01A036I

Project Number: 46-JMW-MSAT

PRACTICING AND RE-TAKING THE SAT BY MBTI TYPE

An Interactive Qualifying Report

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in the partial fulfillment of the requirements for the

Degree of Bachelor of Science

By

Ben Dean-Kawamura

Date: October 24, 2001

Approved:

Professor John M. Wilkes, Major Advisor

Table of Contents

Executive Summary

The Scholastic Assessment Test (SAT) is one of the most prominent tests given in the United States. It's goal seems simple: predicting how well a high school senior will perform in their first year of college using a test of verbal and mathematical reasoning. However, the SAT has been a long-standing subject of controversy. Critics of the test question the validity of using a standardized test to predict academic potential and furthermore they claim the SAT is biased against minorities, women and the working class, hence the significance given to it unfairly disadvantages these groups.

This project is a continuation of the work of about 8 WPI teams that have studied another possible source of bias in the SAT: one based on personality type. These teams administered the Myers-Briggs Type Indicator (MBTI) to high school students and compared the SAT and PSAT scores of different types. They found that certain "Personality Types" consistently scored higher than their counterparts. Since personality type is not a measure of native intelligence, this is evidence that the SAT may indeed be biased for and against different types of learners

This continuing research initiative is important because the SAT has an impact on virtually every high school student who has aspirations of going to college. It is the most commonly used measure that college admissions offices employ to estimate a student's "potential." Moreover, students use the SAT themselves in order to gauge what colleges they have a chance of being accepted into and therefore what range of colleges they can reasonably consider. This makes the SAT incredibly influential in the transition from high school to college in the United States. Furthermore, it is a fine example of a technology with direct social implications that is often not recognized as such because it is not a machine. The SAT is technology in the same sense that contour farming is technology: the effect comes from skillful use of existing implements. Since colleges use the SAT as a means of sorting and assessing prospective students, learning what one can and cannot legitimately say about someone based on SAT scores, critically assessing the claims of the college board, is crucial in promoting a fair and effective higher educational system.

The previous teams who studied the effects of personality type on SAT scores had done a large amount of work before this project began. Two pilot studies and one large-scale study had been conducted over the span of several years, resulting in the collection of personality type data and SAT/PSAT scores for about 75% of the high school students in the Worcester Public Schools (Class of 1996-1999). Several surrounding school districts have also been involved in this series of studies, some to provide suburban contrast to the urban data, some because they administer the ACT as well as the SAT with regularity. In addition, data was collected on the students' socio-economic status and academic transcripts. The transcript data had been analyzed in some depth, by the teams themselves and also outside contributors to correlate student academic program with PSAT performance and to study the makup of the different programs. Analysis of the social class data (coded by occupational prestige of the parent or guardian) was more uneven. Still, the major effects of personality type on SAT/PSAT test taking had been fairly well established.

Hence, when this project started, there were several problems with the collected data that were preventing the full scale conclusive analysis the data set could in principle support - but there was a good basis for theorizing about what the results would be based on partial studies and pilot analyses. Different teams had gathered and assembled the different data sets, which led to the information being distributed among several files instead of in a central location. Also, many students had taken their PSAT and SAT tests in the years after the studies had collected personality type data. These students were new cases that could be included in the analysis if their SAT and PSAT scores could be located. Lastly, many of the data sets had gaps - missing data for certain variables and the follow-ups of cases absent on the day that the data were collected at their school (the follow-up had been done, but the new data had not been not merged into the original data set). Fixing these holes would mean many more cases could be used in the analyses, and the resulting data set would be more representative.

This was the state of the things when this project began. A large amount of data had been collected, but the data had never been merged into a single data set; recent SAT and PSAT test scores had not been collected; and the data sets contained various holes that could be filled by getting transcript data. If these problems could be fixed then a very robust data set could be created to study the effects of personality type on the SAT. One of the main focuses of this project was to do the work needed in order to create that data set so that this project and future ones could make use of it, and then to pick a worthwhile topic to explore personally.

After the data set was created, I used it to examine the how personality type related to the practice effect, test-retest improvement on the SAT. The team of Jennifer Sauvron and Emily Ballinger first raised this question of test-retest, but at the time they did not have a data set capable of supporting a thorough analysis.

An important finding was that “Intuitives” (students who prefer to think abstractly) were significantly more likely to take the SAT, and PSAT, than “Sensors” (students who prefer to think concretely). It seems to me that Sensors were being systematically discouraged from aspiring to try for a college education, a position that Isabel Myers, author of the MBTI, came to after studying other kinds of high school data in a Philadelphia suburb in 1958. When one takes into account previous findings that Intuitives score considerably higher on the SAT, ACT and MCAS tests, an explanation for the Sensors’ lack of SAT participation becomes apparent: they come to perceive themselves as not being “college material” because many measures of academic ability are biased against them. Also, it seems Sensors lower test scores become a self-fulfilling prophecy. They take less challenging high school programs in the junior and senior year and, because of this, really are less well prepared for college.

Another finding was that while "Introverts" (people who prefer the inner-world of thought and reflection) score higher on the initial PSAT than "Extraverts" (people who prefer the outer-world of things, people and action), they don't improve as much on retests. Extraverts were able to substantially raise their scores through retesting more often than "Introverts". This means that the practice of allowing students to use their highest scores gives more advantage to Extraverts than to Introverts, and if students are encouraged to try again, the Extravert-Introvert difference can be minimized.

These findings, along with others detailed in my analysis, indicate that personality type not only affects one's final score, but also many parts of the process that one goes though to get that score. The experience of SAT test taking varies considerably for different types of learners.

This paper consists of three main parts: An overview of the research that preceded this project; a description of the work done to merge together and supplement the data sets created by previous teams; and finally an analysis of the question of test-retest issue using the resulting data set. I hope that it impresses upon the reader how many factors affect SAT test taking and, in general, how complex it is to judge a student's ability to benefit from a college education using a standardized test that gets at qualities that come naturally for some and only with difficulty to others. A one test-fits-all-types approach is questionable on its face and becomes more so when you look into the details.

Introduction

Finding a college student who has not been affected by the Scholastic Assessment Test (SAT) is a near impossible task. A student's SAT score is one of the most important numbers he or she will receive, since it is one of the most commonly used measures by college admissions offices to determine acceptance. Colleges use the SAT because it allows them a standardized way of comparing students as opposed to high school grades, for instance, which are affected by standards that vary from school to school.

Although the SAT is not a machine, it is a technology since it is a device that is used to assess academic potential, and the significance given to the SAT means that it has direct social consequences. Thousands of high school students every year to spend a significant amount of their time preparing for, taking and retaking this highly consequential test. Furthermore, it represents an attempt to set up a "meritocracy" in the United States, where people are rewarded in accordance with their intelligence. High scorers on the SAT are able to get into more selective colleges and universities, are more likely to be offered scholarships and hence they are better able to secure higher paying jobs and higher social status.[1]

This project is a study of how different "personality types" (among the high school class of 1996-1999 in Worcester) went through the process of SAT practice, test and retest. However, before the study can be explained some background must be covered.

The SAT has been the center of much controversy over the years. Many critics have argued that the SAT is biased against minorities, women and the working class and the importance that is placed on it by admissions offices in the most selective schools puts these groups at a disadvantage when applying

[1] Lemann, Nicholas, The Big Test (New York, NY: Farrar, Straus and Giroux)

to prestigious colleges. The debate over bias based on race, gender and social class has been a long-standing one in the United States.

This study is the continuation of a series of eight projects done at WPI to examine another potential source of bias in the SAT, a bias against certain personality types and learning styles. These prior researchers administered the Myers-Briggs Type Indicator (MBTI), a standard measure of personality type, to over 1000 high school students in the Worcester area. Then the SAT and PSAT (Preliminary SAT, a test taken by sophomores and juniors to prepare for the SAT) scores of different Psychological Types were compared. The teams consistently found that there were significant score differences between different learning styles. The largest of these differences correlated with the Sensing-Intuition dimension of the MBTI, with "iNtuitives" scoring over 100 points higher than "Sensors," on average

The past WPI projects have collected data on many variables for the high school classes of 1996-1999 in the Worcester Public Schools (WPS). Different teams have collected archive data on SAT/PSAT score; administered instruments to obtain MBTI type, GCSI type and socio-economic status; and gleaned academic preparation (the average level of difficulty of a student's courses) from student transcripts. One problem however, was that these projects had different methods of data collection, so it has been difficult to link the data collected by the various research teams.

A large part of this project was presenting an overview of the past findings. I reviewed and summarized the contributions of eight WPI teams (and outside contributor Keith McCormick), covering their hypotheses, collected data and results. The story behind these projects is complicated at times, however it is necessary in order to fully understand their findings, which were the basis for this project and some of the vagaries of the data set.

The second major part of this project was to perform the linking needed between the past team's data sets in order to put together a single data set that contained all the relevant information collected so far. This would allow future teams analyze data collected as part of different projects. Also, since all of the classes have graduated it was possible to get completed versions of their SAT scores and transcripts. This allowed the original data sets (which only contained the information available in 1996) to be supplemented with much more data from the Worcester Public School records. The main prior analyses was done on the class of 1997 when they were juniors, though transcripts for them covered classes selected for their senior year. Now the class of 1998 and 1999 records were complete as well.

The final part of this project was to study how different types went through the process of practice, testing and retesting with the SAT. Specifically, how does personality type affect practicing for the SAT using the PSAT? Which types are the most likely to take multiple SAT tests? Which types see the most gain in score though taking multiple tests? The team of Jennifer Sauvron and Emily Ballinger first raised these questions, however they lacked a data set capable of supporting definitive analysis on it, and really never used the right statistical tools to fully examine what data they had.

This project also includes an implicit examination of how different personality types react to a change in school policy. The Worcester Public School system was beginning to implement a new policy of encouraging both sophomores and juniors to prepare for the SAT by taking the PSAT. This gave me an opportunity to examine what types responded most to advice from their school system, presumably represented by their guidance counselors.

I hope that this project is of use to future researchers, educational policy makers, and anybody wishing to understand the complexities of the SAT test taking as technology shaping a social institution.

Background

The SAT

The Scholastic Assessment Test (SAT) is one of the most widely used measures by selective colleges and universities to determine which students are to be considered their best candidates for admittance. The SAT's main purpose is to help predict the academic success of high school students in their first year of college. It consists of two sections, math and verbal, with each score in the range of 200-800 points. This makes the combined score in the range of 400-1600 points. This SAT is developed by the College Board and administered by Educational Testing Services (ETS).

It is hard to overstate the significance of the SAT in the transition from high school to college. Many college admissions offices use SAT score to make their initial cut in selecting from prospective students. Students with an SAT score below a certain level are taken out of the pool, regardless of their performance on other indicators. In turn, many students to do not even consider colleges that have a high average SAT score when compared to theirs, since they see it as a waste of time. This gives SAT score a very important role, not only determining what students a school will consider accepting but what schools a student will consider applying for.

The SAT was designed to be an aptitude test rather than an achievement test, meaning it is intended to reflect a student's potential as reflected in their reasoning ability, not the knowledge that he or she acquired in high school (the ACT purports to do that). This is an effort to allow students from different academic programs, that vary in coverage and rigor, compete on a level playing field as well as to identify high school underachievers with considerable latent potential.

Controversy over the fairness of the SAT has been a long-standing issue in America. Proponents of the SAT claim that it is a fair and accurate predictor of success in the first year of college, especially in combination with high school grades (Most College Board sources claim that it adds 10% more variance explained to the predictive power of high school transcripts alone, and that together they explain 40-50% of the variance), however there are many critics who disagree. They argue that the SAT is biased based on race, gender and socioeconomic status, and they point to statistics that show consistently lower scores for minorities, women and the working class, though it is unlikely that these groups have a different distribution of talent, compared to that of white males. They charge that the lower average scores of these groups reveals a bias in the questions, caused by the writers who have a tendency to pick vocabulary words and math problems that reflect their experience, giving an advantage to students with middle class and professional families.

The College Board, however, does not agree that lower scores for certain groups means the SAT is biased. For instance, minorities and students from the working class get lower grades on average in the first year of college (although women perform better than men) so they claim that the SAT is correct in giving lower scores to those groups[2]. In their view, the lower scores does not represent bias in the SAT, but rather a bias in the American education system that is giving a worse education to minorities, women and the working class. They must mimic the bias in the larger system in order to be useful and are just a messenger with the unwelcome news.

Even so, the College Board has not ignored complaints that the SAT might be biased. In March 1994, a new version of the SAT was introduced, specifically designed to mitigate, if not eliminate, racial, gender and class bias in the SAT. This new test made several structural changes in the SAT including the removal of the antonyms section from the verbal section and the addition of a student designed non-multiple-choice math section, called "grid-in" items.

---

[2] Batey, Daniel, Brezniak, Paula, and Purohit, Ashwin, Cognitive Bias in the SAT (Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1995) Pg 5.

The purpose of this study is to examine another possible bias in the SAT. This bias is not based on race, gender or social class, but rather learning style and/or personality type. Previous studies done by WPI students have documented sizeable differences in the scores of different personality types (greater differences than are found between blacks and whites, woman and men, or the working and middle classes, in fact). However before these studies can be discussed, it is important to discuss the nature of the main measure they used, the MBTI.

The MBTI

The Myers-Briggs Type Indicator (MBTI) was developed by Isabel Briggs-Myers and Katherine Briggs to measure personality type, based on Carl Jung's psychological type theory. It is one of the most widely used measures of personality type today.

MBTI type theory is connected with the way a person perceives the world and the way he or she makes judgments based on those perceptions. It is based on the idea that there are two methods of perceiving: *sensing*, using your senses to gather facts about a situation and *intuition*, using the meanings and relationships in those facts to understand a situation and to see future possibilities. During childhood, a person will develop a small a preference for either sensing or intuition to a very great one representing greater comfort, practice, experience and confidence with one or the other. Of course, daily life still requires the use of both modes of thinking and everyone uses both on a regular basis. Furthermore, later type development will start to redress the balance and should lead to competent usage of the less preferred mode later in life. However, a person's preferred method of perceiving will still come more naturally throughout life. A person will tend to trust his or her preferred method more, use it more often and often develop greater capability with it, unless forced to stay in an environment that demands the use of the less preferred mode.

The preference for either sensing or intuition will have many consequences in a person's life. People with a preference for one tend to work well in different task environments and find enjoyment from different sources than people with a preference for the other. For instance, Sensors enjoy working with facts and applying thoroughly known knowledge while iNtuitives enjoy working with their imagination and using their skills in changing situations and new ways.

Similar to perception, there are two methods of judgment: *thinking*, judgment based on impersonal analysis in search of objective and just decisions and *feeling*, judgment based on personal values in the sense that one engages empathy and seeks harmony, taking into account how those involved will be affected as the starting point in a decision. Both methods are required for daily life, but people start with a preference for one over the other and develop their preferred mode first. It is important to understand that both thinking and feeling are both rational processes and are both valid methods of reaching decisions. Thinkers work best with things that tend to behave in a logical and consistent manner where they are not forced to use subjective reasoning, for instance, mechanical systems. Feelers work best with people, who are not always logical nor always consistent, but Feelers are able to use empathy and value harmony so they are able to "read" people well. Thinkers typically have a strong desire for justice, while Feelers have a strong desire for harmony producing win-win situation, where everyone can come out ahead. Feelers feel no compulsion to treat everyone alike, they help people in accordance to how much help they need.

One of the consequences of an individual's preference for thinking or feeling is the kinds of motivators that lead the person to act. Generally, Feelers are motivated to perform a task when they believe that they will be emotionally supported as they go through it. Therefore to motivate a Feeler, you must make them believe, before they begin, that you care about them as a person and are going support them through it (using phrases such as “I know you can do it”, “I believe in you”, and the like). Thinkers, on the other hand, are motivated to perform a task when they believe that there will be some kind of commensurate reward after they are through. To motivate a Thinker you must make them believe that, once they are through, there will a reward worth their effort. That reward can be recognition, attention, pride in their accomplishments and other subjective rewards, but the point is that the important kinds of support come and the end, not the start.

Not only will a person have a preferred method of perceiving and judging, one of the two will be the *dominant* process, the process most used and trusted. The dominant process is the first to develop in a person and becomes a governing force in that person’s life. The other preferred process is called the *auxiliary* process, since it mostly serves in a support role to the dominant process and when conflict occurs the dominant process will win out. The auxiliary process also helps balance the individual since a person who never developed a perceiving process would be terribly closed-minded and a person who never developed a judging process would have little self-control. Life requires one to both take in information and make decisions in some fashion.

For example, people who prefer intuition and feeling, but whose dominant process is intuition will organize their lives around pursuing their iNtuitive goals. They will readily use feeling if it helps them pursue an activity that interests their intuition, but they will ignore feeling when the two conflict. On the other had people who prefer feeling over intuition will use their feeling to establish goals and employ their intuition when it helps accomplish those goals. Normally the dominant process is given first, so the first group would be called iNtuitives with a preference for feeling while the second group would be called Feelers with a preference for intuition.

People rely mostly on their dominant and auxiliary processes, but at times will use the remaining two. The third most used process is the one opposite to the auxiliary process and is called the *tertiary* process. The least used process is the one opposite to the dominant process and is called the *inferior* process. For instance a Thinker with a preference for sensing would sometimes use intuition (the tertiary process) and on rare occasions use feeling (the inferior process). Very well balanced individuals during adulthood will develop their tertiary process enough to use it very competently, however they are the exception, not the rule. Normally it is relatively underdeveloped compared to the dominate and auxiliary processes. The strength of the preference for the dominant process causes the inferior process to be used very little and thus is very rarely developed beyond a minimal level of competence, and sometimes not even developed to that extent.

Sensing-Intuition and Thinking-Feeling are called dimensions of the MBTI. There are two other dimensions that the MBTI uses: Introversion-Extraversion and Judging-Perceiving. Like Sensing-Intuition and Thinking-Feeling, these dimensions represent preferences that affect information gathering and processing as well as the context that one is likely to feel natural and at ease.

Introversion is a preference for the inner world of thoughts and ideas. *Extraversion* is a preference for the outer world of action, involving both people and things. The original Latin meanings for introversion and extraversion, inward-turning and outward-turning, are much closer to the MBTI meanings than the current usages for the words. Introverts are most comfortable when they are reflecting and tend to think things over thoroughly before they act. Extraverts are most comfortable when they are doing things and tend to take quick action. Introverts tend to choose careers where the work mostly takes place in their heads, for instance as lawyers, scientists or computer programmers. Extraverts enjoy work that mostly takes place outside of them, for instance as managers or actors.

Finally, *Judging* is a preference to face the world using a judging perspective (thinking or feeling), while *perceiving* is a preference to face the world using a perceiving perspective (sensing or intuition). Judgers like order and planning in their lives and have a strong drive towards closure. Judgers like things to be over and done with. Perceivers enjoy flexibility and change and try to keep planning to a minimum in order to keep their options open.

The Extraversion-Introversion and Judging-Perceiving dimensions are strongly related to a person's dominant process. If an Extravert's dominant process is a perceiving one (sensing or intuition) then he or she will want to have a steady stream of information for that process to perceive and will therefore face the world with a perceiving attitude. Likewise, if an Extravert's dominant process is a judging one (thinking or feeling), then he or she will enjoy exercising it and will have a more controlling attitude towards the world to keep things well structured and to get things done.

The opposite is true for Introverts, whose dominant process is focused on their inner world. For them, the outside world is less important and therefore is handled by their auxiliary process. This means Introverts whose dominant process is a perceiving one will be Judgers while Introverts whose dominant process is a judging one will be Perceivers.

A note on terminology: An individual's four preferences determine his or her MBTI type. Convention orders them Extraversion-Introversion then Sensing-Intuition then Thinking-Feeling then Judging-Perceiving and uses the first letter of each word to abbreviate a person's type. For example the last example a person who preferred Extraversion, Sensing, Thinking and Perceiving (an Extraverted Sensor with a preference for Thinking) would be shortened to ESTP. The one exception is intuition, which is represented by an N, since I is used for introversion. I am an INFP for example since I prefer Introversion, iNtuition, Feeling and Perceiving (an Introverted Feeler with a preference for iNtuition).

ISTJ 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ISFJ** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**INFJ** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**INTJ** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I
ISTP 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ISFP** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**INFP** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**INTP** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I
ESTP 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ESFP** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ENFP** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ENTP** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I
ESTJ 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ESFJ** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ENFJ** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I	**ENTJ** 1) Sensing (Dominant) - I 2) Thinking (Auxiliary) - E 3) Feeling (Tertiary) - E/I 4) Intuition (Inferior) - I

I=Introverted E=Extraverted I/E = Type theorists disagree on the orientation of the tertiary

Priorities of Functions for the 16 types[3]

The distribution of types over the general population is far from uniform. It is estimated that Extraverts make up 70% of the population, as do Sensors. Judgers account for 55% of the population. 60% of males are Thinkers while 60% of females are Feelers.[4] Since different personality types enjoy different activities, most self-selected activities will be skewed towards one or more groups. A football team would most likely be largely made up of Sensing Extraverts; a math team would be made of largely Introverts and iNtuitives. It is important to always keep in mind the non-uniform distribution of types in self-selected activities such as colleges, careers and companions.

The Sensing-Intuition dimension of the MBTI should be commented on further since it is by far the most relevant to learning styles and the SAT. Sensing and Intuition are the processes that people use

[3] Martin, Charles, Looking at Type: The Fundamentals (Gainesville, FL: Center for Applications of Psychological Type, Inc., 1997) Pg 9.

[4] Lawrence, Gordon, People Types and Tiger Stripes (Gainesville, FL: Center for Applications of Psychological Type, Inc., 1993) Pg 39.

to process information and understand their world, so it is not surprising that it is strongly related to learning style.

Sensing types learn best from direct experience and like to use a patterned, step-by-step method for problem solving. Intuitives learn best when they first understand the theory behind a subject and tend to gain knowledge by early flashes of inspiration, filling in the details later. Sensors accumulate facts first and later generalize their knowledge and put it into context. Neither Sensors nor iNtuitives should be thought of as more intelligent than the other, the types simply have different preferences that can develop into different strengths and weaknesses with practice.

Sensors are usually more observant of practical or immediate details than iNtuitives. On the other hand, iNtuitives are better at recognizing large scale patterns since they are less bogged down by minutia. Intuitives are better at working with abstract subjects while Sensors are better at finding practical solutions to real world problems. Intuitives tend to think that "true intelligence" is embodied by a quick understanding and a vivid imagination. Sensors tend to think that "true intelligence" is embodied by a deep understanding and good common sense.[5]

Many type theorists, including Isabel Briggs-Myers, have said that the American school system favors iNtuitive understanding at the expense its sensing students (who make up the majority of the population). Textbooks and tests are written from an iNtuitive perspective, and most classes are taught in an abstract way that favors iNtuitives. This is especially true of math and science. Subjects are taught by introducing the theory and definitions first, then by the giving examples and lastly by students using the skills themselves. This order is very good for iNtuitives who want to understand things theoretically, but is difficult for Sensors who like to first understand how a subject relates to the real world in a concrete way. Gordon Lawrence has said, "The record of American education in the twentieth century is a record of neglect for the sensing intelligence, the intelligence typified by the

[5] Lawrence, Pg 7.

young diesel mechanics student who bypassed the prescribed instruction process and solved his problem his way."[6]

Interestingly, one of the first studies done using the MBTI was a study on the distribution of the different types among National Merit Scholars (students with exceptional scores on the PSAT) by Isabel Briggs-Myers and Mary McCaulley. It was conducted while the MBTI was in its infancy and so was a study to validate the MBTI rather then to examine the SAT. Briggs-Myers considered the SAT to be like an IQ test where students that did well in school would score the highest. She reasoned that the types that were best suited for academic learning, INTJs, would score the highest. The dimensions that would make the most difference were Sensing-Intuition and Introversion-Extraversion, in that order.

Briggs-Myers' predictions generally agreed with the findings. The National Merit Scholars were made up of 83% iNtuitives and INTJs were the highest scorers.[7] The point of this study was not to look for possible biases in the SAT however, so the gap between iNtuitives and Sensors was not investigated further at that time. The finding was replicated several times, but the subject was largely left alone until 1992 when a group of students from WPI examined the relationship between MBTI and SAT score in greater detail.

The Initial Pilot Study

The first team of students from WPI to study the relationship between MBTI type and SAT score were David Kingsland, Charles McTague and Benjamin Kibler. They wanted to see if a case could be made for a score bias based on psychological type, similar to the cases for bias based on race, gender and social class. To do this they would test the theory that the certain personality types scored significantly higher on the SAT.

---

[6] Lawrence, Pg 43.
[7] Isabel Myers With Peter Myers, Gifts Differing (Palo Alto: Consulting Psychologists Press, 1980) Pg 37-39.

Of course, even if the team could document a difference in SAT score between types, supporters of the SAT could reply in the same way as against other accusations of bias. If one type scores higher on the SAT then that type must be better suited for academic learning, or at least better suited for the type of learning stressed in most American colleges and universities, which have faculty dominated by intuitive types in most fields. Establishing that certain types scored higher on the SAT would only be the first step in establishing a case for bias. In order to fully make the case, it would also be necessary to demonstrate that the types that scored worse on the SAT performed just as well as their counterparts in the first year of college.[8]

As the first students to study the relationship between SAT score and personality type, Kingsland, McTague and Kibler laid most of the theoretical groundwork for future projects. Their hypotheses were generally on the mark, and the later teams did not deviate far from them, although the differences Kingsland, McTague and Kibler observed between MBTI types were not as large as those in later studies.

The team argued that iNtuitives and Introverts would score higher than their counterparts. The largest difference would be seen between the iNtuitives and Sensors. This was based on the previous study of National Merit Scholars reproduced by Briggs-Myers in her book Gifts Differing and because of their own reasoning that the nature of the problems on the SAT would favor certain types. Myers had hypothesized that Intuitives would hold the largest advantage on the test because of speed, which is a critical factor to timed, multiple-choice tests like as the SAT, Kingsland, Kibler and McTague came to the same conclusion by a different path. "The preference for intuition should have the strongest impact on SAT performance in this study. Intuitives are adept at symbol and word recognition, giving them a speed advantage. (...)Also, they should perform well on reading comprehension sections because they

[8] McCormick, Private Communication

are faster readers and recognize sentence meanings quicker.[9]" Kingsland, McTague and Kibler predicted iNtuitives would score higher on both the math and verbal sections of the SAT, but the difference would be greater on the verbal section since maturity of the symbol system involved in mathematics was more widespread. Math reasoning would be affected more by training and familiarity than verbal reasoning

Introverts would also have an advantage on the test, since they tend to read more and take their academics more seriously than Extraverts. The team regarded the SAT as at least partly an achievement test, so this would give the Introverts an edge.

Kingsland, Kibler and McTague did not make a definite prediction about performance based on the judging-perceiving dimension. Judgers would make decisions quicker so they would have a speed advantage, but this could be offset by other factors. Perceivers tend put off judgment until they have as much information as possible. This would make Perceivers less likely to fall for trick questions and overlook pertinent facts than Judgers.

Thinking-Feeling would account for almost no difference, the group predicted. Verbal scores would be the same between the types though math scores might be slightly higher for Thinkers since they would be more adept at logical problem solving.

When Kingsland, McTague and Kibler turned their attention to data collection they ran into the problems. The team estimated the size of the Worcester high school senior class to be about 1000 students, and they believed they would be able to get 500 students who had taken the SAT. However, that turned out to not be the case. The team ended up including the junior class as well and still only 346 students could be found who had taken the PSAT or the SAT. Of these 346 students, only 276

[9] Kingsland, David, McTague, Kibler, Learning Style Biases in the S.A.T. (Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute 1995) Pg 36.

provided usable data. In order to get these cases, volunteers were recruited from auditoriums full of eligible students, gathered to hear their pitch, so self-selection could have played a role in the results. Also there was a large (.63) correlation between being an iNtuitive and being a Perceiver in their sample. Since the team expected iNtuitives to score significantly higher than Sensors, this could skew the results on the J-P dimension as well.

Dimension	N	Verbal Advantage	Math Advantage	Combined Advantage
I-E	I - 121 E- 155	+13 (I)	+21 (I)	+34 (I)
S-N	S - 136 N- 140	+62 (N)	+46 (N)	+108 (N)
T-F	T-128 F-148	+17 (F)	+6 (T)	+11 (F)
J-P	J-108 P-168	+59 (P)	+43 (P)	+102 (P)

Results from the Kingsland, McTague and Kibler Study

The results of study were mostly consistent with the group's hypotheses, though there were some differences. Intuitives outscored Sensors by 62 points on the verbal section and 46 points on the math section for a total of 108 points. As predicted, intuition provided the largest advantage in the sample. The advantage was more for the verbal section than math, but this was by a statistically negligible margin. Introverts outscored Extraverts by 13 points on the verbal section and 21 points on the math section. This was consistent with the group's hypothesis, but the margin was not large enough to make a very convincing case. Perceivers outscored Judgers by 59 points on the verbal section and 43 points on the math section. The team believed that Perceivers could have an advantage, but this was a much larger margin than they anticipated. The result was probably caused by the high correlation between a preference for perception and a preference for intuition. Feelers outscored Thinkers by 17 points on the verbal section while Thinkers outscored Feelers by 6 points on the math section. This difference was in the opposite direction as the group's predictions, but it was too small to draw any kind of conclusions on, except that the T-F dimension accounted for little difference in SAT score.

Overall, the results of the study showed that Kingsland, McTague and Kibler probably were on the right track with their predictions. However their small sample size and problems with self-selection mean more data would have to be collected before any definite conclusions could be made.

The Second Pilot Study

The study by Kingsland, McTague and Kibler was done the last year that the SAT was using the old SAT format, before the College Board made changes to mitigate or eliminate cultural biases. The obvious question after the format change was, would Kingsland, McTague and Kibler's findings replicate with the new format? This question would be answered by Daniel Batey, Paula Brezniak and Ashwin Purohit's follow-up study, conducted the year after Kingsland, McTague and Kibler's.

Though the test had undergone several changes, Batey, Brezniak and Purohit saw it as being fundamentally the same. They predicted that the same personality types that scored high before would continue to score high with the new format. In fact, they expected the results to be clearer than in the Kingsland, McTague and Kibler study, because they believed their new method of coding SAT scores was an improvement over that used by the previous team. This team used a student's first SAT score, rather than his or her highest score, since they believed cognitive bias would show up most clearly when a student was experiencing the test for the first time. They also disapproved of Kingsland, McTague and Kibler's mixing of PSAT and SAT scores, and ignoring whether it was the first or second time one had taken the SAT.

Another new feature in this study was that Batey, Brezniak and Purohit collected data on the socio-economic status of each student (SES). Students were asked to supply their parent or guardian's occupation and it was coded into a "prestige" score of 1-100. It would have been interesting to see how social class interacted with personality type and SAT scores, but Batey, Brezniak and Purohit decided they did not have time to code and examine this variable. Since only about half of their sample gave them codeable answers, they were discouraged about the possibility of putting considerable effort into looking up occupation and still not being able to do an adequate analysis. A paid assistant would later o this analysis, breaking the prestige scores for father's occupation at 1-49 and 50-100, and report some striking relationship between cognitive style and SES, but also some powerful correlations between social class and what school one had attended among the 120 responses.

Data collection would again prove to be difficult, with the team only gathering 229 complete cases. A major problem was that at Doherty and Nashoba high schools another study had already administered the MBTI to all students enrolled in a physics class. Batey, Brezniak and Purohit decided to get the data from the other research team rather than try to re-administer the MBTI at these schools. However, the data turned out to be coded with ID numbers specific to the study. It would be impossible to link this data without having the keys that were used. These keys were left with the classroom teachers, but by the time Batey, Brezniak and Purohit requested them the information had been lost. Doherty and Nashoba were the largest sources of cases for the first study, so low returns from these schools seriously affected the comparability of the data from each study. Also, since only students who were enrolled in a physics class were lost the data set was skewed away from an identifiable group of students.

However in comparison, many more cases were collected from South, Burncoat and North high schools, schools. Where data collection was weakest in the first study they did much better. This almost made up for the cases lost at Doherty and Nashoba, so the total number of cases was about the same as in the Kingsland et al. study, but the sample was better balanced and at least as representative of the whole student body as the first study had been.

It would have been nice to combine the cases from the first two studies. This would have resulted in a 500 case data set, with good representation for each school. However, it was not possible. The SAT changed it's format and the second team of researchers changed their dependent variable from highest SAT score to first SAT score. The cases could not be legitimately pooled.

Dimension	N	Verbal Advantage	Math Advantage	Combined Advantage
I-E	I - 98 E- 148	+21 (E)	+5 (E)	+26 (E)
S-N	S - 107 N- 139	+91 (N)	+70 (N)	+161 (N)
T-F	T-114 F-132	+51 (F)	+16 (F)	+67 (F)
J-P	J-84 P-162	+64 (P)	+47 (P)	+111 (P)

Results from the Batey, Brezniak and Purohit study

Again, the S-N dimension was the most significant factor, this time by an even larger margin. Intuitives outscored Sensors by over 150 points in the sample. The J-P dimension was also an important factor, with Perceivers again scoring 100 points more than Judgers. This was explained in the Kingsland, McTague and Kibler study by the high correlation between a preference for perceiving and a preference for intuition, but that correlation did not exist this time around. The J-P dimension was emerging as another important factor in predicting SAT scores, with P's having the advantage.

The I-E dimension switched from being a slight advantage for Introverts to a slight advantage for Extraverts. In both studies though, the I-E dimension was accounting for very little difference in SAT score. The T-F dimension changed from causing almost no difference in the first study, to a 67-point difference in favor of Feelers. There was concern that there was a gender difference in disguise, but if so, it was a strange one since men outperform women by 40-50 points on average at the national level.

At this point, it was clear that iNtuitives had a significant advantage over Sensors on the SAT. It seemed likely that Perceivers also had an advantage, though a smaller one. The difference the I-E and T-F dimensions seemed small and the direction was not clear, with score differences going in both directions for the first two studies. The best prediction for these two dimensions was not to make one. The need was clear for a large-scale study, where small sample size and self-selection would not be a problem.

Pieper's Large Scale Study

The next study, done by John Pieper, would be an effort to replicate the results of the previous two pilot studies in a large-scale study encompassing all four main Worcester Public high schools and several others. This study would get even larger than Pieper expected and he would gather a considerable amount of data.

Like the team before him, Pieper decided to collect data on a different test outcome score than the previous studies. This time instead of highest SAT score or first SAT score, Pieper used first PSAT score. His reasoning was that if differences by personality type were better seen by coding the first experience on the SAT they would be clearest in PSAT scores, an even earlier experience and the real first contact with an SAT type test.

Another reason for using PSAT scores was a new school policy that put Pieper in a very good position to collect PSAT data. In an effort to encourage practice using the PSAT, the WPS schools were strongly encouraging sophomores and juniors to take the test, even offering to pay the test fees for the sophomores. This meant that the bulk of both the sophomore and junior classes would be taking the PSAT for the first time that year and a large number of cases could be collected.

During data collection, the principals of the high schools decided that it would be easier to administer the MBTI to all students rather than to find a scheme that would allow half the school to take the MBTI while the other half continued with class. They selected a Monday and Tuesday just before Thanksgiving break, a short week anyway, for data collection. Burncoat and Doherty collected data on Monday and South and North collected data on Tuesday. This led to the MBTI being administered to every WPS High School student. Not every student would show up to take the measure (indeed absences and tardy students were fairly common, especially on Tuesday), but a large portion of the 9^{th}-12^{th} grade student body did spend the 90 minutes to take the MBTI and GCSI those two days. MBTI scores for 2813 students were recorded out of about 4000 students in the system's four high schools, the coverage was about 70%. However, only the sophomores and juniors would be taking the PSAT that year, so many of these cases could not be used immediately. In the end Pieper collected 920 cases where the student had taken both the MBTI and the PSAT and this in a school system where traditionally half of the student body had taken the SAT or PSAT. Presumably those college bound students who had taken the PSAT were more likely to be there to take the MBTI than those who were not considering college, which is not surprising.

In addition to the 4 main WPS High Schools, Pieper also collected data from the suburban Nashoba Regional High School and an experimental public high school located at WPI called the Massachusetts Academy of Math and Science (Mass Academy). Participation rates at both of these schools was very high and Pieper was able to get PSAT and MBTI data for 319 out of the 329 sophomores and juniors at Nashoba High School and for all of the 35 juniors at Mass Academy.

In total, Pieper's study gathered MBTI forms and PSAT scores for 1263 students. This would be the first time that the relationship between personality type and SAT score could be analyzed with a large data set, free from the problems of under-representation and self-selection.

Pieper intended to study the affect of social class in relation to personality type, but never got the chance. He administered a questionnaire to the students in order to create a social class variable. Using the same measure of social class as Batey, Brezniak and Purohit: parental occupation. However the sheet wouldn't scan electronically and had to be hand coded. Pieper was already dealing with the analysis of a very large data set and decided to drop the social class variable at this point.

Pieper was somewhat vague in his report when discussing the results. The numbers of students in each type was not given. Neither were the results for the I-E or T-F dimensions, however there was never a clear difference in the past with these dimensions and one may assume that these were statistically insignificant in Pieper's study.

Dimension	Verbal Advantage	Math Advantage	Combined Advantage
S-N	+73 (N)	+49 (N)	+122 (N)
J-P	+8 (P)	+37 (P)	+45 (P)

Results from the Pieper study

Again the S-N dimension accounts for a large point differential, this time at about 120 points. By now it was clear that this dimension had a large impact on SAT and PSAT scores. The J-P advantage is diminished, mostly from a much smaller Verbal advantage for the P's. The smaller verbal advantage could be the result of Pieper's sample not having a large correlation between being a Perceiver and being an INtuitive. However the P's still held a significant advantage, especially in the math section.

McCormick's Analysis

At the same time as Pieper was doing his analysis, there was another analysis being done using a smaller but more complete data subset by Keith McCormick.

McCormick used data from the same group of students as Pieper, however he took a different stance on data collection. Pieper wanted to replicate the previous findings with as large a sample as he could get. If there was partial, but not complete, information on a student then Pieper kept the case and used it whenever possible. If a student's GPA was missing then the case would still be used whenever GPA was not needed. McCormick, on the other hand, wanted to do his analysis in as much depth as possible and wanted a stable data set, not one in which sample size changed with every analysis. He dropped all cases where complete information was not available. Also, McCormick only wanted to examine students that had full transcript data, where it was known what courses they would be taking senior year. This was possible for the juniors but not for the sophomores, so McCormick therefore decided to drop the sophomores from his data set. All of this caused McCormick to have a much smaller sample size than Pieper, 448 juniors instead of 1263 sophomores and juniors, however each of these cases contained complete data on high school program and MBTI type.

One of McCormick's ideas was to take into account the impact of academic preparation on SAT scores. McCormick had seen SAT literature from the College Board that alluded to score differences connected with core program challenge, but it was not clear how large a difference it made. This made sense and he decided to look into the matter since Isabel Briggs-Myers had already predicted and documented a tendency for IN's to take more challenging programs of study in high school than ES's. He reasoned that this variable could be the variable to explain the "INtuitive advantage" if it was large and consistent enough of a factor. He found that academic preparation was highly correlated with SAT score when he analyzed the data in these terms. McCormick then worked to see how academic preparation interacted with intuition, the MBTI preference that made the greatest difference on the SAT.

McCormick divided students in the WPS into 5 categories; based on the level of challenge of the classes they had taken and were planning to take. McCormick then compared the proportion of iNtuitives to Sensors at each level of difficulty and compared how Sensors and iNtuitives from programs of the same level of difficulty performed on the PSAT.

(While reading the following chart, keep in mind that the level of difficulty for Worcester High School classes, from least to most challenging, is General, College, Honors then Advanced Placement (AP). It is impossible to fill a schedule entirely with AP classes in Worcester, so the highest level of academic preparation was students who had taken at least one AP class in a basically Honors program.)

Program Challenge	Mean PSAT	S-N Dimension	# of cases	Mean PSAT	Difference
Some AP	1107	N	49	1117	
		S	20	1083	34 pts
Only Honors	973	N	46	1017	
		S	54	936	81 pts
Some Honors	841	N	59	883	
		S	61	801	82 pts
Only College	700	N	42	742	
		S	88	680	62 pts
Some College or All General	643	N	3	610	
		S	23	648	-38 pts

McCormick's Class of '97 study

The results of McCormick's analysis suggest that intuition gives an advantage on the PSAT and SAT in two ways. The first is that the more difficult academic programs were made up of mostly iNtuitives even though the general population was mostly Sensors. For example, 71% who had taken at least one AP class were iNtuitives, even though iNtuitives only made up 43% of the total students in the sample. Also, when comparing equally well prepared students, iNtuitives score higher on the SAT at all levels (the one exception was at the General level which only had 3 iNtuitive students, not enough to draw a conclusion from).

A comparison of the scores between the difference academic preparation groups shows the score differential is the highest in the middle levels of academic preparation. When comparing the middle levels of academic preparation, where students have taken some or all honors classes, iNtuitives outscored Sensors by over 80 points. This margin drops to 62 points when comparing students at the low end where students have only taken college level courses. The amount drops further to 34 points when comparing students that have taken an AP course. McCormick's explanation for this goes as follows:

The experience that students have on an SAT section can be split into three broad categories. At the start of the test where the questions are the easiest, he or she will be able to move through them with little to no thought. At this stage the student is answering questions mostly from memory and practiced algorithms. This could be called the mastery zone.

As the student progresses through the test and the questions get progressively harder, he or she will get to a new type of question. This type of question will use concepts that the student has been exposed to, but not yet mastered. These questions will still be possible to answer, but the student will have to take conscious thought and the possibility of errors will be much greater. This could be called the "attractor" zone or zone of uncertainty.

Even farther along the test, the student will get to a new level of question. This last category will consist of questions where the student has not had a formal exposure and probably does not even understand the basic concepts. At this stage the student will either have to take an uneducated guess at the answer or skip the question. This could be called the wild guessing zone.

The length of time spent in each zone will vary with each student of course. It is conceivable that a very well prepared student would spend the entire test in the mastery zone and a very poorly prepared student would spend the entire test in the guessing zone. In real life most students will spend some time in all three zones, with better prepared students spending more time in the mastery zone and less well prepared students spending more time in the guessing zone.

While students are in the mastery zone then the difference between Sensors and iNtuitives should disappear. At this stage both types will be able to answer the questions easily and should score close to perfect. The same effect will occur at the guessing zone since both types will be guessing randomly and scores should be close to 0, given the penalty points the ETS uses to compensate for true guessing.

The zone where iNtuitives outscore Sensors is in the zone of uncertainty. In this stage students are struggling interpret the problems correctly and understand what is the question. Intuitive student's speed edge, use of trial and error, reading between the lines for cues, and other advantages will all benefit them at this stage.

In the zone of uncertainty it can be said that iNtuitives score above their level of mastery or true competency, while Sensors can only score at their level of mastery. Intuitives score higher because they are able to choose the correct multiple-choice answer better without fully understanding the subject, it is enough for them to have the basic concepts down and figure out the rest. It is doubtful, though, that the iNtuitives would be able to perform a procedure that related to the subject in real life - when they had to have the matter really mastered in detail, not just understand it in general terms. Intuitives were able to score points because they could rule out at least 2 of the 4 answers before they guessing based on a hunch. By contrast, McCormick found that the sensing students in the zone of uncertainty went for the "attractor" items – what the College Board calls "distracters." This would be the right answer if you used the wrong algorithm – and they tended to put it first among the choices (i.e. if the question called for division between 2 numbers, the "attractor" would be those 2 numbers multiplied)

Students that had taken some AP classes were well prepared for the SAT and spent most of their time in the mastery zone, while students some or all Honors classes spent much of their time in the mastery zone and students who were taking had not taken any Honors classes spent the largest portion of their time in the guessing zone. This would explain why the score differences between iNtuitives and Sensors was the least in the "Some AP" level, greatest in the "Only Honors" and "Some Honors" levels and beginning to lessen at the "Only College" level. Again, no conclusion should be from the difference in the "Some College/General" level because it contained only 3 cases that were iNtuitives.

McCormick's theory also held for the Mass. Academy students. The program these students were in corresponds to a level of difficulty above the highest in the main Worcester High Schools. By McCormick's reasoning there should be little difference between Sensors and iNtuitives, since both types would not be challenged for and forced to read between the lines until the final 15% or so of the SAT. Indeed, the total difference in scores was only a 1-point advantage for iNtuitives. It is worth noting this was the result of a 36-point verbal advantage for the iNtuitives and at 35-point math advantage for the Sensors, which canceled each other out.

McCormick was not the only other person to study Pieper's data set. Pieper had collected enough data to support many studies afterward (including this one). Like McCormick's, these studies took for granted an iNtuitive advantage on the SAT and went on to examine more subtle aspects.

Test and Retest

Pieper's data set was collected just as the Worcester Public School's began a new policy of encouraging the its students to take the PSAT twice before the SAT, as sophomores and again as juniors. Their rationale was that the extra practice would raise SAT scores in the city. A natural question to ask was, what was the effect of practice on the PSAT when personality type was considered?

Emily Ballinger and Jennifer Sauvron would raise this question in 1997, two years after Pieper's study. They hoped that their timing would allow them to get a large number of students that had taken the PSAT and SAT multiple times. The class of 1997, the juniors in Pieper's study, had taken the PSAT and then the SAT. The class of 1998 would have already taken the PSAT for the second time as juniors and now would be taking the SAT. The class of 1999 would just be taking the PSAT for the second time. Of course, there were also some of students that ignored their guidance counselor's advice and only took the SAT.

Sauvron and Ballinger had two basic questions to answer. Would practice benefit some types more than others? They predicted that it would, and identified which types. Also, would certain types choose to retake the test more than others? They weren't so sure about this question, since they were not clear if a high or a low initial score would motivate a student more to retake the test. It depended on the personality of the test taker in their view.

Unfortunately, the team had to deal with many unexpected problems in data collection and never got to a position where they could answer those questions adequately. The PSAT data for 1997 would not be available in time for the team to use. Also, the data from previous years was not organized in any sort of convenient form. They were forced to enter the data by hand, from a 163-page paper printout provided by the WPS. The team also found that a low proportion of the members of the class of 1998 retook the PSAT as juniors. They were being asked to pay for a test that they had taken for free as sophomores and many decided against retaking it.

In the end, Sauvron and Ballinger turned their attention to proposing the creation of a database that would be able to support the analysis they had hoped to do. They performed an analysis on the data set that they were able to put together, but could not go into much depth. The team looked at what proportion of each type saw an increase in their score as a result of retaking the SAT. The results were that INs saw a score increase most often, 83% of the time. The other types all increased their score about 70% of the time. It is hard to assign a meaning to this statistic however, since it could be the result of many factors and it does not take into account the amount that the scores increased. The team had to be satisfied with setting the stage for future research and documenting that most students improve on the SAT with practice, but some are more likely to retake it - those who were encouraged by high initial scores.

The Impact of Social Class

Throughout these studies socio-ecomonic class (SES) was considered an important variable, but it was never the main focus of a study. The Batey, Brezniak and Purohit team and Pieper team both recorded a SES variable based on parent occupation, but neither coded it. It was Gerald Noble that took up this task. This would involve following up the people, especially at North High, who were absent on the day of the testing or did not finish. The SES items were the last section, so those not finishing were a special problem for him.

Noble's plan was to hand code the questionnaire that Pieper gave up on when it would not scan electronically. He would then calculate a prestige score based on parental occupation ranging from 1 to 100 points. Students would be split into a low prestige group (1-50 points) and a high prestige group (51-100 points) and the groups would be compared together.

Noble hypothesized that academic preparation would be the mechanism that created differences by social class. Those students in the high prestige group would be more likely to take AP and honors level classes and McCormick's work showed students who took more challenging courses scored significantly higher on the SAT. To test this theory, Noble planned to add the academic preparation variable from McCormick's data set to his own and examine how this interacted with social class and SAT score.

When Noble turned to data collection he ran into problems though. To begin with the questionnaire that he wanted to code was the form least likely to be filled out correctly in Pieper's study. To deal with this Noble gave a follow-up administration of the questionnaire at North and Doherty high schools and may have gotten to South and Burncoat too, but did not code those data. This follow-up would be given to the sophomores and juniors of Pieper's study since Pieper focused his attention on these classes; they were the only students to have taken the PSAT at that point. North and Doherty were chosen because they had the smallest number of usable cases when Pieper collected the data. Noble decided to focus his study on these two schools, giving much less attention to the two other Worcester schools in Pieper's study, South and Burncoat.

Even after two attempts to collect the SES data there were still many students unaccounted for, due to lapses in reporting. However Noble managed to get SES and MBTI data for 182 students from Doherty and 187 students from North, for a total of 369 students. The total numbers of sophomores and juniors enrolled was 838 so this was, not quite half of the total, a relatively small, but still statistically valid number of cases to use. He would have to see how representative it was, but samples of 25%-50% are commonly reported in the literature and can be valid representation of the universe as a whole. In this case, he was interested less in describing the universe of students in the sampling pool, than establishing a relationship between two variables, so total sample size was more important than how representative the samples were. As long as he had good variation, a fair number of cases at both ends of the student body that had taken the PSAT, he was okay.

Noble ran into his largest problem when trying to include McCormick's academic preparation variable. McCormick had not received the data that Noble collected in his follow-up trips and was forced to do his analysis without them, discarding any cases where he did not have complete information. Because of this, few cases were common to both Noble's and McCormick's data sets. This problem was worsened by the fact that Noble had focused on getting scores for both sophomores and juniors, but for only North and Doherty high school while McCormick collected data for all schools, but only the junior class. In the end, there were only 96 cases common to both McCormick's and Noble's data sets. To try analyzing this by dividing it into 16 types and 5 academic tracks would have been ridiculous. Noble simply couldn't do an in depth analysis that involved MBTI type, academic preparation and social class for lack of the academic preparation variable, and he discovered the problem too late to repair it. However he could examine the straight SES to SAT relationship and see if SES correlated with PSAT score (as expected from national data). In this regard, Worcester was like the nation as a whole; SES did correlate with PSAT score, but the question of why could not be addressed adequately.

Summary

By the time this project started the study of SAT scores and personality type was well underway. The teams that had worked on the topic had already created a theoretical framework, collected a sizeable amount of data and done quite a bit of analysis on that data. The major finding, that iNtuitives had a significant advantage on the SAT, had been replicated enough times that a very convincing case could be made for it. However there were still many problems to deal with.

A large amount of data had been collected, but it had never been organized into a single data set. Also, the test-retest question raised by Sauvron and Ballinger had not been answered as well as the issue of social class that was worked on by Noble. McCormick's work was very detailed and had many theoretical implications, however his results had not been replicated on the full data set, which would be at least three times as large as his sample when Noble's follow-up cases were included – and thus well over 1000 cases.

These factors determined what the present project would consist of as I tied up loose ends. I would have to merge the many data sets that were collected together in order to get a single one and fill the holes in the data that had accumulated. Using that data set I could examine the test-retest question with care, settle the social class question and attempt to replicate McCormick's findings on a more statistically reliable data set.

Methodology

Overview

This project can be loosely separated into three parts. The first is the overview presented in chapter two that describes the efforts of the eight prior WPI research teams interested in the PSAT or SAT. The second was the assembly of a data set that built upon the data collected by Pieper and Noble and to include complete PSAT and SAT scores and academic transcripts for the WPS classes of 1996-1999. Transcript data was used by Keith McCormick to construct the academic track variable that was used to very good affect in his study of the class of 1997. The last part was to use that data set to examine the test-retest question raised by Sauvron and Ballinger. My contribution to the data was to create at SES variable building on Gerald Noble's aborted effort to add a reliable SES variable and to convert the PSAT/SAT data from the WPS to usable variables. Presented here is the methodology for the last two parts: the procedures to assemble the data set, the hypotheses for the test-retest question and the variables that were used to analyze it.

Data Assembly

By using data already collected it was possible to put together a fairly comprehensive data set in a reasonable time, which included all the variables that were analyzed in previous projects. The data set would contain, for each student:

- MBTI type, collected by Pieper or in Noble's follow-up.
- GCSI type, collected by Pieper or in Noble's follow-up.
- SES data, collected first by Pieper or in Noble's follow-up.
- SAT and PSAT scores for each date the students took the exam from ETS, collected by the WPS.

- Academic program, using transcript data collected by the WPS to ascertain the proportion of AP, Honors, College or General courses that each student took. (Numerous other variables could be created out of the transcript data including grades and electives, but the class of 1997 study suggested that a key variable was program challenge.)

These variables were collected by different teams with different goals. In order to create a data sets that contained all of them it would be necessary to do quite a bit of assembly work.

Since Noble focused his follow-up on North and Doherty high schools and concentrated his coding of SES exclusively on those samples, he never entered the SES data on South and Burncoat and these would have to be entered by hand. Entering the data was fairly time consuming, but fortunately this data was also useful to two other projects done at the same time. I was able split the work with the authors of those projects, Bai Lan Zhu (the SAT ethnic bias study) and Matthew Marino (study of the class of 1999's transition to college).

Academic transcript data as well as SAT and PSAT scores had already been collected for Pieper's study, however both had to be collected again. Many of the students were sophomores or freshman when Pieper collected the data and had taken classes and SAT exams since. Sauvron and Ballinger's made an effort to get the PSAT scores for the class of 1998 and SAT scores for the class of 1997 from after Pieper's study, but there was still data that was more recent. Getting these data would require the cooperation of the WPS.

The WPS had an archive of SAT score reports that they received from the College Board. Each test date was stored in a text file with information on each student who took the test that date. This was a very inconvenient form to use. The WPS was willing to give this data to me if they could receive back a file that was organized by student and contained every test that the student took, in a form that could be linked to other data. This form was very similar to what was needed in order to link the PSAT/SAT data with the other data sets I needed, so the arrangement worked out nicely. It would also be necessary to add WPS ID numbers to the data set in order to link the PSAT/SAT data, since they were not included in the files from the College Board.

To merge the SAT score reports together I first tried to parse and merge them using SPSS. However parsing them into SPSS did not work and I could not find a clear reason why. It may be that SPSS had a difficult time with the size of the files that contained SAT data.

In the end, I decided to write four programs myself to do the work. The SAT score reports were in text format so it was not too much work to write C++ programs to parse them. I wrote one program to extract the useful variables from the College Board's files and output them it comma separated values format (CSV, a standard file format that many programs can read). Another program read those files and merged them into one file that contained the information from the last time a student took the test. The SAT score reports contained SAT scores from previous test dates, so this file included every score the student received. A third program converted this file into a format suitable for the WPS. What the score reports did not contain were analyses from previous dates (the number of correct, incorrect and omitted questions for each section). I wrote a fourth program that merged all the files as well, but the output contained for each student and each date, the analysis of that student's score. Using SPSS is was possible to merge these files together and get a single file that contained both the scores and analysis for each student.

One problem along the way was that SPSS still could not read the files when they were put into CSV format. This is more evidence that SPSS has difficulties reading large text files. Reading the files into Microsoft Excel and then saving them in Excel format solved this problem. SPSS had no problem reading the files after they were in Excel format.

Once the SES data were entered and SAT/PSAT scores were merged, then all of the data was stored in one of the four data sets and all of the data sets were organized using the student as the case reference. The final step was to link these data sets together to create a single one that contained all of the information. Most of the data sets contained high school id number as a variable, which made linking them together fairly easy. ID number is unique for all students and can be used to as a key to combine two data sets.

The data set most difficult to link was SAT and PSAT scores. These came from the College Board and did not contain student id number, which I would need to link the data to my other data sets. To add the student id number to this data set, I used the transcript data set, which contained every student's first and last name, date of birth, and student id number.

Initially I tried to get the id by matching a student's first name, last name and date of birth from the transcript data set to the SAT and PSAT data sets and then adding the student id from the transcript database to the SAT and PSAT data sets. The problem with this method was that relatively few matches were made. If a student bubbled in one character of their name or date of birth differently for the PSAT or SAT than was on their school records, then the match would fail. More matches would have been made if I had only used first name and last name to match on, however this would have created many incorrect matches since there were many students that shared the same first and last name. Similar problems would have occurred if I had used the other reasonable combination: last name and date of birth.

To solve those problems, I created two temporary student id variables, one based on matching a student's first name and last name and the other based on matching a student's last name and date of birth. Then a final student id variable was created based on the two variables. If the two variables had the same value then I would use that value for the final student id variable. If one value was missing then I would use the value I had. If both values were missing, then I would leave the final student id variable blank. If the variables had different values (this was caused by two student's sharing the same name or having the same last name and date of birth) then I would inspect the data by hand and assign the correct value. This method had the advantage of making a large number of matches and also having a double check for students who shared the same first and last name, or the same last name and date of birth.

Bumps

The previous section gives the picture that the data collection was done at the start of this project and that there were few bumps along the road. This was done in order to give a clearer view of the assembly process. In truth, however, the initial data collection effort did not fix all of the holes in the existing data set and more data had to be collected late in the project. There are two reasons for this: I was not able to collect all of the data that I wanted in the initial collection effort and more holes were discovered afterwards.

Both Matthew Marino's Project and Bai Lan Zhu's project were running late, so the SES data that they were in charge of was not available at the time of my analysis. This was not a major setback though, since I was planning to focus on the test-retest question, which could be answered without SES data. However it did mean that again an analysis based on the elusive SES variable would have to wait for another project.

The SAT data collected from Dr. Mostue of the WPS contained data from the 1996-1997 school year through October 2000. One problem from this was that though the data contained 1996 scores, it did not contain all of the scores from the class of 1996 (who also took SATs in 1995 and 1994). This could be solved using the data collected by Sauvron and Ballinger, however this data was in a much different form than the rest, so it would be inconvenient to use that. A larger problem was in the 1999

SAT data, which was given to Matthew Marino because his project focused on the class of 1999. This was a bad decision though, since the class of '99 would be taking the SAT mostly in late 1998 and early 1999, but not at all in late 1999. As a result the data was useless to Marino and he discarded it. The problem was this data *was* useful to me and I expected to be able to get it from him at any time, but by the time that I requested the data from Marino, he did not have it anymore. The last problem was that Dr. Mostue believed one of the test dates was missing from the set of diskettes that I received, since the College Board sent it to her blank. The scores would only exist in paper form and would be scattered at the six High Schools in Worcester. From what I could figure out the missing date was April 1997, but I had not confirmed this with the WPS.

Another area where data was missing was the academic transcripts. Dr. Mostue sent class of 1999 and 1998 data quickly after we requested them, however the class of 1997 and 1996 data was still not in. McCormick's database of the class of 1997 could be used to get the level of academic preparation variable, but again the data would be in a different form than the rest of the project. It was much more preferable to get the data from the WPS in the same form as before.

Lastly, a search through the data set created by Pieper and McCormick and the one created by Sauvron and Ballinger showed that neither contained any of the follow-up data that Noble collected. It was expected that at least a large portion of the data would be in one of the data sets, however this was not the case. We did however have these forms in paper form. The forms did not contain scores on them, but they did have batch numbers that could be sent to the Center for Applications of Psychological Type (CAPT), which scores the MBTI and would have the scores in their records.

Most of these deficiencies in the data set were discovered late in my project when I was hoping to move to the analysis stage. It was clear that another wave of data collection would have to be done.

It would have to be done quickly since the time needed to do a proper analysis and write up was running out.

At this point it was not clear that I had enough usable cases so I was forced to consider adding later classes to my data set. I made a list of the data that I had in my possession and the data that was available for the 2000 and 2001 classes:

CLASS	Number of SAT cases	Transcript Data
1996	0, (but Sauvron + Ballinger's had ~450 cases)	No
1997	453 (missing dates before the 96-97 year)	No (but McCormick's data set contained academic preparation)
1998	527 (missing 4/97)	Yes
1999	493 (missing 1999 dates)	Yes
2000	268 (missing 1999 dates)	Yes
2001	469 (missing 2001 dates, since they hadn't happened)	Yes

State of the data set after first wave of collection

1998 was the class year with the least amount of missing data. It was short the SAT cases from the April, 1997 date but no more. 1999 had the next least amount, missing the SAT cases from tests taken in 1999 however it seemed that most seniors took at least one test before then. 2000 had very few SAT cases, since most of that class took their SAT during 1999. 2001 had a fair amount of data, but it would be somewhat strange to use it without including 2000. 1997 was missing a good amount of SAT dates, the ones before the start of the 1996-97 school year and I had not received their transcript data from the WPS. 1996 was missing all of the SAT data and also had no transcript data.

If the second wave of data collection brought in no new data, then I would still be able to get a fairly sizeable amount of cases. In order to do this I would have to use the data sets assembled by Sauvron and Ballinger and by McCormick. This would give me an academic preparation variable for the class of 1997, SAT scores for the class of 1996 and more SAT scores for the class of 1997. However it would be much more difficult to put together a unified data set, since both of these were in different formats than I was using. In fact, it was not even clear what the format of Sauvron and

Ballinger's data set was, since the variables were not labeled. Also, the 1996-1999 classes would be missing the follow-up MBTI data that Noble collected. This would be quite disappointing, so I focused my attention on collecting more data.

The three main areas of data that were missing were the 1999 SAT test dates, the 1997 and 1996 transcripts and the follow-up MBTI data. I sent out a request to CAPT for the archive versions of the follow-up data and a request to Dr. Mostue of the WPS for the SAT and transcript data. I also sent a request to Dr. Mostue to confirm or deny if April 1997 was indeed the missing case, and for SAT data from the 1995-1996 school year, however this was less of a priority than the SAT and transcript data. At this point, all I could do was wait and hope.

The Second Wave of Data Collection

The second wave of data collection turned out to be a large success. All of the extra pieces of data that I requested were given to me and I was able to build a fairly complete data set. This was very good news. One new problem appeared though: after I had collected all of the new data, some of the reliability of my old data began to seem suspect. Even so, the data set that I finished with was definitely adequate for my purposes, and should serve future projects as well.

After some correspondence with CAPT they were able to locate the electronic copies of the follow-up MBTI data and sent it to me. These files did not contain student id number, however the paper reports did have the number hand written on them. I was able to add the student id number by hand copying it from the paper versions to the electronic versions. There were approximately 450 cases of MBTI data in these files. Not all of these cases were usable, but I was able to add a large portion of these cases to my data set.

Dr. Mostue was very helpful in getting me the data that I requested from the WPS. She created a access database containing the transcript data for the classes of 1996 and 1997 and gave me the disks containing the 1999 sat data. She also confirmed that the missing test data was April 1997, but by this time it was too late to collect the data from the schools on paper and enter it. This was not a major loss, since most students would have taken the test again in early 1998 and I could get their earlier scores from that test date. The one piece of data that I couldn't get was the SAT scores for the 1995-1996 school year. The district had not started collecting the scores in electronic form at that time so the data simply didn't exist. I didn't expect this to be a problem, since Sauvron and Ballinger's data set contained these scores from when they collected them by hand.

At this point I thought that I had complete data in all the areas that would go into my data set and I could move on to analyzing it. Unfortunately that wasn't the case. Once I began analysis, I discovered that some of the data that I had just added seemed very different from the data I had.

When I tried to use the 1996 and 1997 transcript data to group students by academic preparation I found that there was a large difference in the distribution between those years and the 1998 and 1999 years. This seemed to imply that transcript data that I had just received (1996 and 1997) was somehow created differently than the transcript data that I already had (1998 and 1999).

The transcript files were organized so that each class that a student took was one case. Since a student takes approximately 32 classes in their 4 years at school, I would need at least 16 cases for that student in the transcript file to classify their academic preparation reasonably. It would be better if these classes came from the junior and senior years. When I looked into the transcript data for 1998 and 1999 I found many fewer classes for a student than in the 1996 and 1997 data. In some cases there were no classes listed for a student for a given year. It is not clear what happened, but the suspicion is that the 1998 and 1999 files were somehow truncated when they were emailed to us (each file was around 60 megabytes). The 1996 and 1997 were retrieved in person using 2 zip disks so they would not be affected. Because of this I decided that the 1998 and 1999 data might not be valid and that for my analysis I would only use 1996 and 1997 data when I needed the academic preparation variable.

During the analysis of SAT score increase I noticed that the pattern in the 1996 data was much different than that in other years. These were the scores that I collected from Sauvron and Ballinger's data set, instead of directly from the College Board's files. When I examined Sauvron and Ballinger's data set I found the data was much different than other years. Almost all the students in it took the SAT twice (this number was only about 50% for the later years) and that almost all of them raised their score by over 100 points (for the later years only about 40% raised their score 50 points). This led me to believe that Sauvron and Ballinger's data set was for some reason incompatible with mine and I decided that I would have to drop the 1996 data from my analysis.

After discussing the matter with Professor John Wilkes, the advisor of this project, he convinced me that the data could have been used, and the differences were caused by differences in WPS policy. In 1996, very few students were taking the PSAT since it was before the WPS's policy of paying for the test in their sophomore year. This meant that most students were unfamiliar the format of the SAT when they first took it and would probably score low. The WPS considered the low test scores an embarrassment, and concluded that the students needed more practice. Hence the pressure was on the guidance officers to encourage retaking the SAT. This would explain why so many students in the class decided to take the test for a second time.

After they gained so many points by taking the test over, it is easy to see how the WPS policy of encouraging sophomores to practice with the PSAT emerged in the next year. It is also clear why the WPS was convinced that a practice round was worthwhile, enough to pay for poorer students taking the PSAT as juniors. However I discovered the differences very far into my analysis and I did not have time to consult with Professor Wilkes before I finished it, so my analysis makes very light use of the 1996

SAT scores. Using the scores would have distorted the later, more normal patterns of test taking in any case. However, it is worth noting that the rise in scores from first SAT to second SAT for the class of 1996 and the rise from PSAT to SAT for the later classes was comparable; the practice policy of the WPS had the intended effect of raising the average SAT scores for the district.

Dropping those two variables hurt my analysis, but not by very much. For most analyses I could use the classes of 1997, 1998 and 1999, which is a fairly large sample. I was most affected when I wanted to look at academic preparation since I was only able to use the class of 1997 data set, but most aspects of test-retest could be answered without this variable.

Overall, the data set was in good shape after this wave of data collection. I was fortunate to have almost all of my requests for extra data granted and I was able to use all the variables I was interested in once I moved on to my analysis. However, before the analysis can be discussed I must give the hypothesis that I used going into it.

Sauvron and Ballinger's Test-Retest Hypotheses

The test-retest portion of this project was able to leverage heavily on Sauvron and Ballinger's work. This was quite helpful while writing my hypotheses, since it allowed me to compare my ideas to Sauvron and Ballinger's who first hypothesized on the test-retest subject. Their hypotheses will be given before mine with a comparison of the two so the reader can contrast them.

One of the main differences between the two hypotheses was that Sauvron and Ballinger hypothesized based on only the S-N dimension of the MBTI, since it had the strongest relationship to SAT score. However I thought that it would be better to examine all of the dimensions, since it would provide a more complete picture of the situation.

Sauvron and Ballinger predicted that Sensors would see a higher average score gain than iNtuitives. This is partly because it is easier to improve a low score than a high one, and Sensors score lower on average than iNtuitives and partly due to their preference to learn “hands on”, by raw experience. In my hypotheses, I took the same stance as Sauvron and Ballinger and also extended it to the other dimensions. Both Sensors and Judgers should all see a higher average score gain than their counterparts because they score lower initially and are most likely to fall prey to attractor items, due to a need for closure. I hypothesized that they “close” early on the first reasonable answer without reading the rest of the possibilities, and are usually tripped up by “attractor” items.

The team also predicted that Sensors would be more likely to retake the test since iNtuitives would be more satisfied with their scores. They were not totally confident on this point though, and also gave the counter-hypothesis that low initial scores could discourage Sensors. Here I disagreed with Sauvron and Ballinger. Instead of either satisfaction or discouragement with one’s initial score, I considered the most important factor here the student's level of academic interest. Since iNtuitive students are more likely to have aspirations of going to a selective college (and college in the first place), they will be the most likely to want to improve their score. This will cause iNtuitives to take more PSAT tests to prepare and more SAT tests to try to improve their score if the first score is low enough to be an impediment to their plans, hopes and desires. A low score could also threaten their confidence and self-image if it is radically out of line with their academic records in high school.

My Test-Retest Hypotheses

A person's MBTI type represents the way they prefer to view the world and the method of reasoning they prefer to use to deal with it. It is only reasonable that personality type would have a significant effect on the probability that a person would take multiple SAT/PSAT tests and the relative score gain that they would receive as a result. How they will react to “low” or “high” scores is probably a characteristic of type, assuming equal level of support from academic achievement across school environments.

In order to best comprehend the effect of personality type on SAT practice, it was helpful to me to split the issue into three related, but distinct components: the likelihood of each type to practice for the SAT by taking a PSAT test in a setting where this is encouraged, but not required; the likelihood of each type to take an additional SAT test to improve their score; and the score gain that each type gets after practice with an initial SAT/PSAT test. These questions of practice, retest and gain were addressed separately.

Another question that could be addressed was which students were most likely to prepare for the SAT using methods other than the PSAT. This would include preparation from books, workshops, full-length courses and computer programs. Part of the form that SES data was taken from included a self-reported answer on which of those methods were used. I decided the question to be beyond the scope of this project, however it would be a good topic for a future research team, when the SES data from all 4 schools is finally assembled.

In deciding if a student is both likely to practice with the PSAT and likely to retest on the SAT, the most predictive factor should be their level of interest in academics and the degree of investment they have in academic success as part of their self-image. Students who aspire to go to a "highly selective" college have much more at stake when they take the SAT than students who are planning to go to "less selective" colleges, vocational schools or the workplace. My model is more of a "rational choice" model than Sauvron and Ballinger's "subjective discouragement" model. In my view, those students with more to gain are much more likely to spend the time, money and effort it takes to take multiple SAT/PSAT tests. This means that the types more likely to retake the test are the ones that normally do well in school, the iNtuitives and Introverts. These types also score higher on average than their classmates so they are less likely to be discouraged by a low initial score. In fact, they will likely be encouraged by guidance counselors to try to maximize their scores. Since Matthew Marino has learned that the WPS records the future plans of graduating seniors and literally records the college they say they have gained acceptance to, whether they are joining the military, taking a job, becoming a homemaker, etc. The possibility of seeing how well retesting patterns align with college selection patterns is a promising topic for a future team to explore.

If academic interest is controlled by looking at students who have taken programs consisting of classes of a similar level of difficulty then this difference will disappear and another factor will become visible. I predict that Sensors will be more likely to practice using the PSAT than iNtuitives. Sensors prefer to learn by direct experience, while iNtuitives prefer to learn the general concepts behind things. Sensors will want firsthand experience with the kind of questions that will be asked, which is exactly the purpose of the PSAT. Intuitives, on the other hand, will want an understanding of the subjects that are on the SAT and would more likely study using books on SAT strategy or similar means.

The thinking-feeling dimension normally should not account for much difference in academic interest since it is rarely associated with grades or aspiration levels, although it does impact which fields a person will choose to study and what careers they will choose. However, because of the special circumstances in the Worcester school system, the Thinking-Feeling dimension should have an impact on test taking.

The WPS was encouraging students to take the PSAT as both sophomores and juniors with the goal of making it easier to practice for the SAT, though there was a risk of students being discouraged from low scores. MBTI theory suggests that Feelers would be more responsive to this official encouragement from a counselor, therefore would be more likely to take the PSAT multiple times than Thinkers. This effect should be amplified if the student is also an Extravert, since Extraverts are more likely to respond to encouragement from outside sources.

The situation will be different when the SAT is considered. The WPS did not set up a special policy to encourage students to take the SAT like they did the PSAT. High schools generally encourage everyone to take the SAT once, even if they aren't sure about wanting to attend college, so this may cause slightly more Feelers to take the SAT at least once, however this won't be a factor in taking the SAT multiple times. Instead, Thinkers will retake the SAT more than Feelers. This is because the different preferences will cause students to take different views of the question, "should I take the SAT again?" The Thinkers will want to look at it from a numerical point of view. Their score can only go up, so they have nothing to lose. The Feelers will want to look at it from a personal point of view. They will see drawbacks to taking multiple tests: the effort it will take them to study and the hours they will spend sitting in a room on a Saturday morning in order to take the test.

To summarize, I thought that the most predictive factor in determining if a student practices with the PSAT or retakes the SAT is academic interest. This will causes Introverts and iNtuitives to be more likely to do both. If students are only compared to students who have taken similar courses, then this difference will disappear and instead Sensors will practice more with the PSAT than iNtuitives. Also, Feelers will be more responsive to any encouragement that they get to take the PSAT and will take it more, but Thinkers will see more benefit to retaking the SAT if it will get them where they want to go than Feelers so they will retake the SAT more often.

The types that score lowest on the SAT test, Sensors and Judgers, will benefit the most from practice. I predict that Sensors will be less likely to try again in the first place. These types will have an easier time increasing their scores, since they will have lower scores to start with and will have more room to improve. By the same reasoning, the improvement for Sensors and Judgers will be seen the most in the verbal section, where they score lowest in comparison to their iNtuitives and Perceivers.

Test-Retest: Identifying the Pertinent Data and Variables

In order to test those hypotheses, several variables were developed and analyzed. The rationale for choosing these variables follows:

To determine which types were most likely to practice with the PSAT or retake the SAT there was an obvious choice for the variable: the number of times each student took the PSAT and the SAT. This was easily calculated and it measured exactly what was wanted.

The choice of variables to measure the gain that the different types get from practice was more difficult since gain can be looked at in several different ways. Three variables were used, each reflecting a different aspect of the situation.

To compare different classes I looked at the total increase in their combined SAT score (the difference between a student's first score and his or her highest score when they took the measure more than once). This variable was easy to calculate and showed the general effects of the WPS policy on its student's total score.

To analyze which personality types were gaining the most from practice, I examined percent of each type saw a "large" gain as a result of retaking the SAT. A large gain was defined as one that increased the student's score by 50 or more points. The reasoning was that gains lower than 50 probably would not be seen as much different in the eyes of a college recruiter. Also, the college board estimates that students see an average increase of 50 points, so students who saw greater increases than this could be seen as having a higher than average gain.

Lastly, when I wanted to examine differences in the math and verbal scores separately, I used the initial gain for each section (the difference in score between the first and second tests). Initial gain was chosen over average gain, the total gain divided by the number times the student retook the test, because average gain would have been unfair to students who retook the SAT many times. The gain a student gets from retaking the SAT is highest initially and then goes down the more that he or she retakes the test, unless he or she is in a test preparation program of a very high caliber. Students who retook the test more times would see a lower average gain than initial gain. This lower gain would not be reflecting a real difference in the amount that their score went up, simply that the student decided to retake the test more times afterwards.

Throughout all of the analysis, academic preparation was used as a control variable. Academic preparation has already been shown by McCormick to be an important factor in a student's SAT score and I believe it is a large factor in amount of encouragement a student receives to take the PSAT and to retake the SAT. This led to it being reasoned that academic preparation would positively influence both the amount of times a student would take the PSAT/SAT and the gain that the student would receive.

Analysis of the Data

Introduction

Once the data collection process was finished, the data set contained a fairly large number of cases with SAT, PSAT, MBTI and academic preparation data. The subset of cases that contained combinations of these (students for whom we had SAT, MBTI and academic preparation data, for instance) was of course smaller, however a fairly in depth analysis was still possible.

Year	Total N	1 PSAT	2 PSATs	1 SAT	2 SATs	3 SATs	MBTI	Prep.
1996	816	87	1	450	440	0	388	799
1997	962	548	3	435	255	72	663	782
1998	986	558	165	453	263	66	634	768
1999	1113	416	132	406	181	44	749	662
Total	4324	1610	301	1743	1138	182	2327	3121

Number of Cases Collected

General Analysis

From the previous chart, one can see the effects of the WPS's policy regarding PSATs. For the class of 1996, before any policy went into effect, very few students took even one PSAT test. For the class of 1997 the students were encouraged to take the test as juniors and thus have a high number of students taking the test one time. The classes '98 and '99 were encouraged to take the test as sophomores and juniors and thus have a higher number of students taking the test once and a sizeable portion of students taking the test a second time. This led to interesting results when the classes' scores are compared.

Year	Avg. 1st PSAT score	Avg. 1st SAT score	Avg. Final SAT score
1996	94	796	898
1997	85	864	901
1998	79	882	916
1999	79	896	914

The classes of '98 and '99 saw their average score increase from a first PSAT score that was just under an 80 to the high 800s first SAT score to the low 900 final SAT score. The class of '96 for the most part did not take a PSAT, instead choosing to begin with the SAT (the ones who did take it seem to be the more serious students, scoring an average of 94). The average first SAT for the class of '96 was a 796 and their average final average score was an 898.

The similarities are striking between the class of 96's first SAT score and the class of 98 and 99's first PSAT score, as well as the class of 96's second SAT score and the classes of 98 and 99's first SAT score. One possible explanation for this is that these classes went through the same process, just using different tests. The classes of '98 and '99 took the PSAT then the SAT while the class of '96 went through a similar process taking two SATs.

One difference was the class of '98 and '99 was able to raise their score by about 25 points, by retaking the SAT a second time. However a large part of this could have been caused by variance in their scores between tests and the fact that students are allowed to use the highest verbal and math scores they receive. Very few of the class of '96 took the SAT three times, so they would not have seen this gain.

The class of '97 is a third case, who went through a different process from both the class of '96 and the classes of '98 and '99. For the most part the class began the testing process with the PSAT in their junior year, which gave them an initial advantage over the later classes, who took the PSAT their sophomore year and the class of '96 who took the more difficult SAT their junior year. The average first PSAT score for the class of '97 was 85, significantly higher than the other classes starting scores. The class's score then improved only slightly for their first SAT, averaging an 864, however they improved more than other classes the third time around and ended with a average score of 901. Though the class of '97 went through a much different test taking process, their ending score was very close to the other classes.

Overall, it appears that the WPS classes were able to increase their scores by about 110 points between their first encounter with the test and their second. What portion of this increase came from an extra year of study and what portion came from added familiarity with the test? The unique situation of the class of '97 is useful in answering this question.

The WPS began encouraging practice with the PSAT while the members of the class of '97 were juniors and the members of the class of '98 were sophomores. This means that these two classes took the PSAT for the first time together, the main difference being that the class of '97 had an extra year of schooling. The class of '97 outscored the class of '98 by an average of 6 points (equivalent to 60 SAT points). It seems that approximately half of the 110-point increase in score came from maturation and the other half came from familiarity with the test.

On a side note, during this period of time the WPS came under criticism for low average SAT scores. However, much of this criticism was unwarranted as it was based on misreported information. The Worcester Telegram and Gazette reported, on April 20, 1997, that the average SAT score for the WPS class of '96 was 793 (although the class year was not actually specified, this number almost certainly referred to the class of '96, since the class of '97 had not graduated nor had they ever averaged a score below 850). The Telegram and Gazette used the class's average initial score, a number that was lower than the class's final SAT score (the score they would report to colleges) by the very substantial margin of 100 points. It is not clear who is at fault, the Telegram and Gazette for sloppy reporting or the College Board for hard to understand score reports, but the lesson is clear: the average final SAT score for a class can be very different from the initial score, and care must be taken to ensure that the correct number is used when reporting the scores.

Analysis of PSAT/SAT Participation

I began my analysis by examining what types were taking the SAT and PSAT most often. For this, I looked at what percent each type took an initial PSAT, a second PSAT, an initial SAT and a second SAT. Below are the results for the years of 1997, 1998 and 1999. Unfortunately, 1996 was left out because of the reasons mentioned in the methodology.

Note that the percentages for students taking 2 SATs or 2 PSATs is the percent of those that already took 1 test, not of the entire group. Also the percentages of students taking one PSAT and 1 SAT is much lower for 1999 than the other years. Part of this is a reflection of fewer students taking the PSAT in that class (and slightly fewer taking the SAT in that year), but the numbers are exaggerated because of the method I used to determine class year. I used the class year that the student was expected to graduate in the 1995-1996 school year, when Pieper's study was being done. This resulted in a large number of students considered to be in the class of 1999, since it included students who would later transfer to other schools, fail a grade or drop out of school (the drop out rate is about 8% for the WPS). The low percentages are a combination of a real difference in students taking the tests and an artificially high number of students counted as the "base" for the class of 1999.

	1997		1998		1999	
	E (N)	I (N)	E (N)	I (N)	E (N)	I (N)
1 PSAT	**73.7% (283)**	**74.2% (207)**	69.8% (261)	75.4% (196)	**44.9% (208)**	**45.8% (131)**
2 PSATs*	**0.7% (2)**	**0.5% (1)**	26.8% (70)	34.7% (68)	**36.5% (76)**	**31.3% (41)**
1 SAT	**57.8% (222)**	**53.4% (149)**	53.7% (201)	54.6% (142)	**48.8% (226)**	**45.1% (129)**
2 SATs*	**64.9% (144)**	**55.0% (82)**	59.2% (119)	59.9% (85)	**42.9% (97)**	**45.7% (59)**
Total N	**384**	**279**	374	260	**463**	**286**

SAT/PSAT participation by Extraversion-Introversion

	1997		1998		1999	
	S (N)	N (N)	S (N)	N (N)	S (N)	N (N)
1 PSAT	**71.4% (279)**	**77.6% (211)**	67.8% (257)	78.4% (200)	**40.7% (186)**	**52.4% (153)**
2 PSATs*	**1.1% (3)**	**0.0% (0)**	31.9% (82)	28.0% (56)	**32.3% (60)**	**37.3% (57)**
1 SAT	**52.9% (207)**	**60.3% (164)**	49.6% (188)	60.8% (155)	**41.6% (190)**	**56.5% (165)**
2 SATs*	**60.4% (125)**	**61.6% (101)**	59.0% (111)	60.0% (93)	**44.7% (85)**	**43.0% (71)**
Total N	**391**	**272**	379	255	**457**	**292**

SAT/PSAT participation by Sensing-Intuition

	1997		1998		1999	
	T (N)	F (N)	T (N)	F (N)	T (N)	F (N)
1 PSAT	**68.1% (276)**	**82.9% (214)**	69.5% (264)	76.0% (193)	**46.2% (214)**	**43.7% (125)**
2 PSATs*	**0.7% (2)**	**0.5% (1)**	28.8% (76)	32.1% (62)	**31.8% (68)**	**39.2% (49)**
1 SAT	**53.1% (215)**	**60.5% (156)**	52.9% (201)	55.9% (142)	**44.7% (207)**	**51.7% (148)**
2 SATs*	**52.6% (113)**	**72.4% (113)**	58.2% (117)	61.3% (87)	**44.9% (93)**	**42.6% (63)**
Total N	**405**	**258**	380	254	**463**	**286**

SAT/PSAT participation by Thinking-Feeling

	1997		1998		1999	
	J (N)	P (N)	J (N)	P (N)	J (N)	P (N)
1 PSAT	**74.1% (212)**	**73.7% (278)**	70.6% (185)	73.1% (272)	**41.4% (121)**	**47.7% (218)**
2 PSATs*	**0.9% (2)**	**0.4% (1)**	31.9% (59)	29.9% (79)	**34.7% (42)**	**34.4% (75)**
1 SAT	**57.3% (164)**	**54.9% (207)**	50.8% (133)	56.5% (210)	**45.9% (134)**	**48.4% (221)**
2 SATs*	**57.9% (95)**	**63.3% (131)**	67.7% (90)	54.3% (114)	**53.0% (71)**	**38.5% (85)**
Total N	**286**	**377**	262	372	**292**	**457**

SAT/PSAT participation by Judging-Perceiving

*** These percentages represent the number of students who took the test twice out of those who already took it once.**

I expected Introverts to take both the PSAT and the SAT more often, however this wasn't the case. There were some differences in the proportion of Introverts and Extraverts taking test, but there was no consistent notable pattern. This happened for the class of '98, here the Introverts did take more first and second PSATs, but there was essentially no difference for the classes of ’97 and ’99 between Extraverts and Introverts. There was essentially no difference in taking multiple PSATs for those years either. Only in the class of '98, the year that the new policy of practice with the PSAT was introduced (while the class of '98 was sophomores), did the expected pattern emerge, where Introverts took and retook the PSAT more often.

I made no prediction for the J-P dimension and there was little difference when comparing the two groups. Apparently this dimension does not have very much impact in determining if a student will take or retake the PSAT and SAT.

In the S-N and T-F dimensions there was a type that was clearly taking the tests more often, the iNtuitives and the Feelers, respectively. Hence, I examined the S-N and T-F dimensions in more detail.

Analysis of the S-N Dimension

The analysis of the S-N dimension was consistent with the predictions that iNtuitives would be more likely to take the PSAT and SAT. The hypothesized reason was that iNtuitives tend work better in academic settings and are more likely to aspire to go to college. This means that they have more at stake in the test and so would both be more likely to practice with the PSAT and retake the SAT.

Following that logic, if the students with similar levels of academic interest (preparation) were compared, then the larger amount of iNtuitives should disappear. In fact, I hypothesized that when comparing groups of similar academic interest, the reverse would happen and Sensors would be more likely to take the tests, since they prefer to learn through first hand practice. I performed this analysis on the class of '97. The reason for excluding classes of '98 and '99 was that they had less reliable transcript data. This was somewhat limiting since there were not enough cases to analyze students taking 2 PSATs in that class year; it was before the WPS policy of encouraging PSATs came into full effect. However this does not seem to be of great concern, since the results were closely in line with predictions and there is no reason to believe that they wouldn't be for the other classes.

(Again, note that the percent for students taking 2 SATs is out of the number off students who took 1 SAT to begin with)

WPS Class of 1997

Academic Interest	**Mostly General/ Some College**		**Mostly College/ Some Honors**		**Mostly Honors/ Some College**		**Some AP/ All Honors**	
S-N	**S**	**N**	**S**	**N**	**S**	**N**	**S**	**N**
Took 1 PSAT (N)	51.2% (43)	43.5% (10)	80.9% (110)	74.2% (66)	91.0% (71)	95.1% (77)	96.8% (30)	100% (46)
Took 1 SAT (N)	22.6% (19)	26.1% (6)	68.4% (93)	57.3% (51)	75.6% (59)	72.8% (59)	87.1% (27)	80.4% (37)
Took 2 SATs (N)	26.3% (5)	0.0% (0)	53.8% (50)	60.8% (31)	83.1% (49)	64.4% (38)	63.0% (17)	81.1% (30)
Total	84	23	136	89	78	81	31	46

SAT/PSAT Participation by Sensing-Intuition and Academic Interest

Once academic interest is controlled Sensors are indeed taking the tests at least as much as the iNtuitives. There is also good evidence to support the claim that Sensors would be taking more PSAT tests than iNtuitives, but this pattern is limited to the two lower academic levels. Sensors took one PSAT less often except at the higher academic interest levels, but at this level over 90% of the students of both types were taking the test. Interestingly, Sensors took the first SAT more often than iNtuitives in every interest level except Mostly General/Some College, and there they trailed by less than 4 percentage points. The situation was not at all clear when it came to students taking a second SAT though, with large the percentage going both ways as one moves through the different interest levels. Among the students in the toughest program the iNtuitives were most likely to be disappointed and 81% retook it. At the next level the Sensors were more likely to retake it (83%) while the iNtuitives at this level and the next settled into a pattern of nearly two thirds retaking it (64% and 61% respectively). The proportion of Sensors retaking dropped in steps from the high of 83% to 53% and to 26% as one gets into successively less challenging academic programs.

Analysis of the T-F Dimension

In my hypothesis I predicted that Feelers would be more responsive to the WPS's policy on PSATs and would therefore take the PSAT in greater numbers than Thinkers. I also reasoned that Thinkers would be more likely to retake the SAT, since they would focus on the numerical argument of "my score can only go up," instead of focusing on the emotional drain of sitting in a testing room for hours on a Saturday morning. Hence, they would be less affected by encouragement, official and otherwise.

As it turns out, Feelers were not only more likely take one PSAT and a second PSAT, they were also more likely to take the first SAT and to take a second SAT for almost all of the years. One explanation for this is that females are more likely to be Feelers than males, and the WPS believed the female students to be more serious students than their male students[10]. In addition, females generally score lower on the SAT than males so they might have been motivated to try again to prove that they could score well by retaking the SAT. To test this theory I controlled by gender and analyzed the T-F dimension again.

---

[10] Wilkes, reporting a private communication from the Deputy Superintendant James Caradonion, who was discussing grades, attendance and awards. He found equal performance by males and females on the PSAT surprising, since by every other measure the females were outperforming the males.

FEMALES

Test	**Took 1 PSAT**		**Took 2 PSATs**		**Took 1 SAT**		**Took 2 SATs**	
T-F	**T**	**F**	**T**	**F**	**T**	**F**	**T**	**F**
1997 *(N)*	70.8% (119)	87.6% (162)	0.8% (1)	0.0% (0)	63.1% (106)	64.3% (119)	53.8% (57)	77.3% (92)
1998 *(N)*	66.9% (107)	76.3% (148)	31.8% (34)	35.8% (53)	56.3% (90)	57.7% (112)	66.7% (60)	65.2% (73)
1999 *(N)*	42.3% (88)	44.9% (96)	29.5% (26)	39.6% (38)	43.3% (90)	54.2% (116)	47.8% (43)	44.0% (51)

Female SAT/PSAT Participation by Thinking-Feeling

MALES

Test	**Took 1 PSAT**		**Took 2 PSATs**		**Took 1 SAT**		**Took 2 SATs**	
T-F	**T**	**F**	**T**	**F**	**T**	**F**	**T**	**F**
1997 *(N)*	66.2% (157)	71.2% (52)	.6% (1)	1.9% (1)	46.0% (109)	50.7% (37)	51.4% (56)	56.8% (21)
1998 *(N)*	71.4% (157)	75% (45)	26.8% (42)	20.0% (9)	50.5% (111)	50.0% (30)	51.4% (57)	46.7% (14)
1999 *(N)*	49.4% (126)	40.2% (29)	33.3% (42)	37.9% (11)	45.9% (117)	44.4% (32)	42.7% (50)	37.5% (12)

Male SAT/PSAT Participation by Thinking-Feeling

It seems that the main reason that Feelers were taking more SAT tests in the class of 1997 was indeed gender. The females were more likely to take the PSAT that year and females tended to be Feelers more often then males. In the later years (the classes of '98 and '99), the relationship breaks down for the PSAT, but the women are still more likely to take the SAT in the class of 1998 and especially more likely to retake it. For the class of 1999, the relationship weakens, especially among the Thinkers. Feeling females were still more likely to take the SAT in the class of 1999, by a 10% margin.

Once gender is taken into account there was little difference in SAT participation between the Thinkers and the Feelers. Virtually all of the differences between the two groups were not significant at a .1 level (A .1 significance means that the difference in participation levels would appear at random 10% of the time). The only exception was the female Feelers taking more second SATs in the class of '97. More investigation might explain these differences, however I cannot give a theoretical reason for them currently. In general though, the T-F dimension had little independent impact on SAT participation, though gender did.

When gender was taken into account for first PSAT participation, the T-F dimension accounted much less difference, however there were still some significant differences. The prediction that Feelers would take more first PSATs than Thinkers held for females in the class of '97 and '98, however the class of '99 had little difference between the types. For males, none of the differences were significant. Sample size reduction due to my unwillingness to pool the 97-99 data sets hamper our ability to look at the impact of type, gender and program all at once, and make it harder to produce statistically significant findings.

For the second PSAT, the female Feelers also took the test slightly more than Thinkers, however this differences were not significant. For males, the Feelers took more tests in the class of '99, but Thinkers took more in the class of '98 and neither of the differences was significant.

There is an interesting interpretation of the finding that, for females (and for the class of '97 and '98 especially), the Feelers were more likely to comply with their school system's advice of practice for the SAT using the PSAT. When one contrasts the types of encouragement that are effective for Thinkers and Feelers (reward after the fact for Thinkers and emotional support beforehand for Feelers), it seems the Feelers want encouragement generally thought of as more "feminine". It is possible that gender stereotypes caused the WPS guidance counselors to give a more Feeling-type encouragement to women and this motivated the Feelers more than it did the Thinkers. This difference would be especially apparent for the classes of '97 and '98, since the policy first went into effect when they took the PSAT. This theory does not explain, however, the mixed results for the male classes.

To summarize, the WPS females were taking both PSAT and SAT tests more often than the males. This resulted in more Feelers taking the tests, since females are more likely to be Feelers. Once gender is controlled for this difference largely disappears, however the female Feelers seem to have complied with the WPS policy of practice with the PSAT more than the female Thinkers.

Analysis of Gain from Practice

To analyze which personality types were gaining the most from practice, I looked at which students were seeing a large gain (greater than 50 points) as a result of taking the SAT twice. For the first analysis, I grouped the classes of '97, '98 and '99, since the point was to identify which dimensions should be examined in more detail.

Type	**E**	**I**	**S**	**N**	**T**	**F**	**J**	**P**
gained >= 50 points (N)	42.8% (154)	34.1% (77)	38.3% (123)	40.8% (108)	38.7% (125)	40.3% (106)	37.9% (97)	40.6 (134)
gained < 50 points (N)	57.2% (206)	65.9% (149)	61.7% (198)	59.2% (157)	61.3% (198)	59.7% (157)	62.1% (159)	59.4% (196)
Total N	360	226	321	265	323	263	256	330

% Seeing Large SAT Gains Between Tests 1 and 2 by Type

My prediction was that the types that do tend to score worse on the SAT, the Sensors and Judgers, would be the most likely to increase their score. The reasoning was that these types would have the easiest time improving since they were starting from lower scores. However, this turned out not to be the case at all. Intuitives, Feelers and Judgers tended to see an increase more than their counterparts, however the difference was very small and none were significant at the .1 level. The only dimension that was making significant difference was I-E, with the Extraverts getting big gains more often than Introverts.

Since the I-E dimension was the only one to have a significant impact on a student's gain from practice, the rest of this section will be devoted to it.

Analysis of the I-E Dimension

Having identified a difference in the scores of Introverts and Extraverts I examined the difference between the two groups for the different class years.

Type	**Extravert**	**Introvert**
1997 - gained >= 50 points (N)	38.2%(55)	32.9% (27)
1998 - gained >= 50 points (N)	42.9% (51)	36.5% (31)
1999 - gained >= 50 points (N)	49.5% (48)	32.2% (19)

Extraverts vs. Introverts - % Seeing Large SAT Gains by Class Year

In every class year the Extraverts were more likely to see a large score increase than the Introverts. This is convincing evidence that Extraverts are better able to improve through practice than Introverts. I now examined how the Extraverts were gaining more than the Introverts, by examining the math and verbal section separately and by controlling based on the student's starting score.

Type	**Extravert**	**Introvert**
Avg. Math increase	13.75	13.81
Avg. Verbal increase	17.83	8.54

Average Verbal and Math Score Increase by Extraversion-Introversion

Type	**Extravert**	**Introvert**
400-599 - gained >= 50 points (N)	90.9% (9)	62.5% (5)
600-799 - gained >= 50 points (N)	47.1% (48)	45.8% (22)
800-899 - gained >= 50 points (N)	39.0% (30)	37.5% (18)
900-999 - gained >= 50 points (N)	41.4% (29)	31.6% (12)
1000-1099 - gained >= 50 points (N)	42.9% (24)	26.5% (9)
1100-1199 - gained >= 50 points (N)	32.3% (10)	11.1% (3)
1200-1399 - gained >= 50 points (N)	28.6% (4)	36.4% (8)
1400-1600 - gained >= 50 points (N)	0.0% (0)	0.0% (0)

% Seeing Large SAT Gains by Initial Score and Extraversion-Introversion

The verbal section had a large difference in the score increase between Extraverts and Introverts while the math section saw almost no difference in score increase between the two types. Also, the largest differences occurred when the student scored either abysmally or relatively well on his or her first attempt, not in the middle range of 600-899. While the Introverts have a more difficult time sharply improving from scores above 900, the Extraverts are able to improve about the same percentage of the time as those starting with lower scores. Above 1200 points was the exception, but very few students scored that high so it is difficult to draw any good conclusions based on the small number of cases in these categories.

Conclusions

For this study, I examined how personality type affects a high school student's journey through the process of PSAT and SAT test taking and retaking. It found that psychological type affects the amount of times that a student is likely to practice with the PSAT, retake the SAT and the amount of gain that the student is likely to see as a result. Evidently personality type is intimately connected with the entire process of SAT test taking, and hence is a background influence that shapes college admission credentials.

One important finding was that iNtuitive students are more likely to take both the PSAT and the SAT than sensing students. This probably happens because iNtuitives were more likely to be in more challenging academic programs and therefore more likely aspirations of going to a more prestigious college. On the other hand, when the students that were in similar academic programs were compared Sensors were more likely to take both the PSAT and SAT. This is evidence supporting Isabel Myer's view that the American educational system serves Sensors less well than iNtuitives, since the curriculum is biased against their preferred way of learning – at least in the "higher-level" courses.

Another finding was that the female Feelers were more likely to follow the advice of their school system and practice for the SAT by taking the PSAT. It makes some theoretical sense that Feelers would respond to advice more, but why only the WPS females responded to this advice is unclear. One possible explanation is that the WPS guidance counselors' advice was influenced by gender stereotypes, causing female Feelers to receive advice more suitable for their psychological type than female Thinkers. This seems like an interesting topic for future research.

The last finding was that Extraverts had an easier time improving their scores from practice than Introverts. This is counter to my hypothesis was that types that scored lower would find it easier to improve their score since they more room to grow. For instance, Sensors score significantly lower on average than iNtuitives, but the two types were able to increase their score equally well. Further analysis into the issue showed that Extravert's advantage came mostly from the verbal section and from fairly high initial scores. I cannot give an explanation for this phenomenon. Again, I leave must the issue to future researchers. The only clue available is from a WPI student study documenting higher levels of confidence and self perceived academic ability for Extraverts on arrival at college. This did not lead to Extraverts outperforming their Introvert classmates in terms of freshman year grades, but it might reflect a pattern of greater confidence in the SAT test retake situation when seeking admission.

The one major disappointment of this study was that the social class variable (still) could not be analyzed. Noble's hypothesis that SAT score differences by social class are caused by differences in academic preparation is very interesting and it would be quite nice to test it. It is also likely that social class has an effect on the test/retest issue. In addition, there are many more interesting topics could be explored once a social class variable was entered into the data set. I hope that the social class variable can be entered into the data set in the near future, as it is long overdue to be studied.

In addition to the leftover questions from this study and the issue of social class, there are many more topics to explore in the study of personality type and the SAT. The data set created by this project should allow future teams to study these topics without spending too much time on data assembly.

I wish these teams luck in their efforts. The topic of SAT test taking is a complex one, but the enormous impact of the SAT on students' lives makes its study an important task. I hope that this research aids to our understanding of the SAT and standardized tests in general and allows college admissions officers to make more informed decisions on the use of SAT scores to predict academic potential.

Bibliography

Briggs-Myers, Isabel, McCaulley, Mary, Quenk, Naomi, and Hammer, Allen. MBTI Manual: A Guide to the Development and the Use of the Myers-Briggs Type Indicator. Palo Alto, CA: Consulting Psychologists Press, 1985.

Lawrence, Gordon. People Types and Tiger Stripes. Gainesville, FL: Center for Applications of Psychological Type, Inc., 1993.

Briggs-Myers, Isabel with Myers, Peter. Gifts Differing. Palo Alto, CA: Consulting Psychologists Press, 1980.

Martin, Charles. Looking at Type: The Fundamentals. Gainesville, FL: Center for Applications of Psychological Type, Inc., 1997.

Lemann, Nicholas. The Big Test. New York, NY: Farrer, Straus and Giroux, 1999.

Kingsland, David, McTague, Kibler. Learning Style Biases in the S.A.T. Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute 1995.

Batey, Daniel, Brezniak, Paula, and Purohit, Ashwin. Cognitive Bias in the SAT. Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1995.

Pieper, John. Cognitive and Learning Style Biases in the SAT/PSAT. Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1996.

Ballinger, Emily and Sauvron, Jennifer. The PSAT/SAT Test-Retest Demonstration Study: Administrative Reactions to a Learning Styles Student Information System. Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 1997.

Noble, Gerald. How Much Does Social Class Affect PSAT Performance? Unpublished Interactive Qualifying Project, Worcester Polytechnic Institute, 2000.

Mostue, Patty. 1997 SAT Fact Sheet. 1997.

Appendix A: SRTT Analyses

Source of data	Group tabulated:	MBTI Type Table
		Center for Applications of Psychological Type
	Portion Taking 1 PSAT	Legend: % = percent of total choosing this group who fall into this type.
(null)		I = Self-selection index:
	N = 1286	Ratio of percent of type in group to % in sample.

SENSING types with THINKING	SENSING types with FEELING	INTUITIVE types with FEELING	INTUITIVE types with THINKING
ISTJ	ISFJ	INFJ	INTJ
N= 122	N= 62	N= 26	N= 40
%= 9.49	%= 4.82	%= 2.02	%= 3.11
I= 0.96	I= 1.07	I= 1.22	I= 1.08
ISTP	ISFP	INFP	INTP #
N= 90	N= 51	N= 57	N= 86
%= 7.00	%= 3.97	%= 4.43	%= 6.69
I= 0.89	I= 1.03	I= 1.04	I= 1.23
ESTP "	ESFP	ENFP #	ENTP
N= 122	N= 83	N= 156	N= 123
%= 9.49	%= 6.45	%= 12.13	%= 9.56
I= 0.88	I= 0.99	I= 1.14	I= 0.99
ESTJ #	ESFJ	ENFJ	ENTJ
N= 121	N= 71	N= 26	N= 50
%= 9.41	%= 5.52	%= 2.02	%= 3.89
I= 0.85	I= 1.00	I= 0.96	I= 1.14

			N	%	I	
J		E	752	58.48	0.98	
U		I	534	41.52	1.03	
D	I	S	722	56.14	0.94	*
G	N	N	564	43.86	1.10	*
I	T	T	754	58.63	0.96	#
N	R	F	532	41.37	1.06	#
G	O	J	518	40.28	0.98	
	V	P	768	59.72	1.01	
P	E	IJ	250	19.44	1.03	
E	R	IP	284	22.08	1.03	
R	T	EP	484	37.64	1.00	
C	S	EJ	268	20.84	0.94	
E		ST	455	35.38	0.89	*
P		SF	267	20.76	1.02	
T		NF	265	20.61	1.11	#
I	E	NT	299	23.25	1.09	#
V	X	SJ	376	29.24	0.94	"
E	T	SP	346	26.91	0.93	#
S	R	NP	422	32.81	1.10	*
	A	NJ	142	11.04	1.10	
J	V	TJ	333	25.89	0.95	
U	E	TP	421	32.74	0.97	
D	R	FP	347	26.98	1.07	"
G	T	FJ	185	14.39	1.04	
I	S	IN	209	16.25	1.14	*
N		EN	355	27.60	1.07	"
G		IS	325	25.27	0.97	
		ES	397	30.87	0.91	*

Note concerning symbols following the selection ratios:
" implies significance at the .05 level, i.e., Chi-square >3.8;
implies significance at the .01 level, i.e., Chi-square > 6.6;
* implies significance at the .001 level, i.e., Chi-square > 10.8.
_ (underscore) indicates Fisher's exact probability used instead Chi-square.

Base population used in calculating selection ratios:
WPS Population (Class of 1997-1999)
Base total N = 2046. Sample and base are dependent.

* * * * Calculated values of Chi-square or Fisher's exact probability * * * *
Type table order

```
                                          E  2.0770 IJ  0.5111 SJ  4.9542 IN 11.6834
0.7331    0.8493    2.7455    0.6356      I  2.0770 IP  1.0841 SP  7.2647 EN  5.8489
                                          S 21.1285 EP  0.0038 NP 13.4386 IS  1.2292
3.2427    0.1020    0.2760   10.7494      N 21.1285 EJ  3.1534 NJ  3.6240 ES 13.9114
                                          T  8.1438 ST 25.6376 TJ  3.3167
6.2115    0.0122    8.4873    0.0505      F  8.1438 SF  0.3094 TP  1.5095
                                          J  0.8610 NF  9.0002 FP  5.7054
9.4411    0.0000    0.1074    2.2822      P  0.8610 NT  6.9875 FJ  1.0583
```

Source of data

(null)

Group tabulated:

Portion taking 2 PSATS

N = 258

MBTI Type Table
Center for Applications of Psychological Type

Legend: % = percent of total choosing this group who fall into this type.
I = Self-selection index: Ratio of percent of type in group to % in sample.

SENSING types with THINKING	SENSING types with FEELING	INTUITIVE types with FEELING	INTUITIVE types with THINKING
ISTJ N= 27 %= 10.47 I= 1.10	ISFJ N= 14 %= 5.43 I= 1.13	INFJ N= 7 %= 2.71 I= 1.34	INTJ N= 7 %= 2.71 I= 0.87
ISTP N= 19 %= 7.36 I= 1.05	ISFP N= 11 %= 4.26 I= 1.08	INFP N= 12 %= 4.65 I= 1.05	INTP N= 13 %= 5.04 I= 0.75
ESTP N= 23 %= 8.91 I= 0.94	ESFP N= 19 %= 7.36 I= 1.14	ENFP N= 25 %= 9.69 I= 0.80	ENTP " N= 33 %= 12.79 I= 1.34
ESTJ " N= 15 %= 5.81 I= 0.62	ESFJ N= 17 %= 6.59 I= 1.19	ENFJ N= 7 %= 2.71 I= 1.34	ENTJ N= 9 %= 3.49 I= 0.90

		N	%	I
J	E	148	57.36	0.98
U	I	110	42.64	1.03
D I	S	145	56.20	1.00
G N	N	113	43.80	1.00
I T	T	146	56.59	0.97
N R	F	112	43.41	1.05
G O	J	103	39.92	0.99
V	P	155	60.08	1.01
P E	IJ	55	21.32	1.10
E R	IP	55	21.32	0.97
R T	EP	100	38.76	1.03
C S	EJ	48	18.60	0.89
E	ST	84	32.56	0.92
P	SF	61	23.64	1.14
T	NF	51	19.77	0.96
I E	NT	62	24.03	1.03
V X	SJ	73	28.29	0.97
E T	SP	72	27.91	1.04
S R	NP	83	32.17	0.98
A	NJ	30	11.63	1.05
J V	TJ	58	22.48	0.87
U E	TP	88	34.11	1.04
D R	FP	67	25.97	0.96
G T	FJ	45	17.44	1.21
I S	IN	39	15.12	0.93
N	EN	74	28.68	1.04
G	IS	71	27.52	1.09
	ES	74	28.68	0.93

Note concerning symbols following the selection ratios:
" implies significance at the .05 level, i.e., Chi-square >3.8;
implies significance at the .01 level, i.e., Chi-square > 6.6;
* implies significance at the .001 level, i.e., Chi-square > 10.8.
_ (underscore) indicates Fisher's exact probability used instead Chi-square.

Base population used in calculating selection ratios:
WPS Population who took 1 PSAT (Class of 1997-1999)
Base total N = 1286. Sample and base are dependent.

* * * * Calculated values of Chi-square or Fisher's exact probability * * * *
Type table order

				E	0.1642	IJ	0.7266	SJ	0.1388	IN	0.3058
0.3598	0.2576	0.7789	0.1690	I	0.1642	IP	0.1101	SP	0.1647	EN	0.1874
				S	0.0004	EP	0.1736	NP	0.0608	IS	0.8630
0.0664	0.0751	0.0365	1.4058	N	0.0004	EJ	0.9774	NJ	0.1128	ES	0.7245
				T	0.5550	ST	1.1249	TJ	1.9599		
0.1230	0.4429	1.8038	3.8836	F	0.5550	SF	1.6288	TP	0.2756		
				J	0.0171	NF	0.1389	FP	0.1684		
4.8939	0.7060	0.7789	0.1380	P	0.0171	NT	0.1102	FJ	2.4476		

Source of data	Group tabulated:	MBTI Type Table
(null)	**Portion taking 1 SAT**	Center for Applications of Psychological Type
	N = 1069	Legend: % = percent of total choosing this group who fall into this type. I = Self-selection index: Ratio of percent of type in group to % in sample.

SENSING types with THINKING	SENSING types with FEELING	INTUITIVE types with FEELING	INTUITIVE types with THINKING
ISTJ	ISFJ	INFJ	INTJ
N= 102	N= 48	N= 22	N= 30
%= 9.54	%= 4.49	%= 2.06	%= 2.81
I= 0.96	I= 1.00	I= 1.24	I= 0.97
ISTP "	ISFP	INFP	INTP
N= 69	N= 42	N= 43	N= 64
%= 6.45	%= 3.93	%= 4.02	%= 5.99
I= 0.83	I= 1.02	I= 0.95	I= 1.10
ESTP *	ESFP	ENFP "	ENTP #
N= 92	N= 75	N= 131	N= 122
%= 8.61	%= 7.02	%= 12.25	%= 11.41
I= 0.80	I= 1.08	I= 1.16	I= 1.18
ESTJ "	ESFJ	ENFJ "	ENTJ
N= 101	N= 56	N= 29	N= 43
%= 9.45	%= 5.24	%= 2.71	%= 4.02
I= 0.86	I= 0.95	I= 1.29	I= 1.18

		N	%	I	
J	E	649	60.71	1.02	
U	I	420	39.29	0.97	
D I	S	585	54.72	0.91	*
G N	N	484	45.28	1.13	*
I T	T	623	58.28	0.96	#
N R	F	446	41.72	1.07	#
G O	J	431	40.32	0.98	
V	P	638	59.68	1.01	
P E	IJ	202	18.90	1.00	
E R	IP	218	20.39	0.95	
R T	EP	420	39.29	1.05	
C S	EJ	229	21.42	0.97	
E	ST	364	34.05	0.86	*
P	SF	221	20.67	1.01	
T	NF	225	21.05	1.13	#
I E	NT	259	24.23	1.13	#
V X	SJ	307	28.72	0.93	"
E T	SP	278	26.01	0.90	#
S R	NP	360	33.68	1.12	*
A	NJ	124	11.60	1.15	"
J V	TJ	276	25.82	0.95	
U E	TP	347	32.46	0.96	
D R	FP	291	27.22	1.08	"
G T	FJ	155	14.50	1.05	
I S	IN	159	14.87	1.05	
N	EN	325	30.40	1.18	*
G	IS	261	24.42	0.94	
	ES	324	30.31	0.89	*

Note concerning symbols following the selection ratios:
" implies significance at the .05 level, i.e., Chi-square >3.8;
implies significance at the .01 level, i.e., Chi-square > 6.6;
* implies significance at the .001 level, i.e., Chi-square > 10.8.
_ (underscore) indicates Fisher's exact probability used instead Chi-square.

Base population used in calculating selection ratios:
WPS Population (Class of 1997-1999)
Base total N = 2046. Sample and base are dependent.

* * * * Calculated values of Chi-square or Fisher's exact probability * * * *
Type table order

				E	0.9937	IJ	0.0067	SJ	5.3888	IN	0.7773		
0.3620	0.0002	2.1506	0.0478	I	0.9937	IP	1.2433	SP	9.6441	EN	24.6954		
				S	25.6704	EP	2.7694	NP	14.7269	IS	3.2929		
5.7906	0.0277	0.2902	1.3765	N	25.6704	EJ	0.5839	NJ	5.7967	ES	12.6830		
				T	6.9531	ST	28.7176	TJ	2.3868				
11.1987	0.9785	6.4157	7.7106	F	6.9531	SF	0.1179	TP	1.6005				
				J	0.5034	NF	8.6945	FP	4.7565				
5.8170	0.3471	4.0640	2.4483	P	0.5034	NT	10.5860	FJ	0.9672				

Source of data

(null)

Group tabulated:

Portion Taking 2 SATs

N = 586

MBTI Type Table
Center for Applications of Psychological Type

Legend: % = percent of total choosing this group who fall into this type.
I = Self-selection index: Ratio of percent of type in group to % in sample.

SENSING types with THINKING	SENSING types with FEELING	INTUITIVE types with FEELING	INTUITIVE types with THINKING
ISTJ N= 52 %= 8.87 I= 0.93	ISFJ # N= 35 %= 5.97 I= 1.33	INFJ N= 11 %= 1.88 I= 0.91	INTJ N= 17 %= 2.90 I= 1.03
ISTP N= 32 %= 5.46 I= 0.85	ISFP N= 25 %= 4.27 I= 1.09	INFP N= 24 %= 4.10 I= 1.02	INTP N= 30 %= 5.12 I= 0.86
ESTP N= 49 %= 8.36 I= 0.97	ESFP N= 36 %= 6.14 I= 0.88	ENFP N= 79 %= 13.48 I= 1.10	ENTP " N= 55 %= 9.39 I= 0.82
ESTJ N= 60 %= 10.24 I= 1.08	ESFJ N= 32 %= 5.46 I= 1.04	ENFJ N= 21 %= 3.58 I= 1.32	ENTJ N= 28 %= 4.78 I= 1.19

			N	%	I	
J		E	360	61.43	1.01	
U		I	226	38.57	0.98	
D	I	S	321	54.78	1.00	
G	N	N	265	45.22	1.00	
I	T	T	323	55.12	0.95	"
N	R	F	263	44.88	1.08	"
G	O	J	256	43.69	1.08	"
	V	P	330	56.31	0.94	"
P	E	IJ	115	19.62	1.04	
E	R	IP	111	18.94	0.93	
R	T	EP	219	37.37	0.95	
C	S	EJ	141	24.06	1.12	"
E		ST	193	32.94	0.97	
P		SF	128	21.84	1.06	
T		NF	135	23.04	1.09	
I	E	NT	130	22.18	0.92	
V	X	SJ	179	30.55	1.06	
E	T	SP	142	24.23	0.93	
S	R	NP	188	32.08	0.95	
	A	NJ	77	13.14	1.13	
J	V	TJ	157	26.79	1.04	
U	E	TP	166	28.33	0.87	#
D	R	FP	164	27.99	1.03	
G	T	FJ	99	16.89	1.17	"
I	S	IN	82	13.99	0.94	
N		EN	183	31.23	1.03	
G		IS	144	24.57	1.01	
		ES	177	30.20	1.00	

Note concerning symbols following the selection ratios:
" implies significance at the .05 level, i.e., Chi-square >3.8;
implies significance at the .01 level, i.e., Chi-square > 6.6;
* implies significance at the .001 level, i.e., Chi-square > 10.8.
_ (underscore) indicates Fisher's exact probability used instead Chi-square.

Base population used in calculating selection ratios:
WPS Population who took 1 SAT (Class of 1997-1999)
Base total N = 1069. Sample and base are dependent.

* * * * Calculated values of Chi-square or Fisher's exact probability * * * *
Type table order

				E	0.2838	IJ	0.4490	SJ	2.1163	IN	0.7942
0.6703	6.6469	0.2105	0.0426	I	0.2838	IP	1.6818	SP	2.1200	EN	0.4186
				S	0.0015	EP	1.9983	NP	1.4762	IS	0.0176
2.1218	0.3909	0.0180	1.7339	N	0.0015	EJ	5.3682	NJ	3.0008	ES	0.0066
				T	5.3241	ST	0.7185	TJ	0.6415		
0.0985	1.5136	1.8153	5.2702	F	5.3241	SF	1.0816	TP	10.1033		
				J	6.1139	NF	3.0902	FP	0.3828		
0.9481	0.1290	3.7264	1.9186	P	6.1139	NT	2.9515	FJ	5.9992		

Source of data: (null)

Group tabulated: **Portion who gained >= 50 Points** (Combined)

N = 231

MBTI Type Table
Center for Applications of Psychological Type

Legend: % = percent of total choosing this group who fall into this type.
I = Self-selection index: Ratio of percent of type in group to % in sample.

SENSING types with THINKING	SENSING types with FEELING	INTUITIVE types with FEELING	INTUITIVE types with THINKING
ISTJ N= 17 %= 7.36 I= 0.83	ISFJ N= 10 %= 4.33 I= 0.72	INFJ N= 4 %= 1.73 I= 0.92	INTJ N= 7 %= 3.03 I= 1.04
ISTP N= 10 %= 4.33 I= 0.79	ISFP N= 11 %= 4.76 I= 1.12	INFP N= 9 %= 3.90 I= 0.95	INTP N= 9 %= 3.90 I= 0.76
ESTP N= 20 %= 8.66 I= 1.04	ESFP N= 14 %= 6.06 I= 0.99	ENFP N= 37 %= 16.02 I= 1.19	ENTP N= 24 %= 10.39 I= 1.11
ESTJ N= 27 %= 11.69 I= 1.14	ESFJ N= 14 %= 6.06 I= 1.11	ENFJ N= 7 %= 3.03 I= 0.85	ENTJ N= 11 %= 4.76 I= 1.00

		N	%	I
JUDGING	E	154	66.67	1.09 "
	I	77	33.33	0.86 "
INTUITION	S	123	53.25	0.97
	N	108	46.75	1.03
	T	125	54.11	0.98
	F	106	45.89	1.02
	J	97	41.99	0.96
	P	134	58.01	1.03
PERCEPTIVE / EXTRAVERTS	IJ	38	16.45	0.84
	IP	39	16.88	0.89
	EP	95	41.13	1.10
	EJ	59	25.54	1.06
	ST	74	32.03	0.97
	SF	49	21.21	0.97
	NF	57	24.68	1.07
	NT	51	22.08	1.00
INTROVERTS / EXTRAVERTS	SJ	68	29.44	0.96
	SP	55	23.81	0.98
	NP	79	34.20	1.07
	NJ	29	12.55	0.96
JUDGING / EXTRAVERTS	TJ	62	26.84	1.00
	TP	63	27.27	0.96
	FP	71	30.74	1.10
	FJ	35	15.15	0.90
	IN	29	12.55	0.90
	EN	79	34.20	1.10
	IS	48	20.78	0.85
	ES	75	32.47	1.07

Note concerning symbols following the selection ratios:
" implies significance at the .05 level, i.e., Chi-square >3.8;
implies significance at the .01 level, i.e., Chi-square > 6.6;
* implies significance at the .001 level, i.e., Chi-square > 10.8.
_ (underscore) indicates Fisher's exact probability used instead Chi-square.

Base population used in calculating selection ratios:
WPS Population who took 2 SATs (Class of 1997-1999)
Base total N = 586. Sample and base are dependent.

* * * * Calculated values of Chi-square or Fisher's exact probability * * * *
Type table order

				E	4.4076	IJ	2.4360	SJ	0.2210	IN	0.6561
1.0815	1.8344	1.0000	0.0226	I	4.4076	IP	1.0527	SP	0.0371	EN	1.5666
				S	0.3610	EP	2.2953	NP	0.7845	IS	2.9616
0.9461	0.2294	0.0386	1.1748	N	0.3610	EJ	0.4569	NJ	0.1147	ES	0.9261
				T	0.1563	ST	0.1400	TJ	0.0004		
0.0437	0.0045	2.1027	0.4519	F	0.1563	SF	0.0889	TP	0.2090		
				J	0.4451	NF	0.5769	FP	1.4304		
0.8716	0.2658	0.3379	0.0002	P	0.4451	NT	0.0025	FJ	0.8248		

Apendix B: Source Code for Programs Used

newconverter.cc

This program converts files in the College Board's fixed-width format to CSV. It uses a control file that specifies for each field is wanted the name of the field and it's location on each line. Three control files were made, one for PSAT score reports up to 1996, one for PSAT score reports after 1996 and one for SAT score reports.

```
#include <stdio.h>
#include <vector>
#include <string.h>

struct Field {
      char name[32];
      int start_pos;
      int length;
};


int main(int argc, char **argv)
{
      if(argc < 3) {
            printf("usage:\n");
            printf("newconverter <input file> <control file>\n");
            return -1;
      }

      FILE *input_file = fopen(argv[1], "rt");
      FILE *control_file = fopen(argv[2], "rt");

      if(input_file == NULL || control_file == NULL) {
            printf("couldn't open the files\n");
            return -1;
      }

      vector<Field> fields;
      char buf[4096], buf2[4096];

      while(true) {
            fgets(buf, 4096, control_file);
            if(feof(control_file)) break;

            int first_pos, length, last_pos;
            char name[32];

            if(buf[0] == '#') {
                  sscanf(buf, "# %d %d", &first_pos, &last_pos);

                  length = last_pos - first_pos + 1;
            } else {
                  sscanf(buf, "%s", name);
                  Field f;
```

```
                strcpy(f.name, name);
                f.start_pos = first_pos - 1;
                f.length = length;

                first_pos += length;

                fields.push_back(f);
            }
        }
        fclose(control_file);
        vector<Field>::iterator i;
        i = fields.begin();
        printf("%s", i->name);

        for(i++; i != fields.end(); i++) {
            printf(",%s", i->name);
        }
        printf("\n");


        while(true) {
            fgets(buf, 4096, input_file);
            if(feof(input_file)) break;

            i = fields.begin();
            strncpy(buf2, buf + i->start_pos, i->length);
            buf2[i->length] = '\0';
            printf("%s", buf2, i->start_pos, i->length);

            for(i++; i != fields.end(); i++) {
                strncpy(buf2, buf + i->start_pos, i->length);
                buf2[i->length] = '\0';
                printf(",%s", buf2, i->start_pos, i->length);
            }
            printf("\n");
        }
}
```

<u>group.cc</u>

This program takes many CSV files and prints them all to standard output. Duplicate students are only printed the first time that they are seen. This ensures that if the input files are specified in reverse chronological order then the outputted file only contains the information from the last time each student took the test.

```
#include <stdio.h>
#include <string.h>
#include <vector>
#include <algorithm>

class student {
private:
	char f_name[13];
	char l_name[16];
	long dob;
	long ssn;

public:
	student() {
		dob = ssn = 0;
		strcpy(f_name, "");
		strcpy(l_name, "");
	}

	student(char *row) {

		row = strchr(row, ',') + 1; // skip over hs_code

		strncpy(l_name, row, 15);  // get the last name
		l_name[15] = '\0';
		row = strchr(row, ',') + 1; // skip over last name

		strncpy(f_name, row, 12);  // get the first name
		f_name[12] = '\0';
		row = strchr(row, ',') + 1; // skip over first name

		row = strchr(row, ',') + 1; // skip over initial
		row = strchr(row, ',') + 1; // skip over sex

		dob = ssn = 0;
		sscanf(row, "%ld,", &dob); // get the date of birth
		row = strchr(row, ',') + 1; // skip over dob

		sscanf(row, "%ld,", &ssn); // get the social security number
	}

	student(const student &other) {
		dob = other.dob;
		ssn = other.ssn;
		strcpy(l_name, other.l_name);
		strcpy(f_name, other.f_name);
	}
```

```
	bool operator==(const student &other) {

		// 1st try to use ssn to match students, but if this was
		// left blank then use first name, last name and dob

		if(ssn != 0 && other.ssn != 0) return ssn == other.ssn;

		return (strcmp(f_name, other.f_name) == 0) &&
		      (strcmp(l_name, other.l_name) == 0) &&
		      (dob == other.dob);
	}

	bool operator!=(const student &other) { return !(*this == other); }

	student const &operator=(const student &other) {
		if(this != &other) {
			dob = other.dob;
			ssn = other.ssn;
			strcpy(l_name, other.l_name);
			strcpy(f_name, other.f_name);
		}
		return *this;
	}
};

int main(int argc, char **argv)
{
	char buf[4096];

	FILE *f = fopen(argv[1], "rt");

	fgets(buf, 4096, f);
	buf[strlen(buf) - 1] = '\0';
	printf("%s\n", buf);

	fclose(f);

	vector<student> included;

	for(int i = 1; i < argc; i++) {
		fprintf(stderr, "reading %s...\n", argv[i]);
		FILE *f = fopen(argv[i], "rt");
		fgets(buf, 4096, f);

		while(true) {
			fgets(buf, 4096, f);
			if(feof(f)) break;

			buf[strlen(buf) - 1] = '\0';

			student current = student(buf);

			// If the student hasn't been included yet then print
			// his/her info
			if(find(included.begin(), included.end(), current) ==
			   included.end()) {
```

```
                    printf("%s\n", buf);
                    included.push_back(current);
                }
            }
            fclose(f);
        }

        return 0;
}
```

group_by_test.cc:

This program is similar to group.cc, but it outputs all of the score reports, including duplicates. I used SPSS to convert this file into one that contained the score analyses for each student. The output of this program was also given to group_for_wps.cc in order to create a file in the format that the WPS wanted it in.

```
#include <stdio.h>
#include <string.h>
#include <vector>
#include <algorithm>

int main(int argc, char **argv)
{
	char buf[4096];

	FILE *f = fopen(argv[1], "rt");

	fgets(buf, 4096, f);
	buf[strlen(buf) - 1] = '\0';
	printf("%s\n", buf);

	fclose(f);

	for(int i = 1; i < argc; i++) {
		fprintf(stderr, "reading %s...\n", argv[i]);
		FILE *f = fopen(argv[i], "rt");
		fgets(buf, 4096, f);

		while(true) {
			fgets(buf, 4096, f);
			if(feof(f)) break;

			buf[strlen(buf) - 1] = '\0';

			printf("%s\n", buf);
		}
		fclose(f);
	}

	return 0;
}
```

group_for_wps.cc:

This program takes the output of group_by_test.cc and creates a CSV file where the columns represent a SAT 1 test date and the rows represent a student.

```
#include<stdio.h>
#include<assert.h>
#include<string.h>
#include<vector.h>

struct test_score {
	char date[5];
	char verbal[4];
	char math[4];
};

struct row {
	char buf[4096];
	test_score scores[6];
	int score_count;
};

void read_row(row *row, const char *text)
{
	// read the header info
	const char *pos;

	pos = text;
	for(int i = 0; i < 9; i++) pos = strchr(pos, ',') + 1;
	strncpy(row->buf, text, pos - text - 1);

	int i = 0;
	while(strncmp(pos, "    ", 4) != 0) {
		strncpy(row->scores[i].date, pos, 4);
		row->scores[i].date[4] = '\0';
		pos += 7;
		strncpy(row->scores[i].verbal, pos, 3);
		row->scores[i].verbal[3] = '\0';
		pos += 4;
		strncpy(row->scores[i].math, pos, 3);
		row->scores[i].math[3] = '\0';
		pos += 4;

		i++;
		if(i == 6) break;
	}
	row->score_count = i;

}

void add_test(row *r, char *date)
{
	for(int i = 0; i < r->score_count; i++) {
		if(strcmp(r->scores[i].date, date) == 0) {
```

```
                        char to_add[1024];
                        sprintf(to_add, ",%s,%s",
                                r->scores[i].verbal, r->scores[i].math);
                        strcat(r->buf, to_add);
                        return;
                }
        }

        strcat(r->buf, ",    ,    ");
}

int main(int argc, char **argv)
{
        char buf[4096];
        char header[4096] =
                "hs_code,l_name,f_name,ititial,sex,dob,ssn,grad_m,grad_y";

        char dates[][5] = { "1000", "0600", "0500", "0400", "0200",
                "1299", "1298", "1198", "1098", "0698", "0598", "0498",
                "0398", "0298", "1297", "1197", "1097", "0697", "0597",
                "0397", "0297", "1296", "1196", "1096", "DONE" };

        vector<row> cases;

        FILE *f = fopen(argv[1], "rt");
        assert(f != NULL);

        fgets(buf, 4096, f); // skip over the header line

        while(true) {
                fgets(buf, 4096, f);
                if(feof(f)) break;
                row r;
                read_row(&r, buf);
                cases.push_back(r);
        }

        vector<row>::iterator i;
        for(int cur = 0; strcmp(dates[cur], "DONE") != 0; cur++) {
                char add[1024];
                sprintf(add, ",%s-verbal,%s-math", dates[cur], dates[cur]);
                strcat(header, add);
                for(i = cases.begin(); i != cases.end(); i++) {
                        add_test(i, dates[cur]);
                }
        }

        printf("%s\n", header);
        for(i = cases.begin(); i != cases.end(); i++) {
                printf("%s\n", i->buf);
        }


}
```

group_for_wps_sat2.cc:

This program takes the output of group_by_test.cc and creates a CSV file where the columns represent a SAT 2 test date and the rows represent a student.

```
#include<stdio.h>
#include<assert.h>
#include<string.h>
#include<vector.h>

struct test_score {
	char date[5];
	char code[3][3];
	char score[3][4];
};

struct row {
	char buf[4096];
	test_score scores[6];
	int score_count;
};

void read_row(row *row, const char *text)
{
	// read the header info
	const char *pos;

	pos = text;
	for(int i = 0; i < 9; i++) pos = strchr(pos, ',') + 1;
	strncpy(row->buf, text, pos - text - 1);

	pos = text;
	for(int i = 0; i < 53; i++) pos = strchr(pos, ',') + 1;

	int i = 0;
	while(strncmp(pos, "    ", 4) != 0) {
		strncpy(row->scores[i].date, pos, 4);
		row->scores[i].date[4] = '\0';
		pos += 5;
		for(int j = 0; j < 3; j++) {
			strncpy(row->scores[i].code[j], pos, 2);
			row->scores[i].code[j][2] = '\0';
			pos += 3;
			strncpy(row->scores[i].score[j], pos, 3);
			row->scores[i].score[j][3] = '\0';
			pos += 4;
		}

		i++;
		if(i == 6) break;
	}
	row->score_count = i;

}
```

```
void add_test(row *r, char *date)
{
	for(int i = 0; i < r->score_count; i++) {
		if(strcmp(r->scores[i].date, date) == 0) {
			char to_add[1024];
			sprintf(to_add, ",%s,%s,%s,%s,%s,%s",
				r->scores[i].code[0], r->scores[i].score[0],
				r->scores[i].code[1], r->scores[i].score[1],
				r->scores[i].code[2], r->scores[i].score[2]);
			strcat(r->buf, to_add);
			return;
		}
	}

	strcat(r->buf, ",  ,   ,  ,   ,  ,    ");
}

int main(int argc, char **argv)
{
	char buf[4096];
	char header[4096] =
		"hs_code,l_name,f_name,ititial,sex,dob,ssn,grad_m,grad_y";

	char dates[][5] = { "1000", "0600", "0500", "0400", "0200",
		"1299", "1298", "1198", "1098", "0698", "0598", "0498",
		"0398", "0298", "1297", "1197", "1097", "0697", "0597",
		"0397", "0297", "1296", "1196", "1096", "DONE" };

	vector<row> cases;

	FILE *f = fopen(argv[1], "rt");
	assert(f != NULL);

	fgets(buf, 4096, f); // skip over the header line

	while(true) {
		fgets(buf, 4096, f);
		if(feof(f)) break;
		row r;
		read_row(&r, buf);
		cases.push_back(r);
	}

	vector<row>::iterator i;
	for(int cur = 0; strcmp(dates[cur], "DONE") != 0; cur++) {
		char add[1024];
		sprintf(add, ",%s-c1,%s-s1,%s-c2,%s-s2,%s-c3,%s-s3",
			dates[cur], dates[cur], dates[cur], dates[cur],
			dates[cur], dates[cur]);
		strcat(header, add);
		for(i = cases.begin(); i != cases.end(); i++) {
			add_test(i, dates[cur]);
		}
	}

	printf("%s\n", header);
```

```
	for(i = cases.begin(); i != cases.end(); i++) {
		printf("%s\n", i->buf);
	}


}
```