Increasing Your ACT Composite Score: Should You Test a Second Time?

Matt Harmston, MA

Out of over 1.9 million students in the 2018 ACT-tested graduating class, 44% took the full ACT® test at least twice in hopes of improving their scores. Camara and Allen (2017) support this practice, confirming that factors such as time in the classroom between test administrations are associated with increases in ACT Composite scores from the first to second testing. Obvious benefits of increased scores include improved chances of meeting college admission and merit-based scholarship criteria (Doyle, 2006). Given motivation for and potential benefits of retesting with the ACT, an important question centers on the likelihood of students improving their Composite score between the first and second test. This paper and its associated web application are intended to provide student-specific information that assists students, families, counselors, and educators in deciding whether a student's chances of increasing his/her Composite score warrants taking the full ACT test for a second time.

Background

In order to explore first-to-second ACT Composite score gains, a **web-based application** was created to depict expected gains associated with levels of student academic and testing characteristics. Among others, some of these characteristics include high school GPA, previous ACT Composite score, and the expected wait time between first and second test.

The multitude of factors commonly associated with improving test scores are often grouped into three general categories (Lievens, et al., 2007). These include enhancement in cognitive proficiency, reduction in construct-irrelevant variability, and development of test-specific skills. Reported score gains due to any one or more of these factors assume that a student is taking identical or equated test forms designed to have parallel content and psychometric-based specifications.

Student development and application of test-taking strategies and/or test-specific skills between sessions represent additional factors underlying retest score increases. For instance, additional exposure to high school curricula can be considered an essential intervention targeting college and career readiness skill development as assessed by the ACT. Given that assessment of readiness is an important purpose of the ACT (ACT, 2017), it comes as no surprise that participants in Camara and Allen's (2017) study demonstrated gains of approximately 0.20 to 0.25 points per month of classroom time between tests. While this estimate reflects the overall results of their study, slightly greater gains were observed in the first three months of students retesting. The authors speculate that this increase could be due in part to practice effects.

Because performance on the ACT differs between levels of student academic characteristics and behaviors, it is important for students, counselors, and educators to consider how these factors work together in predicting a student's second ACT Composite score. This paper and the associated web app are intended to help explore this relationship.

Source Data

The base population for this study is the 2018 ACT-tested US graduating class (*N*=1,914,817). These students reported graduating in the 2018 academic year, took their latest ACT exam in 10th, 11th, or 12th grades under standard- or extended-time conditions, and attained a college-reportable score. First and second test records including scores for all four ACT test sections were selected for all graduates taking the ACT two or more times with at least one of the first two instances occurring on a national test date (*N*=794,632). Because students testing on a national test date are not required to do so as part of a contract, this criterion insured all students were potentially college-bound. After screening out records missing one or more predictors, the final sample included N=677,315 records.

After evaluating multiple models, student educational performance and behavioral variables with the strongest relationship to score changes were identified. These variables are listed in Table 1.

Table 1. Variables by Which Student Score Gains were Disaggregated.

Variable List
• Most recent ACT Composite score
• Number of times the ACT has already been taken
• Grade-level at time of first testing
• Grouped time-between-test (i.e., 0-3, 4-6, 7-12, and more than 12 months)
• Squared time between test indicator
• Interaction term with ACT Composite score and grouped time between test
• High school cumulative GPA category (i.e., 0.00 – 2.50, 2.51 – 3.50, over 3.50)
• Planning to take physics before graduation
• Planning to take calculus before graduation
• Plans for taking one or more accelerated, honors, or Advanced Placement courses before graduation

Method

With Composite gain scores grouped into categories of no score gain (which includes negative prediction) and gains of one, two, and three or more points, predictive modeling required incorporation of this variable's ordinal nature. Gains exceeding three points were grouped into the highest category because such gains occur with low frequency. As such, the relationship between various predictors and the aforementioned gain categories was modeled using ordinal logistic regression. These predictors were selected after balancing educational relevance and model parsimony with changes in Akaike Information Criterion

(AIC), pseudo R^2, and percent concordant/discordant pair classification. While each evaluator criterion has limitations, a parsimonious model on which the criteria aligned was selected. High-level details of this selection are discussed in the Results section below, with finer-distilled detail shared in the technical appendix.

With model coefficients in hand, a dataset was created containing all possible combinations of the predictor variables and associated probabilities of no gain and gains of one, two, and three or more points. This dataset serves as the source lookup table for the web app. In building the web app, an adjustment for the top end of the score scale was made. Because the models didn't control for maximum values, gains corresponding to scores above 36 (beyond the score scale maximum) can result. To keep predictions meaningful, out-of-bounds predictions are omitted from the graph. For instance, students with a score of 35 will not see chances of gaining two or more points.

Results

As referenced above, multiple models were considered for use with the score gain prediction tool, and a balance was struck between parsimony and accurate prediction. Model accuracy was determined using three criteria. After arriving at a superset of theory-based possible predictors to include in the model, a final set was selected reflecting a model that had (a) the lowest AIC value; (b) a pseudo R^2 among the highest available (0.04); and (c) among the highest/lowest percentages of concordant/discordant pairs (58%/41%). Each of these selection criteria has its own associated limitations, but all three converged on selecting the final model while keeping predictors focused on student achievement and academic behavior.

When considering indicators like the pseudo R^2 and concordant pairs, one would hope to see numbers higher than those observed here. However, the model is predicting gain scores. As such, predicted values contain an abundance of random error simply by virtue of reflecting differences, thereby resulting in lower values for these indicators. Regardless, the model helps students more accurately predict their chances of Composite score gains upon retest than basing prediction purely on chance.

Beyond overall descriptions of model efficacy, the importance of each predictor had to be considered. And, because multiple models were evaluated, the family-wise Type-I error rate had to be monitored. In the final model, all predictors and intercepts were statistically significant at $p<0.0001$, thereby negating Type-I error rate concerns. Among these predictors, the three with the highest logistic pseudo partial correlation (LPPC) were GPA ($r=0.058$), first Composite score ($r=0.051$), and grade of first test ($r=0.046$). As LPPC values approach 1 or -1, they indicate stronger relationships between the variable and the outcome. Though LPPC can be impacted by small or moderate N-counts, this limitation was not an issue given this study's large N-count.

Interpretation of these parameters can be approached using a wide range of specificity. In the end, however, there is a practical interpretation that can be applied. First off, high school GPA range had a negative coefficient (-0.345), such that moving up one level in GPA translates into a lower probability of no score gain. Conversely, first ACT Composite score and grade at time of first test (coefficients of 0.080 and 0.336, respectively) correspond with increased probabilities of no score gain.

Using the Model

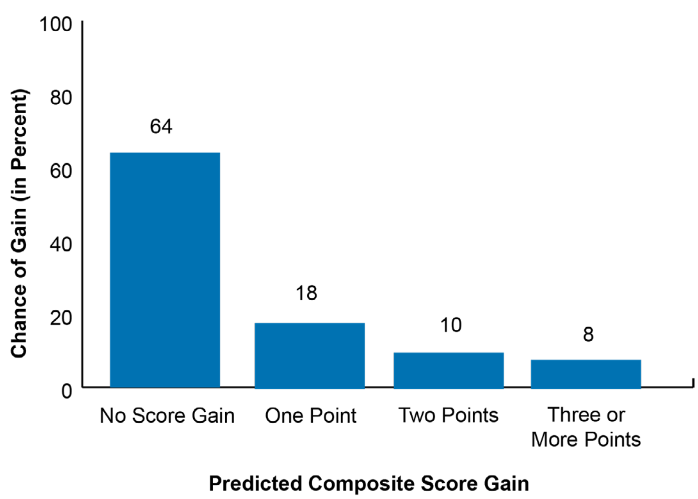
For the purposes of the web application, outcomes from parameter estimates/predictor values inserted into the model, including intercepts specific to each level, were converted into cumulative probabilities (CPs) of each gain level occurring. With cumulative probabilities in hand, differences between those in adjacent levels produced probabilities of attaining a specific gain level. For instance, the probability of no gain was equal to that level's cumulative probability. The probability of gaining two points was obtained by subtracting the CP for a gain of one point from that for a gain of two points. Similarly, the probability of gaining three or more points was derived by subtracting the CP of gaining two points from 1.00.

After creating a dataset containing all unique combinations of predictor values and gain levels, the probabilities for each gain level were computed and appended to the source lookup table. It is this table that is used as a lookup table in the web application.

Using the Tool

When accessing the **prediction web app**, a student's most recent ACT Composite score and subsequent fields are specified via dropdown menus. As each variable is entered, the tool refreshes with the percent chance of no score gain and gains of one, two, or three or more points (see Figure 1).

Figure 1. Web Application Predictions Using Default Values.



As seen in Figure 1, entering the tool presents the user with default settings. Here, the default parameters correspond with a 64 percent chance of no Composite score gain. Conversely, the chance of gaining one, two, or three (or more) Composite points is 18, 10, and 8 percent, respectively. The outcomes are reiterated in subsequent interpretive text, which also summarizes the outcome in terms of there being a 64 percent chance of no score gain juxtaposed with a 36 percent chance of gaining one or more points.

Conclusion

Whenever the question of taking the ACT a second time arises, the fact that a subsequent Composite score may go up or may go down must be taken into account. This is a risk that comes at the cost of sitting for another test session. Predictions of Composite score gain found in this paper's associated web application are meant to aid in the decision to retest, both from the perspective of what a student's chance of Composite score gain might be and by increased understanding of some of the factors associated with score gains. As part of that decision-making process, it is helpful to note that students can enhance their chances of increasing subsequent scores through the use of free test preparation tools such as **ACT Academy** (Payne & Allen, 2019). Beginning in September 2020, ACT will further enhance the retest experience by providing opportunities for students to take only specific sections upon retest. Updates to this tool are anticipated as such data becomes available.

References

ACT. (2017). *ACT technical Manual*. Iowa City, IA: ACT. Retrieved from http://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf

Camara, W., & Allen, J. (2017). *Does testing date impact student scores on the ACT*? Iowa City, IA: ACT.

Doyle, W. R. (2006). Adoption of merit-based student grant programs: An event history analysis. *Educational Evaluation and Policy Analysis, 28*(3), 259-285.

Lievens, F., Reeve, C. L., & Heggestad, E. D. (2007). An examination of psychometric bias due to retesting on cognitive ability tests in selection settings. *Journal of Applied Psychology, 92*(6), 1672–1682.

Payne, J. S., & Allen, J. (2019) *An early look at ACT Academy usage and effectiveness.* Iowa City, IA: ACT.

Technical Appendix

Cohort details are captured in Table 2 below. While some of the student groups are broken out fairly equitably (e.g., Plans to Take Calculus), that is not always the case. In these extreme instances, substantive reasoning justified how students were distributed. For instance, only three percent of the cohort waited more than 12 months between testing. Though a seeming outlier, this accurately depicts ACT testing practices, as over half of these students first tested in their sophomore year, which represents a small percentage of any graduating class. Similarly, only 5% of students fell into the GPA category of 0.00 to 2.50. This was expected because data cleansing was designed to ensure only potentially college-bound students would be included, and such students are less likely to report GPAs of 2.50 and below.

Table 2. Cohort Details

			First Composite Score		Second Composite Score	
Student Cohort	**N**	**PCT**	**Avg**	**STD**	**Avg**	**STD**
All Students	677,315	100	21.9	5.1	22.8	5.3
Score Change Group						
No Gain	288,131	43	22.1	5.1	21.3	5.1
One Point	149,490	22	22.0	5.2	23.0	5.2
Two Points	117,232	17	21.8	5.1	23.8	5.1
Three or More Points	122,462	18	21.3	4.8	25.1	4.8
Grade at Time of First Testing						
10	96,267	14	22.3	4.6	23.8	4.9
11	537,629	79	22.0	5.1	22.8	5.3
12	43,419	6	19.9	5.0	20.6	5.1
Months Between First and Second Test						
Zero to Three	222,692	33	22.9	5.0	23.6	5.2
Four to Six	238,245	35	22.1	5.0	22.9	5.2
Seven to 12	192,925	28	20.8	4.9	21.7	5.3
More than 12	23,453	3	20.1	5.0	21.7	5.5
High School GPA Range						
0.00 to 2.50	34,791	5	16.3	3.2	16.8	3.4
2.51 to 3.50	257,151	38	19.5	4.2	20.2	4.3
3.51 or Higher	385,373	57	24.0	4.6	25.0	4.8
Plans to Take Physics						
No Plans to Take	232,717	34	20.2	4.5	21.0	4.7
Have Taken/Plan to Take	444,598	66	22.8	5.1	23.7	5.3
Plans to Take Calculus						
No Plans to Take	317,179	47	19.9	4.2	20.7	4.4
Have Taken/Plan to Take	360,136	53	23.6	5.1	24.6	5.3
Plans to Take Accelerated, Honors, or Advanced Placement Courses						
No Plans to Take	216,713	32	19.0	4.3	19.7	4.6
Have Taken/Plan to Take	460,602	68	23.3	4.8	24.2	5.0

Beyond distributions, another seeming anomaly was the group demonstrating no gains. Here, the difference between average Composite scores for first and second testing is -0.8 points. It is important to remember that not everyone who tests more than once has a score increase. Thus, while ACT recommends that students test more than once, there are no guarantees that a second test will result in a higher score.

In Table 3 below, model outcomes from the ordinal logistic regression analyses are tabled.

Table 3. OLR Results

Variable	Parameter Est.	Std. Err.	Wald Chi-Sq.	Prob. Chi-Sq.	Standardized Est.	Logistic Pseudo Partial Correlation
Intercept 1	-4.61035	0.06970	4376.5140	<.0001	_	0.050
Intercept 2	-3.68080	0.06960	2796.3220	<.0001	_	0.040
Intercept 3	-2.74750	0.06960	1560.4400	<.0001	_	0.030
First Composite Score	**0.07979**	**0.00118**	**4569.9210**	**<.0001**	**0.2225**	**0.051**
Grade at First Testing	**0.33568**	**0.00551**	**3717.5740**	**<.0001**	**0.0828**	**0.046**
Time Between Tests	0.45564	0.01860	598.6006	<.0001	0.2178	0.018
First Comp x Time Between	-0.01163	0.00052	505.3765	<.0001	-0.1293	0.017
$(\text{Time Between})^2$	-0.06849	0.00306	501.3065	<.0001	-0.1430	0.017
HS GPA	**-0.34498**	**0.00449**	**5892.3710**	**<.0001**	**-0.1129**	**0.058**
Physics Plans	-0.12734	0.00510	624.0246	<.0001	-0.0333	0.019
Calculus Plans	-0.19286	0.00513	1416.0890	<.0001	-0.0531	0.028
Acc/Adv/AP Course Plans	-0.21315	0.00531	1611.5910	<.0001	-0.0548	0.030

Note. Items with highest logistic pseudo partial correlations (LPPCs) are presented in bold.

Highlights from this table are discussed in the text above, and those items with highest logistic pseudo partial correlations (LPPCs) are presented in bold font. These LPPCs aren't very high, though the full model produces 58% concordant pairs. Via multiple indicators of model efficacy, the use of this model by the web application contributes important information to students' retake decision-making processes.

Matt Harmston, MA

Matt Harmston is a senior research scientist in Research Technology, Data Science, and Analytics specializing in secondary research and report development.