

Juliette Bellinson

An Elite Based on Merit: The Origins of the SAT

In May of 2019, College Board announced the introduction of a new and controversial portion of the SAT called the “Adversity Score.”¹ The Adversity Score was a measurement of the hardships a student had faced in their high school career, which would be presented to colleges alongside a student’s verbal and quantitative score. The score was based on factors such as the tester’s socioeconomic background, the amount of crime in their neighborhood, and the relative quality of their high school. By providing this context, College Board hoped that the SAT could provide a profile of a student’s merit that superseded the limitations of their background. However, the Adversity Score was met with criticism from students and parents alike, who claimed that “a student’s qualitative experiences cannot be successfully expressed in a single number.”² To them, the framework of the Adversity Score did not allow for the nuance that was needed to properly contextualize a student’s performance on the SAT. After these criticisms were heard, the Adversity Score was quickly abandoned the next month. In a comment to the New York Times, College Board executive David Coleman apologized for his audacious proposal, stating “we have acknowledged that we have perhaps overstepped, and we’ve adopted a humbler position. That’s admitting that the College Board should keep its focus on scoring achievement.”³

While Coleman’s words were intended to address a single controversy, they speak to a century-long pattern exhibited throughout the test’s history. The pattern in question is the

¹ Hartocollis, Anemona. “SAT’s New ‘Adversity Score’ Will Take Students’ Hardships Into Account.” *The New York Times*, May 17, 2019.

² Hartocollis, Anemona. “SAT ‘Adversity Score’ Is Abandoned in Wake of Criticism.” *The New York Times*, August 28, 2019.

³ Ibid.

dissonance between the lofty goals and the limited framework for execution that has defined the SAT since its very inception.

The SAT was designed as a means to reform the university admissions process. During the late 1880's, administrators of several elite, private universities concluded that it was in the best interest for their institutions, and American society at large, for them to admit students on the basis of merit, as opposed to the circumstances of their birth. In order to achieve this goal, these universities sponsored the creation of a standardized college entrance examination that would select students with outstanding abilities for admission to university. The original SAT, that lasted from 1901 to 1915 measured a student's merit through their attainment of secondary school concepts. However, from 1915 onwards, designers of the exam began to change how they defined merit. Test designers came to believe that a student's merit was best represented not by their prior achievement, but by their demonstrated potential. Thanks to the advocacy of several psychologists such as Carl Brigham, they came to the conclusion that the test should seek to measure a student's natural intelligence.

However, they soon realized that this approach to the SAT was built upon faulty preconceptions. By the 1930's, new research came to light which found that intelligence tests like the SAT were predominantly a reflection of a student's socioeconomic background. With this realization, it became clear to test designers that the "merit" that the SAT sought to measure was inextricable from privilege. Ultimately, in its quest to admit students to elite universities on merit rather than birthright, the SAT reinforced the same stratifications it sought to to topple.

Part 1 - Beginnings

To discuss the aims and impact of the SAT, it is important to note the conditions that led to its inception. Before the turn of the 20th century, the world of private higher education was dominated by the wealthy. The student body of elite universities consisted of well-to-do young men who had matriculated from private academies or studied with a tutor.⁴ Upon completion of their secondary school studies they would sit for a school's entrance examination. Each university administered their own examinations, covering subjects from Greek and Latin to English history and Chemistry.⁵ To attend an elite university, one had to possess a wide enough range of knowledge to meet the standards of one or more individual colleges. Unsurprisingly, these examinations constituted an obstacle to many prospective students. However, these concerns would not be addressed by university administrators until the late 19th century.

In the latter half of the 1800's, it became clear to university administrators that demand for post-secondary education was exploding. In 1880 16 U.S. states had instituted compulsory schooling laws for children ages 6 to 14.⁶ By 1900, this number had grown to 34.⁷ A significant percentage of these children went on to matriculate from a public high school, causing the number of high school graduates in the U.S. to swell. From 1869 to 1900 the number of American 17 year olds with high school diplomas grew from 2% to 9%.⁸ Naturally, quite a few members of this new crop of public school graduates possessed both the qualifications and the desire to attend an elite university. However, the desultory nature of the college entrance exams

⁴ Fuess, Claude. *The College Board. Its First Fifty Years*. Amsterdam University Press, 1950, 2

⁵ Ibid., 3-4

⁶ Lingwall, Jeff. "Compulsory Schooling, the Family, and the 'Foreign Element,' Evidence from the United States, 1880-1900." *Carnegie Mellon University Press*, December 31, 2009, 6.

⁷ Ibid.

⁸ Snyder, Thomas D. "120 Years of American Education: A Statistical Portrait." *U.S. Dept. Of Education, Office of Educational Research And Improvement*, December 31, 1992, 31.

was an insurmountable hurdle to many. For one thing, these entrance examinations were exclusively administered on-site.⁹ This automatically excluded many poor and rural students for whom long-distance travel was impracticable. In a similar vein, each school's entrance exam put great emphasis on subjects which were not taught at public high schools, such as Greek and Latin.¹⁰ The overpresence of these subjects ensured that most public school graduates, including those who had demonstrated extraordinary merit in their secondary school careers, were weeded out by college entrance exams. This practice continued until the 1890's when many elite universities began to consider adapting their admissions practices to be more accessible to this growing group of students.

The first signs of change began with Harvard's eminent president, Charles William Eliot. Before his tenure at Harvard, Eliot had spent several years studying the operations of Europe's universities, and the role of higher education in their societies. He was particularly impressed with French and German polytechnic schools, of which he wrote,

I have given special attention to the schools here provided for the education of young men for those arts and trades which require some knowledge of scientific principles and their applications, the schools which turn out master workmen, superintendents, and designers for the numerous French industries which demand taste, skill, and special technical instruction. Such schools we need at home.¹¹

Eliot was astonished at the quality of the European Polytechnic school graduates. He noticed that, in Europe, the primary purpose of their universities was to produce the nation's future innovators and leaders. In order to do this, these Polytechnic schools made sure to impart

⁹ Fuess, Claude., *The College Board. Its First Fifty Years*, 1950 7

¹⁰ *Ibid.*, 8

¹¹ James, Henry. *Charles W. Eliot, President of Harvard University 1869-1909*. Volume 1. AMS Press, 1930, 130

tangible skills onto their students that directly translated to their flourishing in the workplace. However, these schools' curriculum was not the only factor which contributed to their success.

While in Europe, Eliot took note of the polytechnic schools' admissions practices. The schools which Eliot visited, such as the French Ecole Polytechnique, held rigorous admissions examinations which only a small portion of talented students were able to pass.¹² The school would select a small group of high performers for admission, without taking into account their family's wealth. If a student without the means to fund their education did well on the exam, the Ecole Polytechnique would admit them on a scholarship.¹³ Overall, the mission of these polytechnic schools was to admit the most qualified and capable students from all socioeconomic backgrounds. In doing so, they ensured that their universities would educate those with the potential to lead and revolutionize their country's industries. As such, one of the principal lessons which Eliot learned from his time in Europe was that the success of a university was contingent not just on what it taught, but also on whom it admitted.

In response to these observations, Eliot became inspired to reform the American university to mirror their European counterparts' aims and practices. He believed that the key to Europe's success in industry and government lay in the organization of their universities.¹⁴ If American universities could emulate the European model, he thought, then U.S. society would progress as a result. Thus, when he was appointed as president of Harvard University, he set out to reform the character, curriculum, and mission of the institution, in order to pave the way for

¹² Grattan-Guinness, Ivor. "The Ecole Polytechnique, 1794-1850: Differences over Educational Purpose and Teaching Practice." *The American Mathematical Monthly* 112, no. 3, 2005, 237-238

¹³ Ibid.

¹⁴ Eliot, Charles. "The New Education." *The Atlantic*, February 19, 1869

change in the American university system at large. As his first step toward this goal, Eliot turned his attention to completely restructuring Harvard's admissions practices.

Eliot was dismayed at the low academic caliber of the students whom Harvard admitted. In a letter to a Harvard trustee, Eliot wrote, "I want to have the College open equally to men with much money, little money, or no money, provided they all have brains... I am inclined to think that you would be no more tolerant than I of the presence of stupid sons of the rich."¹⁵ Eliot bluntly stated that if Harvard were to become an eminent institution for innovation, it would have to start by rethinking the way in which it selected its student body. By admitting students on the basis of birthright, Harvard was sacrificing an immense opportunity to consolidate the country's most talented, capable students into its ranks. To ameliorate this, Eliot offered a revolutionary suggestion. He proposed that Harvard should instead seek to admit students of all socioeconomic backgrounds by virtue of their merit.

In an 1890 essay entitled *The Gap Between Common Schools and Colleges*, Eliot suggested a way for meritocratic admissions to become a reality. He proposed that Harvard's entrance exam, along with all other college-specific exams, should be abolished and replaced with a standardized entrance examination.¹⁶ Eliot promised that his new exam would be utterly different from its predecessors, in that it would put far less emphasis on antiquated subjects such as Greek and Latin. To compensate for this, Eliot suggested that his exam would test more applicable skills such as English grammar, the natural sciences, and mathematics. These changes would accomplish two main things. First, they would ensure that the new entrance exam would

¹⁵ Eliot, Charles William. Letter to Charles Francis Adams Jr. June 9, 1904, in *Hoover Institution Digital Collections*.

¹⁶ Eliot, Charles William. *Educational Reform Essays and Addresses*. The Century Co, 1901, 194-219

favor students who demonstrated merit in areas that were more conducive to future success. By testing students on their grasp of these tangible skills, elite universities would be poised to admit students with the greatest potential to lead and innovate. Alongside this, these changes would also alleviate some of the barriers for public school students applying to college. By minimizing the influence of classical studies, this new test would open doors to public school students whose education revolved around other areas. If the standardized entrance exam were implemented along the lines of Eliot's suggestions, it would ensure that America's elite colleges were comprised of the most talented students from all walks of life. Or as Eliot himself put it in his inaugural address, "the poorest and the richest students will be equally welcome here, provided that with their poverty or their wealth they bring capacity, ambition, and purity."¹⁷

Eliot's advocacy inspired many other prominent university presidents to take action. In 1893, faculty and administrators from 10 universities formed the "College Entrance Examination Board," which sought to administer a standardized entrance exam among all American colleges.¹⁸ With Charles Eliot as its president, the Board began work on convincing other universities to join in their mission.

In 1899 the Board held a meeting of faculty members and administrators from private colleges throughout the Northeast. At this meeting, Board members articulated Eliot's philosophy behind admissions reform, and their plan for a standardized entrance exam. This proposal was met with contention from several of the faculty members present. Namely, President Warfield of Lafayette College was hesitant to change the parameters on which students

¹⁷ Ibid., 23-24

¹⁸ Fuess, Claude. *The College Board. Its First Fifty Years*. 1950, 9

were admitted his college. He protested that “Lafayette College does not intend to be told by any Board whom to admit and whom not to admit. If we wish to admit the son of a benefactor, or of a Trustee, or of a member of the Faculty, and such action will benefit the institution, we are not going to be prevented from taking it.”¹⁹ Following President Warfield’s statement, Eliot retorted,

The President of Lafayette College has misunderstood the proposal. The College Entrance Examination Board is not to admit students to any college, but to define the subjects of admission that they will be uniform, to conduct examinations in these subjects at uniform times throughout the world, and to issue to those who take the examinations certificates of performance, good, bad, or indifferent. And, President Warfield, it will be perfectly practicable under this plan for Lafayette College to say, if it chooses, that it will admit only such students as cannot pass these examinations. No one proposes to deprive Lafayette College of that privilege.²⁰

As these words left president Eliot’s lips the room was reported to have erupted in laughter. Eliot’s response, while delivered in jest, succeeded in selling the vision of the board to those in the audience. Eliot argued that it was in the best interest of any self-respecting university to reserve admission to the most academically qualified students. With this witty retort, the administrators in attendance left the meeting with a sense that Eliot’s proposals, and the philosophies behind them, were nothing short of common sense.

Following this meeting, several new schools joined the ranks of the College Entrance Examination Board.²¹ As membership in the Board grew, it became clear to both Eliot and his detractors that the character of the American university was changing. Following the introduction of a standardized entrance exam, it would be the new norm for elite universities to

¹⁹ College Entrance Examination Board. *The Work of the College Entrance Examination Board, 1901-1925*. Ginn and Company, 1926, 4-5

²⁰ Ibid.

²¹ Ibid., 6

admit students not on the basis of birthright, but instead on merit. With this guiding philosophy in mind, the Board set out to design and administer their very first entrance exam.

Part 2 - The First Attempt

The first Board examination was administered in 1901 at Columbia University.²² This test was the precursor to what is now known as the SAT. However, at the time it was simply referred to as the “Standardized College Entrance Examination,” which could be shortened to the SCEE. The SCEE was designed by a collection of faculty members from the universities involved with the Board. It took the form of a curriculum test which measured students’ proficiency in English, Foreign Language, Science, History, and Mathematics. The results of this preliminary experiment were met with both excitement and hesitation. Nicholas Butler, then president of Columbia University and member of the Board, wrote that the SCEE “[will] make an immense contribution to American education,” and “break down untold barriers to sound collegiate education by carrying high and well defined standards of teaching and testing into secondary schools, both public and private, in every part of the United States.”²³ Butler did believe, on some level, that the current form of the test was sufficient to achieve Eliot’s lofty goals of meritocratic admissions. However, in the same report Butler expressed worry about the long-term implications of an “achievement test” being standard. He warned that if the SCEE were to keep its current form, private schools would soon employ special coaches to prompt students to cram the information needed for the test.²⁴ This, he believed, would lead to a

²² College Entrance Examination Board. *First Annual Report of the Secretary*. Pennsylvania State University Press, 1901, 2

²³ *Ibid.*, 33

²⁴ *Ibid.*, 30

continued monopoly of private schools on elite colleges, and defeat the purpose of meritocratic admissions. With this in mind, the Board began to rethink the structure of the SCEE.

One potential solution for improving the SCEE lay in a relatively new testing philosophy called the “comprehensive examination.” The comprehensive examination claimed to measure students’ ability to not only memorize information, but to apply that knowledge through independent reasoning and problem solving. In doing so, it aimed to paint a picture of a student’s capabilities in academia and beyond. The comprehensive exam was championed by many education activists throughout the early 20th century, as well as several notable members of the Board. President Lowell of Harvard, the successor of Charles Eliot, was a staunch advocate for the SCEE taking a comprehensive approach. As the newfound president of Harvard, Lowell felt obliged to pursue the vision set forth by his predecessor. He believed that by adopting a comprehensive approach, the SCEE could be better poised to facilitate Eliot’s goal of meritocratic admissions. In an article for the *Atlantic Monthly* he argued that the SCEE should change from a curriculum-based test to a comprehensive one, stating,

A knowledge of the facts is a small part of education. The important thing is to understand their relation to one another; to be able to correlate them, as the current expression goes; not merely to grasp and retain the relations one has been taught. The pupil must learn to apply principles to new and unexpected situations, and the extent to which he can do so will largely determine the degree of his future effectiveness.²⁵

Lowell believed that by measuring students on the basis of attainment, the SCEE provided an incomplete picture of a student’s merit. Instead, he advocated for the comprehensive approach, which tested not only knowledge, but also skills that could directly indicate a student’s potential

²⁵ Lowell, Lawrence. “The Art of Examination.” *The Atlantic*, January 28, 1926

for success. All in all, In Lowell's mind, the comprehensive approach was far better suited to the philosophies on which the SCEE was originally created.

In April of 1915, the College Entrance Examination Board voted to change the conduct of their examinations from an achievement-based approach to a comprehensive one.²⁶ This change marked a noticeable shift in both the content of the SCEE, and the skillsets which it was testing. These differences can be noted when comparing the content of the 1901 test to its 1915 counterpart. In the 1901 exam, in its section on Latin grammatical constructions, students were asked to do the following.

Write the rules for the following constructions and illustrate each by a Latin sentence.²⁷

- a) Two uses of the dative.
- b) The cases used to indicate the relations of place.
- c) The cases used with verbs of remembering.
- d) The hortatory (or jussive) subjunctive.
- e) The supine in **um**.

Meanwhile, in a similar section on English grammatical constructions, the 1915 exam instructed students to do the following.²⁸

Copy the following sentences, making such grammatical changes as you think necessary.

Explain your changes.

- a) Between you and I, I think I would prefer not to acknowledge the mistake
- b) Tell me the circumstances both pleasant and otherwise
- c) Each one said goodbye in their own way

²⁶ College Entrance Examination Board. *Annual Report of the Secretary, 1915*. Pennsylvania State University Press, 1915.

²⁷ College Entrance Examination Board. *Examination Questions in Latin and Greek 1901-1915 Third Series #1B*. University of California Press, 1919, 8

²⁸ College Entrance Examination Board. *Entrance English Questions Set By the College Entrance Board 1901-1928*. University of California Press, 1929, 175

To answer the question in the first approach, all a student would have needed was knowledge of Latin cases, rules, and definitions. However, the second question touched upon something very different. While it did require that a student possessed background knowledge of English grammar conventions, it also demanded that a student apply this knowledge to an unprecedented circumstance. By reframing the SCEE from a curriculum based test to a comprehensive one, the Board intended to ensure one thing. No matter how much raw preparation each student had been provided by their secondary institution, they would only score high if they could properly utilize their knowledge in a hypothetical real world situation. From 1915 onward, the SCEE measured a student's merit not only through attainment, but also through their demonstrated potential.

The Board's decision to modify the SCEE into a comprehensive examination demonstrated their flexibility and dedication to Eliot's philosophy. However, the changes would not stop there. As the SCEE's developers embraced comprehensive examinations, its developers began to flirt with an even more radical approach. In the decade following 1915 a series of psychologists published research which would fundamentally change the character of the SCEE. This burgeoning field of study came to be known as "psychometrics."

Part 3 - Psychometrics

Throughout the 1910s, as the SCEE was gradually taking shape, so too was the field of psychometrics. Psychometrics was a field of psychology concerned with the theory and technique of measuring human intelligence. The psychologists who practiced psychometrics were focused on developing a means of measuring human intellectual capacity through a written test. The first person to attempt this was French Psychologist Alfred Binet. In 1904 Binet was

commissioned by the French government to design a test which would identify students with learning disabilities so that public schools could provide them with individual attention.²⁹ He developed a test called the “Binet Scales” which measured children’s verbal, spatial reasoning, and pattern recognition skills. From the onset of his research, Binet made it clear that his test was not a “general device for ranking all pupils according to mental worth.”³⁰ He believed that intelligence was a multifaceted phenomenon which could change over time with proper instruction. Despite Binet’s warning, when the Binet Scales were brought to America by psychologist Henry H. Goddard they took on a vastly different role. Goddard was fascinated with the Binet scales, and believed that they could have far-reaching applications beyond their original purpose. Goddard’s interpretation of the Binet test, however, was clouded by his own preconceptions about what intelligence was.

In 1919 Goddard published a book entitled *Human Efficiency and Levels of Intelligence* which stated that the chief determinant of a person’s merit was “the unitary mental process called intelligence.”³¹ In Goddard’s mind, intelligence was a singular measurement composed of an amalgamation of different skills, such as the ones tested in the Binet Scales. A person’s proficiency in each of these disparate skills could be averaged into their “Intelligence Quotient,” or IQ for short. A person’s IQ, to Goddard, had a direct relationship to their capacity for exemplary behavior and success in their life.³² Throughout the book, Goddard also argued, mainly through anecdotal evidence, that intelligence was innate and could not be augmented nor reduced by a person’s circumstances of birth.³³ With this in mind, Goddard perceived Binet’s test

²⁹ Siegler, Robert S. “The Other Alfred Binet.” *Developmental Psychology* 28, no. 2, 1992, 179–180

³⁰ Binet, Alfred, and Th. Simon. “The Development of Intelligence in Children (The Binet-Simon Scale).” Translated by Elizabeth S. Kite. *L’Année Psychologique*, 1916, 274.

³¹ Goddard, Henry Herbert. *Human Efficiency and Levels of Intelligence*. Princeton University Press, 1920, 1

³² *Ibid.*, 15

³³ *Ibid.*, 8

as a goldmine of opportunity. In the preface to his book, Goddard addressed the possibility of a real-world application of Binet's intelligence test, writing,

If mental level plays anything like the role it seems to, and if in each human being it is the fixed quantity that many believe it is, then it is no useless speculation that tries to see what would happen if society were organized so as to recognize and make use of the doctrine of mental levels...Testing intelligence is no longer an experiment or of doubted value. It is fast becoming an exact science. Greater efficiency, we are always working for. Can these new facts be used to increase our efficiency? No question! We only await the Human Engineer who will undertake the work.³⁴

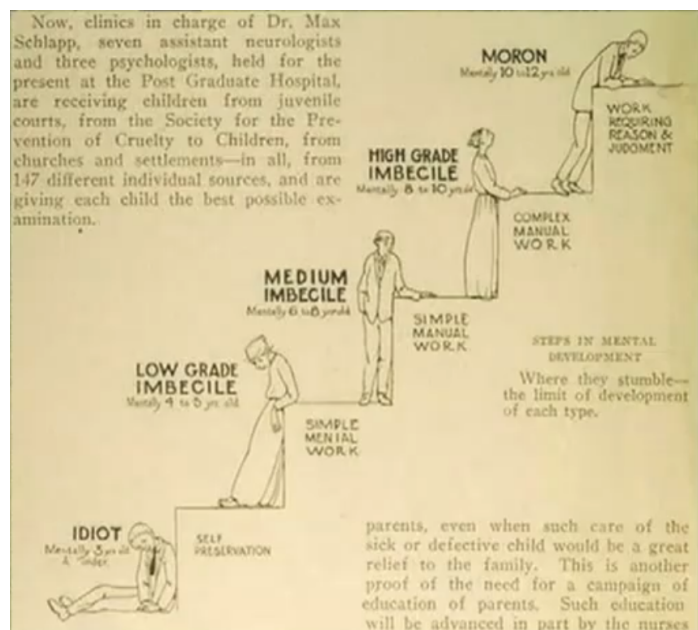
Having clearly ignored Binet's initial warning, Goddard saw intelligence tests as a potential basis for restructuring American society. He desired for intelligence tests to be perfected and implemented across the country. If society were stratified based on the results of these tests then, as implied by Goddard, the country could be more productive and orderly, and all would benefit as a result. Later on in his book, Goddard elaborated on what this society would look like.

Goddard's ideal society was one where vocations were allocated to citizens based on their intellectual promise. Jobs which required reasoning capabilities would be relegated to those who exhibited high intelligence, while manual labor jobs would be given to the less intelligent.³⁵

Goddard's proposal was no doubt a radical one. At the time, Goddard and other proponents of his utopian vision thought of themselves as akin to revolutionaries. In an article from "The Survey Magazine" praising intelligence testing, Goddard's

³⁴ Ibid., iv

³⁵ Ibid.



proposal was caricatured in a very poignant manner.³⁶ This drawing appeared quite reminiscent of contemporary depictions of class stratification, such as those drawn by Marxists depicting Capitalism. However, instead of descriptive roles such as “the Bourgeoisie” or “Proletarian laborers” this depiction marked each class with a description of their intellectual capabilities. This drawing spoke to the fact that Goddard and his followers believed that they were fundamentally restructuring the class system of 20th century America. In this new order, one’s role in society would be delegated not on the basis of inheritance or capital, but on intellectual promise. Notably, rather than depicting this new class stratification as a pyramid or tower, as the Marxists often did, the author depicted it as a staircase. The implication of this depiction was that any person, rich or poor, could climb the social hierarchy so long as they possessed the mental facilities to do so. Goddard promised to optimize American society by replacing the traditional class structure with one where anyone, no matter their present station, could rise only as far as their intelligence allowed them. Through his advocacy for intelligence tests, Goddard had birthed the idea of an “elite based on merit.”

Goddard’s philosophical ideals overlapped strikingly with those of the Charles W. Eliot and the College Entrance Examination Board. Both believed that it was in the best interest for American society if the educated “elite” were selected based on merit rather than birthright, and both sought to use written tests to achieve this goal. However, Goddard firmly believed that the best way to quantify a person’s merit was through the supposedly “exact” science of intelligence testing. Despite the remarkably similar aims and worldviews between the Psychometricians and the Board, they did not come into contact until the early 1920’s. Thanks to Princeton University

³⁶ Virginia State Board of Charities and Corrections. *Mental defectives in Virginia: a special report of the State board of charities and corrections to the General Assembly*. In Survey Magazine. Brandeis University Libraries, 1919.

professor and psychologist Carl Brigham, many Psychometricians would begin to see the SCEE as an opportunity to implement their goal of systematized intelligence-testing.

Carl Brigham's fascination with psychometrics began with his time in the army. During World War 1, Brigham was commissioned by the U.S. army to conduct an experiment on the predictive abilities of intelligence testing.³⁷ He and his team were instructed to administer a modified version of Binet's test - called the Army Alpha Test - to both regular soldiers and officers. Brigham's team was wary of the fact that each soldier's educational background could influence their scores. As a means to prevent this, according to Brigham, each question was designed as a puzzle where no knowledge outside the instructions could affect a tester's score.³⁸ As they administered the Army Alpha, Brigham's team kept track of each soldier's scores alongside their job performance. At the end of their experiment, the team found a remarkable positive correlation between job performance and test scores in officer roles. Candidates who were in the top 20% of scorers exhibited high competency in their jobs, while those below 20% tended to perform worse.³⁹ However, among regular recruits, nearly every candidate above the 80% threshold performed well.⁴⁰

These results of the Army Alpha tests constituted the first compelling evidence of the predictive value of intelligence testing. By presenting a clear correlation between one's IQ and their capability to perform high-skilled leadership work, they had seemingly confirmed Goddard's presuppositions about the determinative nature of intelligence. Following the results of

³⁷ Downey, Matthew T. 1961. *Carl Campbell Brigham Scientist and Educator*. Princeton N.J: Educational Testing Service, 1961, 3-4

³⁸ United States War Department et al. *Army Mental Tests: Methods, Typical Results and Practical Applications*, Henry Holt and Company, 1920, 43-44

³⁹ *Ibid.*, 13

⁴⁰ *Ibid.*

the Army Alpha Test, Brigham began to vocally sympathize with Goddard's philosophical ideals. He became immersed in the question of how intelligence tests could be implemented and utilized in wider American society.⁴¹ As it turned out, his pedagogical background provided him with the perfect answer as to when, where, and how these intelligence tests could be administered.

Carl Brigham concluded from the Army Alpha experiments that it was definitively possible for a test to measure either "pure intelligence, or at least something that makes for educational and industrial success."⁴² Knowing this, Brigham concluded that a version of the Army Alpha should be administered in intellectually rigorous fields in order to select candidates with the most potential to perform well.⁴³ As he formulated a plan to administer these tests, Brigham likely noticed that nearly all of the vocations that required a high degree of mental aptitude were only accessible with a university degree. He also paid attention to the fact that the College Entrance Examination Board had already been experimenting with comprehensive exams, which were similar in aims and design to the Army Alpha. With all of this in mind, Brigham concluded that the best and most feasible place to utilize an intelligence test was in university admissions. In 1922, Brigham crafted a statement on college admissions that was likely intended to be read by the Board. In it he wrote,

This question of admission is one that suggests desirability of making improvements in two directions. Is it not possible to admit a larger number of students who are now excluded but who could profit by a college course? Is it not possible to exclude a larger number of students who come to college and fail, wasting their own time and father's money? The use of psychological methods ought to help in solving both of these problems.⁴⁴

⁴¹ Downey Matthew T. *Carl Campbell Brigham Scientist and Educator*. 1961, 6

⁴²Brigham, Carl. Letter to E.G. Boring. 3 January, 1923 in *Harvard University Archives: E.G. Boring Correspondence*.

⁴³ Brigham, Carl. *A study of American Intelligence*. Princeton University Press, 1923, 210

⁴⁴ United States War Department, et al. *Army Mental Tests: Methods, Typical Results and Practical Applications*, 1920, 176

Brigham's statement demonstrated that he was acutely familiar with the aims of the College Entrance Examination Board. As president of the Board, Charles Eliot had hoped that the SCEE would bring about change in the admissions practices of private American universities. More specifically, he wanted Harvard's freshman class to be composed of the most promising students from all walks of American life, rather than the "stupid sons of the rich." By drawing a distinction between the talented poor student and the unremarkable rich one, Brigham addressed that desire directly. He relayed to the board that if they wanted to admit students purely on the basis of merit, the best way to do so was through intelligence testing. With this, Brigham sent an unspoken message to the Board. He emphasized that the future of the SCEE was as an intelligence test.

Part 4 - The New Plan

Brigham's pitch to the College Entrance Examination Board prompted lengthy debate inside their ranks. The Board had been paying careful attention to the psychometric research of Brigham and his counterparts. When the results of the Army Alpha test were released to the public in 1922, many members of the board began to view Brigham's suggestions as a serious possibility. One Board member, Professor John J. Coss of Columbia University, remarked that after reading the results of the Army Alpha test he, "return[ed] to my teaching job full of hope, believing that at last we had a means of prediction which, when perfected, would be superior to any previously known type of examination."⁴⁵ Though he did not outright reference the SCEE, Coss strongly implied that an intelligence test would be a far superior metric of a student's merit than the present test. The first official call for a change in the SCEE came from Dr. Fiske of

⁴⁵ Coss, John "John J. Coss Diary." September, 1919 in *The Selected Papers of John J. Coss 1908-1952*. Quoted in Claude Fuess, *The College Board its First Fifty Years*. Amsterdam University Press, 1950, 103

Dartmouth College, who wrote a report to the board entitled “Comprehensive Examinations as Intelligence Tests.” In this report he stated,

It should be the purpose of the College Entrance Examination Board not only to ascertain whether the candidates have acquired the information and methods of thought necessary for successful work in college, but also to determine whether they possess certain important intellectual qualities. Up to the present time the Board's endeavors in this direction are exhibited most conspicuously in the comprehensive examinations. If these examinations continue to respond to the demands of the times they should eventually become the best possible tests combining the necessary elements of an informational and of a mental, or psychological, character.⁴⁶

In his report, Dr Fiske argued that in order for the board to admit the most capable students, they should keep the SCEE up to date by employing the techniques of intelligence tests. Notably, nowhere in his statement did Fiske question the accuracy or predictive value of intelligence tests. Rather, he accepted their purported merits as hard, scientific fact. This sentiment appeared to have resonated with the rest of the College Entrance Examination Board. In November of 1922, the Board voted to adopt a resolution stating that it looked with “favorable interest upon the use of general intelligence testing”⁴⁷ and stood ready to integrate them into the SCEE as soon as possible. Thus, they set out on their path to change the character of the exam.

In 1924 the Board appointed an advisory committee of experts on intelligence testing to redesign the SCEE.⁴⁸ At the head of this committee was Carl Brigham himself. From 1924 to 1926 Brigham and his team were tasked with designing the new SCEE which would come to be known as the Scholastic Aptitude Test (SAT). In a 1925 report to the Board, Brigham outlined the design of his new exam. The principal aim of the new SAT was to “put as little premium as

⁴⁶ College Entrance Examination Board. *The Work of the College Entrance Examination Board, 1901-1925*. Ginn and Company, 1926, 163

⁴⁷ *Ibid.*, 189

⁴⁸ *Ibid.*, 210

possible on specific training, and more emphasis on potential promise as distinguished from prior accomplishment.”⁴⁹ In other words, Brigham hoped to negate the perceived flaws of the 1901 curriculum tests by measuring intellectual capacity as opposed to attainment. If Brigham's SAT achieved what it set out to do, it would ensure that every child, whether from a private preparatory school or an underserved public school, would receive a score which demonstrated their promise. Thus, the elite universities of the Board would have their pick of the brightest minds from all walks of American society.

After two years of work, Brigham completed the first SAT. This test diverged radically from the SCEE administered in 1915. In order to ensure that it would be an accurate measure of academic potential and devoid of any circumstantial bias, the SAT was designed to mirror the Army Alpha Tests. As such, like the Army Alpha it was formatted in a similar manner to a puzzle, where no outside information was required to solve it. In an excerpt from the language section, the test asked the following question.

SUB-TEST FOUR

CO

FORM A

Read the vocabulary and rules of the artificial language given below. Then study the sample sentences. Do not try to memorize the vocabulary or forms, but consult them freely while translating the sentences below.

<p>I — ot me — oteb he — ol him — oleb that — ren</p>	<p>VOCABULARY</p>	<p>is — pue please — thanto satisfy — borpo live — homo forever — thos</p>
---	-------------------	--

RULES

1. Plurals are formed by adding "a." Only nouns and pronouns have plurals.
Examples:
we — ota
them — oleba
2. Past time is expressed by placing "erp" before the verb.
Example:
pleased — erpthanto
3. Future time is expressed by placing "bel" before the verb.
Example:
will please — belthanto
4. Nouns are formed by substituting the ending "ac" for the "o" ending of the verb.
Example:
pleasure — thantac
5. Adjectives are formed by substituting the ending "em" for the "o" ending of the verb.
Example:
pleasant — thantem
6. Adverbs are formed by substituting the ending "id" for the "o" ending of the verb.
Example:
pleasantly — thantid

SAMPLES

- | | |
|--|--|
| <p>(a) He pleases me.
Ol thanto oteb</p> | <p>(e) He is living.
Ol pue thanto</p> |
|--|--|

In this question, Brigham and his team aimed to test a student's ability to express themselves in a foreign language. However, they understood that by testing a student's knowledge of a real foreign language, they were favoring only those who had prior training in that language. In order to circumvent this, they asked students to learn the principles of a fake language and tested them on their ability to construct sentences with the information given. In doing this, Brigham and the College Entrance Examination Board believed that they had leveled the playing field for all testers. To them, this new SAT was sure to produce a measure of pure ingenuity and aptitude which superseded their prior degree of education.

On June 23, 1926 the first official SAT was administered to a group of 24,403 candidates.⁵⁰ The change from the SCEE to the SAT was a huge accomplishment in the eyes of the Board. However, for Brigham and the other test developers, it was only the beginning of a long journey toward perfecting the new college entrance exam. After the 1926 administration of the SAT, Brigham kept careful track of each student's performance in order to identify areas where the test could be improved. From 1926 to 1930, Brigham embarked on a series of analyses to make sure that the SAT lived up to its lofty expectations.

Part 5 - Results

In his book about the history of the SAT, author Nicholas Lemann stated that if the SAT were to exist as an accurate quantification of human intelligence, it needed to demonstrate two things.⁵¹ The first of these was the question of reliability. The SAT claimed to be a measure of the supposedly innate and constant phenomenon of intelligence. If this were true, SAT scores should

⁵⁰ Fuess, Claude. *The College Board. Its First Fifty Years*. 1950, 120

⁵¹ Lemann, Nicholas. *The Big Test: The Secret History of American Meritocracy*. Farrar Strauss and Girot, 1999, 32

demonstrate similar patterns each year that the test was administered. Ideally, they should fall on a relatively consistent curve, and remain constant for each test taker from one administration period to the next. Alongside reliability, the SAT needed to demonstrate some degree of validity. By measuring intelligence, the SAT should be able to predict a pupil's future academic and vocational performance to a reasonable degree. As Brigham and his team analyzed the data from the first several administrations of the SAT, they kept these principles in the back of their mind. However, the trends that they found in their analysis would ultimately shed an unfavorable light upon their test.

Brigham first noted that there was a discrepancy between testers' scores on the verbal and quantitative sections of the SAT. In the 1926 administration of the test, nearly all students who had performed well on the verbal section underperformed on the math section.⁵² Seeing as Brigham believed intelligence to be a unitary, comprehensive metric, he considered the discrepancy between verbal and math scores to be an error in reliability.⁵³ After much thought, Brigham concluded that the math section "measured a different component of intelligence"⁵⁴ than the verbal one. This statement, though unclear, likely implied that the math section was measuring secondary school attainment rather than one's innate intellectual capacity. To remedy this, the Board removed the math section from the 1927 SAT.⁵⁵ However, the reliability issues with the SAT did not stop there.

⁵² College Entrance Examination Board. *Twenty Seventh Annual Report of the Secretary*. Pennsylvania State University Press, 1927, 4

⁵³ Brigham, Carl. *A Study of Error: A Summary and Evaluation of Methods Used in Six Years of Study of the Scholastic Aptitude Test*. College Entrance Examination Board, 1932, 351-355

⁵⁴ *Ibid.*, 353

⁵⁵ Lawrence, Ida, et al. "A Historical Perspective on the Content of the SAT" *College Entrance Examination Board*, 2003, 6

From 1927 onward, Brigham's team continued to collect data on students' performance on the SAT. However, as more tests were administered to the public, new issues began to appear. The most stark of these problems were the fluctuation of test scores from year to year. In the 1926 administration of the SAT only 52.8 percent of students passed the Physics portion of the exam.⁵⁶ In 1927 this number rose to a staggering 81.9 percent.⁵⁷ Another issue which was illuminated with subsequent examinations was the variability of scores for repeat test takers. Among students who had taken the SAT more than once, Brigham noticed that their scores often increased far beyond the standard error.⁵⁸

With these facts, it became clear to Brigham and his team that the SAT had failed to demonstrate proper reliability. Upon realizing this, the Board concluded that these "eccentric results" were the product of a design flaw.⁵⁹ To them, Brigham's team had made the simple mistake of administering tests which varied in difficulty. If these flaws were remedied, the SAT would eventually be able to properly quantify a student's natural intelligence. Brigham himself, however, viewed these results with a bit more pessimism. Brigham had carefully and deliberately designed the SAT to resemble the Army Alpha test, which had been supposedly proven to measure intelligence. If the SAT could not demonstrate reliability, he thought, then perhaps it was indicative of a larger problem with intelligence testing in general. With this in mind, Brigham and his team turned their attention to the demonstrated validity of the SAT.

⁵⁶ College Entrance Examination Board. *Twenty Seventh Annual Report of the Secretary*. Pennsylvania State University Press, 1927, 188

⁵⁷ *Ibid.*

⁵⁸ *Ibid.*, 189

⁵⁹ *Ibid.*, 3

Brigham and his team set out to determine the degree to which the SAT could predict a student's potential for success. Unfortunately, given that his study was published in 1927, the researchers could not definitively measure each student's achievement in their post-graduate career. Thus, they settled on using freshman year grades as a broad metric of "success." In comparing students' SAT scores with their freshman year GPA, they found a relatively weak positive correlation. The correlation coefficient, which ranges from 0.0 (no relationship) to 1.0 (complete linearity) was found to be 0.6.⁶⁰ This indicated that the SAT did not have anywhere near the predictive abilities they had anticipated. Brigham first explained this discrepancy as a product of environmental factors. At the end of his study, he wrote that freshman year grades were a flawed metric of success, given the fact that "social distractions" prohibited students from "working at their full capacity."⁶¹ However, on some level, Brigham may have begun to question the effectiveness of his SAT. If the SAT could not demonstrate proper reliability nor validity, then it was clear that it was not measuring any sort of innate intelligence. Knowing this, Brigham had only one question in mind. That was, if the SAT could not quantify a person's intelligence, what was it truly measuring? Brigham set out to answer this question in his 1932 book *A Study of Error*.

After almost 5 years of deliberation, Brigham wrote a scathing critique of his SAT. He admitted that the SAT's failure to demonstrate reliability and validity was not a remittable design flaw. Rather, it was proof that the notion of intelligence could not be successfully quantified using a test. At the end of his book Brigham wrote,

The test movement came to this country some twenty five or thirty years ago accompanied by one of the most glorious fallacies in the history of science,

⁶⁰ Ibid., 200-202

⁶¹ Ibid., 211

namely, that the tests measured native intelligence purely and simply without regard to training or schooling. I hope nobody believes that now. The test scores very definitely are a composite including schooling, family background, familiarity with English, and everything else, relevant and irrelevant. The native intelligence hypothesis is dead.⁶²

Brigham's scathing rebuttal of the SAT noted that the test had failed to differentiate itself much from the attainment tests of the 1910's. He noted that a student's performance on an intelligence test, such as the SAT, was contingent upon the degree of training they had received in secondary school. This revelation provided a conclusive explanation for the increase in scores among repeat test takers who had received further schooling in the time between their two tests. However, to Brigham, the SAT was a reflection of more than just a student's attainment. He also conceded that intelligence tests reflected the social and environmental circumstances a student had been reared in. By admitting this, Carl Brigham demonstrated that the SAT had rebuked the principles on which it was based. The SAT was designed to facilitate the admission of students on the basis of their merit rather than privilege. However, in actuality, it reflected the very social stratifications it sought to negate. In other words, the supposed "merit" which the SAT quantified was inextricable from privilege.

Conclusion

Despite Brigham's warnings, the Board continued to use the SAT to pursue their goal of admissions reform. During the 1930's and 40's, the SAT exploded in popularity among elite universities.⁶³ This was due, in part, due to the advocacy of Harvard's president James Conant. Conant promoted the widespread use of the SAT as a means to foster "equality of opportunity" in

⁶² Brigham, Carl. *A Study of Error: A Summary and Evaluation of Methods Used in Six Years of Study of the Scholastic Aptitude Test, 1932*, 380

⁶³ Fuess, Claude. *The College Board. Its First Fifty Years*. 1950, 125

university admissions.⁶⁴ He was an ardent believer that by measuring intelligence, the SAT provided a picture of a student's merit independent of their circumstance. However, as the SAT became more widespread in American universities, the Board began to notice some of the patterns that Brigham had originally called to attention.

In the 1960's and 70's, the SAT became mandatory for most American universities. With more colleges outside the original group of elite, private universities requiring the SAT, the number of test-takers skyrocketed from around 24,000 in 1926 to 1,000,000 in 1975.⁶⁵ As more students began to take the SAT, the average score declined from 1,060 in 1960 to 1,010 in 1975.⁶⁶ This sharp decline was addressed in a 1990 study by Lawrence Stedman of Yale University. Lawrence's study concluded that around 40% of the score changes from 1960 to 1975 could be attributed to "demographic changes in test takers."⁶⁷ In other words, as more students from lower socioeconomic backgrounds took the SAT, the average score plummeted. This highly publicized study proved to the Board, and the American public, that the SAT did not transcend circumstance. Rather, each student's score reflected their background, and the educational opportunities that had been allocated to them.

Following this discovery, the College Entrance Examination Board, which by the 1990's had become a private company known as the "College Board," began to distance themselves from discussions about equity in admissions. In 1993 the College Board renamed the SAT from

⁶⁴ Conant, James Bryant. *Our Fighting Faith: Five Addresses to College Students*. Harvard University Press, 2013, 28

⁶⁵ College Entrance Examination Board. *Annual Report of the Secretary 1975*. Pennsylvania State University Press, 1976, 10

⁶⁶ *Ibid.*, 22

⁶⁷ Stedman, Lawrence. Kaestle, Carl. "The Great Test Score Decline: A Closer Look" *Yale University Press*, 1990, 132.

the Scholastic Aptitude Test to the Scholastic Assessment Test.⁶⁸ When asked about this change, Board president Donald Stewart stated that it was done "to correct the impression among some people that the SAT measures something that is innate and impervious to change regardless of instruction."⁶⁹ By 1993, it was clear that the College Board had given up on the intelligence-approach to the SAT. They realized, just as Brigham did 60 years before, that it was impossible for a test to measure a student's merit in a way which negated their prior instruction and circumstances of birth. After reconciling with this fact, the Board determined that the SAT would continue to be used. However, from 1993 onwards, the Board framed the SAT as a test which measured merit through attainment. After years of trial and error, the SAT ended up right back where it started.

All in all, the story of the SAT is ultimately one of hubris and miscalculation. The test was originally designed with a lofty goal in mind. It aimed to erode many of the barriers faced by disadvantaged students applying to college by providing a means for universities to judge applicants based on merit. The test was originally designed to measure a student's attainment of certain secondary school concepts that were taught in public and private schools alike. However, many Board members believed that attainment could not truly paint a picture of a student's promise in a way that overlooked their privilege. As the 1920's approached, the Board made the bold and fallacious decision to measure a student's merit through their perceived intelligence. However, this intelligence-based approach was a remarkable failure. Despite claims made by psychologists that intelligence tests provided an unbiased glimpse into a student's innate

⁶⁸ Jordan, Mary. "SAT Changes Name, but It Won't Score 1,600 with Critics." *The Washington Post*, March 27, 1993.

⁶⁹ Ibid.

capacity, studies would eventually demonstrate that it did the opposite. The new SAT, in actuality, measured a combination of attainment and a student's wealth. With this realization, the Board decided decades later to give up on their quest for equity in admissions. They recognized that, despite their good intentions, true and unbiased meritocracy could not be achieved through the framework of a test. While Eliot's dream of admissions based on merit rather than circumstance has lived on, it is evident that this illustrious goal cannot be achieved singlehandedly through the SAT.