



**Inter-rater Reliability of Holistic Measures Used in the Freshman Admissions Process of
The University of Texas at Austin
Summer/Fall 2005**

Abstract

For applicants not automatically admitted via the Texas Top 10% Law, the freshman admissions process at The University of Texas at Austin considers two computed numerical values: The Academic Index (AI) and the Personal Achievement Index (PAI). The PAI is the result of an holistic review by readers of an applicant's resume and two required essays. This report is an investigation into the inter-rater reliability of PAI scores assigned by those readers.

The UT Austin Admissions Process

The admissions process of any university is an exercise in both selecting qualified students with a high probability of success, and crafting an entering class that also meets the university's mission. In 1996, as a result of the *Hopwood* case, UT Austin looked carefully at its admissions process. At the time, merit, when applied to an admission decision, consisted of high standardized test scores and high class rank. In 1997, the Texas State Legislature exerted its own definition of merit through the "Texas Top 10% Automatic Admissions Law." Today, any student graduating in the top 10% of his/her Texas high school class is automatically admissible to any public college or university in the state—including The University of Texas at Austin. From 1997 through 2002, automatically admitted students accounted for approximately one-half of all admits and entering freshmen at UT Austin. In 2003, the percentage jumped to nearly two-thirds of the entire class and approximately 70% of the entering class from Texas high schools.

The UT Austin admissions routine for students not automatically admitted is elaborate and entails a broad concept of merit. Beginning with the entering class of 1997, for those not automatically admitted, the idea of merit was expanded from class rank and test scores exclusively to the inclusion of the following factors:

The Academic Index (AI)

High School Record:

- Class rank
- Completion of UT required high school curriculum
- Extent to which students exceed the UT required units
- SAT/ACT score

The Personal Achievement Index (PAI)

- Scores on two essays
- Leadership
- Extracurricular Activities
- Awards/honors
- Work experience
- Service to school or community
- Special circumstances:
 - Socio-economic status of family
 - Single parent home
 - Language spoken at home
 - Family responsibilities
 - Socio-economic status of school attended
 - Average SAT/ACT of school attended in relation to student's own SAT/ACT
 - Race/Ethnicity (beginning with fall 2005)

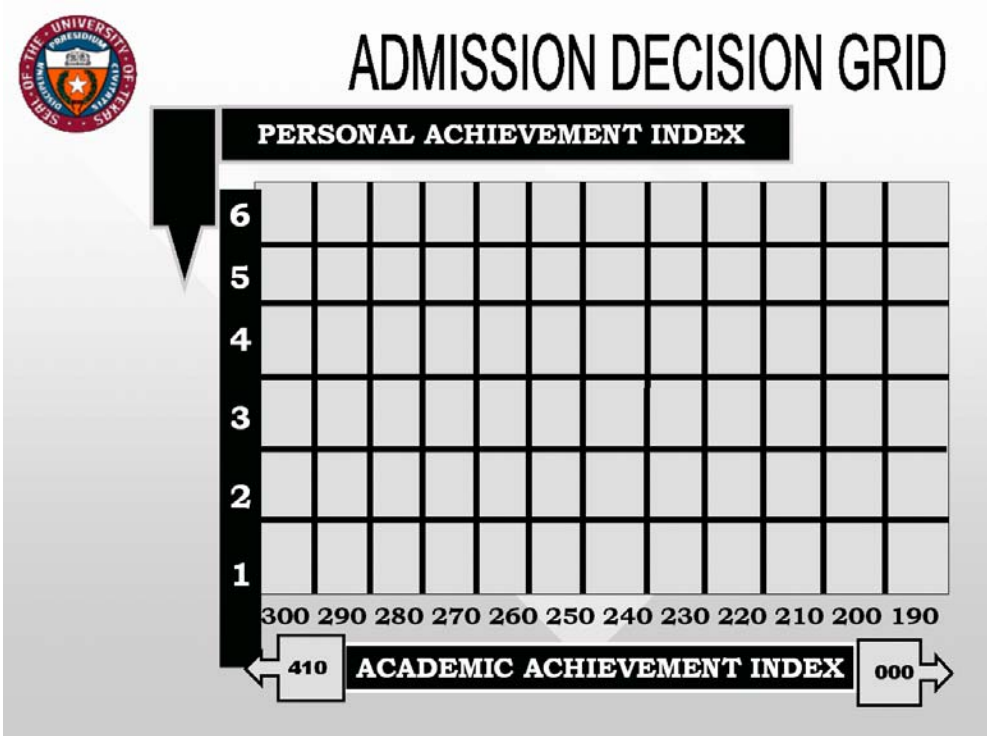
Thus, merit includes the ambition to tackle rigorous high school coursework, the production of quality prose, and the desire to make a difference in one's school, home, or community. Evidence of employability (work), and some sense of having excelled in any number of areas are also considered. Moreover, admissions officials place these attributes in the context of the circumstances under which the student lived.

The Academic Index (AI) is determined by a multiple regression equation utilizing a high school percentile derived from an explicit class rank $[1-(\text{class rank}/\text{class size})]*100$, and verbal and math test scores from the ACT Assessment or the SAT I: Reasoning Test. The equation produces a predicted freshman year grade point average. After a review of the high school transcript, an applicant can be "awarded" a tenth of a point if he/she exceeded UT's required high school curriculum. Thus, AI values range from 0.00-4.10.

The Personal Achievement Index (PAI) is UT Austin's holistic approach to admissions. Admissions officers are trained each year to conduct comprehensive reviews of every application from students not automatically admitted. All applicants are required to submit two essays. Each are read and scored on a scale of 1-6. The application itself, and any attachments an applicant chooses to include, is then reviewed. A "personal achievement" score on a scale of 1-6 is then assigned to the application. From the three scores, two essays and a personal achievement score, a PAI is computed. The equation reflects a 1997 faculty decision to give slightly more weight to the personal achievement score than the essays: $PAI = [(personal\ achievement\ score * 4) + (mean\ essay * 3)] / 7$.

AIs and PAIs of applicants not automatically admitted are then plotted on an admissions decision grid. (See Figure 1 below.) The most-qualified candidates are located in the cells closest to the upper left corner. Admissions liaisons, and/or representatives of Deans' offices or faculty, then make decisions as to which cells to select as admitted students. Applicants who are Texas residents are either admitted, "cascaded" to their second choice of major, offered Summer Freshman Admission, or offered the Coordinated Admission Program (CAP) at a UT System component school. Thus, Texas residents submitting a completed entering freshman application by published deadlines are not denied admission to UT Austin. All non-residents compete with one another in accordance with the routine described above and are either admitted or denied.

Figure 1



There are few exceptions to the routine described above. In some colleges there are more top 10% student applicants than available spaces. In those cases automatic admissions is limited to 80% of the admitted class. Thus, in some cases, top 10% students, while guaranteed a place at UT Austin, will nonetheless compete for a slot in their first-choice major. These students compete as described above.

Holistic Review and the Personal Achievement Index

Figure 2 below represents the score scale of the elements of the Personal Achievement Index.

Figure 2
Score Scale for the Personal Achievement Index
Since 1997

Personal Achievement Score	Essay Score
<p>Score of 6 A score in this category means that the applicant demonstrates superior achievement in terms of leadership, service, or special honors. This achievement can be demonstrated in any field: academics, athletics, the arts, or extracurricular activities. Or it can be demonstrated through extraordinary circumstances: working to help support a family, developing a life skill, or facing a personal crisis. The applicant's achievement shows an unusually high level of maturity, commitment, and perseverance.</p> <p>Score of 5 A score in this category means that the applicant demonstrates commendable achievement in terms of leadership, service, or special honors. This achievement can be demonstrated in any field or through unusual circumstances. The applicant's achievement shows a high level of maturity and commitment.</p> <p>Score of 4 A score in this category means that the applicant demonstrates creditable achievement in terms of leadership, service, or special honors. This achievement can be demonstrated in any field or through unique circumstances. The applicant's achievement shows a good level of maturity and commitment, though the achievement itself might not be extraordinary.</p> <p>Score of 3 A score in this category means that the applicant demonstrates slightly below average achievement in terms of leadership, service, or special honors. This achievement can be demonstrated in any field or through personal circumstances. The applicant's achievement shows some level of maturity but an inconsistent level of commitment. There may also be only slight evidence of meeting a challenge or pursuing an opportunity.</p> <p>Score of 2 A score in this category means that the applicant demonstrates below average achievement in terms of leadership, service, or special honors. The applicant's achievement shows some level of maturity, but the level of commitment is either short-term, superficial, or merely active participation. There may also be little evidence of meeting a challenge or pursuing opportunity.</p> <p>Score of 1 A score in this category means that the applicant demonstrates well below average achievement in terms of leadership, service, or special honors. The applicant's achievement shows little maturity or commitment, and there may be no evidence of interest beyond classroom requirements. There may also be no evidence of meeting a challenge or pursuing opportunity.</p>	<p>Score of 6 An essay in this category demonstrates <u>clear and consistent competence</u> though it may have occasional errors. Such an essay:</p> <ul style="list-style-type: none"> • effectively and insightfully addresses the writing task • is well organized and fully developed, using clearly appropriate examples to support ideas • displays consistent facility in the use of language, demonstrating variety in sentence structure and range of vocabulary <p>Score of 5 An essay in this category demonstrates <u>reasonably consistent competence</u> though it will have occasional errors or lapses in quality. Such an essay:</p> <ul style="list-style-type: none"> • effectively addresses the writing task • is generally well organized and adequately developed, using appropriate examples to support ideas • displays facility in the use of language, demonstrating some syntactic variety and range of vocabulary <p>Score of 4 An essay in this category demonstrates <u>adequate competence</u> with occasional errors and lapses in quality. Such an essay:</p> <ul style="list-style-type: none"> • addresses the writing task • is organized and somewhat developed, using examples to support ideas • presents minimal sentence variety <p>Score of 3 An essay in this category demonstrates <u>developing competence</u>. Such an essay may contain one or more of the following weaknesses:</p> <ul style="list-style-type: none"> • inadequate organization or development • inappropriate or insufficient details to support ideas • an accumulation of awkward expressions <p>Score of 2 An essay in this category demonstrates <u>some incompetence</u>. Such an essay is flawed by one or more of the following weaknesses:</p> <ul style="list-style-type: none"> • poor organization • thin development • little or inappropriate details to support ideas • frequent awkward expressions <p>Score of 1 An essay in this category demonstrates <u>incompetence</u>. Such an essay is seriously flawed by one or more of the following weaknesses:</p> <ul style="list-style-type: none"> • very poor organization • very thin development • expressions so awkward that meaning is somewhat obscured

Methodology

Each year all readers of admissions files are required to attend approximately two days of intense training. This is conducted by a university faculty member who is also nationally recognized as an expert in holistic reading and scoring. The training serves the purpose of calibrating scores to assure fairness and consistency. For this study the trainer was asked to select two examples of each personal achievement score, and three more examples selected at random. (For purposes of this report, the trainer's scores were considered true scores.) The resulting packet of fifteen (15) applications was then sent to all (n=44) of the readers of freshman applications. (The primary author of this report is a reader, but was excluded.) The scores were entered into a dataset for analysis.

Results

Inter-rater reliability is traditionally measured in three ways:

1. Exact Score Consistency is the frequency, expressed in percentage, of an exact score being assigned by different readers.
2. One-Point Score Consistency is the frequency in which readers were within one point of each other.
3. Pearson Correlations between the true score, in this case the trainer's score, and the raters' scores.

Table 1 below illustrates the first two inter-rater reliability measures described above. The resulting degrees of consistencies are similar to goals set by ACT, Inc. and the College Board for inter-rater reliability of the writing samples of their standardized tests.

Table 1
Consistency Measures Between Raters
2005 Study

	Exact Score Consistency	+/- One-Point Score Consistency
Personal Achievement Score	47%	88%
Essay Scores	51%	91%
Overall	49%	90%

Table 2 below examines the correlations between the raters and the trainer. The relationships are strong and consistent ($r=.80--.82$).

Table 2
Pearson Correlations Between Readers and Trainer
2005 Study

	Pearson Correlation
Personal Achievement Score	.82
Essay Scores	.80
Overall	.81

The results of this study represent considerable improvement from a similar study conducted by UT Student Affairs Research in 1997. Exact Score Consistency of essay readers at that time ranged from 35-42% and for the PAI it was 34%. The One-Point Score Consistency for the essays was 88-89% and 83% for the PAI.

The evidence is that inter-rater reliability for holistic measures in UT's admissions process is typical of industry standards, acceptably high, and has improved over time.