Facts and Fantasies Regarding Admission Standards

By Ted Micceri, Ph.D.

Budgets, Human Resources and Information Technology
Budgets and Policy Analysis
Office of Institutional Research

Paper Presented at the AIR Forum 2001
Long Beach, CA, June 3-6
Facts and Fantasies Regarding Admission Standards

Summary

This study sought to determine which, if any, of the possible predictor variables available at freshman entry would prove useful in selecting students for admission based on relationships with college performance. A study of relationships between four college performance variables and eight admission variables was conducted using seven freshmen cohorts (summer and fall) at the University of South Florida (USF). Supporting an extensive literature, meaningful relationships with all four outcomes occurred for both high school G.P.A. and class rank. Although simple relationships with outcome variables occurred for other variables such as sex, race/ethnicity and test scores, most of their predictive capacity was already included in the G.P.A.-based measures, and they therefore add little to predictions. Even the very strongest relationships between predictors and outcomes fail to identify meaningful performance differences between any two adjacent scale points. Using a standardized 15-point scale (each scale point includes roughly 7% of all students) across G.P.A., Class Rank, ACT and SAT scores, attrition differences between adjacent scale points at the maximum, were about 2.4% (that between high school G.P.A. and first year grades) and with meaningful outcomes (e.g. graduation, retention to second year) were generally between 0.5% and 2%. Even avoiding the issue of the measurement error inherent in all such measures, these small differences in outcomes from scale point to scale point show the fantasy of setting specific admission cut-offs. When one sets a cut-off, no difference in outcomes will occur for the students immediately above and below that cut-off one in at least 97.6% of cases and almost always in 98-99% of cases. Mortenson (1999) indicates that affluence relates far more with graduation and retention, and could reduce the incorrect decision rate from 98% to perhaps 70% of rejected applicants.

Introduction and Background

Because of recent increases in FTIC enrollment and projected future increases, and to provide improved information for admissions decisions, this study sought to determine what short- and long-term performance and attrition effects relate to student characteristics that are available to inform selection during the admissions process. 1

An extensive literature on predictors of college success exists. Generally, the research indicates that high school grades are the best predictor, with such as race/ethnicity, sex, and standardized test scores providing some information. Test scores are commonly used as an admission requirement, however, the literature provides very limited support for their use. Elert (1992) summarizes the findings of many studies by stating that overall, previous grades are about twice as good as the SAT at predicting first semester grades.2 The main justification for requiring the tests for admissions is that although the SAT is an inferior predictor relative to high school grades, it can increase the accuracy of prediction when used in combination with them. However, research indicates that inclusion of the SAT increases early grade prediction by an average of only 5%. The major reason that the benefits are so low is that the SAT provides redundant information. He notes:

Marginal as they are, the predictions of first year grades are the test’s most accurate forecasts. Correlations between scores and grades in later years, and

---

1 Note that full-time enrollment and particularly full-time completion of hours (a proxy for affluence) is about as strong a predictor of outcomes as high school G.P.A., but is not available as a selection criterion. This is true for Graduate Students also.

2 Even test makers themselves do not claim that standardized tests measure either achievement or school outcomes (Bracey, 1997). Test makers such as ETS and ACT do claim a low-level relationship with first semester college grades (Murphy, 2000; Elert, 1992).
overall college average, are lower still. One study found that the ability of college admission tests to predict grades declined consistently from one semester to the next throughout eight semesters (Humphreys). The virtual disappearance of the aptitude tests' ability to predict beyond the freshman year has been explained by some commentators as a result of the nature of advanced study. Multiple choice testing predominates introductory courses, they argue, but intermediate and advanced courses demand a broader range of performance.

The Florida State University System (Florida SUS, 1995) conducted extensive research on traditional predictors of college success across the then nine SUS institutions, and concluded:

The greatest single predictor of success in College is High School Grades.

Two methods are most frequently used to quantify high school performance: High School G.P.A. and class standing or rank. Class standing is reported as a standard more often than is G.P.A.

To give some idea of how extensive research conducted on standardized tests such as the SAT and ACT is, one of the highest relationships in the testing literature between the SAT and anything other than another test is a correlation of $r = .66$ reported by French & John (1967) between SAT scores and uric acid levels in the blood. Such ETS-sponsored studies clearly show just how widely ETS has searched for something their product predicts.

Methods

The following predictor variables (admissions characteristics) were submitted to analysis (Acronym used for each specific variable is in parenthesis): (1) High School G.P.A. (HS G.P.A.); (2) Graduating Class Rank (Rank) – Percentile Rank in High School Graduating Class (available for cohorts: 92 to 99); (3) Test Scores– Either the ACT or the combined SAT; (4) Minimum Test Scores – Scores at or below 440 on either the Quantitative or Verbal portion of the SAT and below 18 on the ACT; (5) Alternative Admits – Students who were either admitted below the BOR Sliding Scale or who lacked adequate academic credits for admission. (6) Race/Ethnicity; (7) Sex; (8) Age.3

The following outcome variables were submitted to analyses: first semester G.P.A., first year G.P.A., returning for the second year, long-term retention/graduation and 1-7 year graduation. For regression purposes, the retention/graduation variable was expanded to three ranks, with 0=lost in 1st year, 1=retained two or more years, and 2=graduated.

Variable Sources – All data were derived from official USF or BOR sources.

Sample – The sample included all fall 1991 to fall 1999, summer and fall freshmen (FTIC).

Analysis Issues: Parametric (OLS) correlations (Pearson $r$) consistently showed lower relationships between USF G.P.A. and incoming items (HS G.P.A. TEST Scores, etc.) than non-parametric (Spearman Ranks). Therefore, when attempting to determine whether additive effects occurred among any of the variables, I therefore ran backward elimination, robust regression analyses using rank transformed data.

Limitations - Numbers reported in this study will not specifically match with those of any official IR reports, because only students listed in both the Admissions Files and Student

---

3 Note that in this report, only measures 1, 2, 3, 6 and 7 are reported. For a more complete copy of the study, contact Ted Micceri: tmicceri@admin.usf.edu.
Data Course Files (SDCF) for a given semester were retained for analysis. Recent cohort data for both transfers and stop-outs are incomplete due to time limitations.

Results

Sample Characteristics
Because approximately 97% of USF’s average annual new FTICs enroll in either the fall or the summer semester, this study included data from those two sources. The sample involved 21,138 students of whom 12.8% matriculated in summer semesters.

When looking at combined summer and fall cohorts the percentages of females and various racial/ethnic groups have remained comparatively steady since about 1993, with females representing about 58% of the average cohort, and whites about 65%. In the average cohort, about 93% of the students were either 18 or 19 years old.

Differences occurred between semesters with the following consistent patterns occurring in the summer: greater percentages of females enroll, smaller percentages of whites.

Regarding scores on high school academic variables, fall cohorts had consistently higher scores, while comparatively complex and interactive relationships tended to occur between sexes and among racial/ethnic groups. Whites, Asians and Males score higher on tests (ACT, SAT). Whites, Asians and Females score higher on G.P.A. variables (e.g. Class Rank).

Regarding course loads, the average student in the summer semester took two courses, compared with four for those in the fall. Fall students average a drop of 16% between hours attempted and hours completed, compared with only 3% for the summer cohort.

Method Note – Two approaches used to evaluate possible relationships between predictor and outcome measures are reported here: (1) Simple relationships were analyzed using Correlations and Equivalency Scores, and (2) Complex Relationships were evaluated using robust Multiple Regression techniques.

Simple Relationships - Correlations
Table 1 shows median Spearman $r$ correlations\(^4\) between four academic variables from high school and four USF outcomes for cohorts from 1991 to 1997\(^5\). Clear findings include:

- None of the relationships are particularly strong (e.g. maximum correlation is .45 between high school G.P.A. and first year USF G.P.A.).
- Correlations of all entry variables with short-term variables are substantially larger than are correlations with long-term variables.
- All variables show greater correlation with first year G.P.A. than with first semester G.P.A.
- All variables show low simple correlations with Graduation/Retention variables.
- High School Performance (G.P.A./Rank) shows a greater relationship than do test scores with all performance variables.
- Correlations between the SAT and ACT test and first semester G.P.A. are consistent with historical relationships from other universities, which occasionally reach $r = .20$\(^6\).

---

\(^4\) These were higher than Pearson correlations.
\(^5\) 1998 and 1999 lacked adequate performance data.
Table 1
Summary Relationships Between High School Source Variables and Performance at USF

<table>
<thead>
<tr>
<th></th>
<th>HS G.P.A.</th>
<th>HS Class Rank</th>
<th>SAT</th>
<th>ACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>first Semester G.P.A.</td>
<td>0.33</td>
<td>0.31</td>
<td>0.19</td>
<td>0.15</td>
</tr>
<tr>
<td>first Year G.P.A.</td>
<td>0.45</td>
<td>0.41</td>
<td>0.26</td>
<td>0.30</td>
</tr>
<tr>
<td>Graduation</td>
<td>0.18</td>
<td>0.18</td>
<td>0.11</td>
<td>0.13</td>
</tr>
<tr>
<td>Graduation/Retention</td>
<td>0.20</td>
<td>0.19</td>
<td>0.12</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Simple Relationships - Equivalencies
In order to directly compare the four academic source variables with outcomes, 15-point equivalency scores with each scale point representing 7% of USF’s population were created for each of the predictor measures based on over 15,000 historical students, as shown in Table 2. An equivalency score of 3 indicates that approximately 20-22% of all USF FTIC enrollees scored a lower value, and each score represents approximately the same percentage for each of the four source variables (HS G.P.A., HS Class Rank, ACT, SAT).

Table 2

<table>
<thead>
<tr>
<th>Scale</th>
<th>Target Percntl</th>
<th>HS G.P.A.</th>
<th>Nearest Percntl</th>
<th>Class Rank</th>
<th>Nearest Percntl</th>
<th>SAT</th>
<th>Nearest Percntl</th>
<th>ACT</th>
<th>Nearest Percntl</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-7</td>
<td>&lt; 2.33</td>
<td>&lt; 40</td>
<td>&lt; 800</td>
<td>&lt; 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>2.33</td>
<td>5.6</td>
<td>40.0</td>
<td>6.7</td>
<td>800</td>
<td>6.5</td>
<td>16</td>
<td>6.8</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>2.50</td>
<td>12.9</td>
<td>51.0</td>
<td>13.7</td>
<td>864</td>
<td>14.5</td>
<td>18</td>
<td>17.6</td>
</tr>
<tr>
<td>4</td>
<td>28</td>
<td>2.75</td>
<td>22.2</td>
<td>58.0</td>
<td>20.6</td>
<td>908</td>
<td>21.3</td>
<td>19</td>
<td>24.3</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>2.85</td>
<td>27.3</td>
<td>63.0</td>
<td>26.6</td>
<td>930</td>
<td>28.3</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>42</td>
<td>2.99</td>
<td>34.6</td>
<td>69.0</td>
<td>35.1</td>
<td>960</td>
<td>35.8</td>
<td>20</td>
<td>33.2</td>
</tr>
<tr>
<td>7</td>
<td>49</td>
<td>3.10</td>
<td>44.6</td>
<td>73.0</td>
<td>42.1</td>
<td>992</td>
<td>43.5</td>
<td>21</td>
<td>43.6</td>
</tr>
<tr>
<td>8</td>
<td>56</td>
<td>3.20</td>
<td>50.4</td>
<td>77.0</td>
<td>49.3</td>
<td>1024</td>
<td>51.7</td>
<td>22</td>
<td>53.8</td>
</tr>
<tr>
<td>9</td>
<td>63</td>
<td>3.30</td>
<td>56.2</td>
<td>80.0</td>
<td>55.5</td>
<td>1040</td>
<td>55.2</td>
<td>23</td>
<td>62.4</td>
</tr>
<tr>
<td>10</td>
<td>70</td>
<td>3.50</td>
<td>66.1</td>
<td>84.0</td>
<td>64.0</td>
<td>1072</td>
<td>62.4</td>
<td>24</td>
<td>70.4</td>
</tr>
<tr>
<td>11</td>
<td>76</td>
<td>3.60</td>
<td>70.7</td>
<td>87.0</td>
<td>70.9</td>
<td>1104</td>
<td>69.7</td>
<td>25</td>
<td>76.5</td>
</tr>
<tr>
<td>12</td>
<td>82</td>
<td>3.74</td>
<td>75.0</td>
<td>89.0</td>
<td>75.0</td>
<td>1136</td>
<td>75.8</td>
<td>26</td>
<td>82.8</td>
</tr>
<tr>
<td>13</td>
<td>88</td>
<td>3.96</td>
<td>82.8</td>
<td>92.0</td>
<td>82.4</td>
<td>1184</td>
<td>83.2</td>
<td>27</td>
<td>88.1</td>
</tr>
<tr>
<td>14</td>
<td>94</td>
<td>4.14</td>
<td>89.2</td>
<td>95.0</td>
<td>89.2</td>
<td>1232</td>
<td>88.8</td>
<td>28</td>
<td>91.7</td>
</tr>
<tr>
<td>15</td>
<td>95-100</td>
<td>4.30</td>
<td>94.1</td>
<td>97.0</td>
<td>93.4</td>
<td>1296</td>
<td>94.2</td>
<td>29</td>
<td>94.6</td>
</tr>
</tbody>
</table>

Outcomes were defined as percentages of populations at a scale point accomplishing an outcome. If a variable is a good predictor of performance, then the percentage accomplishing a specific task (e.g. attaining a first year G.P.A. of 2.0 or above) should be less for the lower scale points (e.g. 1-7) than for the higher ones (e.g. 8-15). The best predictor variables will exhibit a monotonic increase (steady increase from lower scale point to next higher scale point) in percentages accomplishing an outcome (e.g. returning for the second year). To better show trends, all percentages represent smoothed estimates computed by averaging a scale point’s percentage with the percentage above and below. This smoothing technique eliminates idiosyncratic fluctuations that tend to occur for

Note that simple correlations between test scores and first year GPA tended to be more highly positive for whites than for all, but zero-order to negative for each minority group when analyzed separately. This almost surely reflects language bias against minorities.
specific scale points and provides a far clearer picture of trends. The sample sizes used in these comparisons range from 9,000 for the ACT (600 per scale point) to 15,800 for G.P.A. (1,050 per scale point).

**Short-Term - First Year G.P.A. of 2.0 or higher**

Figure 1 shows that a comparatively monotonic relationship occurs between both High School class rank and High School G.P.A. and attaining a USF G.P.A. of 2.0 or higher during one’s first year. Each point on the chart reflects the percentage of students between 1991 and 1997 who accomplished a first-year G.P.A. of at least 2.0 and who scored at that point on G.P.A. or Rank. Scale point 8 represents a 3.2 high school G.P.A. Percentages attaining a 2.0 increase quite steadily from scale point 1 (G.P.A. < 2.33, Rank < 77) to scale point 15 (G.P.A. >= 4.30, Rank >= 97). The bottom panel shows that a far flatter trend line occurs for the tests. A very gradual upward trend occurs from scale point 1 (ACT < 16, SAT < 800) to scale point 5 (ACT = 19, SAT = 940). The trend is then flat until scale point 9 (ACT = 23, SAT = 1040) after which a somewhat steeper increase occurs. Note that the steepest segments for tests are less steep than the flattest segments for G.P.A. variables.

![Figure 1](image_url)

**Figure 1**

2.0 or Higher G.P.A. During First Year and Equivalent Scores – 1991-1997 Cohorts
Short-Term - Returning for Second Year at USF

Figure 2 shows that a comparatively monotonic relationship occurs between both High School class rank and HS G.P.A. and returning to USF for a second year. For Rank, the increase is very small, but steady from scale point 1 to scale point 15, while the G.P.A. variable shows a dip between points 5 and 9. The bottom panel shows that no such increase occurs for either of tests used, but instead an almost flat relationship with values of 1 being near 80% returning, values of 11-12 being at or about that point, and values of 14-15 being only three to four percentage points above.
Long-Term - Retention (includes Graduation, Transfer and Stop-out)

Figure 3 shows that a low-level, but comparatively monotonic relationship occurs between both High School class rank and HS G.P.A. and retention/graduation at USF over the long term. The increase is small, but steady from scale point 1 to scale point 15 for Rank, while G.P.A. shows flatness between points 3 and 8. The bottom panel of Figure 3 shows that no such increase occurs for either of tests. Again, as in Figure 2 approximately the same percentage of students score at scale points 1-5 either graduate or return, as do those with very high scores on the tests. Note that the curves for Figure 2 and Figure 3 are quite similar.

Figure 3
Relationship Retention/Graduation and Equivalent Scores
Long-Term - 5-7 Year Graduation Rates
Because only earlier cohorts (1991 to 1993) may be used for these analyses, sample sizes are more limited ranging in size from 3,500 for Class Rank to 6,500 for High School GPA. This is a concern at USF because USF’s graduation rates increase steadily through the eighth year after matriculation. These data, therefore, reflect what one might term short-term graduation rates for a Metropolitan Institution like USF (see Blumberg, et. al, 1997). Figure 4 shows that both GPA variables again show a comparatively monotonic relationship with degree attainment, with the lowest scores graduating at a 28% rate and the highest at a 61-67% rate (Rank is highest and lowest). Test scores also show a relationship with degree attainment, ranging from low score to high score of 31% to 60% for the ACT, and from 37% to 56% for the SAT. A comparatively steep upward trend begins at about scale point 10 for GPA, RANK and ACT (respectively, Rank=84, GPA=3.5, ACT=24). Note, however, that both tests show very flat results in the middle for ACT from 5 to 10, and at the lower end for SAT, from 1 (< 800) through 9 (9=1040).

Figure 4
Relationship Between 5-7 Year Degree Attainment and Equivalent Scores
Complex Relationships – Multiple Regressions Findings

Robust, backward-elimination, rank-transform multiple regression analyses were conducted to determine whether relationships between predictor variables and USF performance might be additive. Five variables High School performance (G.P.A. or Class Rank), Test Scores (SAT or ACT), sex, race and age were regressed to First Year G.P.A., 5-7 Year Graduation and Graduation/Retention. The smallest sample size was 4,800 and the largest, 9,800. Figure 5 and Figure 6 show relative obtained $F$ values for each predictor variable in twelve different regressions. Each pair of columns represents a separate regression, one including SAT scores (lighter bars) and one including ACT scores (darker bars). Clearly, the most important variable for all predictions proved to be High School Performance, represented either by High School G.P.A. (Figure 5) or Class Rank (Figure 6). Only for First-year G.P.A. using Class Rank did either standardized test show a relationship of any relative magnitude (bottom panel of Figure 6). Note that the negative $F$ values shown for the SAT reflect negative parameter estimates since $F$ values are squared and cannot be negative. For the graduation/retention variables, the SAT showed negative relationships (parameter estimates) with the outcome variables. Sex and race tended to show as great, if not greater relationships as the tests in all cases except relationships with first-year G.P.A. (positive values for sex associate with being female, and for race, with being white).

Analyses conducted when both G.P.A. and Class Rank are entered into the regression model, showed two clear effects (The Ns were 6,300 for SAT and 4,850 for ACT; too few had SAT, ACT, G.P.A. and Rank to run regressions using these):

1. Class Rank dominates G.P.A. for long-term (Graduation/Retention) effects.

Of course, almost all of the variance of G.P.A. and Class Rank is shared. This is the reason when both are used in the same regression model, and one takes away its variance from the model, little is left for the second. Overall, the G.P.A.-based variables are clearly the strongest predictors of any outcomes of any variables available in admissions process, and these analyses show that the variables that may be legitimately used for admissions decisions are not additive (e.g. G.P.A., Rank and Tests, assuming that sex and race/ethnicity may not be used as criteria). Thus, as Elert (1992) and Astin (1996) note, tests add little, if any to the prediction of a student’s performance at USF.

---

7 Note that the use of $F$-values is not strictly legitimate when attempting to estimate the comparative magnitude of partial and semi-partial correlations. A far better indicator is $R^2$ changes, but I was using SAS for these purposes, and after seeing the great differences, simply did not take the time to run the many regressions required to estimate $R^2$ changes.
Figure 5

Figure 6

F Values from Regressions Using High School Class Rank – 1991-1997 Cohorts
Conclusions

The results of this study show that G.P.A.-based source variables (High School G.P.A. and Class Rank) are far more powerful predictors of all outcomes at USF than other available measures, none of which add useful information to prediction when G.P.A is considered.

One factor at USF is the substantial percentage of each cohort that transfers to another SUS institution (up to 20%). Such students were treated as retained in the Graduation/Retention outcome variable, which shows different relationships than simple graduation rates.

An Important Consideration

These analyses appear to suggest that one could legitimately use either High School G.P.A. or Class Rank as admittance criteria. However, the greatest simple percentage performance difference of any source variable with any outcome variable occurs between Class Rank and a first year G.P.A. above 2.0 (Figure 1). This success percentage difference ranged for 61% for the lowest scale point to 97% for the highest, and differences between any two scale points were relatively monotonic. Thus, the average difference between any two scale points was 36% divided by 15 = 2.4%. If a cut-off is set between any two scale points, one would be correct in making such a distinction for an average of 2.4% of those rejected and would be incorrect in making such a distinction for 97.6% of those rejected. That is, 2.4% more of the accepted would succeed than the rejected and this, where the greatest relationship occurs. For all other performance outcomes, the increase between scale points is smaller.

Mortensen (1999) shows that affluence is a far more efficient predictor of performance. Six-year graduation rates decline rapidly by income quartile: 70%-80% for the top quartile (most affluent 25% of the population), 22-27% for the third, about 15% for the second and about 5-7% for the bottom. If one sets a criterion of not accepting any student who applies for need-based financial aid, that should eliminate many students in the bottom two quartiles, which, of course, would make such cut-offs far more effective than those derived from other available measures, perhaps reducing the rejection error rate from 98% to 70%.

References

Astin, A. W. (1996). *Degree Attainment Rates at American Colleges and Universities: Effects of Race, Gender, and Institutional Type*. ERIC_NO: ED400749


