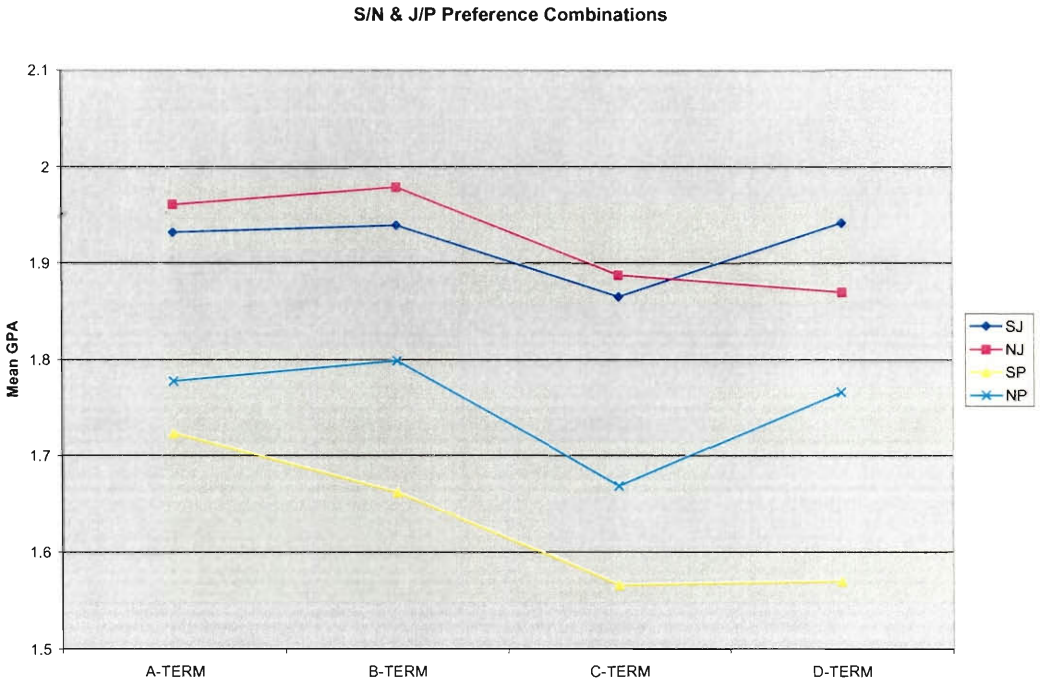
may be worth looking for in future studies. I would also like to point out that, while no results are shown, an analysis was performed to explore the Thinking or Feeling dimension of the MBTI. The results for these two types of students were so nearly identical that no significant distinction could be made. While initially I believed that thinking students, with their penchant for logic and analysis, would have an advantage over the Feelers at an institution like WPI where science and math usually having one "right" answer. Since the skills necessary to succeed in science and math are dominant in Thinkers, it could be assumed that they would have an advantage over Feelers. However, this ultimately proved not to be the case.

Since there was a clear distinction between the preferences of the MBTI, I decided to see if the same was true once all the preferences were put together into the students' true MBTI type. The result of that analysis is the graph shown below. While complex at first, a deeper inspection can yield some interesting information.



The most obvious is the wild variations of the INFPs. This seemingly important result can be explained by the fact that there are only three cases of this type; therefore the average score is not representative enough to be significant. Also important to note is the unfortunate ESTPs who, consistent with the previous analyses, hold the lowest average over the course of the year, an average of 0.3 points behind all the rest.

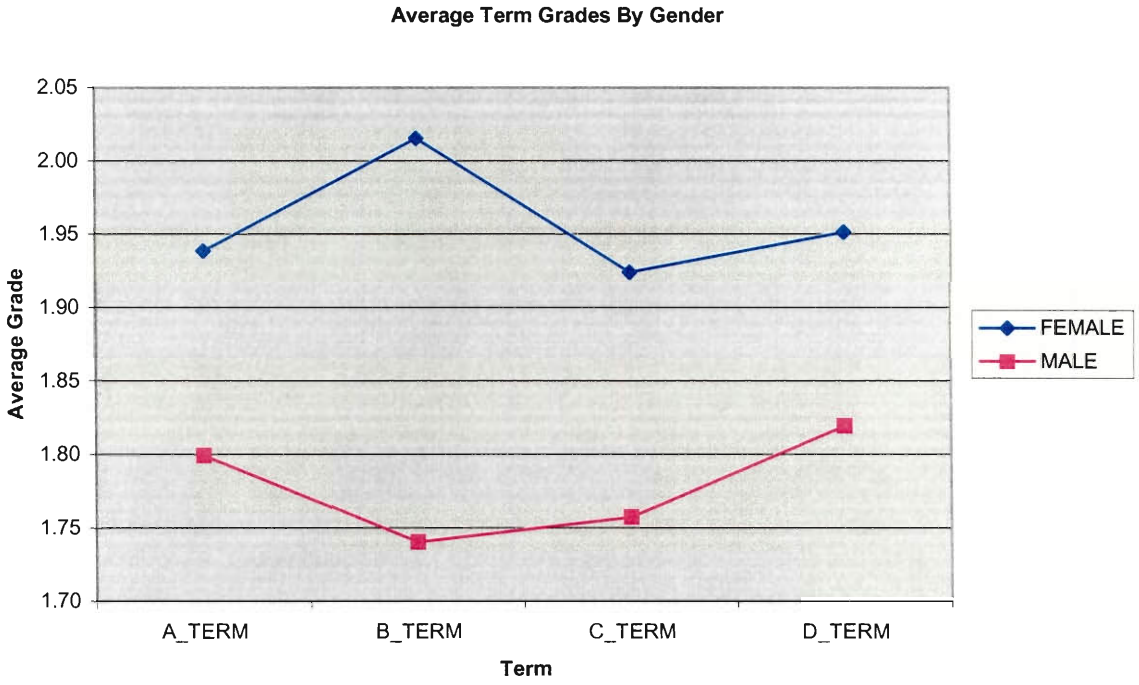
These prior analyses also set me up to replicate the B-term plunge findings. When I combined the SN and JP preferences and ran the analysis the result was a little surprising.



For this year the lowest averages occurred during C-term and while the SP's remained the worst performers there is no sign of improvement as there was in the past. An important analysis would be to see what level math course the 2001 students started with and if indeed it was more advanced than most years. If so this might explain why the "slump" was pushed to later in the year. If their overreaching caused the students to fail their first math course they would essentially be one term behind the students from the previous studies. Unfortunately one of the problems with the original 2001 dataset is that it does not include the information about which courses were taken that would allow me to make that conclusion.

With a greater understanding of the role MBTI type plays in a student's academic achievement I began to examine what other factors could predict a student's level of success. It seemed to me an obvious test to explore the other major

type difference, which is gender. WPI is a school with a predominantly male population, 76% male versus 24% female. I wanted to determine if this carried over into academic performance.

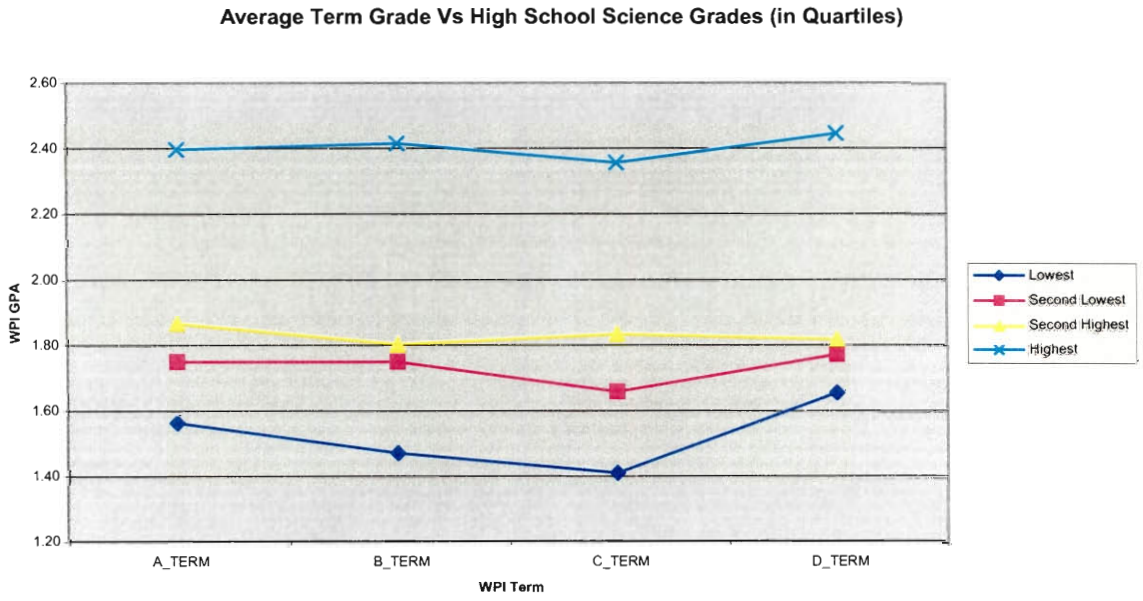


By breaking the dataset down by male and female students and performing the same type of averaging as in the previous studies the above results were generated. It is clearly seen that the female students outperform the males over the course of the year. Explanations for this include the possibility that females feel the need to prove themselves equal in the scientific field that make up the majority of programs at WPI and therefore work harder. It could also be that the male students have a higher level of confidence and a tendency to "coast" through their classes believing that extra effort is not required. Whatever the reason, the female minority appears to hold a distinct advantage over the male population.

I next turned my attention to the students' high school academic performance. Since the curriculum of WPI is based on scientific principles and classes the logical conclusion is to see how a student's background in these courses relates to that student's performance at WPI. Running a correlation analysis using SPSS produced the following results.

	A_TERM	B_TERM	C_TERM	D_TERM
Correlation with High School Science Average	0.464	0.436	0.442	0.348

These values indicate that a student's performance in high school science classes is a relatively significant indicator of their performance at WPI. To see just how well a student's performance in science courses affects their grades at WPI, SPSS was used to split the students in quartiles based on their average high school science grade. The average WPI term grades were then calculated for each of these groups. The graph below shows the relative performance of each grouping.

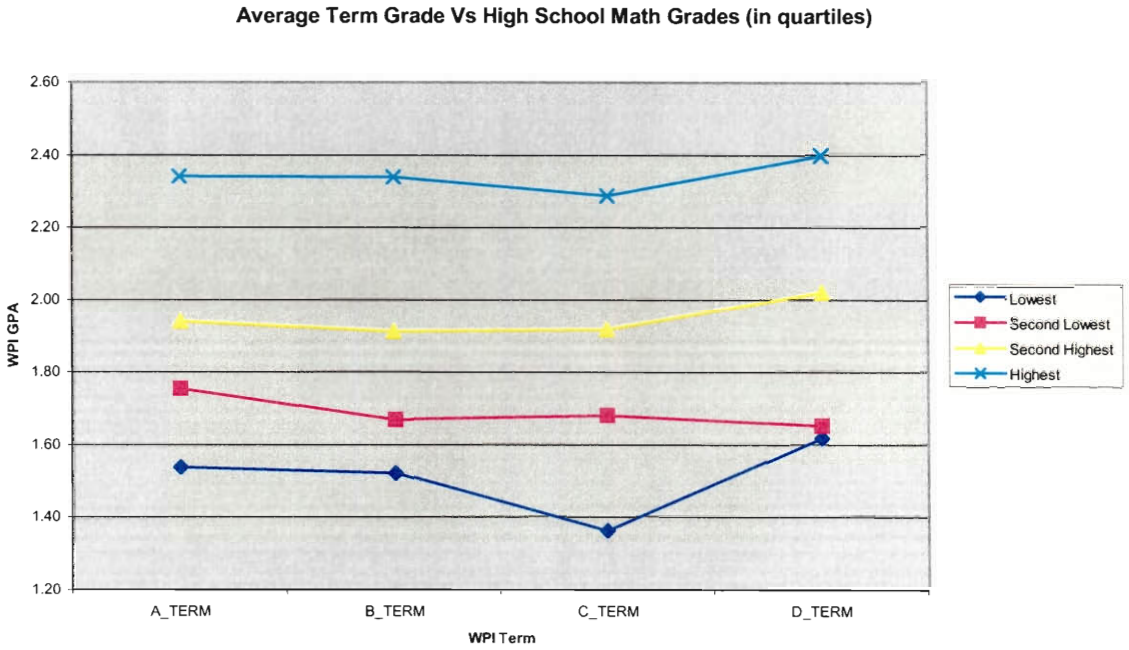


Not surprisingly, the student that performed the best in their high school classes also received the best grades in their WPI courses. What is surprising is the large gap between the highest quartile and the rest of the students. On average this group received almost a full letter grade higher than the lowest group. This is especially interesting given that the two "middle" groups are so close together in terms of performance. Also of interest is the fact that these averages do not take into account the level of science course taken during high school. If a student earned an A in an introductory high school course they appear to have the same ability to perform well in college courses as those students who received A's in advanced level courses.

Of equal, if not greater, importance in the WPI curriculum is mathematics courses. Since science courses can be used to predict a student's academic achievement it would follow that the same could be said for their past experiences with mathematics. Indeed, when looking at the correlation of math to WPI grades there are significant results.

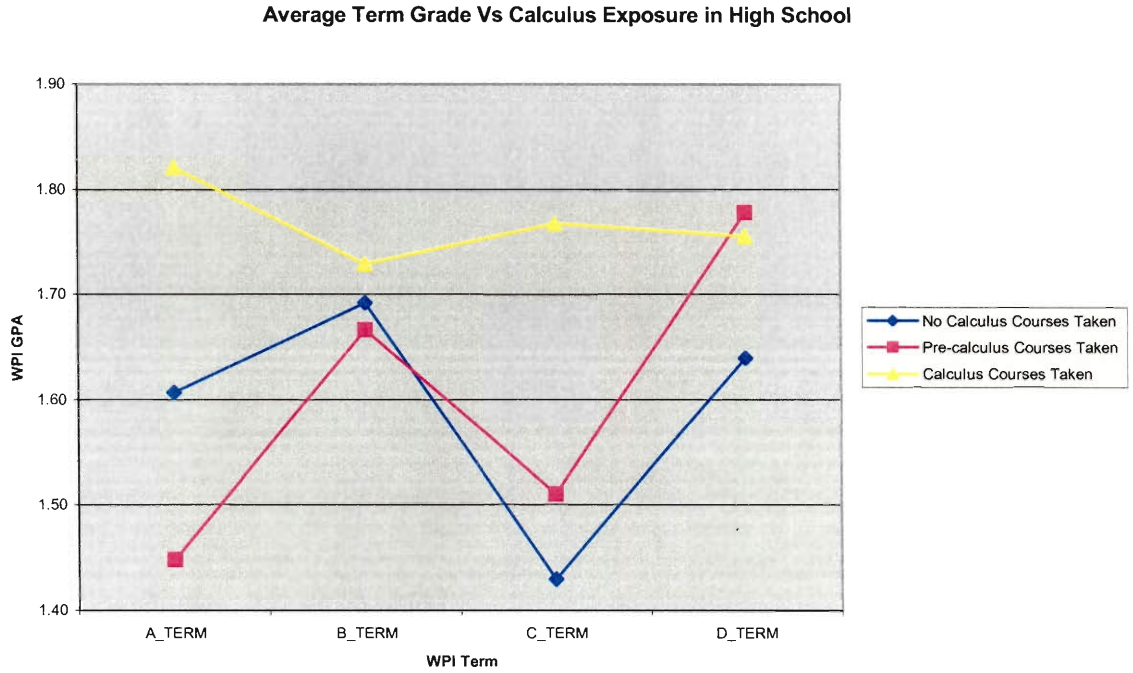
	A_TERM	B_TERM	C_TERM	D_TERM
Correlation with High School Math Average	0.461	0.424	0.438	0.380

Once again there is pretty strong evidence suggesting that if a student does well in high school math courses they will do well in their college work as well. To see how this theory plays out the students were broken into quartiles and the average grade for each group was calculated and plotted as shown below.



As expected, those students who performed well in high school continue the trend once in college. In this analysis, in contrast to that relating to the science grades, there was a larger discrepancy between the groupings but the relative order was maintained. To see if the level of math courses taken has any bearing on the level of student performance a second analysis was run. This second analysis looked at whether or not a student had any exposure to calculus courses while in high school and how that experience affected their level of college achievement.

The dataset was split into three groups, those that had no calculus courses, those who took a pre-calculus course, and those who completed a full calculus course. The average term grade for each of these groups was calculated and the results are shown in the following graph.



As expected, those students who had taken calculus courses were more prepared for the challenges of WPI courses. The results generated for the remainder of the students illustrate an interesting phenomenon. For the first two terms, those students with no exposure to calculus were able to out perform those students with at least some prior knowledge. While there is no clear explanation for this I can speculate that those students with prior exposure to calculus decided to start themselves in a higher level math course upon entering WPI than those with no exposure. This aggressive strategy may have caused these students to end up in a more challenging course than they expected while those who started slowly were able start off well until the courses became too challenging and prior knowledge could be applied. On the basis of this and other findings it should be clear that there are several variables that affect a student's academic success.

Conclusions and Recommendations

I must admit that I am a little disappointed with the findings of this report. When this project was started I was hoping to make some great discovery into what determines a student's success at an institution like WPI and be able to write a paper that would spark a revolution in the way Academic Advising and the Office of Admissions went about making their decisions. But at this time I have found no major revelation to that end.

It seems that the most accurate predictors of a student's academic success are the usual suspects; if you do well on the SATs you will do well in classes. The same could be said for student's who flourish in math and science; the skills they learn in high school are readily adaptable to the college classroom. Not surprisingly, students who take advanced level courses such as calculus are able to achieve a greater amount of academic success upon entering the college environment.

To fully understand the role of high school courses more information would need to be added to the dataset including what level of course was taken (honors, AP, or general studies for example) as well as information relating what courses are taken during the freshman year at WPI.

It does appear that the MBTI type plays a role as well. Average grades by type did show some evidence of an advantage for students with certain preferences, such as Introversion, or Intuition. It also continues to show that SP students are at risk and may benefit from assistance. This revelation underscores the need to determine the students MBTI type before, or soon after they enter WPI.

In the interest of analysis for the class of 2001 I would like to see the dataset extended so that a full comparison to other classes could be completed. I also suggest testing the findings of this paper, mainly the shift of the B-term drop, to see if it can be explained. If the shifting remains unaccounted for future studies must certainly be done to replicate the B-term findings of the previous studies. In this way it can be determined if that phenomenon was a coincidence or if there is something about the second term that causes students to experience such a drastic reduction in the GPA.

As my project shifted gears from the start I would also like to see research done into the sophomore year experience of these and future classes. To truly explore the C-term slump of the 2001 SP's it is necessary to see if they recovered in the later terms. Extending the analysis into later years of study would also shed light onto whether or not the B-term drop is caused by a specific factor in a student's freshman year or if it is the nature of academic life to experience a few pitfalls along the way.

Bibliography

Crouse, James and Dale Trusheim. *The Case Against the SAT*. The University of Chicago Press, Chicago, Illinois. 1988

Demarest, Larry. *Looking at Type in the Workplace*. CAPT, Gainesville, Florida. 1997

Lawrence, Gordon. *People Types and Tiger Stripes.* Center for Applications of Personality Types, Inc., Gainesville, Florida. 1979

Sharp, Daryl. *Personality Types: Jung's Model of Typology*. Inner City Books, Toronto, Canada. 1987

Appendix A

Prior Data on the Classes of 2002 and 2003

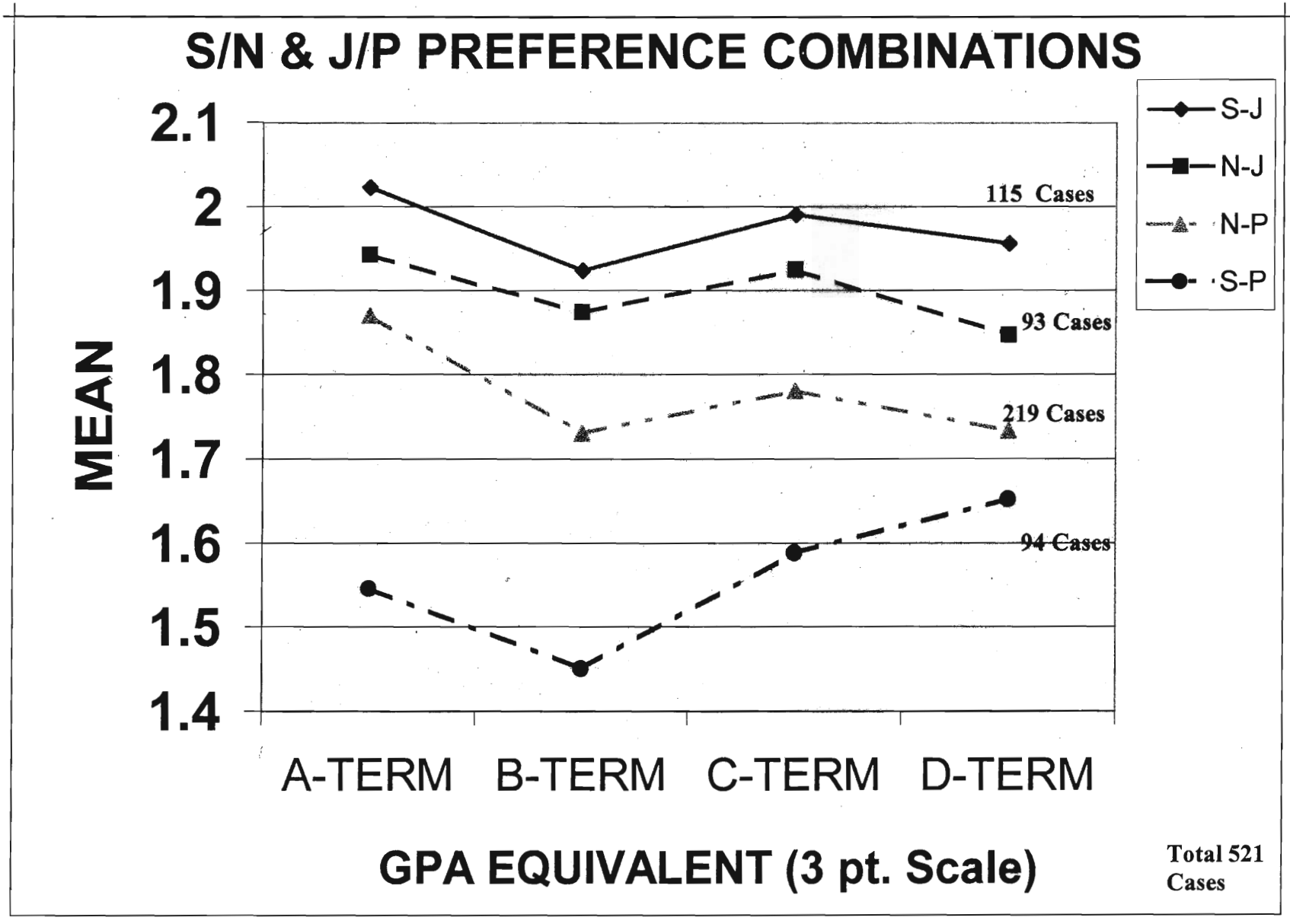

S/N & J/P PREFERENCE COMBINATIONS
MEAN
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
S-J
N-J
N-P
S-P
115 Cases
93 Cases
219 Cases
94 Cases
A-TERM
B-TERM
C-TERM
D-TERM
GPA EQUIVALENT (3 pt. Scale)
Total 521 Cases

TABLE 2

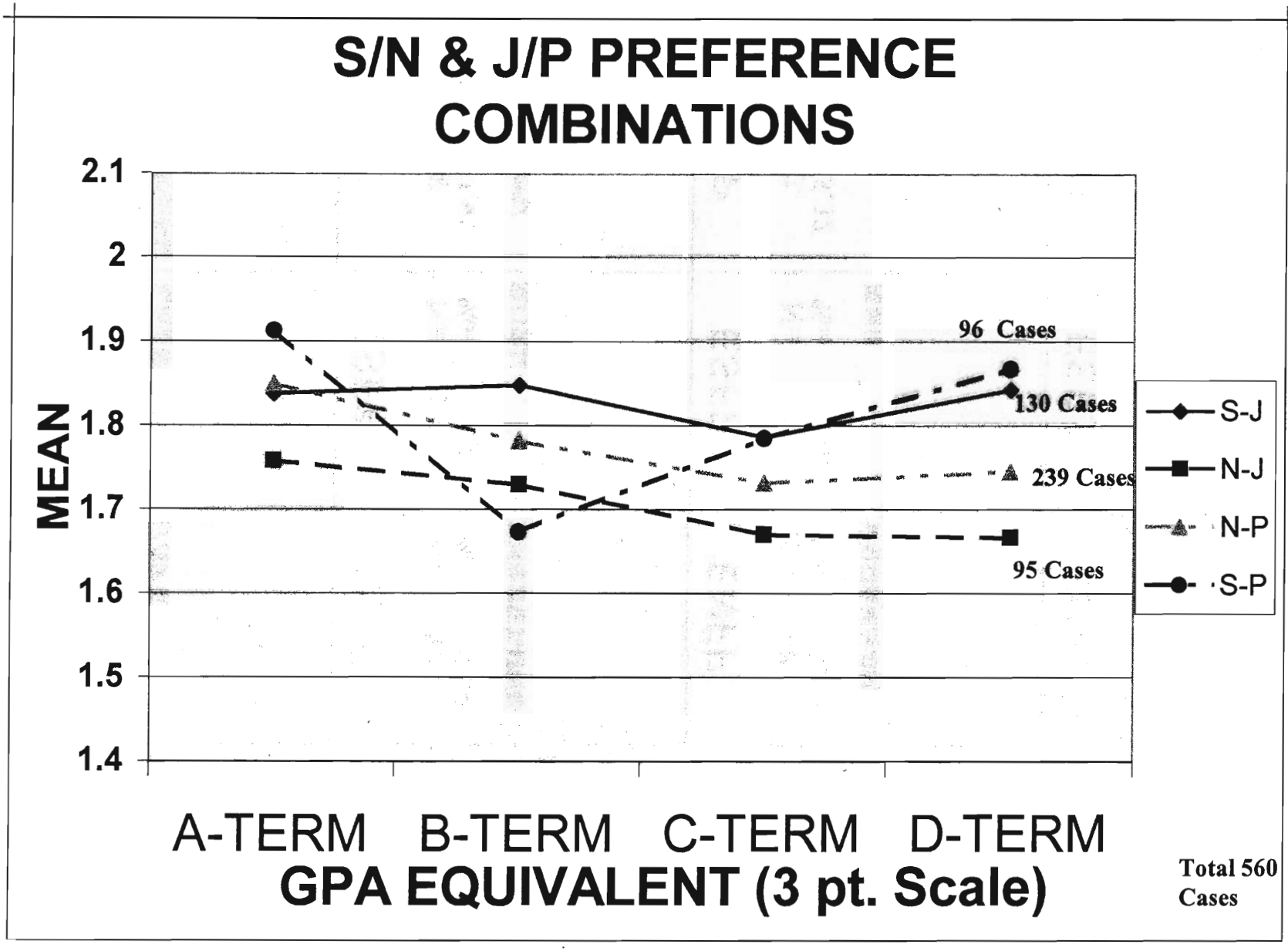
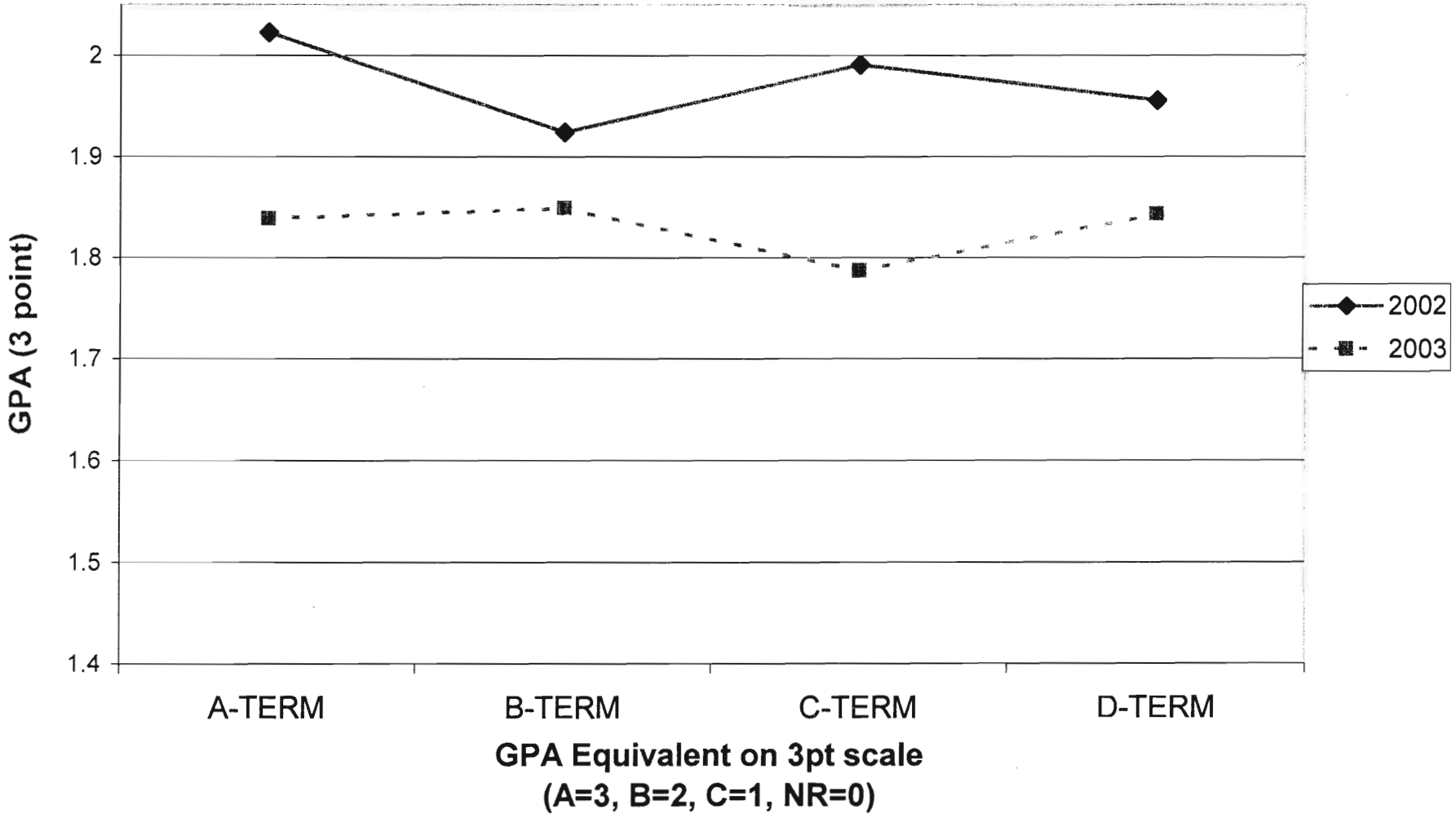
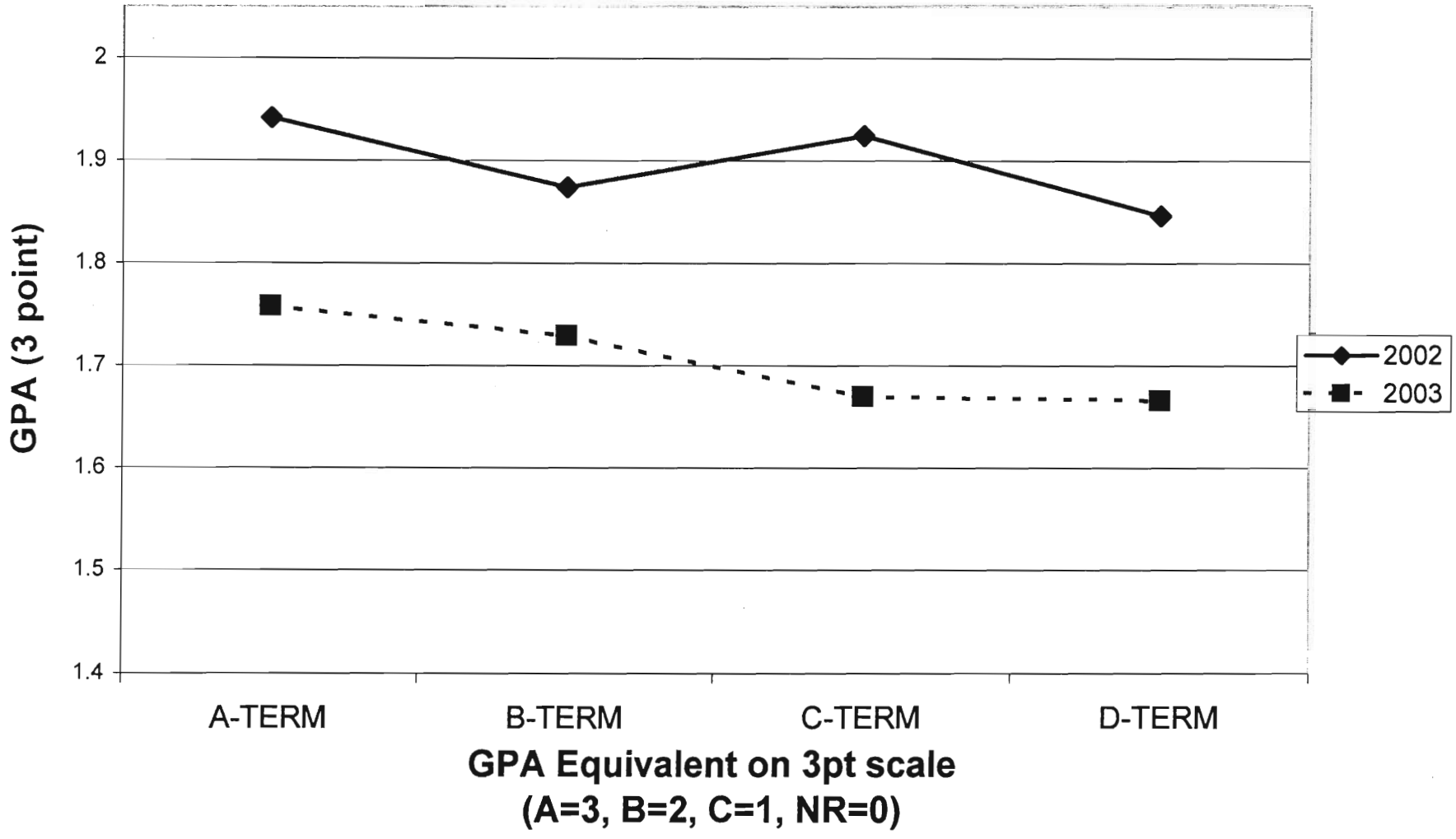

S/N & J/P PREFERENCE COMBINATIONS
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
MEAN
96 Cases
130 Cases
239 Cases
95 Cases
S-J
N-J
N-P
S-P
A-TERM B-TERM C-TERM D-TERM
GPA EQUIVALENT (3 pt. Scale)
Total 560 Cases

TABLE 3b

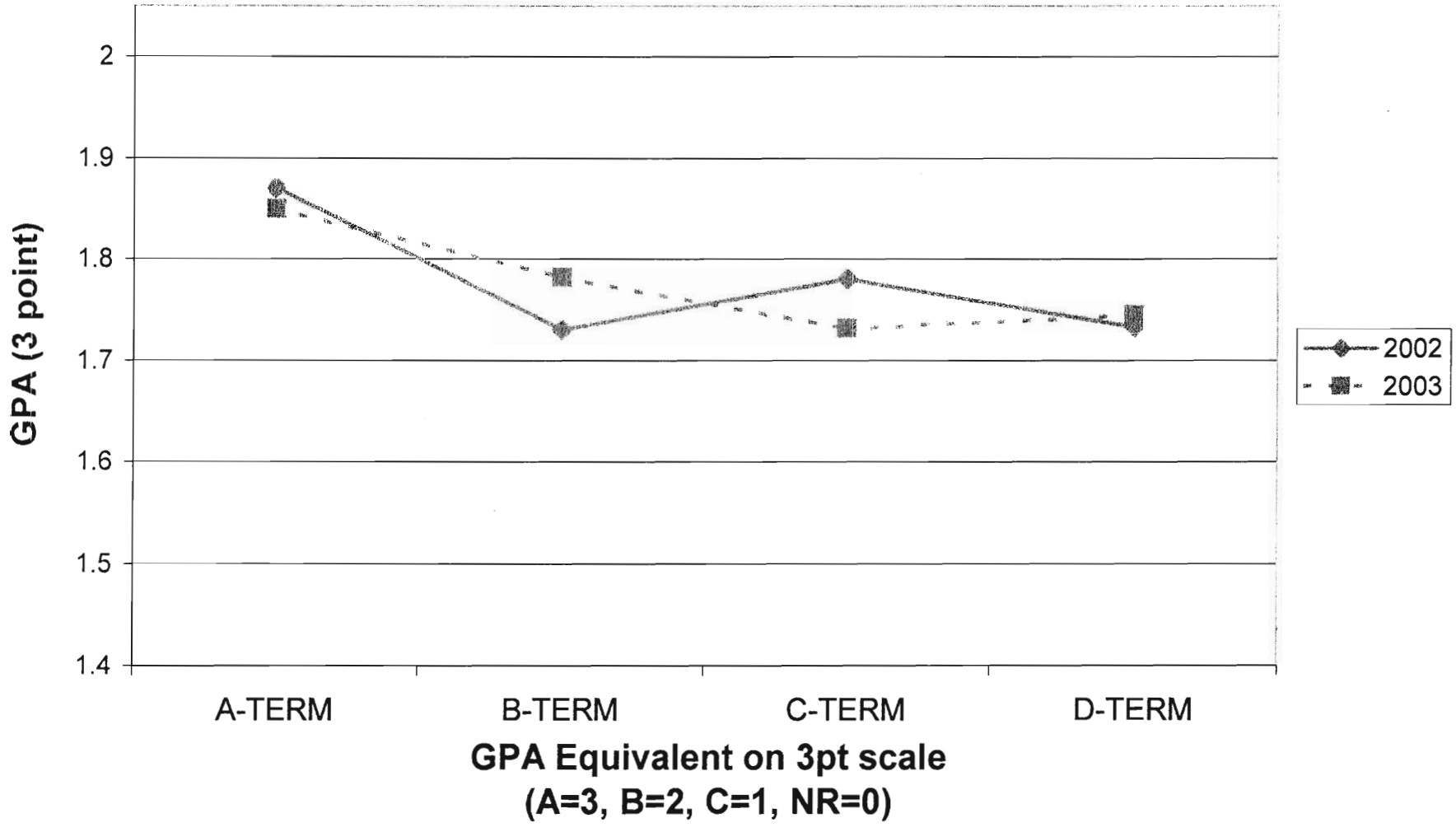
WPI Class of 2002 vs 2003 Freshman Year Grades by MBTI Classification
(Sensing Judging (SJ) Students Only)



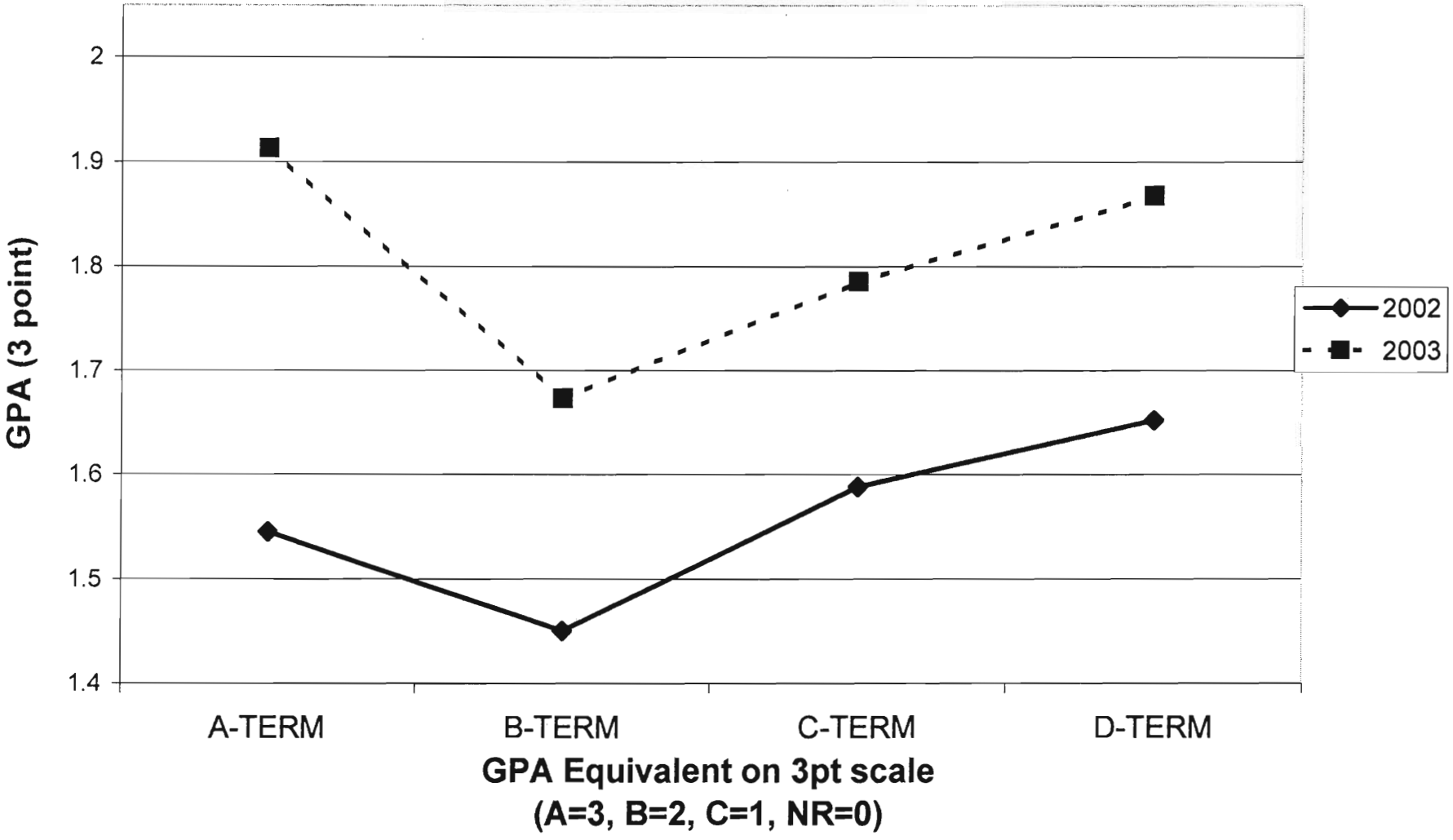
WPI Class of 2002 vs 2003 Freshman Year Grades by MBTI Classification
(Intuitive Judging (NJ) Students Only)



WPI Class of 2002 vs 2003 Freshman Year Grades by MBTI Classification
(Intuitive Perceiving (NP) Students Only)



WPI Class of 2002 vs 2003 Freshman Year Grades by MBTI Classification
(Sensing Perceiving (SP) Students Only)



Appendix B

Graduation Rates (Class of 2001)

MBTI (F) and "On Time" College Graduation (in 4 years) at WPI
WPI Class of 2001 (57% officially graduate "On Time")

SJ (166 cases)	70%	(116)
SP (102 cases)	62%	(63)
NJ (103 cases)	66%	(68)
NP (234 cases)	52%	(122)
Total (605 cases)	61% (369) of those who filled out the MBTI	

The extreme types	
ESFJ	78%
(18)	
ENFP	40%
(47)	

Appendix C

Leading Academic Performance Predictors - The McCormick Study

Summary of Major Class of 2001 Study Findings Relevant to the Proposed Admissions-Advising Office Collaboration

by John Wilkes and Keith McCormick

Based on the study planned by John Wilkes and carried out by two analysis teams under his general direction, it seems that a strong case for the collaborative sharing of student assessment data between Admissions and Academic Advising can be made. The study dealt with the academic performance of the members of the Class of 2001 during their Freshman Year.

Ann Garvin of Student Advising has long sought a way to identify the students most at risk of failing to perform well enough to stay at WPI; i.e. those who fail to demonstrate the promise that led to their admission in the first place. Her task is to identify as early as possible the 70-80 students (about 10% of the class) that would benefit the most from the intervention of her limited resources. She wants to find the people for whom a timely intervention can make all the difference and allow them to go on to succeed at WPI. She does not want to spend her time on the people who will spontaneously recover from a rough start, nor does she wish to temporarily prop up those who will not or really cannot do college level work independently and will leave later anyway.

Ann estimates the target pool to be about half of the bottom 20% of students in terms of end of year grades. While she has found that this group emerges for the most part by the end of B term, many of them are already discouraged and in very serious shape by that time. At least some attrition among people we could have helped is inevitable at that point. The goal is to take action on a group which we were already monitoring as having lead indicators of being at risk before the end of A term, and when their first grades are in hand, to decide which ones to work with at that time. This involves the ability to differentiate grades which are truly ominous from those which are an aberation and transition adjustment problem that needs no concerted response.

The data currently coded by the college during the admissions process, basically SAT 1 and 2 scores and HS class rank; were inadequate to the task of even identifying the bottom 20% of the class. We were recruited to try to find a new set of lead indicators that would be in hand by the 6th week of A term. In effect, this limits us to data collected during the admissions process and New Student Orientation.

Wilkes led a team which took on the task of creating a data base of MBTI and GMCS learning styles data which, were then linked to the freshman grade, SAT and HS rank data by adding in several variables from the HS transcript (for those 300 students who gave consent) and housing data covering dorm, number of roommates and gender mix of residence hall. The teams working with him also verified grade data and created sub-grades by subject. The major goal was to improve the predictive value of the SAT by controlling those variables which were obscuring the predictive value that the SAT is supposed to have (as claimed in the College Board literature).

McCormick was on a separate mission to assess the predictive value of the MBTI and see whether in combination any of the variables in the data set could be used to create an equation that would identify the bulk of the bottom 20% before any WPI grades were turned in. He did an analysis in which all four dimensions of the MBTI (but especially the MBTI Sensing and

Intuition (S-N) Dimension), the Average Math grade, Average Level of Challenge in the Science program (AP, Honors, Non-Honors), Level of Challenge in the Humanities program (based on History or English courses), and Average H.S. Science Course grade were all under consideration. It seems that one can do pretty well predicting Freshman GPA with HS Math Grade, Science Level and History Level in combination. However, the best equation included Sensing-Intuition, HS Science Grade and Science Level. (Obviating the need to look at HS Math or Humanities performance providing you know their learning styles was a shock to all of us.)

How strong a predictor was the "best" equation? Keith found that a sensing student with science grades one standard deviation below average in a program one standard deviation less challenging than average would be 16 times more likely to be at risk by the end of the year than the average member of the class. It cuts the other way too. An Intuitive student with science grades one standard deviation above average in a high school class one standard deviation more challenging than average would be only 6% as likely to be at risk at the end of the year as the average student in the Class of 2001. Between these extremes, the three factors can be used to calculate tradeoffs in order to array the students in terms of probability of being at risk by the end of the year.

Ann is looking for those with several factors placing them at risk; making them long odds against a spontaneous rebound. She also wants more time to work with them before things reach crisis proportions. Hence, Keith was told to err on the side of tagging someone to be monitored who might not turn out to need help rather than to err on the side of missing those who will need help. Again we were leaving to her the problem of identifying who is beyond help and really in over their heads based on a closer look, but not giving up on anyone right away

Using a cutoff designed along these lines, i.e. try not to miss any, throw a broader net, Keith predicted that 68 students would be at risk by the end of the year which he figured was Ann's offices, maximum capacity. Nineteen of those students turned out to be at risk and forty-seven of them were not. (The large majority of the "false positives" were in the 4th rather than the 5th quintile of the class in terms of grades, but still would likely have benefited from Ann's attention). On the other hand, Keith's formula found 19 of the 24 students who really were at risk by year end; missing only 5. This was done with information in hand on the first day of classes. Not bad. He wants to know how many more of those at-risk we are willing to miss to narrow the pool that we monitor through their first year? I am wondering if we should allow the next run of predictions to include the first terms grades at WPI, since no one is talking about intervening before the end of A term.

MBTI data are not available to admissions staff, and the other HS transcript data is available, so his equation based entirely on HS transcript data is still of interest. Further, the only reason Math grades and programs are not predictive of performance among WPI students is that the reputation of the school and admissions screening has narrowed the range of variation in that area for matriculating students that they differ more among themselves in other ways. Statisticians call this a "restriction of range" problem. The same problem is evidenced in our analysis of SAT scores, but there was still value in those numbers if you learned to look at them in the right way.