

A Quantitative Analysis on the Differences Between the SAT and the ACT:

Evaluating the Disparities in Yield Curves

Matthew Rudolph Larriva

Abstract

A common and pressing question in the pre-college preparation realm asks which test should a student take: the SAT or the ACT. It can be reasoned that this question is more appropriately phrased, “which test will a student score better on?” While a survey of current research and publications yields a plethora of information about the differences between the tests, there is little in the way of quantitative research on the *implications* of these differences. One instantly finds all the disparities between the tests, but is left to hearsay to decide which to prepare for. Given the availability of quantitative data on the results of the tests, this question should be answerable.

This analysis will start by reframing a commonly asked question, “which is easier?” and instead pursue a more appropriate question, “which test allows faster point gains?”. Once this question is edified, it needs to be shown how the differences between the tests are not nullified by the standardized nature of the tests themselves. Once the question is reformatted and the nullification argument is satisfied, the analysis will present a statistical summary of findings illustrating that the SAT rewards preparation disproportionately to the ACT up to about 37 hours of prep, after which the ACT rewards preparation to a greater degree than does the SAT. Based on these differing yield curves, the analysis will suggest three reasons for the differences and make recommendations on how to evaluate a student’s potential success in preparation for the SAT versus the ACT.

Not “Which Test Is Easier” but “Which Test Better Rewards Preparation?”

The zeitgeist needs to move away from asking which of the two tests is easier. If we define “easier” as which test has more perfect scorers, then the answer is the ACT.ⁱ In 2012, .04% of ACT takers scored a 36ⁱⁱ, while The Collegeboard reports approximately .02% perfect scoresⁱⁱⁱ. Thus it is twice as likely to score perfectly on the ACT than on the SAT. But this result is unsatisfying and myopic. The results of the ACT are skewed right, a phenomenon having to do with the nature of averages, to be discussed in greater detail later. More importantly, most test takers are not realistically targeting perfect scores. So this definition of “easy” is unimportant to the vast majority.

If a student were targeting a perfect score, and had the starting-score and intellectual capacity to make that a reality, then the ACT would be the preferred test—more people obtain perfect scores, and that number is increasing over time^{iv}. But, again, this is irrelevant to most.

The right question to ask is, which test is easier to change one’s score on, because for the vast majority of test takers, this is the question of importance.

Standardization Does Not Nullify the Differences

Now that the question is properly framed it is important to discuss the nature of standardization to establish that the standardization component of the tests does not in fact nullify their differences. A lazy sophism states, “as both the SAT and ACT are graded on a curve would not any differences simply be muted?”

First, these tests are not “curved.” Let us be clear on what a curve would look like. A “curve” in the nature of which most people inquire would mean that there would need to be the same percentage of perfect scorers on every sitting of every SAT and ACT. This is not the case. Quite the opposite: if the entire population suddenly became brilliant, then both tests would allow 100% of test takers to receive perfect scores.

Although the process that the Collegeboard (makers of the SAT) and ACT (makers of the ACT) use to standardize their results is vague and nebulous, we are perpetually assured by researchers and test-makers that it is not through a “curve” in the abovementioned sense.

So now let us hone the initial question: if one test is easier to change one’s score in, would not that effect benefit all test-takers equally and thus be nullified?

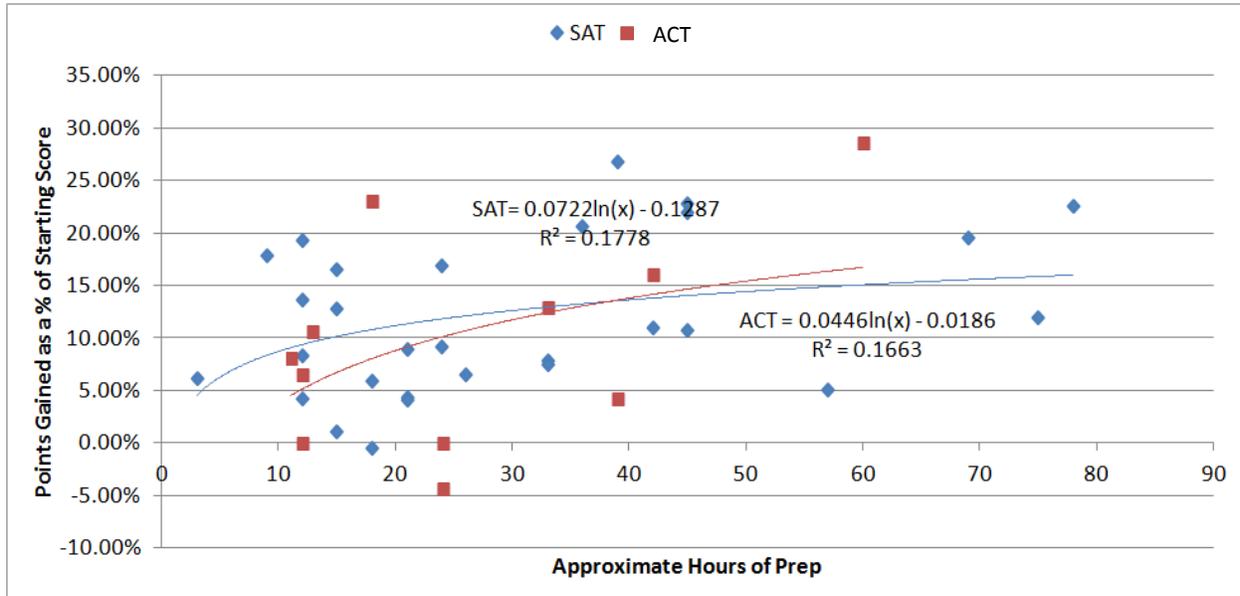
Again: not so. If the entire pool of test-takers learned the definition to the word *pugilist*, and that word were explicitly tested on the SAT, then we would expect about a 10-point increase on average SAT Critical Reading sections. The test makers standardize their tests so that iterations are comparable, not so that variations in intelligences are neutralized. The goal is to ensure that an October 2012 sitting is directly comparable to a December 2013 sitting by standardizing the difficulty of the tests. Their goal is not to lower the scores of the October test-takers simply because a larger proportion of them happened to score high.

Tangibly, what this indicates is that the standardizing process of the test makers does not neutralize the differences in the tests; it simply neutralizes variety within different iterations of the tests. By way of example, to say that the SAT and ACT differences are irrelevant because of “the curve” is like saying the differences between the MCAT and the GRE are insignificant because these tests are graded on curves. The tests are not standardized *to each other* but only *to themselves*. So if it could be proven that one test allowed easier travel in points than the other, clearly, that one would be the test to take.

Thesis

According to the data of Powerful Prep, Inc.—a test prep firm out of Southern California—the SAT rewards preparation disproportionately to the ACT up to about 37 hours, at which point the yield curves cross. This analysis will demonstrate the methodology and then suggest three reasons for the differing yield curves.

Yield Curves



Methodology

Using SAT and ACT paired points illustrating hours of preparation and percent gained relative to starting score, the data were plotted.

Two points were dismissed, one from the SAT and one from the ACT where gains were abnormally large.

Two natural log curves were plotted as lines of best fit; Rsquare values are displayed.

The Use of Natural Log Curves

The natural log is used to produce trend lines because its shape matches intuitively the course of preparation—strong initial gains followed by a diminishing rate of marginal return. This theory is substantiated by the Rsquared values, which are strong enough for analysis.

At this point, an analysis of the differences in yield curves is warranted.

This study proposes three reasons for the disparity in yield curves:

- The Impact of Vocabulary
- Scoring Processes
- The Intent of the Tests (Innate Ability versus High-School Review)

The Impact of Vocabulary on the Curve

This section will consider the amount of points a test-taker can gain, on both the SAT and ACT, by expanding his or her knowledge of vocabulary.

On the ACT, this impact is null. That is, the ACT does not test, at least not explicitly, a test-taker's knowledge of non-colloquial vocabulary.

On the SAT, this vocabulary impact is acute. Sources indicate that up to 50% of the Critical Reading section is directly or implicitly vocabulary based^v. A conservative analysis suggests 28% as a lower bound.

There are 67 questions on the Critical Reading section. 19 of these questions are so called Sentence Completion questions—which rely largely on vocabulary knowledge. Even without the inclusion of the five-or-more Vocabulary-In-Context questions, 19 of 67 is already 28%.

Now, it is important to realize how facile the Vocabulary-In-Context and Sentence Completion questions become when the definitions of the words are known.

A sentence completion question is as follows:

The _____ professor became even more cranky when his students would _____ and talk without stopping.

- (A) Old.....opine
- (B) Cantankerous...drone on
- (C) Pious.....dither
- (D) Masochistic....machinate
- (E) Belligerent...abstain

A Vocabulary In Context question is as follows:

In Context, the word *dogged* most nearly means:

He was dogged in his preparation, unyielding and strong willed.

- (A) Canine
- (B) Libelous
- (C) Strong
- (D) Persistent
- (E) Faltering

A reader must allow that knowledge of the definitions of these words would render these questions unequivocal and very easy.

A reader must further allow that memorization of vocabulary words is one of the most linear tasks a test-taker would have to complete. In most cases, contributing 1-2 hours will result in the knowledge of 100 words. Because the SAT recycles its tested vocabulary, and because these words are widely and freely available on the internet, the vocabulary component of the SAT is easily and quickly mastered. Compare this vocabulary aspect of test-prep with learning how to factor quadratic equations. The former is guaranteed to occur in 19 out of 67 Critical Reading questions and is easily conquered through memorization. The latter could occur 3-4 times on one test or as few as once, and its mastery is not guaranteed even after hours of practice. Would you rather memorize the definition to *pugilist* or study the process of how to factor $x^2-x+110$?

Thus, because of the ease of memorizing vocabulary words in comparison to learning the steps and applications of more complex strategies, because of the fixed bank of vocabulary words tested, and because of the large portion of the Critical Reading section that depends on vocabulary, we theorize that the SAT derives its greater yield curve slope—its greater returns to 1 hour of preparation—in part from the impact of vocabulary.

The Scoring Process

The SAT is graded as follows: a student receives a score on each of three sections (Writing, Critical Reading, Math) and these scores are added to reach a composite score. Thus a 700 on Writing, 750 on Critical Reading, and an 800 on Math would produce a 2250.

An ACT grade, however, is produced by averaging the scores of four sections (English, Math, Reading, Science). This number is then *rounded*. That means a 22.25 becomes a 22, and a 22.50 becomes a 23.

Therefore, the SAT composite score is produced via *adding* while the ACT composite score is produced through *averaging*. This process explains both why the ACT has lower points travel than the SAT, and why the ACT is easier to score perfectly on.

First, to illustrate why the ACT has lower points travel than does the SAT, consider the following: if a SAT tester improves by 1 question, his score will most likely improve by between 10 and 30 points depending on where the tester's starting score is. On the ACT, scoring a single extra question correctly would rarely yield a higher score.

2012-13 Official SAT Practice Test Raw Score Conversion Tables									
Critical Reading			Math			Writing			
Raw	Scaled	CR %	Raw	Scaled	M %	Raw	Scaled	Convert	W %
67	800	100%	54	800	100%	49	80	800	100%
66	800	99%	53	790	98%	48	78	780	98%
65	800	97%	52	760	96%	47	75	750	96%
64	790	96%	51	740	94%	46	73	730	94%
63	770	94%	50	720	93%	45	71	710	92%
62	760	93%	49	710	91%	44	70	700	90%
61	740	91%	48	700	89%	43	68	680	88%
60	730	90%	47	690	87%	42	67	670	86%
59	720	88%	46	680	85%	41	66	660	84%
58	700	87%	45	670	83%	40	64	640	82%
57	690	85%	44	660	81%	39	63	630	80%
56	680	84%	43	650	80%	38	62	620	78%
55	680	82%	42	640	78%	37	61	610	76%
54	670	81%	41	640	76%	36	60	600	73%
53	660	79%	40	630	74%	35	59	590	71%
52	650	78%	39	620	72%	34	58	580	69%
51	640	76%	38	610	70%	33	57	570	67%
50	630	75%	37	600	69%	32	56	560	65%
49	620	73%	36	600	67%	31	55	550	63%
48	620	72%	35	590	65%	30	54	540	61%
47	610	70%	34	580	63%	29	53	530	59%
46	600	69%	33	570	61%	28	52	520	57%
45	600	67%	32	560	59%	27	51	510	55%
44	590	66%	31	550	57%	26	50	500	53%
43	580	64%	30	540	56%	25	49	490	51%

64 correct questions and 3 incorrect question creates a raw score of 63.25 which rounds to 63, scaling to 770

65 correct questions and 2 incorrect questions creates a raw score of 64.5 which rounds to 65, scaling to 800

1 additional correct answer produces 30 more points.

52 correct answers and 2 incorrect answers produces a raw score of 51.5 which rounds to 52, which scales to 760

53 correct answers and 1 incorrect answer produces a raw score of 52.75 which rounds to 53, which scales to 790.

1 additional correct answer produces 30 more points

If a student had a 54 in English, a 35 in Math, a 23 in Reading and a 24 in Science, then improving by one question on each section would result in the exact same score

Approximate ACT Scale Score	Conversions			
	English	Math	Reading	Science
29	67-68	50-51	32	36
28	66	48-49	31	35
27	64-65	46-47	30	33-34
26	63	44-45	28-29	32
25	61-62	41-43	27	30-31
24	59-60	39-40	26	29
23	57-58	37-38	5	27-26
22	54-56	35-36	23-24	24-25
21	52-53	33-34	22	22-23

The above situations examined the effects of score changes with regards to additional correct questions, but recall the ACT section scores must then be *averaged*.

Taken to its extreme, there is a situation in which an ACT taker could improve 3 points (referring to the leftmost column in the above table, 3 more points is not necessarily achieved by 3 more correct questions) in one section and have no overall score change:

	Science	English	Reading	Math	Composite	Rounded
Starting	23	23	6	22	18.50	19.00
Finishing	23	23	9	22	19.25	19.00
Gain	0	0	3	0	0.75	0

Remember the numbers in the above table are scaled scores, not questions-correct. It may require up to 5 additional correct questions to move from a scaled score of 6 to a scaled score of 9.

The nature of averages can often serve to constrain an ACT taker by limiting points travel.

However, the converse must also be true: the ACT can hide some weaknesses. Take for example the following scores:

	Science	English	Reading	Math	Composite	Rounded
	36	36	36	34	35.50	36.00

A test taker could miss 2 points (not questions) and still arrive at a perfect score because the ACT is averaged and rounded. This same phenomenon would not happen on the SAT.

If the above examples find the reader lost because of his unfamiliarity with the scaling and averaging process, the argument can be distilled to this: there are only 36 possible scores on the ACT, whereas there are 240 possible scores on the SAT. It is easier to move scores on the SAT because there are more scores to *have*. With regards to the ACT: it's much harder to move scores, because there are fewer scores possible. On the SAT, there are 239 ways to *not* score a perfect 2400, but on the ACT, there are only 35 ways to *not* score perfectly.

This, we posit is one reason the ACT reports more perfect scores annually than does the SAT. This is also why we recommend that if a student is starting at a very high score upon entering the test-prep process, then the student should prepare for the ACT, as a perfect score is more achievable because errors are more permissible.

The Intent of the Tests

This analysis asserts that the SAT is an aptitude test. This is suggested based on the SAT's former title, "The Scholastic Aptitude Test"—a title which was then changed to Scholastic Assessment Test to avoid controversy. Makers of the ACT agree, calling the SAT, "...more of an aptitude test, testing reasoning and verbal abilities"^{vi}. The SAT makers suggest a more broad term, saying the SAT, "lets you show colleges what you know and how well you can apply that knowledge."^{vii}

The ACT is a *curriculum* test designed to assess a student's mastery of high-school level curriculum. It self describes as, "an achievement test, measuring what a student has learned in school."^{viii}. Some suggest the test is "uncoachable" in that a coach could not reproduce a high-school education's worth of curriculum over the course of a few-months' test-prep.

Now the successes in fulfilling their respective goals will not be evaluated here. What is worth assessing, however, is the degree to which a student's score can move (on a percentage basis) within both *types* of tests. If one allows that the SAT is closer to an aptitude test, and the ACT closer to an end-of-year final, then the disparity should become more pronounced. Aptitude tests measure abilities to complete set functions: think career aptitude, computer aptitude. Final exams, however measure reasoning and retention of key concepts, and their applications. Which would you rather try to master?

Allow then, that the SAT is more like the GRE, and the ACT more like the MCAT in that the former is another self-proclaimed ability test^{ix} while the later is a curriculum knowledge exam^x. Looking at the times required to prepare for those two tests: Princeton Review reports the average successful candidate spending 200-300 hours of preparation for the MCAT^{xi}, whereas experts recommend only 96 hours of prep for the GRE^{xii}. From this it is easy to infer that so called aptitude or ability tests require less preparation than curriculum review tests do.

This is why a student's score moves more slowly in ACT preparation: because the test intends to, and does test wide-reaching curriculum—a review of which takes longer than preparing for the SAT, which is an aptitude test—a slightly more narrow and function-completion oriented exercise.

Why Do the Curves Cross?

It is necessary to point out that the yield curves on the data set crossover at approximately 35 hours of preparation. That is, after preparing for roughly 35 hours, according to the data, time spent preparing for the ACT is more rewarding-in points- than time spent preparing for the SAT.

This analysis suggests two reasons for this: as mentioned, there are a greater percentage of perfect ACT scores than SAT scores. That is, on the ACT 36 is a relatively easy score to achieve. This was addressed in the above section on scoring. Similarly, and for whatever reason, there seems to be an anomaly in which the 99th percentile on the ACT is *also* relatively easy to attain for those who are starting near it. Thus

when the study removes the effect of those 99th percentile test takers in the data set, the crossover point moves out to 47 hours, which reaffirms our position.

The second reason is the availability of data past the 50th hour of preparation. There were a disproportionate number of ACT test-takers who prepared for more than 50 hours when compared to SAT test-takers. Removing students who prepared for more than 50 hours prevents the SAT and ACT curves from crossing; the SAT curve lies above the ACT yield curve at all points. Again, this reaffirms our central thesis: that the SAT seems to reward preparation more than the ACT does, for the same hours of prep.

Disclaimers and Potential Errors

This analysis was substantiated in large part because of the data set provided by the author's test prep company. This opens up the data for a myriad of human-interaction errors which include, among others, this particular company's ability to tutor the SAT versus the ACT. If it could be shown that this company's results were unlike those of other companies, then the data would merely suggest that Powerful Prep was more adroit at tutoring to the SAT than to the ACT.

However, the remainder of the quantitative analysis (The Intent of the Tests, Scoring, and Vocabulary) would still be accurate and compelling.

Conclusion

It is the findings of this analysis that the SAT rewards preparation at a greater rate than does the ACT. This can be explained in part by the large percentage of vocabulary-based questions, the scoring process of the tests, and the intent of the tests.

Therefore it stands to reason that most students should, as a default, prepare for the SAT if they are constrained in time or if they are not starting with an exceptionally high ACT score (that is, if they score below 30). However, if time is unlimited, or if a student is beginning very strongly, the ACT seems to be the more rewarding test, as it has a higher percentage of perfect scorers than the SAT does.

ⁱ ^ "The ACT Test® Data". ACT Inc. Retrieved September 15, 2011.

ⁱⁱ ^ "The ACT Test® Data". ACT Inc. Retrieved September 15, 2011.

ⁱⁱⁱ <http://press.collegeboard.org/sat/faq>

^{iv} ^ "The ACT Test® Data". ACT Inc. Retrieved September 15, 2011.

^v Own The SAT; © Anthony-James Green, 2011, www.TestPrepAuthority.com

^{vi} <http://www.actstudent.org/faq/actsat.html>

^{vii} <http://sat.collegeboard.org/about-tests/sat>

^{viii} <http://www.actstudent.org/faq/actsat.html>

^{ix} http://www.ets.org/gre/revised_general/faq/

^x <https://www.aamc.org/students/applying/mcat/about/>

^{xi} <http://www.princetonreview.com/medical/mcat-myths.aspx>

^{xii} http://www.ehow.com/decision_7220885_many-hours-should-study-gre_.html