Project Number: JMW-MAC3

Comparative Analysis of Learning Styles and MCAS Scores in the Worcester and Fitchburg Public School Systems

An Interactive Qualifying Project Report

submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

John L. Stambaugh

Kenneth J. Fountain III

Michael McElhaney

Date: April 27, 2001

Approved:

Professor John M. Wilkes, Advisor

TABLE OF CONTENTS

List of Tables

APPENDIX A

MBTI SAMPLE ANSWER SHEET
MBTI SAMPLE SCORING INFORMATION RESULT REPORT

APPENDIX B

LETTER OF REQUEST FOR DATA FROM FITCHBURG
LETTER OF REQUEST FOR DATA FROM WORCESTER

APPENDIX C

SAT I-ACT™ SCORE COMPARISON CHART

APPENDIX D

FITCHBURG SPSS ANALYSIS RESULTS

APPENDIX E

WORCESTER SPSS ANALYSIS RESULTS

APPENDIX F

ERIC TAPLEY'S FITCHBURG RESULTS FROM ACADEMIC TRACK/SAT/ACT COMPARISON

ABSTRACT

The Massachusetts Comprehensive Assessment System (MCAS) has caused a fiery debate among public school officials due to the fact that many students are at risk of failing this test. The MCAS takes effect starting this year as a high stakes state achievement test. It is supposed to improve standards and provide accountability tied to statewide curriculums. Prior study of the SAT has revealed an "intuitive" advantage using MBTI data from the Worcester and Fitchburg Public school systems. This study seeks to find a similar cognitive bias on the MCAS. Also, if there is a bias, is it the same size and strength involving the same variables? In this socio-economic context, there might be a consequence of the MCAS relating so highly to the SAT in what the exams tell you. Finally, the study seeks to help identify those students "at risk" of failing the MCAS given the findings about what it is correlated with.

ACKNOWLEDGEMENTS

The authors would like to thank Professor John M. Wilkes for his advising and mentoring throughout this project. Also, we would like to thank Patricia Mostue, Ph.D., of the Worcester Public School system, as well as her staff, for her unending advice and in providing us with the data we needed to complete this project. Similarly, we would like to thank Mr. Thomas Lamey and the Fitchburg Public School system, the WPI work-study students who aided in data entry, family, and friends.

1 Introduction

In education, the standards of evaluation are very dependent on the school a student attends. Because of this, it is very difficult to gauge what and how much a student has been taught simply by looking at GPA or even individual grades. To try to solve this problem, the State of Massachusetts has mandated a statewide test. This test, the Massachusetts Comprehensive Assessment System (MCAS), is used as a measuring stick of sorts to determine if the schools systems across the state are effectively teaching a curriculum that meets the state's minimum standard for graduation as defined by the curriculum frameworks published by the State Department of Education.

In an ideal world, the benefit of having this test would be looking at a few numbers can give one an idea of what the student has been taught, and how well the material has been mastered. This information would allow teachers to assess their level of success and adjust their teachings to meet the state's requirements if necessary. It would also allow the state to identify schools that are not routinely meeting the required minimum, and take an appropriate course of action to rectify the situation. However, this test is not just used to help assess the school system in general. The student must pass the test between 10^{th} and 12^{th} grade in order to graduate from high school with a diploma.

The problem that has arisen out of this requirement is the fact that in some schools, especially city schools, large majorities of the student body are expected to fail the test. This can be seen in schools such as Worcester, Fitchburg, and Boston. In the Worcester and Fitchburg public schools, 66% of the students failed the Science and Technology portion of the practice round of the MCAS, and Boston had 71% failing.

These were eighth grade students who took the MCAS in 1999, therefore making them part of the class of 2003, the class that must pass the test as 10^{th} graders, or try again in 11^{th} or 12^{th} grade in order to graduate. When this information is combined with the fact that passing the test is required just to graduate high school, the result not to enter college is potentially devastating. If Worcester could only graduate a third of their students, the school system would be thrown into turmoil.

With this fact known, it becomes apparent that something must be done first to assure that the test is not biased, and if it is not, to help the teachers prepare their students more effectively for this exam. The purpose of this study is to examine the effects of personality type (as determined by the MBTI) on MCAS results. If the MCAS performance of some types of learners is predictably weak and people with certain personality types are found to be at risk of under performing on the test, then a case for bias can be made. Also, a look at the MCAS (an achievement test) itself will be instructive. Our goal is to see if there is a correlation between it and the SAT/PSAT, which are considered aptitude tests. If it is found that there is a correlation between the two tests, then re-evaluation of the MCAS might be necessary. A double check using the ACT Plan (which is administered to 10^{th} graders in the Fitchburg system) will also be undertaken. Though it too is a college admissions indicator, it is considered an achievement test.

This project will combine Worcester and Fitchburg High School students' MCAS scores, personal information, PSAT/SAT data, ACT Plan scores, GPA's, programs of study, and related materials with the results of their respective MBTI (a personality and learning styles indicator) to create a sizeable database. Using this database, we will be

able to test our hypotheses and find out if the expected correlations actually exist. If the correlations are found to exist, then this database of MBTI information will prove very useful to the school system as a whole. It would allow teachers to identify students at risk earlier in their high school programs, and assist them in taking steps that are appropriate to help the students most likely to struggle with the MCAS. Ideally at the completion of this project, an expandable MBTI based database will have been constructed. This database will be given to the participating urban school systems in the hope of their continuing to maintain and update it as part of their efforts to cope with the MCAS crisis unfolding around them.

2 Literature Review

2.1 Myers-Briggs Type Indicator (MBTI)

The Myers-Briggs Type Indicator, which is commonly referred to as "The Indicator", or simply as the MBTI, is a psychological tool developed by Isabel Briggs Myers and her mother, Katharine Cook Briggs. Their goal was to create a test that would reflect C. G. Jung's research on understanding individual differences among people. Jung observed that human behavior is not random, but instead follows identifiable patterns that he felt derive from the structure of the human mind.

Myers and Briggs came upon Jung's work in 1923 and thus, began their own effort, which consisted of nearly twenty years of "type watching." After this extensive research, Myers decided that people might make better decisions concerning their careers and lives if they knew about Jung's types. To that end, Myers began constructing a questionnaire to help assess an individual's specific type. The MBTI was developed through several different iterations (over thirty years) as information was collected on thousands of people. At her death Isabel's forms F and G were the standard research instruments, and included the same core load of 100 items used to classify psychological type.

What the MBTI does is provide a description of a person's preference in a set of four letters that define Psychological Type on four dimensions. Armed with this information about a person's preferred, and sometimes habitual, way of dealing with situations involving data gathering or decision making, then approach to learning and likely reactions to various careers are able to be predicted and understood. It is also a way to study small group dynamics and leadership style. But knowing type does not just

help understand others. It also helps one understand one's own inclinations. The valuable area that MBTI data can be used to illuminate which concerns us here is its ability to identify a person's learning style without implying that one particular type is more able or intelligent than another. The MBTI deals with preferences, not abilities. It does not determine what skills or abilities someone will develop, but it does reveal the best channels by which to reach a learner. What it can do (for a teacher) is indicate circumstances under which given teaching and learning strategies are likely to be most effective by taking individual difference and preferences for handling information and coming to a decision into account.

The four dichotomous MBTI type preference dimensions are as follows:

Extrovert or Introvert

Sensing or Intuition

Thinking or Feeling

Judgment or Perception

By providing answers to multiple-choice questions in the indicator (Appendix B), each person shows that he or she is predisposed to either one or the other of these preferences. Following is a brief description of each of the MBTI's preferences assembled by Charles Martin, Ph.D., in Looking at Type: The Fundamentals, in which he indicated that Jung considered these differences in preference innate, rather than learned. What one learns later is how to function effectively when required to operate in one's less natural or preferred mode.

Extroversion (E): *Key words*:
Outer world, people, action, breadth.
Description:

Extroverts are energized by active involvement in events. They are most excited when they are around people. They often find their understanding of a problem becomes easier if they can talk about it aloud.

Or

Introversion (I): *Key words*:

Inner world, ideas, reflection, depth.

Description:

Introverts are excited when they are involved with ideas, images, and memories that are part of their inner world. They often prefer solitary activities or with one or two others whom they feel an affinity. Introverts truly like the idea of something often better than the something itself.

Sensing (S): *Key words*:

Facts, details, experience, present.

Description:

People with a preference for sensing are more concerned with what is actual, present, current, and real. They are often good at seeing the practical applications of ideas. They learn best when they can first see the pragmatic side of what is being taught.

Or

Intuition (I): *Key words*:

Symbols, pattern, theory, future.

Description:

People who have a preference for intuition would rather gain understanding through insight than hands-on experience. They like concepts, and they learn best when they have an impression of the overall idea first.

Thinking (T): *Key words*:

Impersonal, truth, cool, thought-minded.

Description:

People of this type are concerned with the objective truth in a situation. They act based on truth or principle that is independent of what they or others might want.

Or

Feeling (F): *Key words*:

Personal, value, warm, tender-hearted.

Description:

People of this preference are concerned with whether decisions and actions are worthwhile. They feel that they can make the best decisions by weighing what people care about.

Judgment (J):

Key words:

Structured, decided, organized, scheduled.

Description:

What the judging preference often looks like is that they prefer a planned or orderly way of life, and likes to bring life under control to the degree that is possible.

Or

Perception (P):

Key words:

Adaptable, spontaneous.

Description:

People with the perceiving preference are inclined to a more flexible and spontaneous way of life, like to understand and adapt to the world, and like to stay open for new experiences.

Once a person's predilection for each of the above preferences is found, then a four-letter type pattern (Appendix B), which is descriptive of the overall psychological type, is generated. With this four-letter pattern, it is possible to go into greater depth and analyze all sixteen different type combinations. However, this isn't germane to the fundamental understanding of the MBTI. The ultimate goal of the MBTI is to improve the quality of a person's life by helping them understand themselves and the people around them better in terms of how they prefer to handle data and come to a decision.

2.2 Center for Application of Psychological Type (CAPT)

In 1975, Isabel Myers and Mary McCaulley founded the Center for Application of Psychological Type, known as the CAPT. This non-profit organization was created to continue research on the MBTI and to provide training on type. The CAPT also

publishes and distributes many different publications on type and the application of type in different fields, such as education, personal life, and the workplace. The CAPT's main focus is training people to be qualified in the use of the MBTI. Training takes 3-4 days, and costs roughly $700 - $900 dollars. Training seminars are held 2-3 times a month in a different U.S. City.

The Worcester Public schools became interested in the MBTI starting around 1995-1996, when evidence was gathered by several WPI student teams looking at data from the class of 1997, which connected their PSAT scores to their MBTI types. The three highest and three lowest of the 16 types scored 250 points apart, on average. Worcester made arrangements to profile each class using the MBTI during their 10^{th} grade year, and paid to have a core group of about 50 school psychologists, guidance counselors, teachers, and administrators trained to the point of becoming qualified users of the MBTI. They would, in turn, help others in the future.

2.3 Standardized Testing (Aptitude and Achievement Tests)

Objective testing, the use of objectively scored, paper-and-pencil tests of true-false and multiple-choice items, came into widespread use during World War II, but IQ testing began in World War I as the Army tried to identify the potential officers, assigning the rank and file by democratic means. At that time large numbers of recruits had to be tested for placement in armed forces positions. They needed to decide in a very short time whom would become pilots, cooks, infantrymen, officers, etc. The question of whether it is possible to use objective tests effectively to answer these kinds of questions was found to be a qualified yes. The Army found that within limits, it was possible to use a paper and pencil test to predict a person's aptitude for things they had never even tried

to do before. At least it was better than guessing or asking the recruits what they thought they could do best.

Standardized testing, which added rigidly defined procedures for administering the tests, came into use when the armed forces noticed that individuals from certain areas of the country were scoring much lower on the tests than individuals in other parts of the country. When this was looked into, it was found that the test was given under widely varying conditions. Some were given no instructions for taking the test and some were actually asked to take the test standing up, during their physical examinations. In order to alleviate these issues, standardized procedures to be followed by everyone were developed for administering the tests.

With the success of the armed forces achievement and aptitude tests, (the differences between them will be discussed below), it was only natural that this method would be used in schools to evaluate students. Aptitude tests (such as the SAT) were developed for use as college entrance exams, and achievement tests (such as Iowa) began to be used to measure grade-school performance. Most colleges at the time had a procedure to select eligible students that involved an extensive interview with the student, letters of recommendation and high-school performance information. Most selective colleges and universities also had their own admissions tests, which stressed Latin and the classics. Prep schools had sprung up to prepare students for the test of a given Ivy League university. Harvard proposed a College Board exam to replace this system and help identify working class students or unusual merit worthy of scholarship support. The objective, multiple-choice version of the SAT appeared during World War II. When objective tests started to show up, many colleges wondered if they could use them

effectively to predict a student's college success. The reason that virtually all colleges have gone over to using ACT and SAT test scores in addition to high school records of performance are not solely that they are more cost-effective and efficient than subjective interviews and past grades, although they are. No, they hoped that they would do a better job of predicting school success than the older, less efficient procedures and provide a level playing field (to be fairer than grades). By this measure they are a partial success in that the use of SAT scores and high school GPA is about 10% more predictive of freshman college grades than the use of H.S. GPA alone. Together they correlate 0.4 to 0.5 with freshman grades thus accounting for 16% to 25% of the variance in the early college grades.

The two kinds of tests mentioned above, achievement and aptitude, are fundamentally different. Standardized achievement tests are designed to measure how much a student has already learned about a subject. The results from these tests can be used to help teachers develop programs that suit students' achievement levels in each subject area, such as reading, math, language skills, spelling, or science to take them from where they are to where they need to be.

Standardized aptitude tests are an effort to measure students' ability to learn – their potential, or how well they are likely to do in future school work to which they are not yet exposed, as a general matter. Instead of measuring knowledge of subjects taught in school, these tests measure a broad range of abilities or skills that are considered important to success in school. They can be efforts to measure verbal ability, mechanical ability, creativity, clerical ability, or abstract reasoning. The SAT is designed to tap verbal and mathematical reasoning ability, which together are referred to as a general

(information reasoning) ability by the College Board. The results from aptitude tests help teachers to plan instruction that is appropriate (in terms of pace and presentation) for students, considered to be at different levels of aptitude as well as having different backgrounds. The ACT is different in that it focuses on where one is given the typical high school curriculum offered in the United States.

2.4 Massachusetts Curriculum Framework

Equal opportunity for all students has been a focal point in the media and ongoing public debate over state and national education standards for the better part of 20 years. Even before that, there were racial equality battles being fought, which directly resulted in the integration of all races into the same public schools. This represented an early attempt to offer equality of opportunity to students regardless of their race.

In the past two decades concerns not only over racial equality within schools but the underlying inequalities between men and women and the social classes have become part of the SAT debate. As differences between schools and school districts serving the various ethnic groups and the more and less affluent communities were documented activists went to the legislature, the court and even the streets in search of redress. Many changes in the education system that we have today were efforts to achieve more equal opportunity. Affirmative Action Plans by race were only the most visible of these efforts that came to involve average SAT scores. Reverse discrimination cases began to reach the courts in which SAT scores were the primary evidence that a "stronger" student was discriminated against and did not get a college admission because they were of the wrong (white) race, given efforts to balance admissions alone these lines, especially at publicly funded state colleges and universities. The pressures of public debate involving social

class differences came to a head in 1993 with the resolution of the McDuffy v. Robertson case in the Supreme Judicial Court of the Commonwealth of Massachusetts.

McDuffy v. Robertson represented a case of discrimination associated with social class. It was an outgrowth and extension of a case brought before the court system in 1978, Webby v. Dukakis, in which Webby claimed that townships within the Commonwealth were often unequally funded. Specifically, the poor school districts, possessing little income from property taxes, charged that the wealthier school districts had more money per student, thereby treating students differently depending on what school system they found themselves in. As the State contributed a smaller fraction of funding to school districts than did the local towns, this allowed for a wide range of per student spending, and hence monetary inequality under Commonwealth law.

Over the years various school systems, including poor metropolitan areas such as Worcester and Brockton, joined suit, evolving into the McDuffy case, which reached resolution in 1993. The Court decided in that case that the spending structure for the Commonwealth was inherently unequal and must be changed, with the details of a plan to achieve equity left to the legislature, as well as the allowance for tax increases associated with meeting the equal education standard. This opened the door for the rapid passage of new law to try to fix the system. The new laws, referred to the Education Reform Act of 1993, passed by the legislature and were signed by Governor Weld. They specified a new funding structure where the State contribution to education was raised from 30% to 50% for school districts. This was the central issue in the court case, but there were more sweeping changes in the dictates of the act.

The Supreme Judicial Court of Massachusetts laid down the concept of a Curriculum Framework in its decision when it delineated the seven areas a student would need to be proficient in after passing through the school systems of Massachusetts. The concept was simple. The only way to ensure that the equality sought by fixing financial inequity between school districts was being achieved was to measure the readiness of graduating students for the world somehow. These seven areas of knowledge were used to represent a minimum skill set that could be and should be delivered by an equitable statewide school system without vast discrepancies in resources between districts.

The Reform Act addressed these curriculum areas in a number of ways. The time allotted to the various core curriculum areas was increased greatly by the Act, so as to increase student preparedness in this area. Teacher certification requirements were strengthened, particularly in the sphere of core curriculum familiarity. Charter schools were added to the mix in some districts and while free from many generalized district regulations the core curriculum was set to be stressed therein as well. Lastly, the curriculum frameworks for the various content areas themselves were fleshed out.

The Curriculum Framework for the Commonwealth was set to include the following knowledge or skill areas: English/Language Arts; Mathematics; Science/Technology; History/Social Science; World Languages; Art (Culture); and Health. In order to give the new Curriculum some flesh on those bones, the Commission on the Common Core of Learning was formed, consisting of educators, administrators, and others considered likely to have valuable input. The Commission constructed a rough sketch of what students would need to know after passing through a K-12 school system. This sketch was done by 1994, and the Department of Education accepted the Common

Core of Learning as the model for district education plans in the Commonwealth. By 1995 the first full implementation of the meat of the Framework, specifically in the areas of Mathematics, Science/Technology, Art, World Languages, and Health, was released to the schools.

Using this new set of guidelines was difficult for most districts due to the rapid changes that the Framework underwent in the immediate post 1995 period. The cause for this resided in the first revision document of the Framework itself, which was muddled within the learning processes of the Curriculum Framework rather than the knowledge and skill areas themselves. A 1996 Board of Education Chairman appointee named John Silber (Ex-President of Boston University) assessed that first revision and called for its overhaul using a new investigative commission composed of BOE members and curriculum specialists. These people were intended to focus on the intended outcome of K-12 education rather than the means to get there, as was the stipulation of the McDuffy court decision. Hence, the Curriculum Framework came to be a representation of the minimum standard knowledge base they felt a student graduating from a Massachusetts high school should possess in terms of content.

This commission, working under the direction of Silber (and his successor James Peyser) developed an achievement oriented educational environment. Opposition to this structure was mounted primarily by teachers and learning specialists who overwhelmingly held that focusing on knowledge rather than learning process keeps the students from effectively coping with new problems. Instead it limits them to a given pool of knowledge. This debate remains unresolved, though the court decision and the

Reform Act itself do not explicitly call for or prohibit a focus on learning process rather than content outcomes.

Regardless of which position one takes on the debate over the philosophical nature of the Curriculum Frameworks the result is a seemingly easy way to measure the degree to which equity between the school systems has been achieved. If the Curriculum Framework is based on knowledge areas, it should be possible to define the statewide knowledge minimums in each area of the Framework. Once defined, it becomes a theoretically simple task to test the knowledge of students in these areas and not only assess their level of achievement individually but those of the individual schools and school districts as well. Here enters the Massachusetts Comprehensive Assessment System, a standardized test used to evaluate students and schools for their success in meeting the minimums of the Curriculum Framework areas being tested for. MCAS and its meaning are detailed below. For now assume that the MCAS is a tool designed to illustrate the relative competence of schools and districts in teaching their students the areas targeted by the Curriculum Framework – starting with English Language Arts, Science and Technology, and Mathematics. A first practice run of MCAS covering these fields was administered to 4^{th}, 8^{th}, and 10^{th} graders at the end of the 1998-1999 school year. The Science and Technology test was considered experimental and subject to change at that point so attention focused on English Language Arts and Mathematics..

The importance of the MCAS in terms of the Framework becomes clear immediately. The standard performance expected per district in the Commonwealth reflects the minimum standard set forth in the Curriculum Framework and called for in the Reform Act. Taken to its logical conclusion, by assuring a minimum of knowledge

requisite to survive in the world post graduation to all students in the Commonwealth the Reform Act fixes the inequity described in the McDuffy case. Hence, if the standard is not met, the cause of the failure must be assessed. This is where the fiercest debate resides, specifically in the attribution of blame to an entity in the educational system. If the failure is in the exam, as is contended by opponents of MCAS and the Framework, then the test must be modified to more accurately reflect the Framework. If the failure is in the schools to teach the Curriculum completely, then the reasons must be examined and a remedy applied on the basis of the Reform Act. If the failure resides with the student, that student must be assisted in gaining the required knowledge before graduating from the system.

The interdependence of the Curriculum Framework, MCAS, and the school districts involved forms a tight, complex net of relationships. Sorting this web out involves more than just arguing philosophy. It depends on our understanding of how the Curriculum is assessed, and how students, as learning, thinking individuals, respond to this assessment. This requires knowledge of what MCAS is and it's meaning is to us several years after its creation for these worthy purposes.

2.5 Origins of the Massachusetts Comprehensive Assessment System

The Massachusetts Comprehensive Assessment System, or MCAS, is a standardized test designed to assess student and school performance in meeting the minimum knowledge required within the Curriculum Framework of the Department of Education. As specified in the section above on the Framework, the State needed a means to determine if the school districts across the state were affording students an equitable education. By defining the Curriculum Framework minimal knowledge base as

an equitable education minimum, they were theoretically able to test for a student's preparedness in terms of the Framework as well as test the success of a school in teaching the Framework to the student body as a whole. The test was to be administered to 4th, 8th, and 10th graders as an on-going assessment of their progress in understanding the Core Curriculum, but it was the 10th grade test that they had to pass. If they did not, the school would have 2 years to help them meet this standard in a maximum of 5 more attempts to pass the parts of the tests that were initially failed.

In order to be effective in molding the school districts to the Frameworks, certain responses were keyed to take effect should schools under perform on the test. The school would be studied to see if the cause was structural or due to the qualifications and efforts of personnel then within the school. If the cause was student related, or perhaps if the Curriculum needed more definition the staff in place would deal with it. The under performing students may or may not be flagged for special attention depending on their grade level. State funding may or may not be shifted in order to help assure later success. In brief, many things could result from weak performance on the MCAS by students. It is when the cause is determined to be the school, faculty, or school system that the remedy takes dramatic action such as the State taking over a bankrupt school, administration getting fired and curriculum being imported from elsewhere.

Not all of the 7 Curriculum Framework areas are presently tested on the MCAS, as the implementation of the Framework has been incremental throughout its history. By now – the third round of MCAS testing scheduled for May 2001 – the areas of the Framework incorporated into MCAS currently are Mathematics, Science/Technology, English Language Arts, and History/Social Science. The MCAS evaluation itself tests

students in these areas according to different means of evaluation, specifically a mix of traditional bubble-filled standardized questions as well as write-in portions. The types of questions on the test are multiple choice, found in all sections, as well as short answer write-ins for Mathematics portions of the test. There are open-ended questions for all sections of the test as well, which allow students to offer brief paragraph long answers to a myriad of questions. Lastly, in the English Language Arts section, there are writing prompt questions which allow for extensive student writing opportunities. Some involve creative problems, some are inspirational writing, some involve technical expository, etc.

The whole test comprises an enormous amount of material supposed to center on the content mandated by the Curriculum Framework. Therefore, it represents an incredibly ambitious achievement-style test reminiscent of a hybrid ACT/SAT. The MCAS multiple-choice sections are scored by machine, with all other parts being scored by teams of teachers and scoring professionals. Written answers are scored according to a preset scoring guide that all scoring teams refer to, and a range of points (1-4) are awarded on the basis of the student's answer. The scores awarded on the MCAS correspond to a range of points between 200 and 280, with the performance level cutoffs being every 20 points (hence at 220-239, 240-259, and over 260). Student performance itself is ranked below and within these levels as Failing (219 or less), Needs Improvement, Proficient, and Advanced, with the quality of the student's performance being from lowest to highest. A score of 245 for a given section, for instance, means that the student has performed proficiently on the section. In terms of the Framework, the student has learned that requisite portion of the Framework, and the school has provided that student with an equitable education under the Reform Act.

Different scores therefore mean different, important things. Students who are advanced have mastered the material required, and hence are receiving an equitable education. Students who are proficient, while not masters of the content area, are skilled to the point of having received an equitable education. From the district standpoint, these two upper categories of students are considered total successes, with efforts targeted at assisting the lower two categories in understanding the Framework content.

The students who need improvement have met the minimum standard of the Framework, and hence the district has still succeeded in providing those students with an equitable education, and will not be subjected to punitive action but encouraged to attempt further improvement. Admittedly, the students are performing poorly, and large efforts should be made to improve their understanding. Lastly, the failing students do not understand enough material to meet the minimum standard required by the Framework and are considered unready to function in the broader society and world of work. This last has economic implications for the state that warrant intervention. From the district perspective, this can either be a school or student related problem. If a large number of failing students are in a school or district, it can be assumed that the school or district, for whatever reason, has been unsuccessful in teaching the Core Curriculum. If there are few students, then the failure may lie mostly with the individual student (or a certain type of student) who then needs special attention in order to gain an understanding of that core material.

This point of interpretation over the results is where the greatest debates lie, for particular schools (such as vocational schools), or particular students (such as students who primarily speak Spanish), may have considerable difficulty with the MCAS. This

puts some districts in the difficult position of lacking any obvious remedy to the failing situation their students find themselves in. For the students' part, beginning with the class of 2003 graduating seniors will be required to have passed the MCAS between their 10^{th} and 12^{th} grade years, and for students who have a history of failing tests, the MCAS will prove a daunting requirement to meet.

Into this debate, at this point, enters the idea of learning styles as described in discussion of the MBTI. If the MCAS truly approximates an achievement-style test, then the types which get good grades due to consistent application and succeed at achievement tests will do so for the MCAS. Further, since achievement tests generally are considered to be content, not process oriented, there should be little bias overall toward any one type or certain type of learners. Conversely, if the MCAS is not a measure solely of content, but in some way takes into account process related thinking skills, then the types which perform well on it should appear to more closely resemble those who excel at aptitude tests such as the SAT. Either way this concept of psychological type and its relationship to test performance levels promises to be revealing. The S/N dimension of the MBTI seems to be a real factor in aptitude tests, some achievement tests, and other educational measures used nationally. If the MCAS is designed such that a few groups have the edge in success, then there are definite groups of students, differentiated by their MBTI types, which may be more at risk for failure on the MCAS. Does that mean that the educational system in general or a given school has served them badly, is biased against them, or that the test itself is inherently harder for some type of learners than others? It is this group of "at risk" students that represents an as of yet little considered source of tests failures that cognitively might not reflect on the school or their own efforts. Do these same students

also show up badly on other measures such as grades, SAT performance, etc? The question of whether a "one size fits all" test is the appropriate way to assess educational outcomes is illuminated by the use of a learning style indicator with implications for career choice.

2.6 MCAS and the MBTI

As mentioned above, the test was initially given to fourth, eighth, and tenth graders. The possibility that certain areas might be tested at different grade levels to spread the burden was raised after the 2 experimental rounds of 1999 and 2000. To avoid the fatigue factor and let the students and teachers focus their preparations, there is likely to be a change in policy. For instance, the English Language Arts section has been administered to some third and seventh graders. The Mathematics section has been taken by sixth graders. Finally, the Science and Technology portion has been given to fifth graders, and will be omitted for fourth graders in the 2001 MCAS (http//www.doe.edu.mass/ mcas). The reason for giving the examination to the fourth and eighth graders was to assess the performance of the school district in preparing the students according to the core curriculum before the students themselves face their high stakes moment at the end of the 10^{th} grade. The tenth grade examination focuses more on the individual student, and measures his or her ability to pass the MCAS in order to graduate from high school.

The format of most of the questions that appear on the MCAS are similar to those that one would find on the SAT (Scholastic Aptitude Test) or PSAT (Preliminary Scholastic Aptitude Test). However, the MCAS designers stressed the need to balance multiple-choice with open-ended items and focus on achievement rather than reasoning.

Hence, there are multiple-choice questions (all subject areas), short-answer questions (Mathematics only), open-response questions (all subject areas), and writing prompts (English Language Arts only). The results are reported in one of four ways for each section and are used for several evaluations. These include using the results to make improvements in teaching and learning, to show school and district accountability, and individual student accountability. The four modes of reporting results from the MCAS and their respective scores are listed below with full descriptions:

> **ADVANCED (260-280):** Student demonstrates a comprehensive and in-depth understanding of rigorous subject matter, and provides sophisticated solutions to complex problems.
>
> **PROFICIENT (240-259):** Student demonstrates a solid understanding of challenging subject matter and solves a wide variety of problems.
>
> **NEEDS IMPROVEMENT (220-239):** Student demonstrates a partial understanding of subject matter and solves some simple problems.
>
> **FAILING (200-219):** Student demonstrates a minimal understanding of subject matter and does not even solve simple problems.

There has been some indication that previous tests such as the SAT and PSAT, which are aptitude tests, are easier for students of some learning styles as measured by the MBTI (Myers-Briggs Type Indicator) than others. However, so far we have found no research on how learning style can affect a student's performance on the MCAS, which is viewed as an assessment of achievement. The only basis for speculation is a comment based on research by Isabel Myers in the early 1960's at a number of suburban Philadelphia high schools. She reported then that intuitive students, especially INP's would be expected to excel at IQ style tests and seek out college preparation programs but that they might not do as well as the sensing and judging students when average grades based on day to day

performance were assessed. This was her way of stressing that the INP students were not really smarter, just different in learning style as compared to the ESJ students. If learning style does affect MCAS performance, which personality types are most at risk of failure? In addition, does the MCAS in fact test achievement, and not aptitude? Will the school systems' grades be a better predictor of success on the MCAS than scores on an aptitude test, like the PSAT?

2.6.1 Public Debate

An important aspect of studying the MCAS that cannot be pushed aside by the analysis of the test itself is the public's opinion of the exam. Throughout the recent history of education reform in all states, public arguments for and against new policies have raged heated for decades, often reaching no resolution. MCAS thus far has proved itself to be good fodder for such arguments, with much talk and even mobilization of grass-roots organizations representing thousands of people who want their voices heard. This includes the Massachusetts Teachers Association.

One problem that has arisen out of the debate over the MCAS comes from the location of the primary opponents. A large percentage of the people that want to get rid of (or lower the importance of) the MCAS are in the suburbs, where the schools are generally in better shape than in the inner cities and the state is offering few new resources despite the threat of higher taxes. Public support for MCAS is strongest in the state's largest cities, where students "trapped" in failing schools and have the most to gain from a successful shift to higher standards. Unfortunately for MCAS supporters, there are more suburban districts then urban ones, so the suburbs have more of a voice than the urban districts. Further, the urban teachers are not in agreement with the urban

voters and urban parents who now have a fear that their child, lacking a high school diploma, will not be able to go to college in a state university.

Besides the fact that passing the MCAS will eventually be mandatory to graduate from high school, the fact that it makes teachers radically change their curriculum has caused many people to shy away from it. Protesters say the MCAS has caused a dramatic shift in teaching priorities. Instead of teaching general themes and ideas, schools are trying to cram as many individual facts into students' heads as possible in order to increase test scores. This has caused many teachers to oppose the MCAS on philosophical grounds and in terms of educational practice. They have gone so far as boycotting it and refusing to administer it to their students. But, supporting the MCAS are many of the leaders of the state, and they feel that the MCAS is needed to put teeth into the state's education reform initiative and provide accountability for the use of state funds.

A recent addition to the debate has been the issue of retaking the test if a student fails it the first time. It was recently decided that students who fail the MCAS exam on their first try will have at least five more chances to pass it and can take retests only in areas they did not pass that don't include the toughest questions that would simply tire and discourage them. This is according to a plan approved by the state Board of Education. The plan is an effort to prepare for very high failure rates: forty-five percent of 10th-graders in the state failed English and 34 percent failed math on the 2000 test. Some say allowing retests that don't include the most difficult questions represents a retreat from high standards that the MCAS was intended to create, but most

administrators feel that it is pointless to subject this population to the additional stress of a long exam involving a massive fatigue factor.

One problem that may arise because of the MCAS is policies to encourage the privatization of public education. Because there is already dissatisfaction with public schools, very poor MCAS scores will further discredit the urban public schools. If this happens, the public humiliation may set the stage for them to be taken over by for-profit charter schools, and many feel that this is a problem. The Charter schools will be in a position to pick and choose their student bodies and then take funds from the public schools commeasurate with each pupil they take. In the end, the public schools will have fewer resources to deal with the toughest remedial cases.

Interestingly enough many of the districts across the Commonwealth which originally supported education reform, and the Reform Act of 1993 (with its provision for a representative minimum curriculum that would represent equal education in all districts), now vehemently oppose MCAS. From school district and individual school standpoints the MCAS is increasingly viewed as an invalid measure of what the students know, or too different from the curriculum already in place for the school to rapidly adapt to it. Some schools are ill equipped to meet the minimums on all parts of the test, and some students who have special needs or other difficulties may not pass the test at all given its present level of difficulty. This would exclude them from higher education programs designed to accommodate them. Toss into this mix the factors surrounding the upcoming requirement for 10^{th} graders to pass the MCAS prior to graduation and all the elements are in place for a lively public debate.

A prime opponent to the MCAS for some time has been the vocational/technical schools scattered throughout the state. These schools are focused upon teaching their students the skills of a trade rather than core academics. The students in these schools often are often not academically inclined, and typically do no go on to college. Instead, they go on to become carpenters, electricians, and other such blue-collar professionals. Indeed, employment rates for graduates of these schools are incredibly high compared even to many universities, and drop out rates for the vocational schools tend to be low. They bridle at the idea that they do not prepare their students with the skills they need to live in our society. However, with the pressure now on to pass the MCAS in order to graduate, these schools face the specter of high dropout rates with disillusioned students leaving en masse for public schools geared to MCAS preparation after their students start failing the MCAS and lose the option to go on to higher education beyond community college.

The skill range of the vocational students is often focused on the trade that the student is pursuing for all 4 years of their high school education. With half the attention given to academics that would be found in a normal public high school, the students are especially likely to fail the MCAS. Some schools are seeking exemptions on this basis from the MCAS, while others are attempting to adapt. These adapting schools are toying with wide ranging solutions, like expanding the school day to allow for more academics. Some support restricting trade preparation to the last two years of high school so that the first two years of the vocational student's time will be spent on academics associated with passing the MCAS exam. Many call for changes in the MCAS, including additions of material, that would better measure the vocational students' practical and tangible skills

and those of other students who are better with their hands than their heads. David Driscoll, the Education Commissioner for the Commonwealth, said that the "...standard is going to be the same for all, so we have an expectation that vocational schools will get these kids up over the bar, which I think has been properly set for graduation." Despite complaints by vocational schools, as per the concept of equity described by the court case which originated the Reform Act, no one school or district gets special treatments in failing to provide the State mandated minimum standard of public education, just as no school may be denied the opportunity to do more than is required.

This seemingly immovable position of equity in testing has had a variety of effects, mostly forcing schools all over the State, of all socioeconomic backgrounds, to review the curriculum and scramble to meet the common testing requirements. Despite an overwhelming flow of bad news regularly coming from the assessment results, there has been good news from some districts on occasion. For instance in 1999, there were 23 schools that did not have one failing student on the MCAS, quite an amazing feat given the admitted challenge presented by the test. Many schools showed marked improvement, some showed a little nudge of improvement. Meanwhile, the overall numbers remain somewhat dismal, with tens of thousands of students likely to fail the 10^{th} grade test this year (and hence need to retake it to graduate). Still, many observers are hopeful. They think the struggling districts can improve greatly in just a few years, and will be much the better for it while many currently successful districts may have difficulty hitting or maintaining Proficient or Advanced performance for years on end. Complaisancy is on the decline and attention has been focused on public education and resources provided to reform weak schools.

In any case, the performance from 1998 to the present on the test offers some insight into trends based on school-wide and system-wide scores and rankings about the validity of the MCAS. However, the data are insufficient to really make assertions. Given that this is still only the 6th year of incorporation of the Curriculum Framework into the school system, and many districts have yet to integrate much of the material focused on testing into their classes it is still very likely that the MCAS will not yet be testing the students on what they were actually taught since 4th grade. One thing that is certain is that the MCAS and the Curriculum Framework minimums do not represent the maximum that will be demanded of students or districts.

As part of the Reform Act, districts are required to pursue ever higher standards of evaluation in the hope to make them more saleable in the State, National, and world markets as skilled graduates. As such, the minimum performance expected on the MCAS has been set to be an upward moving bar, set to new heights as old heights are reached and surpassed. The drive toward producing more skilled graduates is supposed to compensate for the risk of defining a statewide minimum that many students surpass, and would regress toward, should no driving force be given to them to achieve. Despite low scores in many areas of the State, Governor Cellucci insists that the State will "...hold the line on standards; I think that's been very clear." The fear that schools will focus on the low end students to such a degree that they don't challenge the high end students and serve the gifted young people increasingly poorly is very real.

With some schools failing overwhelmingly, and the bar set to move ever higher in the near future, many schools have come up with creative ways to deal with the exam. Many schools have restructured their curriculum to cover the new material at an earlier

time, swapping the years material is taught or taking material not on the test out altogether. This approach dismays many teachers who see it as disrupting the fabric of the 4-year high school educational experience they are trying to provide for their students. Others see it as beneficial by reducing repetition in non-critical areas and focusing teachers in different grades to systematically decide who is supposed to cover what.

Some schools now offer tutoring, which is either during school hours or afterward, in areas specific to the MCAS. The concept is all to often to teach to the test, rather than to master the Core Curriculum itself. Still in most places this represents rising standards and a more serious approach. Other schools have stressed team teaching, using two or more teachers to illustrate the unity and connections between subjects. Educational integration that is sometimes left for the latter years of some universities is being incorporated at the high school level, allowing for algebra and geometry to be interwoven in an effort to develop the same thinking skills necessary to achieve on Science and Social Studies portions of the MCAS. The curriculum guidelines stress developing a mastery of processes of "inquiry" as much as endorsing the coverage of any particular concept on subject matter.

In the end the people of the Commonwealth will have to be the ones to decide what stays and what goes, for the MCAS and the Curriculum Frameworks. Lobbying groups are being formed to sway opinion at the State House to get the legislature to back down from using the MCAS as an individual graduation requirement, but rather to stress educational system level accountability. Efforts are being made to expand the numerous special education exemptions already allowed under the Department of Education's

assessment rules. Still, other groups are arguing vehemently in support of MCAS, that despite its failings the test fulfills a useful function for the Commonwealth.

What may be necessary is for the public to step away from MCAS for a time, to step back and see if MCAS is a symptom rather than a cause of their concern over their children as students. The MCAS is an assessment of the equitable education provided to students across the State in terms of the Curriculum Framework, a Framework mandated by the Reform Act. The Act itself was a response to the court decision in favor of school districts which called for sweeping changes to make the education system across the State more equitable and homogenized. So perhaps it is time for the public debate to concern itself with redressing a central question of education rather than perhaps compounding a folly that in another 15 years will be resolved by another lawsuit and a new act calling for reforms.

2.7 Scholastic Aptitude Test (SAT)

Adapted from: The Leicester SAT Experience. Written by Gerard Mangenot and Mario Tongol. April 30, 1998.

The Scholastic Aptitude Test was created in 1926 by the College Entrance Examination Board. The exam was aimed at determining high school students' preparedness for college. It was comprised of several lengthy essays. The test was taken by a small percentage of college-bound students and was used by a few college admissions staffs primarily to identify working class students worthy of scholarship support to attend Ivy League Colleges. It was normalized using preparatory school graduates who, on the whole, did well at Harvard and its sister schools. Since its creation, the SAT has changed considerably. Standardized multiple-choice questions

were added to facilitate grading. Eventually, the essays were dropped altogether leaving only the standardized portion which was influenced by the IQ testing movement of the period (Wirtz et al. 1977).

The Educational Testing Service, created in 1947 by the College Board, is currently responsible for administering the SAT. Many colleges and universities require applicants to submit an SAT or ACT score for consideration. Therefore, most high schools "require", or strongly urge, college-bound students to take the SAT. The test is comprised of a Mathematical and a Verbal part. The score range for both parts is on a scale of 200 to 800. The Mathematical portion is geared at the material taught typically in grades one to nine or up to and including geometry and algebra. The Verbal part determines a student's word relationship and reading skills (Crouse and Trusheim 1979).

High school students as well as college admissions staffs depend on the SAT to help them judge whether an applicant will fit in and flourish at a college. It starts when prospective students compare their SAT scores with the average published SAT scores for a certain college to decide whether to bother to apply. College admission staff members use the SAT to determine who to accept and decline for admissions, but it is not their prime decision guide in most cases. High school program difficulty and grades come first, then SAT and class rank. The SAT is valued as a control on local high school grading standards and the difficulty of the courses offered. However, it is in the "benchmark" decisions that the influence of the SAT is more subtle and insidious.

Those with SAT scores that are very high do not have the rest of their file examined very closely. One reader is considered enough and there is no committee review. Those with scores below a given level will not have the rest of their file

reviewed at all or are deemed too great a risk to be eligible for scholarship support. In combination, SAT scores can have a considerable impact in determining access to competitive schools where the admissions office is looking for a weakness to use as grounds for exclusion. Even where a full committer review of the whole file is undertaken the focus of discussion is often discrepancies between the high school grades and SAT score construed as evidence of "over" and "under" achievement.

However, the whole idea of trying to predict future college performance is both necessary and controversial. Students who have worked hard and done well in high school are expecting their reward in the form of a range of options and possibly scholarship support. For them, the SAT looms as a potential threat to their investment. Other students, who have not applied themselves consistently, can make up for it in one day on one test. It could be a "last chance" to showcase their capabilities, as aptitude or promise.

The reason for the inclusion of SAT background and data in this study is to examine whether the MCAS, like the SAT before it, can be shown to have elements of aptitude as well as achievement. Previous studies, which will be mentioned later, have shown that the SAT, which is viewed as an aptitude test, also has elements of achievement affecting scores and the combination favors certain types of learners as identified by the MBTI. Therefore, core curriculum and course preparation can account for up to a 400-point difference in SAT scores. However, comparably prepared students still differ in average by learning style in ways that are predictable and meaningful given Jungian theory, the basis for the MBTI.

2.8 ACT Assessment™

The ACT Assessment™, or "A-C-T", is a national college admission examination that sometimes replaces the SAT, or is taken in combination with the SAT by college-bound students. This SAT competition is based in Iowa and the ACT is more popular in the mid-West and south than the SAT. ACT results are accepted by almost all U.S. colleges and universities in lieu of the SAT anywhere in the country, but it is the underdog primarily in the Northeast. As noted earlier, the ACT is curriculum based, designed to cover what is taught in typical high schools and text books. By contrast, the SAT, which is an aptitude based test, is designed so as to avoid testing content, or what one has been taught (and hopefully learned) at the level of facts. It is an assessment of reasoning ability, so as not to favor students from stronger secondary schools. The ACT is "designed to assess high school students' general educational development and their ability to complete college-level work" (http://www.act.org/aap/ index.html). The ACT covers four skill areas: English, Mathematics, Reading, and Science Reasoning. The ACT-Plan is given as a 10^{th} grade test and it includes 215 multiple-choice questions and takes approximately 3 hours and 20 minutes to complete, with actual testing time being 2 hours and 55 minutes. One more feature of the examination is that it contains a questionnaire which provides the student with valuable information for career and educational planning. The PSAT is normally administered early in 11^{th} grade. The ACT or SAT is then taken primarily in late 11^{th} or early 12^{th} grade.

The ACT is administered on five national test dates, which occur in October, December, February, April, and June. Some states also offer the test in September. The registration fee for the ACT is $23, and $26 in Florida. The ACT Assessment™ is

prepared according to three standards: Standards for Educational and Psychological Testing, Code of Professional Responsibilities in Educational Measurement, and Code of Fair Testing Practices in Education. The test has been administered since the fall of 1959, and it has been offered in all 50 U.S. states since 1960.

The test is scored differently from both the MCAS and the SAT/PSAT. The highest possible score on the ACT Assessment™ examination is 36, and the odds of this happening are 1 in 8,131. A conversion chart between ACT and SAT scores can be found in Appendix C. It is of interest to us here because of its greater philosophical and content area alignment with the MCAS than the SAT, and the fact that the Fitchburg Public schools encourage its administrations, and are an area testing center.

2.9 Overview of Statistics

Statistics, the mixture of art and science behind determining the relationships between different individuals in groups, is the backbone of research into learning styles as they apply to academic testing. A detailed synopsis of statistical analysis is beyond the scope of this report, though a brief description of the concepts employed in the analysis of data in this project would go far to assist the untrained reader as to the meaning of the numbers, and results, we cite. To that end, a brief overview of the major concepts used in this report follows.

Statistical analysis is based on the scientific application of hypothesis testing to large data sets. Hypothesis testing itself is somewhat esoteric, with the accurate description of the nature of the theory which one wishes to test being difficult to develop. Suffice it to say that with statistics one tries to disprove a hypothesis, and typically the hypothesis is that there is no connection between 2 variables in the data set we are

studying. For instance, if one wished to see if the men in a certain school are taller than the women there, one would compare the heights of as many men and women as possible. The hypothesis would be the null hypothesis that there is no difference between the heights. By disproving the null hypothesis, one would demonstrate that there is a difference, and hence can go on to describe that difference. This method of disproving a hypothesis that claims there are no special features in the data refers to testing the null, or no-difference, hypothesis.

Most hypothesis testing demands that a certain amount of assuredness be achieved in reaching a conclusion. This is represented by the p-value for the hypothesis test, a decimal fraction of the total number of times that the result of our hypothesis test would result in the null hypothesis. Typically, this value is set low, perhaps 0.05 or lower. This value means that 5%, or 1 in every 20 data items, or less behave as in the null hypothesis. If less than 5% of the data conform to the no-difference rule, it is typically accepted that the difference is real, and the differences found in the data is statistically significant. Significance is merely a measure of confidence that the finding we see bears on the hypothesis we are testing. If our p-value was 0.2, then 20% of the data would conform to the null hypothesis but we would consider it invalid even if the test resulted in that whole 20%. Fundamentally, this means that we can set the bar for rejecting the null hypothesis as high or low as we want, but we need it low enough to matter in the real world.

What is done with the data at hand is dependent on the nature of the data at hand. Sometimes one has continuously increasing or decreasing test scores, for instance. Other times a variable consists of a list of distinct, discrete values that may be possible for a

trait, such as people being male of female. Statistics work to compare these continuous and discrete data values in a variety of ways. Different methods are employed to study numerical, non-numerical, ordered and non-ordered, and other types of data. Generally, speaking though all types of analysis begin with getting related data in the first place.

Data collection is often a difficult process especially when large numbers of people or large organizations are involved. The problem is that data must not only be gathered accurately, but it is necessary to gather data from large numbers of cases that were selected in a random fashion. The large numbers are necessary because the more individuals in a sample, the more certain one can be that the null hypothesis has been disproved for any given difference on pattern found in the data. If the sample size is 100, then a p-value of 0.05 may be reached easily if there is a sampling error. Let us say that the sampling is not very random, and we end up with 6 data cases which conform to the null hypothesis. Under the 100 size sample this says that the null hypothesis is true. But let us say that we have a sample size of 200, and this time, even with our not quite random sampling we pick up those same 6 conformists and 2 more, we have 8 null hypothesis data points. For the sample size of 200 and p-value of 0.05, then up to 10 data points may conform to the null hypothesis and yet we still reject it. What this means is that for larger and larger samples, we reduce the probability that an analysis will be so mistaken that reverses the null hypothesis test. Typically, statisticians strive for samples sizes near 1000 to ensure a reliable result of their analysis, but they are all forced to work with what they can get, and gathering even 200 complete cases can often be difficult.

With this concept of p-values and hypothesis testing in mind the present statistical analysis proceeds along one of two major paths, depending on what the data look like.

By "look like", statisticians mean if they conform to some pattern that is easily recognizable. For instance, for many tests in school students' scores are scaled so as to make a bell curve, or normal distribution. This type of distribution indicates behavior of the sample tending toward an average value, or put another way, there are fewer outstanding students, and fewer total failures, than there are average students. Statisticians typically employ some initial tests to describe what distribution of the data look like.

These tests are often referred to as descriptive statistics, and include some common tools such as finding the sample's mean, median, and mode. More importantly, measures are based on the actual spread, or range, of the data values in the sample. Clustering of the data according to its ranking by quarters (4 equal groups) is referred to as interquartile range, or the location of 25% and 75% of the data. That statistic gives an idea of general spread. A better idea is provided by the variance, or variation within the range of the sample. The variance is the sum of the distance of the data values in the sample from the mean. This variation can be compared between different samples to determine if their spread is greater or lesser than one another or so great in absolute terms that it really can't be described very well by the summary statistic. The standard deviation is the square root of the variance, a value illustrating the typical deviation of individual cases from the sample mean. Tied to this is the sample's standard error, the deviation of multiple samples from the same population from the sample mean. Together, this substantial list of statistics provides an idea of the spread of the data, the shape of its distribution, and how accurate any sampling measure has been.

Other descriptive statistics are commonly used, such as the confidence interval, the coefficient of variation (comparison of variation between samples), and skew of the data toward one extreme. Together, from these statistics, one gets a picture necessary to understand basics about the behavior of different samples, if one knows enough, and can be sure of some features such as normality on given variables of the distribution, it is possible to proceed with parametric statistical tests. If one is uncertain about some traits of the data, or if the data are non-normal, it is best to err on the conservative side and employ ranked, or non-parametric, statistical tests. For this project, the important analyses will involve Pearson correlation coefficient and standard regression techniques based on them.

Pearson correlation is the process of estimating the amount of variation in one variable, or set of data, due to the size and variation in another variable within the same sample. The meaning of the data is interpreted from the resulting r value of the test, from -1 to 0, and to 1. This corresponds to perfect negative correlation to perfect positive correlation, with no correlation at all being in the middle at the zero point. A Pearson correlation value of 0.91 results in 83% of the variation in the size of the tested variable is "explained by" (can be predicted in terms of) variation in the other variable. If the p-value is very low, for instance 0.005, then the correlation explains 83% of the variation, and the chance that there really is no difference in the larger population despite the strong difference (correlation) in the sample at hand is only 5 chances out of 1,000.

The second major analytical technique used in this study is standard regression. Standard linear regression tests to determine the strength of the prediction value of a cause, or x, variable in determining the affected variable, or y. The strength of the

relationship is often reported in terms of Pearson correlation values, and the slope and intercept for the best fit line through the x-y paired data provides the predictive mechanism for using the regression technique. This test assumes at the very least complementary variances for the tested variables, a linear fit to the data, and perfect determination of the x values. If the assumptions hold, this tool becomes incredibly useful in going a step beyond just describing patterns and into determining relationships. Its major advantage over the Pearson correlation coefficients on which it is based is the capacity to use multiple independent variables to predict values in a single dependent variable. It is essentially a multiple correlation coefficient.

The statistical analyses used in this report conform to the descriptions above, with some additional tools employed in the study of variance. What the reader may understand about the statistical process employed is that the data were studied, described using basic techniques, their variation studied using a variety of comparative methods, and finally major conclusions drawn on the basis of correlation and regression values. For a further treatment of the methods, or a deeper understanding of the meaning of the numbers, consult any introductory statistics text. We used "Applied Statistics for Engineers and Scientists", by Petruccelli, et al, published by Prentice Hall, Upper Saddle River, NJ, in 1999.

2.10 Previous Studies and Findings

Adapted from: The Leicester SAT Experience. Written by Gerard Mangenot and Mario Tongol. April 30, 1998.

2.10.1 Previous MCAS Studies

The MCAS is still new. Hence, published studies on the MCAS have been very scarce despite the huge effort in the state to disseminate the actual results to the schools in a useful form. In fact, there has been no data to present in this report from any previous WPI study. Newspaper reports are about all that we have seen and these present many numbers and system averages, but do not go into deep analysis of the data. The project being proposed in this document is one of the first, if not the first, study to create a database containing MCAS and MBTI data. Hence, this study will probably be the first to see if learning style as determined by the MBTI can help one identify students who are at risk to under perform on the MCAS, as it did on the PSAT. Finally, this study is also the first, to our knowledge, to make a comparison between an achievement and aptitude test, of which the MCAS claims to be the former, by putting the SAT results in the same file with the MCAS results. However, there have been extensive studies done on standardized tests, such as the SAT and PSAT, in relation to learning style. These studies will be summarized in the following sections since the current proposed project also deals with other standardized tests other than the MCAS. The MBTI manual includes correlations of the MBTI dimensions with various standardized tests (including the PSAT and SAT) as well.

2.10.2 Previous ACT Studies

A statistical analysis done by McElhaney (1998) showed that the correlation between MCAS scores and Plan ACT scores were reasonably strong (0.5-0.6). This makes sense due to the fact that both the ACT and MCAS are achievement tests. The

data for this analysis was collected from 10th graders in the Fitchburg school system that took the MCAS in the 1998-1999 school year. Below is a table of the correlation coefficients between sections of the Plan ACT and sections of the MCAS for these students. Notice that most of the coefficients are above 0.5, which means that the correlations are strong. Therefore, a student's performance on one of the tests (i.e. the MCAS) can be used to predict the same student's performance on the other (i.e. the ACT) at levels considerably better than chance. A 0.5 correlation means that 25% of the variance in the dependent variable (MCAS) is explained by the independent variable (ACT). For a 0.6 correlation, the percent of variance explained is a robust 36%.

Table 1: Correlation coefficients between the Plan ACT and MCAS

	MCAS-10 English Scaled	**MCAS-10 Math Scaled**	**MCAS-10 Science Scaled**
ACT-Plan English	0.613	0.541	0.535
ACT-Plan Reading	0.564	0.442	0.430
ACT-Plan Rhetorical	0.554	0.507	0.522
ACT-Plan Algebra	0.535	0.684	0.588
ACT-Plan Geometry	0.374	0.574	0.397
ACT-Plan Math	0.500	0.673	0.561
ACT-Plan Science	0.506	0.531	0.511
ACT-Plan Usage	0.611	0.512	0.500
ACT-Plan Composite	0.639	0.633	0.591

Further, other research has reported even higher levels of correlation between students' PSAT and ACT Plan scores, in the 0.7-0.8 range. This research was presented in an oral

presentation by Eric Tapley in Gainsville, FL, in March of 2000 and is fully documented in his IQP report submitted on March 13, 2001, but was not available to us at the time our analysis was completed.

2.10.3 Previous SAT Studies

There have been several studies comparing personality type, cognitive learning style, and performance on standardized testing conducted by Worcester Polytechnic Institute students in area schools during the past few years. Each of these studies resulted in strong evidence to support the theory that test scores on the PSAT or SAT vary in ways that favor certain types of learners. Though studies have been performed here since 1995, the majority of our hypotheses will be based on the study by John Pieper, which concluded in early 1997. Pieper's project generated by far the largest, most inclusive (PSAT) data set of all. When Keith McCormick completed his work analyzing these data, it had the most detailed analysis of all of the data sets. Pieper and McCormick's study involved students from four Worcester Public High Schools (WPS), Nashoba Regional High School and Massachusetts Academy of Math and Science. The following sections will summarize the major results and findings of that study.

Previous studies have shown that the "sensing" versus "intuitive" dimension of the MBTI is by far the strongest indicator of performance on the SAT found on the two measures (MBTI and GCSI) utilized in this study. Another common indicator of SAT performance is the "perceiving" versus "judging" section of the MBTI. The "extroversion" versus "introversion" and "thinking" versus "feeling" portions of the MBTI did not result in any significant findings in the previous studies, though the literature includes reports favoring the introverts, and Isabel Myers-Briggs expected that

to be the case. Indeed, she expected E/I to be the second strongest predictor. In Pieper's study, the second strongest indicator was the Remote Associater dimension of the GCSI, but there is no MCAS data available in a data set with the GCSI results, so we will not go into those findings here.

In the Worcester Public School (WPS) System, intuitive (N) types scored, on average, 122 points higher than sensing (S) types. In each of three studies, this trend proved true with the intuitive (N) advantage ranging up to 162 points. In WPS 49 of the 122 point intuitive (N) advantage came in the math section while 73 points came in the verbal section.

The intuitive (N) advantage is probably due to the multiple choice format of the SAT. Students with a strong intuitive (N) sense would be able to pick out the right answer from a list even if they could not have figured it out without clues. The discrepancy between that advantage in the math and verbal sections can probably be attributed to the "grid- in" portion of the math section. Since the answers are not listed in this section, the intuitive advantage in recognizing answers rather than generating them is neutralized when the possible alternatives are not offered.

Previous studies have also found an advantage in SAT performance among perceiving (P) types. In WPS perceiving (P) types scored 45 points higher on the SAT than judging (J) types. This trend was supported by data from other previous studies. In these studies the perceiving advantage ranged up to 111 points.

The perceiving advantage can be attributed to perceiving types taking in more information before coming to a decision. This allows them to more fully understand the question instead of picking the first answer that looks possible or reasonable. As with the

intuitive (N) advantage, the perceiving (P) advantage is predominantly found in the verbal section. This could be in part be attributed to the influence of the "grid- in" style questions as well. Also, this can in part be attributed to the typical "perceivers' " superior reading comprehension on the verbal section (Pieper 1997). They tend to pick up details on the first read before they know what they are looking for, and focus their attention. J's focus early, and look for only what is relevant. They will generally have to go back through once they know what to look for.

Pieper's study in the Worcester Public School system included an analysis of the sixteen possible personality types in the MBTI versus performance on the PSAT. The highest scoring MBTI type was INFJ. The lowest scoring MBTI type was ESTJ In Nashoba Regional High School, the study showed the highest scoring type was INTJ. The lowest was, again, ESTJ. These two types were about 300 points apart in both cases. Pieper's report did not include average PSAT scores for all of the sixteen types. He reported the top three versus the bottom three for Worcester. These two groups were about 250 points apart. Tongol and Mangenot hoped to do that analysis again in Leicester, but the cohort size (110 combined) was too small to support an analysis that would divide the "sample" into sixteen categories (1998).

Previous studies also reported similar findings to Pieper's study for the MBTI. Kibbler and McTague reported a 108 point advantage for Intuitive (N) types over Sensing (S) types. They also reported a 102 point advantage of Perceiving (P) types over Judging (J) types. Batey, Brezniak, and Purohit (1995) reported a striking 162 point advantage for Intuitive (N) types over Sensing (S) types. They also reported a 111 point advantage

of Perceiving (P) types over Judging (J) types. In addition, Batey et al. found a 68 point advantage of Feeling (F) types over Thinking (T) types. See the chart below.

Table 2: MBTI Type Advantages Found in Previous Worcester Area Studies

	Batey et al.	**Kingsland et al.**	**Pieper**
Intuitive Advantage	+162	+108	+122
Feeling Advantage	+68	Unreported	Unreported
Perceiving Advantage	+111	+102	+45
Cohort Sizes	229	276	1267 (Jr + Soph)

2.10.4 Average SAT Scores from Previous Studies

One more piece of information that is vital to our study is the average SAT scores at various schools from former studies. Only Pieper's PSAT findings will be noted here because Kingsland et al's previous research was done on the original SAT before it was changed in 1995. Batey et al. used the first available version of the "New" SAT. Batey et al. used only the first SAT score, not the highest one as Kingsland et al did. Pieper reported average Verbal PSAT scores in the Worcester Public School system ranging between 399 to 443 points depending on the high school. Math scores ranged between 375 and 428 points. In Nashoba the Verbal average PSAT score was 506 points and the Math average was 501 points. At the Massachusetts Academy of Math and Science, the average Verbal PSAT score was 669 points and the Math PSAT score was 706 points. All of these scores are expected to move up about 50 points from the PSAT to later SAT scores. The Worcester Scores were also deflated by the administration's attempt to get all students, even those not college bound, to take the PSAT that year. Unlike the WPS,

at Nashoba High School, the 85%-90% of the students in the junior class who took the PSAT and who participated in the study were planning to attend college or at least considering it as an option.

Table 3: PSAT Averages From Pieper's Study

	WPS	Nashoba	Mass Academy
Verbal PSAT	419	506	669
Math PSAT	406	501	706
Total	825	1007	1375
Cohort Sizes	1267 (Jr + Soph)	308 (Jr + Soph)	35 (Jr Only)

A simple exmaination of these numbers shows that the Mass Academy scores are by far the highest and the intuitive advantage diminishes as the scores rise. This is expected since the students are an exclusive pre-selected group that took the PSAT before as part of their admission process to the academy. Here, the E/I relationship become more important. Nashoba is the second highest overall. Since it is a somewhat suburban/rural public school, the Leicester results were expected to somewhat mimic the Lancaster students in the Nashoba district, but Leicester is more of a working class town than Bolton or Stow, the other two towns in the Nashoba Regional System. The WPS results are, on average, the lowest. Since many students at the Worcester schools are from lower socio-economic backgrounds, and many are not college bound, this is to be expected. Pieper does not report PSAT scores by social class, but the larger literature does so, documenting a strong relationship.

Finally, Tongol and Mangenot made some discoveries that will eventually coincide with this study. They found that in terms of MBTI type, the Sensing (S)/

Intuitive (N) and Judging (J)/ Perceiving (P) parameters gave not strong, but moderate findings in relation to PSAT/SAT performance (1998). Intuitive (N) and Perceiving (P) types can "outscore their counterparts consistently when they are compared only to those of their own class year and academic program" (Mangenot, et al. 1998). This was an extremely interesting finding and was one of the foundations for which the proposed project was built. Can the same Intuitive and Perceving parameters as determined by the MBTI predict strong performance on the MCAS as they did for the PSAT and SAT? In order to answer this question, it first must be determined whether or not the MCAS shows elements of aptitude and not only achievement.

3 General Hypotheses

As mentioned in section 2.10 of this report, certain parameters of the MBTI, especially the Sensing (S)/ Intuitive (N) and Judging (J)/ Perceiving (P) types, have been shown to somewhat predict how a student will perform on the SAT. It is best that the data is analyzed under a control variable, i.e. for those classmates in the same academic program. When this is not done, MBTI relationships can be exaggerated as in Pieper's WPS study, or can be obscured altogether, as in the Leicester study. This is due to the impact of differences in preparation and which learning types have the greatest access to the more challenging Advanced Placement and Honors courses in a given high school. Therefore, the correlations that will be made between learning style via the MBTI and the SAT should also be strong because the students are of the same class.

In relation to the MBTI and individual subject grades or overall GPA relationship, it is hypothesized that those with the higher GPA's will tend to be certain learning types, although it is unknown which ones at this point. Isabel Myers predicted that "judging" students would do better in class day-by-day, but the "perceivers" would outperform them on standardized tests like the PSAT. Students with a higher GPA are often encouraged to take more challenging courses, and therefore, they are more prepared for standardized examinations than other students. However, some decline, so as to keep their grades higher in a less challenging program. It is also thought that classroom performance in certain subjects, such as science and English, will be a predictor for performance on the same sections of the MCAS rather than general GPA predicting average scores.

Since both the ACT and MCAS are achievement tests, it is expected that certain learning types determined by the MBTI will be more prone to underperformance than others. This follows the logic that the ACT and MCAS are highly correlated as shown in the literature review. Also, certain MBTI types have under performed on sections of standardized tests which have elements of achievement in them, indicating that certain MBTI types should be expected to perform similarly on the MCAS. While there has been little to no research in this area, the hypothesis has a foundation. Remember that the SAT has been shown to have elements of achievement in it. Also, previous research has shown correlations between the PSAT/SAT and ACT, as mentioned in section 2.10 of this report. Therefore, if certain learning styles by MBTI standards give some students an advantage on the sections of the PSAT/SAT that are supposedly "achievement sections", the same MBTI types will give the same students an advantage on both the ACT and MCAS assessment exams. It was also shown before in section 2.10 that many sections of the MCAS and ACT show strong correlations. Therefore, performance on sections of one test should predict performance of similar sections on the other test.

Getting back to the main focus of the project, which is the effect of MBTI type on MCAS performance, it is expected that there will be certain types that will have advantages and others that will be at risk. In section 4.5 of this report, there are questions of correlation listed. It is expected that there will be significant correlations between most, if not all, of the relationships that are stated.

4 Methodology

4.1 Objectives

The scope of this project goes well beyond trying to find if there is a relationship between the dimensions of the MBTI and the MCAS. By itself, that is not a complex analysis. However, if there are correlations between these two entities that finding will give rise to new hypotheses, a search for more correlations and replications of existing findings may be necessary. For instance, since there has been no other MCAS/MBTI study, this information should be compared to the findings of prior studies on the SAT/PSAT, grades and level of course difficulty, the Plan ACT (if available). Other analyses may be called for by the school systems involved. Whatever they deem necessary to develop an understanding of which students are at risk to under perform on this standardized test is worth doing, so long as they can provide the necessary data. Ethnicity and social class are likely to concern them.

Therefore, in this study, we seek to find in-depth relationships between MCAS scores, overall and by subject, using all four specific MBTI parameters, including discrete and continuous scores for each. Based on previous studies with the PSAT and SAT, the analysis will be focused on the Sensing (S) and Intuitive (N) aspects of the MBTI. This variable has been shown to be the most correlated with test scores on other standardized examinations.

The second part of this study will involve a brief guide to identifying students who are at risk to under perform on the MCAS so that the Fitchburg and Worcester public school systems can handle any plans to use this lead time as they see fit. It must be stressed that it is not the intention of this study to instruct the school systems how to

deal with these students. Suggestions for such ends will not be a part of this study. However, it is the intention of this project team to aid the school systems in looking at MCAS data help identify at-risk students and explain why the test is especially challenging for them. If time permits and the school systems are willing, we would also like to set up a meeting or presentation in which we present our findings and lend our insight as to why students may be receiving lower scores on the MCAS. We would also like to do this in order to answer any questions that the school system officials may have pertaining to the study in person. However, if asked for advice on how to handle students with risk factors, we prefer to leave that to the professionals who have been trained in that particular field, as we can offer only our opinions.

4.2 Current Status of Worcester and Fitchburg MCAS

Because of the MCAS, the Worcester and Fitchburg public school systems are in dire straights. Both schools are seriously under performing on the statewide test, and if the problem isn't resolved, a large percentage of their students will not be able to graduate from high school. The following table shows some statistics for the Science and Technology scores on the MCAS. These figures are from eighth graders who took the MCAS in 1999. They are the class of 2003, which takes the 10^{th} grade test the first time this year, 2001.

Table 4: 1999 MCAS Scores (8th grade)

School System	Rank	Advanced	Proficient	Needs Improv.	Failing
		(280 - 260)	(259 – 240)	(239 - 220)	(219 - 200)
Westborough	11	21%	44%	22%	13%
Algonquin	15	7%	39%	35%	19%
Nashoba	27	8%	30%	34%	27%
Quabbin	55	6%	33%	36%	26%
Leicester	64	3%	37%	32%	28%
Marlborough	80	6%	24%	27%	43%
Narragansett	?	5%	18%	30%	46%
Leominster	91	3%	20%	28%	48%
Fitchburg	118	1%	10%	21%	66%
Worcester	120	2%	13%	18%	66%
Boston	127	1%	8%	15%	71%
Statewide		5%	23%	27%	45%

It can be seen from the table that in Fitchburg, Worcester and especially Boston, over two thirds of the students are likely to fail the test and thus, not be eligible for graduation. Besides the fact that many whom would be normally eligible to graduate will not, there is a large percentage of students in the urban districts that cannot seem to meet the state's minimum required knowledge base. There is currently a movement afoot attacking the MCAS, supported by educators. It is obvious that something must be done to improve this situation if it really reflects educational performance and levels of learning by the school systems. If not, the test will create great hardship, and many educators comfortable with the idea of evaluating schools want to see the provision that will keep some students from graduating repealed.

4.3 Parameters of Study

The study contained students who attended high school in the Worcester and Fitchburg public school systems. All of these students graduated from high school in the year 1999, 2000, or 2001[1]. The MBTI was administered to them when they were in 10^{th} or 11^{th} grade, which was in 1997, 1998, or 1999 for most of them. The MCAS exam was also administered when they were sophomores. The number of students in the Worcester Public School system that took the MBTI and (presumably the MCAS) was about 1200, and there were 182 who took the MBTI (and presumably both exams) in the Fitchburg school system. Therefore, the total number of students available for study from both school systems was 942. This number was large enough to provide a reliable study and give meaningful results. Of these 942 students, the database contained somewhere in the neighborhood of 750-850 students from both systems for whom we received MCAS and PSAT/SAT scores. A large portion of the data came from the Worcester Public School system, simply because there are four high schools there, and a significantly larger number of 10^{th} graders for which there was MBTI data. Information for every student was not available. Hence, the database contained lapses and we took a loss of up to 200 students that took the MBTI.

4.4 Database Creation

Again, the number of students that were contained in the usable Worcester database from both school systems totaled just fewer than 1200 students. Encouraging responses to our request for MCAS data from Worcester were received and data was provided.

[1] The class of 2001 data from Worcester was added in to reach 1000 cases, as this was more statistically reliable. This cohort was missing data from Doherty High School., which serves the middle class area.

Individual transcript data (subject, grade, course number) was received from both Worcester and Fitchburg, SAT data was also obtained from both school systems for about 50% of the students who take that test, Plan ACT data from Fitchburg, any other well-known achievement tests that students in both systems have taken, and any remaining MBTI data from a few students in the Fitchburg school system were requested as well. Letters were written to contacts in both school systems to request these data[2]. As detailed in the letter, a meeting was scheduled with the school systems before statistical analysis in order to outline what the school system would like the project to focus on in addition to our goals.

Once the database was created, many parameters were included, such as class, year when the test was taken, MBTI type, MCAS scores in all subject areas, verbal, math, and overall SAT scores, individual subject grades, ACT scores (where applicable), are also by subject and composite. Microsoft Excel® was used to build the database, and it was organized well enough so that it could be utilized by the school systems involved to see if they could have identified students at risk to under perform on the MCAS with such a mix of variables had it been created before the students took that test at the end of the 10th grade.

4.5 Statistical Analyses and Correlations

Once the database was created and had all the necessary information in it, analysis of the data was then performed. There were a couple of computer software programs that aided in the correlations that were planned. As mentioned before, Microsoft Excel® was

[2] This was Mr. Lamey from Fitchburg and Dr. Mostue from Worcester- refer to Appendix C for the letters.

the location of the database. This information was then imported to Statistical Program for the Social Sciences (SPSS) in order to make correlations, analyze data, and test the hypotheses presented in this project report. The statistical tools that were used in the study included correlation coefficients and other graphing relationships more available in the SPSS package. All of these analyses enabled one to see the relationship between two entities that cannot normally be inferred from simply looking at the raw data. Following is a list of questions that the study intended to answer. What is the relationship, if any, between:

- Overall MBTI type and MCAS scores?
- Individual MBTI parameters (i.e. S/N) and MCAS scores?
- Individual subject grades and MCAS scores (science with science, etc.)?
- PSAT/SAT scores with MCAS scores?
- ACT Plan and Stanford Nine (if available) scores with MCAS scores?

Also, we intended to explore any correlations between individual subject grades (or overall GPA) and MBTI type, as well as any others that were requested from either the school system or other interested parties. Not only were the existence of correlations reported, but the relative strength of these correlations was also determined and reported so as to see how powerful an index could be created by combining them.

4.6 Identifying and Typing Students in the School System

Once the data was analyzed, the school systems were to be informed of the findings. In order to do this, a meeting is to be set up with both Worcester and Fitchburg so that the data and results can be presented. A copy of this report will also be sent to them. The goal of these meetings is simply to inform the school officials of whether or not the

MBTI provided a way to identify students who are at risk to under perform on the MCAS in terms of their learning style, and how reliable a guide subject grades are for the different types in terms of MCAS, SAT, and ACT performance. Prior findings have indicated that SAT scores are good guides to freshman college performance for some types of students, and tell you nothing about how others will perform. A similar logic using the MBTI as a control variable will be employed here too. The goal of the project does not include, however, instructing officials and guidance counselors on how to deal with these students who are at risk. The meetings were simply for suggestion, and since the hypotheses in this report were proven accurate, the MBTI can be used as a lead indicator of how a student will perform on the MCAS. Especially notable to both the study researchers and the school system officials was the relationships between the Sensing (S)/ Intuitive (N) parameters and MCAS scores, by section and overall. Secondarily, the grade predictions by MBTI type were also shown to be revealing.

In any case, the Worcester Public School System has many qualified users of the MBTI who understand its learning styles implications. Few of them are classroom teachers. Most are in guidance or Educational Psychology, and hence their expertise was not going to be brought to bear on these problems until the case was made that it would be worth doing. Since few of the MBTI qualified staff members are comfortable with large-scale statistical analysis, the Worcester Public Schools do not organize the MBTI data in such a way as to facilitate this type of analysis. We see ourselves as stepping into these gaps and creating the data set configuration that was needed to use MBTI results in order to help the students if there are indeed MBTI-MCAS relationships reminiscent of the prior MBTI-PSAT findings.

5 Data Analysis and Results

5.1 Fitchburg Overview and Summary of Results

When the Fitchburg Public School database was finally created, there were 344 cases that had MBTI data out of the two classes (1999 and 2000). However, the class of 1999 did not take the MCAS as sophomores. Hence, out of these 2 classes of students, there were only 191 students for whom there was both MCAS and MBTI data. Therefore, the analyses that involve MCAS data only contain these 191 cases, most of which came from the class of 2000. Obtaining the data from Fitchburg was less problematic than from Worcester, for us because there were fewer cases, and some of the data was already available at WPI due to a prior project on the ACT. A database had already been composed by Mike Bonczek during a previous project. While this database had data entry errors and some fields were left blank, it served as a good framework for the larger database created in this study. The analysis of this Fitchburg data set also served as a model for the analysis to be performed on the larger and more comprehensive Worcester data sets to be created later on.

5.1.1 MCAS/MBTI Relationship[3]

When analyzing the MCAS/MBTI relationship, it was necessary to use correlation coefficients, compare the means, and also cross-tabulate certain parameters. When a cross-tabulation was done, it was found that the highest quintile (highest 20% of MCAS scores) mostly came from the Intuitive students whereas the lowest quintile

[3] In all analyses done on the Fitchburg data set, the first letter of the MBTI discrete pair was assigned a "1" and the second was assigned a "2". For example, in the table on page 65, S was given a "1" and N a "2".

(lowest 20% of MCAS scores) contained mostly Sensing students. In order to further investigate whether Intuitive students had an advantage over the Sensing students on overall MCAS performance, a table of correlation coefficients was calculated. This also showed that Intuitives had an advantage. While the parametric (Pearson) correlation coefficient was low, it was still significant at 0.163. A final comparison of the means, table in Appendix, was able to distinguish individual sections of the MCAS with the S/N discrete variable. This discrete variable was primarily related to the English and Science sections of the MCAS, and not at all with the Mathematics section. These significant tests were also low, but statistically significant. The combination of these advantages for Intuitives most likely makes up their advantage on the overall test score. The Pearson correlation coefficient table follows for the MBTI/MCAS analysis. The Spearman rho results were almost identical.

Table 5: Pearson Correlation Coefficients for MBTI/MCAS Relationship

	Overall MCAS Score	**E/I**	**S/N**	**T/F**	**J/P**
Overall MCAS Score	1.000	0.016	0.163*	0.026	-0.019
E/I	0.016	1.000	-0.042	-0.049	-0.031
S/N	0.163*	-0.042	1.000	0.202**	0.220**
T/F	0.026	-0.049	0.202**	1.000	0.117*
J/P	-0.019	-0.031	0.220**	0.117*	1.000

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

A positive correlation coefficient means the "2's" have the advantage and a negative coefficient means the "1's" have the advantage.

5.1.2 Overall GPA/MBTI Relationship

A comparison was done with the discrete variables of the MBTI and the students 4 year GPA. It was found from a simple correlation table that overall GPA correlated with the E/I and J/P discrete variables. These correlations were also low, but still were statistically significant. The Pearson correlation table can be found below, and the Spearman rho table yielded similar results.

Table 6: Pearson Correlation Coefficients for MBTI/GPA Relationship

	Overall GPA	E/I	S/N	T/F	J/P
Overall GPA	1.000	0.152*	0.038	0.035	-0.160**
E/I	0.152*	1.000	-0.042	-0.049	-0.031
S/N	0.038	-0.042	1.000	0.202**	0.220**
T/F	0.035	-0.049	0.202**	1.000	0.117*
J/P	-0.160**	-0.031	0.220**	0.117*	1.000

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

In can be seen that Introverts outperformed Extroverts in the area of overall GPA, and Judging students outperformed Perceivers. Cross-tabulation and comparison of the means showed the same results, with the Introverts having a mean overall GPA of 2.85 and Extroverts having one of 2.61. Similarly, Judging students had a mean GPA of about 2.89, and the Perceivers had one at about 2.63. All of the analyses performed on these variables resulted in a slightly higher correlation between the J/P discrete variable and overall GPA than the E/I discrete variable and overall GPA. In this particular analysis, we found a negative correlation coefficient, which simply means that the first parameter

of the two (i.e. J or E) was the one associated with high GPA's and represents the lower MBTI score. A positive correlation coefficient meant that the second parameter (i.e. P or I) was related to getting higher grades.

5.1.3 SAT/MBTI Relationship

Again, correlations, cross-tabulations, and a comparison of means were made with SAT scores (composite and individual sections) and the discrete variables of the MBTI. It was found from the correlation coefficients that the verbal section of the SAT correlated with the E/I and S/N discrete variables, with the higher correlation existing with the S/N variable. The mathematics section of the SAT correlated with the E/I variable only. Finally, the composite, or combined, SAT score of both sections again correlated with both the E/I and S/N variables, with the higher correlation existing with the E/I parameter. This was not consistent with the verbal SAT/MBTI relationship found. Therefore, further analysis was necessary since the Introverts had an advantage on the math section and overall SAT score, but not on the verbal section of the SAT.

Comparison of the means showed about a 40-point advantage on the SAT verbal section for Intuitives, a 20-point advantage for the Intuitives on the math section, and an overall advantage of 60 points (again for Intuitives) on the composite score. As for the E/I discrete variable and how it correlated with the SAT, another comparison of the means displayed about a 30-point advantage for Introverts on the verbal section, a 40-point advantage on the math section (also for Introverts), and an overall 70-point advantage for Introverts on the composite SAT score.

From this information, it was still unclear which discrete variable was more powerful in predicting SAT score, both on individual sections and overall. Therefore, the

median scores were looked at and a cross-tabulation was performed. Again, it was found that there was no real difference, and neither the S/N nor E/I variables were better than the other at predicting SAT score. However, if a decision had to be made, it would be most likely concluded that the S/N variable is more correlated to the verbal section of the SAT than the E/I variable. Conversely, the E/I variable is more highly correlated to the mathematics section and overall composite score of the SAT than S/N. Therefore, at least in this small Fitchburg data set, the E/I variable is the most important MBTI parameter involved with SAT scores, followed by S/N. It must be stated, however, that no correlation coefficient, parametric (Pearson) and non-parametric (rank-order), was above 0.2. Neither the J/P nor T/F variables made any correlation with any section of the SAT. Further, this finding is at variance with the results of the other studies in Worcester, Leicester, and even in Fitchburg for the class of 1999 as reported by Eric Tapley when he studied the SAT – ACT relationship, in all those cases SN was the more predictive variable.

5.1.4 Overall GPA/SAT Relationship

Thus far, we have reported only small correlation coefficients. While they are statistically significant, they do not explain a major portion of the variance. This analysis of the overall GPA/SAT relationship was where the correlation coefficients started to become larger and more significant. Only correlation analyses were used in this instance because we were trying to determine, from the Fitchburg data set, if overall GPA was more correlated with the SAT or the MCAS. It was expected that the MCAS/Overall GPA relationship should be stronger than the overall GPA to SAT relationship.

Hence, a correlation table was calculated (found at the end of the section) and it was found that the overall GPA of the students in the Fitchburg system was highly correlated with all three scores on the SAT (verbal, math, and composite), using both the Pearson and Spearman rho (rank order) correlation coefficients. The highest correlation, using both Pearson and non-parametric correlation coefficients, was between overall GPA and the composite SAT score. The lowest was found when overall GPA was correlated with the verbal section of the SAT. The coefficients for the correlations between overall GPA and the individual SAT sections were about 0.5, and the coefficient for GPA and composite score was slightly higher at 0.56. The grades of Fitchburg students are thus explaining about 25% of the variance on their SAT scores.

Table 7: Pearson Correlation Coefficients for SAT/GPA Relationship

	Overall GPA	**Verbal**	**Math**	**Comp.**
Overall GPA	1.000	0.504**	0.512**	0.558**
Verbal	0.504**	1.000	0.681**	0.921**
Math	0.512**	0.681**	1.000	0.913**
Comp.	0.558**	0.921**	0.913**	1.000

**Correlation is significant at the 0.01 level (2-tailed)

5.1.5 Overall GPA/MCAS Relationship

For this analysis, the same procedure for the GPA/SAT analysis was essentially repeated. However, as expected, the correlation between overall GPA and each section of the MCAS was somewhat higher than that of each SAT section and overall GPA. Below is a table showing overall GPA and the correlation coefficients for each section of

both the MCAS and SAT. Notice the high correlation between overall GPA and the English section of the MCAS. It is even higher than the overall GPA/overall MCAS score correlation.

Table 8: Pearson Correlation Coefficients Between GPA, SAT, and MCAS

	SAT Verbal	SAT Math	SAT Composite	MCAS English	MCAS Math	MCAS Science	MCAS Overall
Overall GPA	0.504	0.512	0.558	0.733	0.584	0.583	0.719

So, about 35% of the variance in MCAS math and science grades and 50% of the variance in English MCAS score is explained by grades – as compiled into a 4 year GPA. Unfortunately all the grades necessary to compute the GPA aren't available until 2 years after the test s taken.

5.1.6 MCAS/SAT Relationship

The final analysis to perform in this data set from Fitchburg was the one between the MCAS and the SAT. Analysis was done on the 150 cases that had both the MCAS and SAT data. In theory, these two tests should not correlate highly or at all. This is due to the fact that the MCAS is a measure of achievement based on curriculum mastery and the SAT is supposedly a measure of aptitude. However, the results of this analysis showed a very different situation. This analysis produced the highest correlation coefficient of all the analyses performed on the Fitchburg data set. The lowest correlation coefficient seen was 0.5 between the science section of the MCAS and the verbal section of the SAT. The highest coefficient seen was 0.827 between the overall MCAS score and the composite SAT score. Both the Pearson and Spearman rho

coefficients were similar to each other, so only the Pearson table is shown below. In terms of the way that these coefficients describe a relationship, a 0.9 coefficient is interpreted as an identity (more than 80% of the variance explained). This means that the two entities being correlated are virtually identical. From the results of our analysis, it seems that the MCAS and SAT are almost identical, that SAT scores explain 69% of the variance in the MCAS score. In fact, the MCAS could be used to predict SAT scores and vice versa with considerable accuracy. However, this data set was small and it will be interesting to see if the larger, more statistically credible Worcester data set produces the same results.

Table 9: Pearson Correlation Coefficients for MCAS/SAT Relationship

	MCAS_ENG	MCAS_MATH	MCAS_SCI	OVERALL_MCAS	SAT_VERB	SAT_MATH	SAT_COMP
MCAS_ENG	1.000	0.635**	0.747**	0.898**	0.625**	0.601**	0.681**
MCAS_MATH	0.635**	1.000	0.670**	0.875**	0.473**	0.527**	0.803**
MCAS_SCI	0.747**	0.670**	1.000	0.891**	0.512**	0.529**	0.648**
OVERALL_MCAS	0.898**	0.875**	0.891**	1.000	0.606**	0.624**	0.827**
SAT_VERB	0.625**	0.473**	0.512**	0.606**	1.000	0.973**	0.921**
SAT_MATH	0.601**	0.527**	0.529**	0.624**	0.973**	1.000	0.913**
SAT_COMP	0.681**	0.803**	0.648**	0.827**	0.921**	0.913**	1.000

**Correlation is significant at the 0.01 level (2-tailed)

5.1.7 Transcript/MCAS/MBTI Relationship

For those students in the database which we had MCAS and MBTI data for, a key was created. This key evaluated the difficulty of each student's academic curriculum. The curriculums were given number designations (1-8), with 1 being the hardest high school curriculum and 8 being the easiest. They were separated as follows:

1= All AP Classes

2= Some AP Classes
3= All Honors Classes
4= Some Honors, Some College Classes
5= All College Classes
6= Some College, Some General Classes
7= All General Classes
8= Dominated by Special Education Classes

With this academic "track" variable, we performed a comparison of the means between the MCAS scores (by subject and overall) in each category. The resulting table follows. In the Fitchburg data set, there were no students who had all AP classes. Hence, no "1's" appeared in the analysis.

Table 10: Mean MCAS scores by average Program Difficulty

Program Track- # cases	MCAS_ENG	MCAS_MATH	MCAS_SCI	OVERALL MCAS
2- 11 cases	257	240	238	734
3- 6 cases	250	234	240	724
4- 41 cases	241	224	232	696
5- 42 cases	230	214	224	669
6- 30 cases	223	207	218	648
7- 2 cases	201	200	204	605
8- 1 case	220	200	236	656

Notice that each average MCAS score and the overall MCAS score decreases as the program track goes from harder to easier. This is to be expected since the more difficult courses cover more material and thus better prepare the student for the MCAS test, which focuses on the core curriculum.

The next analysis that was performed was a comparison of the means for the Sensing and Intuitive students in each curriculum category. The results from this analysis

are also found below[4]. Notice that as the curriculum becomes easier, there is less difference between the Sensors and Intuitives on the MCAS. However, with the more difficult program tracks, the "2's" and "3's" especially, the Intuitives have about a 20-point advantage on the MCAS over the Sensing students.

Table 11: MCAS/SN/Program Difficulty

Program Track S/N Preference- # cases	MCAS_ENG	MCAS_MATH	MCAS_SCI	Overall_MCAS
2 (S)- 3	250	237	231	718
(N)- 8	259	241	241	740
3 (S)- 3	246	225	239	711
(N)- 3	255	243	240	738
4 (S)- 20	240	223	231	695
(N)- 21	242	224	232	698
5 (S)- 22	230	215	224	669
(N)- 20	231	213	225	668
6 (S)- 16	224	206	218	648
(N)- 14	223	208	218	648
7 (S)- 2	201	200	204	605
8 (N)- 1	220	200	236	656

After seeing these results, we decided to do an identical analysis with SAT data, both by subject and overall (table below). Notice that there is a difference as the course difficulty decreases, something that contrasts with Tapley's Fitchburg Class of 1999 averages (Appendix G). Only the two most difficult academic tracks showed an Intuitive

[4] Only averages are found in this table. However, the entire output table, which includes medians, variance, and standard deviations from SPSS can be found in Appendix D.

advantage with SAT scores. After that, the Sensing students seemed to have an advantage in the class of 2000.

Table 12: SAT/SN/Program Difficulty

Program Track S/N Preference- # cases	SAT_VERB	SAT_MATH	SAT_COMP
2 (S)- 4	565	520	1085
(N)- 7	643	563	1206
3 (S)- 3	457	513	970
(N)- 3	603	563	1117
4 (S)- 17	491	501	991
(N)- 19	458	459	918
5 (S)- 15	451	413	864
(N)- 16	433	414	847
6 (S)- 7	439	396	834
(N)- 6	413	400	813

5.1.8 Sensing/Intuition Difference in MCAS and SAT score with Respect to Gender

The final analysis performed on the Fitchburg data set was also inspired by Tapley's study. A comparison of the means was calculated for each parameter of the MBTI and then separated by sex. There are four tables below which show an average comparison of overall MCAS scores and overall SAT scores between each MBTI parameter by sex. The "1's" in the table represent E, S, T, and J while the "2's" represent I, N, F, and P. Notice that there is no difference between males and females in terms of MCAS scores except that female Intuitives have an advantage over female Sensing students. This relationship cannot be found in the males. However, with respect to SAT

scores, there is a major difference between males and females, especially with the E/I and S/N parameters. For instance, from the table below, it can be seen that female Intuitive students have about a 90-point advantage over female Sensors on overall SAT score while male Intuitives only have a 25-point advantage over the male Sensing students. This same trend can be found with E/I, where female Introverts have only a 50-point advantage over the female Extraverts on SAT score while male Introverts have a 100-point advantage.

Table 13: Test Score Difference by Sex for MBTI Type

SEX	Type	SAT_COMP	OVERMCAS		SEX	Type	SAT_COMP	OVERMCAS
F	E	938	681		**F**	S	910	669
	I	987	680			N	1002	693
	Total	960	680			Total	960	680
	Count	107	71			Count	107	71
M	E	932	670		**M**	S	963	668
	I	1038	672			N	989	673
	Total	976	670			Total	976	670
	Count	85	82			Count	85	82

SEX	Type	SAT_COMP	OVERMCAS		SEX	Type	SAT_COMP	OVERMCAS
F	T	961	678		**F**	J	950	680
	F	960	682			P	966	680
	Total	960	680			Total	960	680
	Count	107	71			Count	107	71
M	T	971	672		**M**	J	983	670
	F	985	668			P	973	670
	Total	976	670			Total	976	670
	Count	85	82			Count	85	82

5.2 Worcester Overview and Summary of Results

The Worcester Public School data set was constructed out of a mixed group of measures and demographics covering the classes of 2000 and 2001, linked together in two distinct ways. As described in the methodology section it was necessary to merge data files either on the basis of student ID number or surname. As a result of this two-

pronged merger approach, we were concerned about the overlapping of duplicate names in both files. It would be conceivable to have one instance of a name tied to one student ID in a MBTI file, and then merge that with SAT data by last name, but there could also be two or three or ten students with the same last name. Many of the students with the same last name also share a first name, the simplest solution was to remove all duplicate student ids and last names prior to file merger.

We conducted basic descriptive statistics for all factors studied in both the 2000 and 2001 datasets (data not shown). Our major concerns lay with the MCAS, MBTI, and difficulty of curriculum variable. Because one of the goals of this study is to identify students at risk of under performing on the MCAS it was necessary to concentrate on the students with MBTI data and MCAS scores – figuring that transcript data could be found to describe the program that most of them took. Beyond that we focused on the proportion failing the MCAS. This allowed for a clearer picture of how the different factors affected students at risk of failing. One goal was to describe those at risk who would need help before or after the exam, preferably before. What kind of learners were most of them going to be?

5.2.1 Class of 2000

5.2.1.1 MCAS/MBTI Relationship

When analyzing the MCAS/MBTI relationship it was found that looking at the mean score as well as the percentage of students failing the exam was a convenient way to see if one MBTI type has an advantage over another. The first type that will be looked at is the Sensing/Intuition vs. MCAS relationship.

Table 14: SN vs. MCAS Mean score

Sensing - Intuition		MCAS English	MCAS Math	MCAS Science
Sensing	Mean	226	214	221
	N	238	238	238
	Std. Dev.	15	15	12
Intuition	Mean	236	225	230
	N	194	194	194
	Std. Dev.	16	20	14
Total	Mean	230	219	225
	N	432	432	432
	Std. Dev.	16	18	14

Table 15: SN vs. % of failing students

Thinking - Feeling			MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
	Sensing	Count	74	156	103
		% within SN	31%	66%	43%
	Intuition	Count	22	87	35
		% within SN	11%	45%	18%
Total		Count	96	243	138
		% within SN	22%	56%	32%

By looking at these two tables it can be seen that students preferring intuition have approximately a ten-point advantage over the sensing students on each test. Ten-points is not a large margin on a test of this size, but it happens that those ten points make a large difference in the number of students failing the test since the averages are so close to the 220-point level defining failure. This can be seen in English and Science section of the MCAS and to a lesser degree on the Math section where in the students general seem less well prepared. The proportion of intuitive students failing the English and Science section are less then half that of the sensing who are failing on their practice rounds.

The next MBTI dimension to be examined at will be the Judging/Perceiving variable as it relates to the MCAS score. The J/P dimension will be analyzed with respect to mean MCAS score as well as percentage of students failing the MCAS

Table 16: JP vs. MCAS Mean Score

Judging – Perceiving		MCAS English	MCAS Math	MCAS Science
Judging	Mean	228	218	223
	N	144	144	144
	Std. Dev.	17	17	15
Perceiving	Mean	231	220	226
	N	288	288	288
	Std. Dev.	16	19	13
Total	Mean	230	219	225
	N	432	432	432
	Std. Dev.	16	18	14

Table 17: JP vs. % of Failing Students

Judging – Perceiving			MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
	Judging	Count	42	85	66
		% within JP	29%	59%	46%
	Perceiving	Count	54	158	72
		% within JP	19%	55%	25%
Total		Count	96	243	138
		% within JP	22%	56%	32%

The mean scores show very little bias, roughly about two – four points. But the percentage of failing students shows of a more significant finding. As with the SN the English and Science section shows a large gap between the two types while the Math shows very little difference.

The third type to be looked at is the Extraversion and introversion type.

Table 18: EI vs. MCAS Mean Score

Extraversion – Introversion		MCAS English	MCAS Math	MCAS Science
Extraversion	Mean	229	214	224
	N	277	277	277
	Std. Dev.	16	17	13
Introversion	Mean	233	222	227
	N	155	155	155
	Std. Dev.	16	20	14
Total	Mean	230	219	225
	N	432	432	432
	Std. Dev.	16	18	14

Table 19: EI vs. % of Failing Students

Extraversion – Introversion			MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
	Extraversion	Count	65	166	95
		% within EI	24%	60%	34%
		I			
	Introversion	Count	31	77	43
		% within EI	20%	50%	28%
Total		Count	96	243	138
		% within EI	22%	56%	32%

As can be seen by looking at these tables, there is negligible difference in the mean scores distinguished on this dimension. The percentage of failing students as well shows very little dependence on type, with the largest gap in the Math score being only about 10%.

The final type to be looked at is the Thinking and Feeling types.

Table 20: TF vs. MCAS Mean Score

Thinking – Feeling		MCAS English	MCAS Math	MCAS Science
Thinking	Mean	229	219	225
	N	270	270	270
	Std. Dev.	17	19	14
Feeling	Mean	233	220	226
	N	162	162	162
	Std. Dev.	15	18	13
Total	Mean	230	219	255
	N	432	432	432
	Std. Dev.	16	18	14

Table 21: TF vs. % Failing Students

Thinking - Feeling			MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
	Thinking	Count	66	154	90
		% within TF	24%	57%	33%
	Feeling	Count	30	89	48
		% within TF	19%	55%	30%
Total		Count	96	243	138
		% within TF	22%	56%	32%

As with the EI type there isn't a significant difference in the mean scores of between the Thinking and Feeling preference groups. Also, there isn't a major difference in the percentage of students failing the test between the groups.

When looking at all four types it gives a clearer picture of what types have the most affect over the students score. The first two types discussed SN and JP have the largest affect, SN being the worst correlated out of those two. For the students on the borderline between failing and passing the test, the mean score difference, while small, appears to have a large effect in terms of number of students affected or at risk.

The next table shows the top four and bottom four of the 16 MBTI types and their percentage of failure.

Table 22: Top 4 and Bottom 4 MBTI Types

	English			Math			Science		
	Type	**% Failed**	**Count**	**Type**	**% Failed**	**Count**	**Type**	**% Failed**	**Count**
Top Four	INFP	4%	1	INFP	21%	5	INFJ	0%	out of 5
	ENFP	6%	3	ENTJ	21%	3	INFP	4%	1
	ENTJ	7%	1	INTP	24%	5	INTP	5%	1
	INTP	10%	2	INFJ	40%	2	ENFJ	14%	1
Bottom Four	ESFJ	47%	9	ESFJ	84%	16	ESFJ	74%	14
	ISTJ	46%	12	ISTJ	69%	18	ISTJ	65%	17
	ESTP	36%	20	ESTP	69%	38	ISFJ	47%	8
	ISFJ	35%	6	ISFJ	65%	11	ESTJ	44%	20

5.2.1.2 MCAS/Difficulty level of Classes

In order to do this analysis a variable had to be created to gauge the students' class difficulty. The curriculum levels were given number designations (1-5), with 1 being the hardest high school curriculum and 5 being the easiest. They were separated as follows:

1= Some AP Classes or all Honors (N = 118)
2= Some Honors, Some College Classes (N = 295)
3= All College Classes (N = 107)
4= Some College, Some General Classes (N = 115)
5= All General Classes (N = 35)

Looking at the number of students in each category shows a large portion fell into category 2. This is because students that took mostly college but a only few honors and students that took mostly honors and only a few college classes where lumped into this grouping. In retrospect we probably should not have done that and it could easily have been changed.

Table 23: Average Course Difficulty Level vs. MCAS Failing

			MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
	Honors +	Count	4	21	6
		% within Category	3%	18%	5%
	Honors/ College	Count	5	173	92
		% within Category	17%	59%	32%
	College	Count	47	89	65
		% within Category	44%	83%	61%
	College/ General	Count	69	100	83
		% within Category	60%	87%	72%
	General	Count	21	28	26
		% within Category	60%	80%	74%
Total		Count	191	411	272
		% within Category	29%	61%	41%

Notice that the percentage of failing students increases as the level of challenge (or difficulty) of the students' courses goes down. This is to be expected since the harder curricula prepare a student a for the MCAS test. It could also be argued that the more able + ambitious students take the harder courses, and that also plays a role in the trend shown above. It is not clear what is the relationship - but the resulting pattern is clear and consistent.

5.2.1.3 MBTI/MCAS/Course Difficulty

The final data that will be shown for the Worcester 2000 data set is the relationship between SN / MCAS / Course Difficulty.

Table 24: SN and Difficulty Level vs. MCAS Mean Score

Average Program Level Of Courses			MCAS Eng Mean	MCAS Math Mean	MCAS Sci Mean
Honors +	29	S	241	235	234
	59	N	245	240	241
Honors/College	123	S	229	215	223
	82	N	235	220	227
College	42	S	218	206	216
	22	N	232	221	228
College/General	31	S	214	206	211
	20	N	221	210	219
General	13	S	216	209	214
	9	N	229	216	221

This table shows clearly that even within the same difficulty level the students that preferring Intuition have an advantage over the students that favor Sensing. This rules out the possibility that the advantage seen in the pure MBTI/MCAS results could be a result of certain MBTI types are simply taking higher-level classes and that is giving them the advantage in the MCAS.

The SAT – MCAS relationship was briefly revisited in this larger data set to see if it was consistent with the results from the Fitchburg dataset.

Table 25: SAT vs. MCAS

	MCAS English	**MCAS Math**	**MCAS Science**
SAT Verbal	0.71	NULL	NULL
SAT Math	NULL	0.83	NULL
SAT Composite	0.70	0.83	0.81

These numbers are very highly correlated in the 0.7 – 0.8 range. They are also very similar the results found in the Fitchburg dataset.

5.2.2 Class of 2001 : A Replication Study

5.2.2.1 MCAS/MBTI Relationship

When analyzing the MCAS/MBTI relationship it was found that looking at the mean score as well as the percentage of students failing the MCAS exam was a convenient way to see if one type has an advantage over another. It should be noted that this dataset is slightly different then the 2000 dataset. This one lacks all the students from Doherty High School, and includes 100 Voke students. The first MBTI dimension that will be looked at is the Sensing/Intuition vs. MCAS relationship, since that was the strongest of the earlier findings from the class of 2000 study.

Table 26: SN vs. % of failing students

Sensing – Intuition			MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
	Sensing	Count	153	219	175
		% within SN	52%	74%	59%
	Intuition	Count	55	110	70
		% within SN	26%	52%	33%
Total		Count	208	329	245
		% within SN	41%	65%	48%

By looking at this table, it can be seen that students with a preference for intuition have less of a chance of failing the MCAS in each section. This can be seen in English and Science section of the MCAS and to a slightly lesser degree the Math section. The

intuition students failing each section is just about half that of the sensing, except for the Math section in which the difference isn't as much.

The next type to be looked at will be the Judging/Perceiving type with the MCAS. The JP type will be analyzed with respect to percentage of students failing the MCAS.

Table 27: JP vs. % of Failing Students

Judging – Perceiving		MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
Judging	Count	99	146	116
	% within JP	51%	75%	59%
Perceiving	Count	109	182	128
	% within JP	35%	59%	41%
Total	Count	208	328	244
	% within JP	41%	65%	48%

The percentage of failing students shows a significant finding. As with the SN, the Math and Science section shows a large gap between the two types of learners while the English shows less of a difference. Here, we can see that Perceiving students have an advantage on all three sections of the MCAS over the Judging.

The third Psychological type we looked at was the Extraversion and Introversion dimension.

Table 28: EI vs. % of Failing Students

Extraversion – Introversion		MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
Extraversion	Count	129	205	161
	% within EI	41%	65%	51%
Introversion	Count	79	124	84
	% within EI	41%	64%	43%
Total	Count	208	329	245
	% within EI	41%	65%	48%

As can be seen by looking at this table, the percentage of failing students shows very little relationship to this dimension of the MBTI. Unlike the class of 2000, there is no difference between the two types in each of the three sections of the MCAS. The EI relationship does not replicate.

The final dimension we looked at was the Thinking and Feeling preferences.

Table 29: TF vs. % Failing Students

Thinking - Feeling		MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
Thinking	Count	153	219	162
	% within TF	49%	70%	52%
Feeling	Count	55	109	82
	% within TF	28%	56%	43%
Total	Count	208	328	244
	% within TF	41%	65%	48%

This result shows that there is not a large relationship between type and the percentage of failing students on each section of the MCAS. However, it seems that

there is a certain percentage of Feelers that have less of a chance of failing the English and Math sections of the MCAS than the Thinkers. This finding was not evident in the class of 2000 study and might be due to data irregularity.

When looking at all four types, it gives a clearer picture of what types have the most affect over the student's score. The first two type dimensions discussed, SN and JP, have the strongest relationship to MCAS scores with SN having the larger correlation out of those two. For the students on the borderline between failing and passing the test, the mean score difference, while small, appears to have a large effect.

The next table shows the top four and bottom four MBTI types for the Math section of the MCAS from the class of 2001. An analysis of the overall MCAS scores was not performed for this class (A different data analyst worked on this late arriving dataset). However, the class of 2001 seems to have the most trouble with the Math section, hence the reason for the presentation of these percentages over the English and Science sections. The English and Science sections do show similar results, and the full table of analysis can be found in Appendix E.

Table 30: Top 4 Bottom 4 MBTI Types (Math Section Only)

	Type	% Failed	Count
Top Four	INFJ	20%	1
	ENFP	38%	22
	INTP	44%	16
	INFP	50%	9
Bottom Four	ESTJ	84%	41
	ENFJ	83%	5
	ISTP	81%	21
	ISTJ	80%	39

5.2.2.2 MCAS/Difficulty level of Classes

The same variable that was created to gauge the students' class difficulty with the class of 2000 was used for this class as well. Again, they were separated as follows:

1= Some AP Classes or all Honors
2= Some Honors, Some College Classes
3= All College Classes
4= Some College, Some General Classes
5= All General Classes

It should be noted that a large majority of the students again fell in the number 2 category. This is because students that took mostly college but only a few honors, and students that took mostly honors and only a few college classes were lumped together. If we had to do it again we would base a division on whether the Honors or College courses were in the majority.

Table 31: Difficulty Level vs. MCAS Mean Score

Average Program Difficulty			MCAS Eng Failing	MCAS Math Failing	MCAS Sci Failing
	Honors +	Count	93	93	93
		Mean	242	236	235
	Honors/ College	Count	265	266	266
		Mean	226	216	220
	College	Count	68	68	68
		Mean	213	210	215
	College/ General	Count	62	62	61
		Mean	207	202	211
	General	Count	19	19	19
		Mean	201	200	204
Total		Count	507	508	507
		Mean	224	216	221

Notice that the average score on each section of the MCAS increases as the level of difficulty of the students' courses increases. This is to be expected since the harder curriculums prepare the student more for the MCAS test. It could also be argued that the more intelligent students take the harder courses, and that also plays a role in the trend shown above, so the direction of causality is not clear, but it probably is a matter of better preparation for the achievement test. It should also be noted that this analysis replicated the class of 2000 findings.

5.2.2.3 MBTI/MCAS/Course Difficulty

The final data that will be shown for the Worcester 2001 data set is the relationship between SN/MCAS/Course Difficulty.

Table 32: SN and Difficulty Level vs. MCAS Mean Score

Track Level - Count		MCAS Eng Mean	MCAS Math Mean	MCAS Sci Mean
Honors +	41 S	241	236	234
	52 N	242	236	236
Honors/	150 S	222	213	218
College	115 N	230	220	224
College	45 S	213	208	214
	23 N	213	212	218
College/	48 S	206	202	210
General	14 N	209	202	213
General	13 S	201	200	203
	6 N	203	200	206

This table shows clearly that even within the same difficulty level the Intuitive students have an advantage over the students that prefer Sensing. This rules out the possibility that the advantage seen in the tables which compared MCAS average by MBTI variable could be an artifact – the indirect result of Intuitive and Introvert types taking higher-level classes more often enough to give themselves the advantage on the MCAS.

6 Conclusions

The Fitchburg data set served as an excellent model for planning our analysis of the 2 part Worcester data set. It was found that the Intuitives have a consistent advantage on the MCAS, even though the correlation coefficient was somewhat small. Also, Introverts and Judgers have higher average GPA's than Extroverts and Perceivers, respectively. In terms of the MBTI/SAT relationship, we were able to replicate results from previous studies, showing that Intuitives and Introverts (individually) have a 60-point advantage on overall SAT scores. The analysis of GPA with both the MCAS and SAT yielded expected results. The MCAS correlated with overall GPA in the 0.7 range (both with Pearson and Spearman rho coefficients) and the SAT correlated with GPA in the 0.5 range, also parametrically (Person) and non-parametrically (Spearman). While the SAT correlated with overall GPA more than we expected, the MCAS correlated to GPA even more than the SAT did, which we did predict. Using transcript data, we were able to compare academic track with MCAS scores (by section and overall), and also compared the averages of the Sensing and Intuitive students in each track. We found that as the difficulty of a student's curriculum decreased, so did MCAS scores, both by section and overall. Notably, students with the more difficult tracks had much better MCAS scores than those in the medium to low difficulty range. When Sensing and Intuitive MCAS scores were compared with academic track, it was found that Intuitive students had an advantage, especially at the higher difficulty academic programs (some AP, all Honors, and some Honors and some College). When SAT scores (subject and overall) were compared using academic track as the independent variable, we found the reverse trend of Tapley's results from the class of 1999 (Appendix G). While the two

most difficult academic tracks showed an Intuitive advantage on SAT scores, the other categories did not. In fact, the Sensing students seemed to have the advantage as course difficulty decreased in Fitchburg, this was not the case in Worcester. An analysis of the S/N difference by sex in relation to MCAS and SAT scores revealed an interesting finding. Only the Intuitives had an advantage on the MCAS for both males and females. No other MBTI parameter revealed MCAS differences between the two sexes. However, the SAT analysis between the sexes revealed different results. The female Intuitive students have about a 90-point advantage over female Sensors on overall SAT score while male Intuitives only have a 25-point advantage over the male Sensing students. This same trend can be found with E/I, where female Introverts have only a 50-point advantage over the female Extroverts on SAT score while male Introverts have a 100-point advantage.

Our most revealing finding was found when we did the MCAS/SAT relationship. Each section of the MCAS correlated with all other sections of the SAT (including overall scores) in the 0.7-0.9 range. In some cases, this is considered a statistical identity, accounting for 81% of the variance. Hence, we decided to include sections on policy debate in this report, as this finding was most striking and a significant issue for policy debate due to one test (SAT) being for aptitude and the other (MCAS) for achievement.

The Worcester data set revealed an Intuitive advantage on the MCAS in all subject areas when Intuitives were compared to Sensing students. Also, when academic track was compared with MCAS scores (by subject), it was found that as academic track increases in difficulty, a student's MCAS score in all three areas also increases. When comparing academic track with average MCAS scores in relation to individual MBTI

parameters, we found the same trend, with the Intuitives having less of a chance of failing than the Sensors.

One problem with the MCAS that might not be avoidable is the fact that it favors one group of MBTI types over another. In this case Intuitives have a consistent but small advantage over Sensing students. This poses a problem for the makers of the MCAS as the MCAS is only supposed to test a student's knowledge of the things he/she has learned in school, not their personality type. The problem is that making a totally unbiased standardized test is very difficult if it is possible at all. To compound this it is unknown to us whether the Intuitive students get their advantage from the questions being asked, or possibly they do better in general on any standardized test structured like the MCAS. If it is found that certain people do better on standardized tests then other, will all standardized tests be abolished? I think not. All test have certain biases built in them, the key to improving the MCAS is to identify theses biases and use them to help the students that are affected by them.

Research into the nature of the MCAS will continue for at least a few years, as the districts strive to bring their curricula into alignment with the "accountability" tests. In particular the challenge of predicting which students are most at risk for failure will continue and intensify. As policy changes in the coming years, it will nonetheless remain critical to work with them in a proactive fashion. One will also need to understand where students have difficulty, and if students can be grouped together according to what their learning styles are, it may help in their learning how to overcome their difficulties with the test format. What is more, constant research into the performance of students and schools on the MCAS allows for more critical assessments of the test leading to revisions

in format and policies about the consequence of failure. Better-informed decisions regarding district wide policy on MCAS preparation should be the result of these studies.

Learning style studies done at WPI, including this one, have covered college, high school, and elementary school populations. All of them have used the MBTI (or MMTIC) as a central standard to which the test and grade performance measures are related. In particular, the S/N and P/J preferences have formed the core of the most interesting results at all levels, proving to be excellent indicators of performance on aptitude and achievement based tests, and even on grades. But personality type is not the only factor involved in this research. Comparisons using a variety of demographic variables have been suggested or carried out, and various tests have been compared to each other, similar to our comparing the MCAS with the SAT and Plan ACT. It is also appropriate to use grades as a predictor of test performance.

7 Discussion

7.1 Policy Analysis

There are several different public policies currently being debated concerning the extraordinarily high failure rates of urban Mass high school students taking the MCAS. The most prominent is the issue of students retaking the test, and also there is great concern over the fact that a very high percentage of the minority did poorly on the test. To a lesser degree the question of whether vocational and special education children are to be required to take the test is also in question. Several new policies that are becoming active are a direct result of the high failure rates.

The current method that the state is using to deal with the high failure rates is to allow the students to retake the test at least 5 more times. This in itself is not a major policy change. The fact that when the student retakes the test many of the more difficult questions will be removed to reduce the length of these sessions and the stress associated with them, has caused many to oppose this. The rationale behind this is that the toughest questions would only demoralize students taking the retests, and that for those students who need to take the retest it makes sense to focus on basic skills not the "advanced" themes. While some feel that this cheapens the MCAS, this policy has come into effect and will be implemented in the near future as the class of 2003 takes the exam this year.

Schools, especially in the city, have turned to having teachers tutor students in an effort to increase the number of students that pass the test. City schools have received several million dollars in additional funding that has paid teachers to tutor children after school and work in summer-school programs. These programs are not only limited to

the city schools. Schools throughout the state have set up these programs. Some communities have already seen improvement in students' Math and English skills as a result of these intensive tutoring programs. These programs have become wide spread with 35,000, in 1999, students from 202 Massachusetts school districts participating in some form of remedial instruction. The Legislature has targeted $80 million to help students at risk of failing the MCAS.

One major issue that is currently being addressed is the fact that results by race on the 1999 test showed that Latino and black students lagged behind whites in every category on the test and for every grade. The difference is especially pronounced in high school where about 45 percent of white 10th-graders failed math, 80 percent of blacks did, and 85 percent of Hispanics failed. This information has sparking fears that students who do not pass the exam - required in 2003 to graduate - will become frustrated and drop out of school.

Outside programs as well as the tutoring discussed above are being used to combat this problem. This includes a large coalition of churches which are attempting to set up after school academic centers. There are as many as 40 of these centers planned in Boston. Many Latino and black leaders feel that if the education department really wants to help Latino and black children, it should do away with the MCAS graduation requirement until their scores improve to avoid the crisis and simply reward those who do pass with opportunities rather than punish those who fail.

One final current policy that has been adapted because of the large failure rates allows special education students to finish high school without passing the MCAS. But instead of receiving a diploma, students in this category would receive a "certificate of

completion". Advocates for special education children said the proposed "certificates of completion" send a dangerous message: that special education students aren't as competent, and it might as well be a certificate of "attendance".

7.2 Identifying Students at Risk

One of the main goals of this study was to find a way to identify students at risk of under performing on the MCAS test. Using the results gathered from the analysis from our data set certain correlations where found between the MCAS and other variables such as GPA, SAT score, MBTI, etc. By examining how strong these variables are correlated to the MCAS, the relative effectiveness of them predicting that student's MCAS score can be predicted. The most basic question was whether grades could be trusted as a guide to how the test would go. There was widespread doubt about this even among top administrators in Fitchburg. They had seen enough students with low averages pass and good students fail to wonder. However overall GPA is a good correlate. This was partially due to the impact of program difficulty on the GPA. Grades alone would not be as good. We wondered how well grades would work for students in the same program of study. However, major questions would be raised if grades were unrelated to test scores. Luckily, that was not true, at least where we studied the matter.

Based on our results several factors were found that would predict an MCAS score very well, and others that would be useful but not nearly as affective. When using these proxy tests to predict the MCAS the one that most closely does this would be most effective. The SAT turned out to highly correlate to the MCAS with the MCAS composite and SAT composite having a correlation coefficient of approximately 0.8.

With this said, if only one thing could be used to identify students at risk an early administration of the PSAT should do it. It would produce reliable results and could be very helpful to both the student and the teachers to know which students are at risk of performing poorly on the MCAS or vice versa.

As a practical matter however – the 8th grade MCAS and the students grades in 9th grade could be combined into a powerful predictor of how the 10th grade MCAS will go. With a correlation coefficient of approximately 0.7 for GPA and overall MCAS scores, while not as high as the SAT it is still very significant. GPA could prove to be the most effective predictor simply for the fact that it is readily available to almost any teacher or guidance councilor. The MCAS is first given in 10th grade, and because most students won't have taken the SAT yet, it wouldn't be available as a predictor, but the students GPA would be available and could be used. The value of the MCAS – SAT correlation will actually be its reverse, a warning about who will likely under or over perform (based on expectations grounded on grades) on the PSAT and SAT.

MBTI types also allow one to predict if a student is likely to under perform on the MCAS and in this case provides information on how best to approach those who will need extra help to pass. The MBTI SN -JP relationship is not nearly as large as the two previous variables, but there is a significant correlation in Worcester. This alone would not be a very reliable source to predict a student's performance on the MCAS. Because the MBTI can be given before the MCAS is given in the 10th grade it would be useful to include with the students GPA, and use both predictors to try to gauge a students performance on the MCAS. This would be a good test to use, both the GPA and the

MBTI, both can be found before the student takes the MCAS, unlike the SAT, and together could be a powerful tool for teachers.

7.3 Future of the MCAS

The future of the MCAS for at risk students is still unknown. With as many people campaigning against it, politicians may eventually waver in their support and the MCAS may be scraped as the state's sole or even main measuring stick of school success. But if this is not the case certain things could be done make the test more conducive to the at risk student.

Outside of changing teaching methods and varying after school help, changing the MCAS itself might be necessary. It is expensive, time consuming, given often, and does not provide much new information. Some students might benefit from being given a prep course a couple weeks before the MCAS. Test strategy alone might get a few of the near misses to pass the test. Being similar to SAT prep courses, such a workshop would go over test taking methods specific for the MCAS. While this would not solve many of the central issues of the MCAS, it might help minority students who may not be accustomed to taking a test similar to the MCAS, and especially specific types of learners who are not "natural" at this sort of cognitive task.

One major flaw that teachers seem to point out, this is especially relevant for the history section, is that the content in the MCAS has no depth. One thing that could be done would be to coordinate the test more closely with teachers, to get an understanding of what they are teaching and to modify the test accordingly. While different tests couldn't be made for each individual school, it would allow for a greater understanding of what the teachers feel is important.

Many of the students that fail the MCAS are students whose English isn't their first language, or especially where English isn't spoken in the home. For these students, a modified MCAS English section might be appropriate. Instead of covering specific components of grammar and punctuation, a more general understanding of the fundamentals of the English language could be tested.

There is no simple answer to the question of how to improve the MCAS for students at risk. While simplifying the test may increase certain student's chances for passing, is that really the goal of the MCAS? The state feels that all high school graduates should have the knowledge level tested in the MCAS. The reason why the MCAS is under such fire, despite widespread business support, might not be because it is a poorly made test, but the first test that REALLY matters. If the test itself does not have to change; the schools and teachers might have to change to meet the expectations of the test. Whether this kind of teaching to the test is a good thing or not is being hotly debated.

What our findings add to the debate is a question about whether the MCAS measures what people think it does, and whether it is not harder for some than others based on learning style – not preparation. In short is a one-size fits all test really the level playing field that the proponents of the MCAS claim it is.

8 Future Areas of Research

8.1 Centralization of Database

Throughout the process of data collection, this project team experienced many problems that limited our progress. First, when we asked the school systems for data, it was coded into a form that was unfamiliar to the average lay person, and when we contacted the school system itself, we realized that some of the people involved with sending the data could not always decipher the data either. Fitchburg's documentation had been lost, and they promised little. However, we got what they had, which included paper transcripts, and we created our own databases. Ironically, though its databases were better, this was also the case in the Worcester system. Prior groups struggled with misaligned databases and inconsistent formats, but we had the paper version to fall back on and there were only 150 cases per class year. It should be made clear that there is no placement of blame being made, because the information from CAPT and organizations that make the various standardized tests (College Board, Advanced Systems) is the output code in their own *asci* formats. It simply seems that there should be someone or some organization in charge of centralizing a database with all pertinent information by case every time new data is collected on an individual. It should then be put into a form that can be read by the school systems and various parties involved, particularly organizations and people that perform analysis and reports on that information. This means the case numbers used for each individual must to be preserved and used consistently. This would then make data collection and merging much more efficient and more analysis could be done in a given project. The current study was supposed to stress analysis of data but ended up being a data identification, acquisition, and organization effort for the most part.

The analysis was rushed, incomplete and still late due to delays and formatting problems on the front end. In the end we had to hire help to deal with the problem.

8.2 Automation of Analysis

What can be seen from any interaction with large data sets is that they are unwieldy, and when humans enter the picture mistakes occur frequently. Computers do not make "matches" without perfect or exact matches. Lapses in having last name first, truncated names, and identical names (as well as column alignment problems) all cut into the proportion of cases that could be linked. At one point, we despaired of getting 400 matches out of 800 cases. At about 40 cases per hour (hand-entering), we were able to save half of the cases in the class of 2000 cohort for the study. Somehow with 810 people in the database – only 675 had taken the MCAS, and 567 had filled out the MBTI. Some did not have complete transcripts hence the analysis was carried out on 432 cases. Given this experience we can attest that the problems associated with disparate data, it being in a variety of formats, locations, and keyed without one unifying identifier per student (repeat names happen regularly), are enormous. Getting that data into a centralized database that is easily accessed can be daunting, but once done, it makes the whole process of data merger easier. A greater problem lies on the other side of that database, however, in the form of analytical mountains to overcome.

Often times there are spot analyses which the researcher would like to have access to at will for use in reporting, researcher, policy, etc. These analyses are typically very time consuming, but represent merely replications of analyses done before on other data. For instance, year to year recalculating Pearson correlation coefficients for MCAS aggregate scores versus S/N or P/J preferences is something that is difficult to work out

the first time it is done, but thereafter represents only doing the analysis over with the same process, just new data. In principal, if the variable names were standardized from year to year the analysis could be preserved and doing the new analysis could be automated. Since time, money, and human resources are often not on the researcher's side in being thorough by doing such replication studies, they often go undone in favor of new ground breaking work. The weakness of this strategy is that changes in the assumptions from those original, non-replicated studies may have taken place over the years. If the new ground breaking work is based on those assumptions, the new work may be invalid and time and effort wasted.

Working out a solution to that problem is technically demanding, though infinitely rewarding and useful in its application. Like the database centralization described above, once a way is developed to replicate studies automatically, with little human effort, then that standardized analysis can occur for years without interruption. The suggestion made is conceptually simple. Assign a server, most likely not the same server that handles the actual database caching and serving, to do replication analysis. Develop command scripts for the operating system to call analysis functions if a specific type of data is updated in the database, and then develop command scripts for SPSS, Excel, or even write small dedicated programs to handle analysis. Output the analysis results to data files for long term storage and study, and print pieces that are predetermined to be important. An example follows...

Assume it is January, 2003, and the May/June, 2002 MCAS data is finally available to the schools for their use. The researchers studying MCAS and the MBTI are interested in the same relationships as always, but want to expand on them and go new

places. But doing the old and new analysis would take time. What they do is take the MCAS file, and add it to the centralized database. The database program parses the file, merges it correctly, flags records that failed to merge, and updates the database itself. The researchers then add in the MBTI file to the database. The database program parses the file, determines that it is standard MBTI format, fixes as many names as it can, merges those names in, and flags failed names for human intervention. The database is then updated with less than 10% of the handwork currently required.

The analysis server, which constantly checks the state of the database, finds that the MCAS and MBTI for the year 2002 were added in, and as such one of the functions of the analysis server is to replicate earlier studies involving MCAS and MBTI once these variables exist. It pulls in the 9^{th} and 10^{th} grade transcript data - calculates GPA by subject and overall and the relative difficulty of the courses taken by subject and overall. The server calls SAS or SPSS, loads up the MCAS/MBTI fields it needs from 2002, and proceeds to do descriptive statistics, find correlations, and attempt to form predictions using regression on the basis of the Intuitive advantage. All of this output is stored for future use, and the pieces that wish to be viewed by researchers are printed to a standard format and labeled for when researchers pick them up to add to an in-house paper file. The server then parses the results, finds significant correlations, and emails a predetermined list of researchers the news. Lastly, it passes the location of the new analysis file back to the central database, which stores it as the "N/S – MCAS/MBTI Replication Study" field associated with 2002 data.

This example highlights the benefits of not only the centralization of the database, but also the automation of repetitive analysis tasks. Such a system would pave the way

for more ease of doing longitudinal analysis, as well as the automated update of all analyses that need to be when new data are added. By printing very little of the output, and emailing researchers only when relationships are significant deviation from prior findings – or very robust replication (according to the researchers predetermined definition) no time or effort is wasted, and new work can be done with the researchers time. The sizeable difficulties in implementing this system are apparent in the need for the creators of the system to be knowledgeable about many things. Specialists in Access, or some other database handler such as StarBASE, must design system components, as well as specialists in server development, be it with free systems such as Linux or FreeBSD, or commercial systems such as NT. Applications developers familiar with Excel, SPSS, SAS, or self-written comparable software would be called on to develop the analysis toolkits. Lastly, an overseer familiar with project management would need to make sure everything worked well together. But should such a system be developed, it would not only be useful at WPI for research purposes, but could be exported for a licensing fee to school districts as a data management system. Districts which are understaffed already may welcome such a backup or support system in their own work.

8.3 Future IQP's and Possible MQP's

From our research and brainstorming, this study could give rise to other possible IQP's and possibly MQP's if a student were to come along that was involved in social science and policy studies as a major. For instance, it is obvious that this study could be replicated every year from now on as long as the MCAS and MBTI are used in Worcester since we have been able to compile an extensive and readable database for statistical analysis. Also, the state of Massachusetts will start counting the MCAS this year as a

requirement for graduation. Therefore, it is our opinion that the MCAS will be heavily studied for years to come. Other projects in the very near future could focus on why the MCAS and SAT/PSAT scores are so similar when in theory the tests are so different. These studies would answer many questions that we did not have time to answer. For example, how accurate and precise is the SAT/PSAT as a predictor of MCAS score? Also, are students getting high or low scores on the MCAS for the same reason(s) that they are getting the same relative scores on the SAT/PSAT? Is it that the PSAT is more of an achievement test than people think? Is the MCAS not really testing what the students are studying?

Further, other projects could pursue a more in-depth study of the policy issues that are caused by this project's results. What are the respective institutions doing to improve the quality of the tests? What are Massachusetts's individual school systems doing to improve curriculum and what will they do with type identification? It was actually very surprising to see that many questions arose from the results of our study, when our project covered so many areas and tried to relate the MCAS and MBTI relationships to previous studies involving standardized tests of aptitude. There may always be research to do in this area, since this project has pioneered the MCAS and its relationship to many parameters, and that is part of a continuing debate. Also, there will be research and statistical analysis to be done as long as tests such as the MCAS, SAT, PSAT, Plan ACT, and other standardized examinations involving both aptitude and achievement are administered. Keeping up with the growing database is important, but that must not become the dominant part of the project as it did in our case. We want to see things focused on the growing suspicion that the "intuitive" advantage is generic to that type of

testing – and cannot be avoided by "fixing" the test.

9 Project Evaluation

The Interdisciplinary Qualifying Project (IQP) is designed for students to do research in an area which integrates science and technology in society. It is the opinion of this research team that the IQP differs from the Major Qualifying Project in two major ways. The first is that more times than not, the IQP is done outside of a student's major. The second, and most important, is that the IQP seeks to explore a society-technology issue or social problem in society, which relates to some sort of policy issue. This issue usually has something to do with a social or political system (involves classical or social science) and the educational system, which is the focus of this project. The MQP, on the other hand, seeks to pioneer an area of scientific or engineering research which has had little to no consideration paid to it. This is not always the case, but in this day in age, where the university is trying to keep up with technological advances, it is more efficient and recognizable to be part of a development team than a problem solving (quality) team.

It is with these thoughts in mind that this evaluation of the MCAS/MBTI project that we write down what this project has done to heighten the skills that both do and do not apply to the career paths we have or will choose in the future. All of the members of the team were able to perform all of the relevant skills involved with database creation and statistical analysis. We gained experience in data handling as well as negotiating with school system officials for data that was not necessarily available for public viewing. Also, while the three of us had taken a statistics and math course before, this project was an actual application of that theory to a real-life problem in which many things were at stake for the schools, teachers, and most of all, the students. We see the MCAS and SAT as devices, or instruments (technologies), designed to assess a social

system, which involves design choices, or tradeoffs. Finally, it provided experience in writing an extensive and detailed report of our findings, so that someone who was not familiar with the situation at all would understand what we intended to do, what we did do, and what we suggest others do in the future to make this research even more meaningful.

However, this project was not always a pleasurable experience. It was very hard to obtain data in a useful form from the school systems. Time was an issue as the school officials had other responsibilities and higher priorities than catering to our needs on the spot. Hence, we did not always get the idea that our time and contributions were valued. We had brief windows in which to help out, and were not on the payroll, so waiting was not making us money, nor making progress. It was hard to be patient and then expend our time fixing dozens and dozens of database errors others had made, and not wonder if we would ever get to use our real skills. Other personal responsibilities on the part of the team members also hampered efforts to meet some deadlines. However, with any research, not everything goes according to schedule, and the manner in which we adapted to these obstacles was a large part of the success of the project. In the end, we made progress, contributed, and answered most of our questions. The next project team will hopefully start in a much better place than we did, and can benefit from out missteps and errors as well as our successes.

BIBLIOGRAPHY

An ACT Program For Educational Planning. (2000, September). *ACT Assessment® Homepage* [Online]. Available on the World Wide Web: <http://www.act.org/aap/index.html>

Batey, Danielle et al. "Cognitive Bias In The Scholastic Aptitude Test." Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1995.

Crouse, James et al. The Case Against The SAT. Chicago: The University of Chicago Press, 1988.

Kingsland, David et al. "Learning Style Biases In The Scholastic Aptitude Test." Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1995.

Mangenot, Gerard C. and Mario L. Tongol. "The Leicester SAT Experience." Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1998.

Massachusetts Comprehensive Assessment System. (2000, September). *Massachusetts Department of Education* [Online]. Available on the World Wide Web: <http//www.doe.edu.mass/ mcas>

McElhaney, Michael F. "MBTI Data And Its Application To MCAS." Unpublished Interactive Qualifying Project Report, Worcester Polytechnic Institute, 1998.

Pieper, John B. "Cognitive And Learning Style Biases In The SAT/PSAT." Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1997.

Wirtz, Willard et al. On Further Examination: Report Of The Advisory Panel On The Scholastic Aptitude Test Score Decline. New York: College Entrance Examination Board, 1977.

APPENDIX A

PLEASE TAKE A FEW MINUTES TO ANSWER THE FOLLOWING OPTIONAL QUESTIONS.

WHICH DO YOU LIKE BEST? (PICK ONE)

- Art
- Biology
- Business
- English
- Foreign Languages
- Geography
- History
- Home Economics
- Industrial Arts/Shop
- Mathematics
- Military Science
- Music
- Reading
- Phys Ed and Sports
- Physical Sciences
- Social Sciences
- Theater
- Other
- None

HANDEDNESS

- Left handed
- Right handed

RELIGIOUS ORIENTATION

- Adventist
- Baptist, Southern
- Baptist, Other
- Church of Christ
- Church of Christ, Scientist
- Church of Christ, United
- Episcopal
- Jehovah's Witnesses
- Jewish, Conservative
- Jewish, Orthodox
- Jewish, Reform
- Latter Day Saints
- Lutheran
- Mennonite
- Methodist
- Muslim
- Orthodox
- Pentecostal
- Presbyterian
- Roman Catholic
- Unitarian Universalist
- Other Christian
- Other non-Christian
- None

RACE GROUP (PICK ONE)

- African American/Black
- Asian/Pacific Islander
- Caucasian/White
- Native American/ Alaskan Native
- Other

HISPANIC ETHNICITY

- Non-Hispanic
- Hispanic
  - Mexican/Mexican-American
  - Puerto Rican
  - Cuban
  - Other

INCOME GROUP

- Lower Income
- Middle Income
- Upper Income

MARITAL/ RELATIONSHIP STATUS

- Single/ Never Married
- Separated
- Divorced
- Living Together
- Widowed
- Married

GENDER OF PARTNER

- Male
- Female

NUMBER OF TIMES MARRIED

0 1 2 3 4 5 6 7 8 9

YEARS IN CURRENT MARRIAGE/ RELATIONSHIP

00 11 22 33 44 55 66 77 88 99

AGE AT FIRST MARRIAGE

00 11 22 33 44 55 66 77 88 99

I was raised primarily in the USA	Yes	No
My parents were raised primarily in the USA	Both	One / None
My fluency in English	Not Much	OK / A lot

FREQUENCY OF RIGOROUS PHYSICAL EXERCISE

- Once a day or more
- A few times a week
- Once a week
- Once a month or less

I CONSIDER MYSELF A VEGETARIAN

Yes / No

SMOKING:

- Smoke now
- Quit smoking
- Never smoked

MARK THE POINT ON THE SCALE THAT FITS YOU BEST:

w I am	SAME AS WHEN YOUNG 7 6 5 4 3 2 1		VERY DIFFERENT
ppiness	VERY HAPPY 7 6 5 4 3 2 1		VERY UNHAPPY
uccess	GREAT SUCCESS 7 6 5 4 3 2 1		TOTAL FAILURE
joyment	ENJOY EVERYTHING 7 6 5 4 3 2 1		ENJOY NOTHING
ess	NONE 7 6 5 4 3 2 1		EXTREME DISTRESS

P CODE HERE YOU LIVE

0 0 0 0 / 1 1 1 1 / 2 2 2 / 3 3 3 / 4 4 4 4 / 5 5 5 / 6 6 6 / 7 7 7 7 / 8 8 8 / 9 9 9

AGE AT FIRST JOB

00 / 11 / 22 / 33 / 44 / 55 / 66 / 77 / 88 / 99

NUMBER OF YEARS IN CURRENTJOB

00 / 11 / 22 / 33 / 44 / 55 / 66 / 77 / 88 / 99

NUMBER OF HOURS PER MONTH SPENT VOLUNTEERING

000 / 111 / 222 / 333 / 444 / 555 / 666 / 777 / 888 / 999

IN TIMES OF PERSONAL CRISIS ARE YOU MORE LIKELY TO:

SEE
- just a few things to the exclusion of everything else

OR
- everything from the trivial to the serious as equally important

TELL
- at least one person

OR
- no one

WANT
- to be alone

OR
- to be with someone

SEEK COUNSELING

Yes / No

HOW MANY NEW PEOPLE HAVE YOU MET OUTSIDE OF WORK IN THE PAST MONTH?

00 / 11 / 22 / 33 / 44 / 55 / 66 / 77 / 88 / 99

HOURS PER MONTH SPENT PLEASURE READING?

000 / 111 / 222 / 333 / 444 / 555 / 666 / 777 / 888 / 999

MARK THE POINT ON THE SCALE THAT FITS YOU BEST:

w I am at work	SAME AS I AM AT HOME 7 6 5 4 3 2 1		VERY DIFFERENT FROM HOME
adition	IS VERY IMPORTANT 7 6 5 4 3 2 1		NOT VERY IMPORTANT
ideas are	VERY CONSERVATIVE 7 6 5 4 3 2 1		VERY LIBERAL
behavior is	VERY CONSERVATIVE 7 6 5 4 3 2 1		VERY LIBERAL
y political views are	VERY CONSERVATIVE 7 6 5 4 3 2 1		VERY LIBERAL
awareness of time	KEENLY AWARE 7 6 5 4 3 2 1		OBLIVIOUS
el understood by others	ALWAYS 7 6 5 4 3 2 1		NEVER
mfort with computers	TOTAL 7 6 5 4 3 2 1		NONE
ical instruments	PLAY NOTHING 7 6 5 4 3 2 1		PLAY AT LEAST ONE EXTREMELY WELL
nge	ACTIVELY SEEK IT 7 6 5 4 3 2 1		WORK TO KEEP THINGS THE SAME

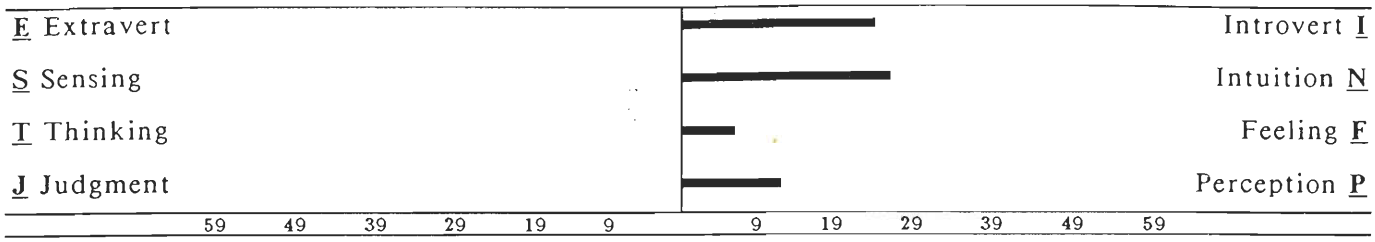
Mark Reflex® by NCS EM-155210-2:65432 ED99 Printed in U.S.A.

Myers-Briggs Type Indicator

Detailed Scoring Information

Professional report for
John Wilkes

MBTI resul
ERICA I



INFP - Introverted Feeling with Intuition

FULL SCALE RESULTS (*Parts I,II and III of the MBTI*)

		E I Direction of Energy	S N Perception Function	T F Judgment Function	J P Extraverted Attitude
INFP	Preference Scores:	I 25	N 27 Auxiliary (Extraverted)	F 7 Dominant (Introverted)	P 13

Total Points			
E	9	I	21
S	6	N	19
T	10	F	13
J	9	P	15

PHRASE QUESTIONS (*Parts I and III of the MBTI*).
These questions ask for responses to everyday events and may therefore be more influenced by the demands of these events.

INTJ Preference Scores: I 15 N 23 T 3 J 1

Phrase Points			
E	8	I	15
S	1	N	12
T	5	F	3
J	8	P	7

WORD PAIR QUESTIONS (*Part II of the MBTI*).
Word pairs are less affected by everyday events and may be nearer to true preferences.

INFP Preference Scores: I 11 N 5 F 11 P 15

Word Pair Points			
E	1	I	6
S	5	N	7
T	5	F	10
J	1	P	8

Answer Sheet Information

Sex	Female
Computed Age is	16
Highest grade completed	11
Are you a student?	
Likes _ best	
Are you working?	
Date of birth	08/04/81
Form G answered on	04/29/98
Total items available	126
Last item answered was	126

Continuous Scores

	Full Scale	Phrase	Word Pair	X Half	Y Half
EI	125	115	111	121	105
SN	127	123	105	115	113
TF	107	97	111	103	105
JP	113	99	115	101	113

Group Code	7QI -98
Case Id	#3999-79
Date Scored	6/ 5/98

Item Omissions

EI	scored items omitted	0
SN	scored items omitted	0
TF	scored items omitted	0
JP	scored items omitted	0
Subtotal of omissions		0
Research items omitted		0
Total omissions		0

APPENDIX B

September 25, 2000

Patricia Mostue, Ph. D.
Testing and Assessment Specialist
Worcester Public Schools
20 Irving Street
Worcester, MA 01609

Dear Dr. Mostue:

This letter is in regards to a project that we are beginning at the Worcester Polytechnic Institute. It is a comparative analysis of learning styles with MCAS scores to see if one MBTI personality type has more or less of an advantage on the MCAS than another. We will need some assistance from your school system if we are going to have a large enough sample to perform a reliable study. We would like to set up a meeting with you and your colleagues to talk about the available data you may possess for an achievement-aptitude comparison which can be measured by grades and tests.

We estimate that there will be a potential pool of about 800-900 students from the class of 2000 who took the MBTI when they were 10th graders in the Worcester and Fitchburg public school systems. Our goal is to produce a working database on at least 700-750 of these cases, and at the same time focus on the relationship of aptitude versus achievement.

We are extremely interested to see if the MCAS is actually an achievement test, or if there are some elements of the exam which suggest it is an aptitude test, like the SAT. Previous studies have shown that both the SAT and PSAT display indications that achievement plays a role. We want to see if there is that same bias built into the MCAS, only that the MCAS tests achievement according to high school curriculum. Are there elements of the MCAS that are geared towards aptitude?

Another goal is to make many statistical correlations with the data. We are not just going to compare learning style data to MCAS scores. We also want to correlate MCAS with SAT score, learning style with SAT score, learning style with overall grade point average, MCAS and SAT with overall grade point average, and if there is enough time and we receive the data, we want to make correlations with the Fitchburg Plan ACT examination. All of these correlations will help us find out which learning styles determined by the MBTI are at risk to underperform on the MCAS. We then want to meet with each school system, discuss our results, and give suggestions as to how to identify these "at-risk" students for the MCAS so that they can receive a high school diploma.

In order to complete our study, we will need some data from both Worcester and Fitchburg. We currently have MBTI data for the class of 2000, which was collected when they were 10th graders for both systems. We need the following:

1. MCAS data for the class of 2000 when they were tenth graders
2. Overall grade point average for these same students.
3. The most recent SAT data before their graduation.
4. If available, ACT data from any students who took this test or data from another well-known achievement test.

These correlations and means of data analysis are not set in stone. However, we are very direct in finding a correlation between achievement and aptitude as well as establishing a working database that can be utilized long after this project is completed. If there is anything that you as a school system would like us to look at, we would be more than happy to do so provided we have enough notice and it is of our time frame. We want to help with this problem that the MCAS is causing the state of Massachusetts. We are especially interested in Worcester and Fitchburg because of the 66% failure rate that the last MCAS predicted. We feel that this project will at least educate the guidance counselors in each school system as to which students are at risk to get lower scores on the MCAS due to their learning style as determined by the MBTI. We want to eventually set up meetings with Worcester and Fitchburg to discuss our findings and also discuss solutions on how to treat the "at-risk" students.

We expect that this project will last from August 31, 2000 until about the middle of January. We will need the data before this to establish a database which contains both MCAS and MBTI data. Previous studies have shown that the SAT and PSAT, which claim to be aptitude tests, contain elements of achievement, especially for the sensing and intuition traits as determined by the MBTI. Therefore, we are persuaded that we will find that there are some students who are at risk to underperform on the MCAS because of their learning style. After this project is complete, you as a school system will know whom these students are and can then have strategies in place to undermine learning style.

We would like to thank you in advance for your cooperation. If you would like to contact us to set up a meeting or to ask some questions, we can be reached at 508-284-9176 (Ken) or 508-831-6451 (John). Again, thank you and we look forward to hearing from you concerning this project.

Sincerely,

Kenneth Fountain

John Stambaugh

September 25, 2000

Thomas Lamey
Assisstant Superintendent
Fitchburg Public Schools
376 South Street
Fitchburg, MA 01420

Dear Mr. Lamey:

This letter is in regards to a project that we are beginning at the Worcester Polytechnic Institute. It is a comparative analysis of learning styles with MCAS scores to see if one MBTI personality type has more or less of an advantage on the MCAS than another. We will need some assistance from your school system if we are going to have a large enough sample to perform a reliable study. We would like to set up a meeting with you and your colleagues to talk about the available data you may possess for an achievement-aptitude comparison which can be measured by grades and tests.

We estimate that there will be a potential pool of about 800-900 students from the class of 2000 who took the MBTI when they were 10th graders in the Worcester and Fitchburg public school systems. Our goal is to produce a working database on at least 700-750 of these cases, and at the same time focus on the relationship of aptitude versus achievement.

We are extremely interested to see if the MCAS is actually an achievement test, or if there are some elements of the exam which suggest it is an aptitude test, like the SAT. Previous studies have shown that both the SAT and PSAT display indications that achievement plays a role. We want to see if there is that same bias built into the MCAS, only that the MCAS tests achievement according to high school curriculum. Are there elements of the MCAS that are geared towards aptitude?

Another goal is to make many statistical correlations with the data. We are not just going to compare learning style data to MCAS scores. We also want to correlate MCAS with SAT score, learning style with SAT score, learning style with overall grade point average, MCAS and SAT with overall grade point average, and if there is enough time and we receive the data, we want to make correlations with the Fitchburg Plan ACT examination. All of these correlations will help us find out which learning styles determined by the MBTI are at risk to underperform on the MCAS. We then want to meet with each school system, discuss our results, and give suggestions as to how to identify these "at-risk" students for the MCAS so that they can receive a high school diploma.

In order to complete our study, we will need some data from both Worcester and Fitchburg. We currently have some MBTI data for the class of 2000, which was collected when they were 10th graders for both systems. We need the following:

1. MCAS data for the class of 2000 when they were tenth graders
2. Overall grade point average for these same students.
3. The most recent SAT data before their graduation.
4. If available, ACT data from any students who took this test or data from another well-known achievement test.

These correlations and means of data analysis are not set in stone. However, we are very direct in finding a correlation between achievement and aptitude as well as establishing a working database that can be utilized long after this project is completed. If there is anything that you as a school system would like us to look at, we would be more than happy to do so provided we have enough notice and it is of our time frame. We want to help with this problem that the MCAS is causing the state of Massachusetts. We are especially interested in Worcester and Fitchburg because of the 66% failure rate that the last MCAS predicted. We feel that this project will at least educate the guidance counselors in each school system as to which students are at risk to get lower scores on the MCAS due to their learning style as determined by the MBTI. We want to eventually set up meetings with Worcester and Fitchburg to discuss our findings and also discuss solutions on how to treat the "at-risk" students.

We expect that this project will last from August 31, 2000 until about the middle of January. We will need the data before this to establish a database which contains both MCAS and MBTI data. Previous studies have shown that the SAT and PSAT, which claim to be aptitude tests, contain elements of achievement, especially for the sensing and intuition traits as determined by the MBTI. Therefore, we are persuaded that we will find that there are some students who are at risk to underperform on the MCAS because of their learning style. After this project is complete, you as a school system will know whom these students are and can then have strategies in place to undermine learning style.

We would like to thank you in advance for your cooperation. If you would like to contact us to set up a meeting or to ask some questions, we can be reached at 508-284-9176 (Ken) or 508-831-6451 (John). Again, thank you and we look forward to hearing from you concerning this project.

Sincerely,

Kenneth Fountain

John Stambaugh

APPENDIX C

ACT™-SAT I SCORE COMPARISONS

ACT COMPOSITE SCORE	RECENTERED SAT I SCORE (VERBAL AND MATH)
36	1600
35	1580
34	1520
33	1470
32	1420
31	1380
30	1340
29	1300
28	1260
27	1220
26	1180
25	1140
24	1110
23	1070
22	1030
21	990
20	950
19	910
18	870
17	830
16	780
15	740
14	680
13	620
12	560
11	500

Source: College Board Online. (2000, September). *SAT I-ACT Score Comparisons* [Online]. Available on the World Wide Web: <http://www.collegeboard.org/index_this/sat/cbsenior/html/stat00f.html>

APPENDIX D

Page missing in original

IQP/MQP SCANNING PROJECT

George C. Gordon Library
WORCESTER POLYTECHNIC INSTITUTE

APPENDIX E

Page missing in original

IQP/MQP SCANNING PROJECT

George C. Gordon Library
WORCESTER POLYTECHNIC INSTITUTE

APPENDIX F

the SAT.

Fitchburg (N = 80)

	Average SAT Score	Average ACT Score
AP or some AP	1226	25.8
Honors or some Honors	999	21.3
College	929	19.0
General	765	15.8

Table 7.7: Overall Fitchburg ACT and SAT Averages, by Track

These Fitchburg high school results are fairly representative of their overall population, and reflect the expected trend in track performance (as seen in earlier studies). At the middle levels (honors and college tracks), there seems to be a greater spread on the ACT than on the SAT, supporting their claims of a curriculum basis for the test.

Leicester (N = 28)

	Average SAT Score	Average ACT Score
AP or some AP	1020	23
Honors or some Honors	1069	24.3
College	942	19.0
General	--	--

Table 7.8: Overall Leicester ACT and SAT Averages, by Track

In Leicester our results were very skewed, both by the small sample size and its selected nature. There were only 2 students in the AP or some AP track, 7 students in the Honors or some Honors track, and 13 students in the college track. While the overall trend of increased track equaling higher test score performance holds true at the middle levels, the small sample for the top track (AP) doesn't show that result. Notice, however, that the SAT trend is also seen in the ACT for these middle tracks.

8.1 SAT Analyses using the MBTI

Fitchburg (N = 73)

TRACK	AVERAGE SAT SCORE	STUDENT TYPE AND NUMBER	SAT SCORE BY TYPE	"INTUITIVE ADVANTAGE"
AP or some AP	1226	Intuitive (8 Students)	1249	82 points
		Sensing (3 Students)	1167	
Honors or some Honors	999	Intuitive (10 Students)	1051	87 points
		Sensing (15 Students)	964	
College	929	Intuitive (16 Students)	995	123 points
		Sensing (19 Students)	872	
General*	765	Intuitive (1 Student)	770	10 points
		Sensing (1 Student)	760	

Table 8.1: Fitchburg SAT by track and S/N dimension (* This cell is underpopulated and was included in the table for completeness; it is disregarded as a finding in the written analysis.)

It's clear that at Fitchburg High School there is an intuitive advantage on the SAT. When compared to the earlier studies we see that there is more of an intuitive advantage at the AP level, this may be due to the larger number of intuitives there. At the general level one can't draw any conclusions because of the small sample size.

In the middle groups at Fitchburg we see the intuitive advantage appear at roughly the same levels as prior studies and a bit stronger than in Worcester. The Worcester findings from the Pieper-McCormick study, in particular, are replicated in this subsample. This is encouraging because there is enough ACT data available for Fitchburg to do a solid comparison between the two tests.