2013

Improving Retention and Fit by Honing an Honors Admissions Model

Patricia Joanne Smith
University of Central Arkansas

John Thomas Vitus Zagurski
University of Central Arkansas

Improving Retention and Fit by Honing an Honors Admissions Model

Patricia Joanne Smith and John Thomas Vitus Zagurski

University of Central Arkansas

For over a century, admissions officers and enrollment managers have relied on external validation of merit in selective admission of undergraduates. A main criterion used for selection is standardized testing, i.e., the SAT and ACT. Since these tests have been long-suspected and then shown to contain class and race biases while not accurately predicting retention (Banerji), the Schedler Honors College at the University of Central Arkansas (UCA) shifted to a holistic, multi-criterion selection process, de-emphasizing standardized tests, and then analyzed the outcomes. The statistical analysis served two goals. The first was to test whether variables in the admissions model, developed in 2007, predicted retention; the results led to changes in the weighting of variables for a revised rubric that we have used since 2010. The second goal was to improve enrollment of a more racially diverse population of students. Our findings demonstrated that most variables used in typical higher education admissions protocols did not accurately predict retention in the Schedler Honors College at UCA. Only one variable correlated to retention in honors, namely, high school grade point average (hsGPA). By increasing the value of hsGPA in the revamped selection rubric, UCA was able to increase rates of retention as well as diversity of incoming students.

Although the ACT and the SAT are widely accepted as indicators of college success by enrollment managers, the College Board states that standardized tests predict only 42% of academic success within the first year of college (Chenowith). Colleges nevertheless continue to base admissions and scholarship decisions on a test with this poor level of reliability. Gilroy claims that the ACT and SAT tests are one of the only ways that colleges can compare students from all over the world on a predetermined scale in a cost-effective manner. A key fallacy in this logic, however, is that the SAT and ACT were not meant to be used interchangeably (Syverson). The two tests measure different characteristics in students. The ACT measures mastery of basic high school material while the SAT tests abstract and critical thinking skills (Syverson).

Because of these inconsistencies, as well as concerns about bias in standardized tests, more than 800 institutions (including Texas Tech University, Central Bible College, Cambridge College, Texas Women's College, University of Arizona, and University of Memphis) have chosen to be test-free institutions, meaning that these colleges do not use the SAT or ACT in their admissions decisions (FairTest). Preliminary research conducted on institutions not using standardized testing has demonstrated that their selection methods have been just as effective (Banerji).

Using grade point average and class rank for selective admissions has its own problems: methodologies used to calculate hsGPA vary from school to school; neither grade point average nor class rank is standardized (Sadler, et. al.); and the scale for reporting hsGPA varies, with some high schools refusing to report class rank altogether. If high schools do not rank students, then the university bears the burden to understand hsGPA in context (Sadler et al.).

Honors programs and colleges with selective admissions typically rely on criteria used more generally in higher education, including standardized tests, despite the fact that honors education in the United States started as a reaction to excessive standardization. Frank Aydelotte, while serving as President at Swarthmore College, noticed that the education system was not challenging top students. Having been a Rhodes Scholar, he was familiar with the Oxford methodology, and used it to begin the first American honors program at Swarthmore College in 1922 (Rinn). Honors programs have broadened teaching and learning practices since then, largely because of shared information among participants in the National Collegiate Honors Council (NCHC). Founded in 1966, NCHC has contributed to the growth of honors education, and, in 2012, had nearly a thousand member institutions nationally and worldwide (NCHC).

Since the mid-1990s, the NCHC's Basic Characteristics of a Fully Developed Honors Program have called for "a clearly articulated set of admissions criteria (e.g., GPA, SAT score, a written essay, satisfactory progress, etc.) [that] identifies the targeted student population served" (Madden). Furthermore, the NCHC's Basic Characteristics of a Fully Developed Honors College, developed in the early 2000s, maintains that the honors unit should "exercise considerable control . . . over honors admissions," which may include a "separate application" process (Sederberg). Where honors administrators control their own admissions protocols, selection criteria, arguably, should be free of bias and, when possible, give the program the best measure of applicants' likelihood of success specifically as honors students.

Patricia Joanne Smith and John Thomas Vitus Zagurski

TRANSITION TO MULTI-CRITERION INDICATORS AT UCA

In 2005, the Schedler Honors College at UCA was using a somewhat typical admissions method to select 150 incoming freshman from approximately 500 applications. Then something unprecedented happened; the honors college suffered its greatest-ever first-semester attrition rate, with twelve students leaving honors by the end of the 2005 fall term (8%). This dropout rate triggered administrators to rethink how they had been admitting students. Exit interviews produced one common theme: students were leaving because they did not feel that the program was right for them. They simply "didn't fit."

The administration set out in spring 2006 to design a selection process that better measured "fit." After convening student focus groups and getting faculty feedback, the administration developed a list of characteristics to better identify a good fit for the honors college: intellectual promise, maturity, motivation, and initiative, all of which recommenders had been asked to discuss in their letters; interest in learning and willingness to "keep the conversation going"; resourcefulness and adventurousness without a constant need for right answers; willingness to talk about unusual topics; and a reaching toward new ways to conceptualize or verbalize thoughts. In light of this consensus, faculty were asked to describe specific traits of ideal honors students, to suggest how these traits would be demonstrated, and to rank the traits' order of importance. Faculty were far more concerned with writing, conversation skills, curiosity, and critical thinking than they were with the standardized measures of ACT and hsGPA. They were also more interested in knowing about students' leadership, service, and collaborative work than about their class rank. Once the faculty had identified and ranked these characteristics, the administration designed a process to measure them.

The administration began by examining the purpose of each part of the admissions process. Students applying in 2005 were asked to submit (1) demographic and background information, (2) a high school transcript with ACT scores, (3) a letter of recommendation from a high school counselor or teacher, (4) an essay describing their interest in the honors college, and (5) a paper written for a high school class. Several weaknesses were evident in this application packet. First, the essays expressing interest in honors echoed information advertised in recruiting brochures and the website. They wanted what they were told to want: scholarships, private rooms in an honors residence hall, small classes, and grants for study abroad. They wanted perks instead of scholarship, citizenship and leadership. A second weakness was that papers written for high school classes varied widely and revealed too little about writing strengths. Reviewers of recommendation letters searched for (rare) red flags to deny admissions rather than assessment of admissibility. In addition to

the application materials, interviews with faculty were part of the admissions process, but they served more as an orientation role than a screening. Given that enrollment was capped, demand for admissions relative to supply of openings was so great that, without more consistent data to go on, ACT played too great a role in selection so that students with a 30 or above ACT score were virtually guaranteed admission. This reliance on ACT as a primary selection criterion resulted in limited diversity of the honors student population.

Aiming to redesign the admissions process for the entering class of 2007, the administration addressed all the concerns that had emerged. The first step, focused on student writing, was asking students not for one new essay and a previously written high school paper but for two new essays, each with a specific purpose. One essay question asked students to read an excerpt from Peter Elbow's 1973 essay "The Doubting Game and the Believing Game," which contrasts ways of knowing. The responses allowed reviewers to assess writing skill, reading comprehension, and critical thinking. The second essay focused on students' interest in honors and did not just ask what appealed to them about honors at UCA but instead led them to write about being leaders in the "public square." Faculty, who had not only been uninterested in perks but had found them disturbing as part of a culture of entitlement, now read essays where students examined the honors website and specifically discussed how an honors education at UCA could affect their leadership, civic service, and collaborative work for a cause important to them. Faculty wanted to cultivate in the admissions process a culture that reflected the values embedded in the program's mission, specifically the integration of self-reflection, civic engagement, and scholarship. The goal was to admit students who grasped and desired this experience.

Assessment of student writing became part of a larger rubric where, along with hsGPA and class rank, it had greater value than in the past. Students who were screened with the new rubric and received the highest scores were then invited to interview. The interview process also changed. Rather than serving merely to orient, the new structure evaluated students *in situ* as they participated in a mock small-group discussion similar to a freshman class experience. This evaluation was separated from orientation, which preceded it. Applicants who passed initial screening spent an "Inform and Interview Day," accompanied by their parents, with honors faculty and administrators. The "inform" part began with a presentation of the vision and the mission of the Schedler Honors College followed by a discussion of the "nuts-and-bolts" of the curriculum, living-learning community, scholarships, and grant opportunities for travel abroad and undergraduate research. Prospects then met with a panel of current students for a "Q & A" about the experience of honors education. After lunch in the cafeteria, parents met with administrators

of financial aid, registrar services, admissions, student services, housing, and campus police while students moved to the "interview" part of their day. They attended a large-group lecture related to the Elbow Essay and delivered by an honors faculty member, and then they participated in small group discussions with a faculty member and three or four fellow applicants. This large-group/small-group experience presented a model of the freshman year curriculum and allowed each prospect to have a day in the life of an honors student.

Between the large-group lecture and the small-group discussion, students were given a post-lecture writing prompt and a few minutes to elaborate their new understanding of the topic about which they had already written and had also just been lectured. The on-site writing served two purposes: providing the admissions committee a sample of the student's unpolished writing and giving introverted students time to gather their thoughts before the discussion. The on-site essay and participation in the discussion were then scored as part of the total rubric in the final assessment.

Applicant scores were calculated using quantitative measures of ACT, hsGPA, class rank, transcript evaluation, letter of recommendation, and additional quantitative assessments of the student's Elbow essay, honors interest essay, on-site essay, and small-group interaction (see Appendices A and B). Applicants were ranked by quantitative measures as well as by a qualitative assessment of their writing and small-group interaction. In order to test the rubric's general comprehensiveness, the honors administration strategically added an overall qualitative assessment to see if the items being evaluated and their weights matched the general impression faculty were getting of students. If students with lower scores had faculty advocating their admission, then something might be missing from the rubric.

The honors college used this process for three years and then, in fall 2010, examined two research questions: (1) Is there inter-rater reliability between the qualitative impression and the quantitative scores? (2) Do items being evaluated predict retention (the operational definition of "fit")? Results were then used to adjust the selection items and rubrics to better predict retention.

PROCEDURES

The subjects of analysis were students admitted to the UCA Honors College in spring 2007, 2008, and 2009 through the redesigned admissions process, producing a sample size of 352. The prediction (independent) variables were ACT composite score, raw high school grade point average, participation in college preparation curriculum, participation in advanced placement courses, letters of recommendation, writing assessment, small-group interaction during the campus visit, and overall impression. Outcome (dependent) variables were first-semester college GPA, second-semester college GPA, and retention in the

program at the end of the first year and second years. Context variables of race and age were also examined.

Raw ACT scores ranged from 19–35 with a mean of 29 and a standard deviation of approximately 1.0. Raw high school grade point average (HsGPA) was unstandardized, meaning that the grade point average could be either a 4.0 or 5.0 scale. Raw high school rank (HsRank) was collapsed into four categories of importance, ranging from 1(lowest) to 4 (highest). Student participation in college preparatory curriculum (CPC) was coded 1 for no and 2 for yes. Student participation in advanced placement classes (AP) was coded 1 for no and 2 for yes. Student demonstration of a pattern of strong grades over time (SG) was coded 1 for no and 2 for yes. Where the recommendation letter (RL) reported that the student demonstrated intellectual promise, motivation, maturity, independence, initiative, writing skills, or any special talents and/or enthusiasm, it was coded 2 for yes and 1 for no when these features were absent.

The Elbow essay (EE) had four scoring sections: shows understanding of purpose, shows understanding of main parts of argument, delineates the primary strengths of the essay, and raises questions and issues related to the essay. Each section received a score of 1 (unacceptable or below expectations), 2 (acceptable or marginal), 3 (commendable or very good), or 4 (exceptional or outstanding).

The honors interest essay (HCE) also had four parts: shows understanding of honors purpose, distinguishes self creatively, distinguishes self academically, and distinguishes self in leadership/service/collaborative work. Each section received a value of 1 to 4, using the same scoring criteria as the Elbow essay.

Writing mechanics and organization (MEC) of the Elbow essay and the honors interest essay were graded together in four parts: coherent sentences and appropriate word choices, strong paragraphs and sentence variety, sufficient address of the question, and use of specific detail and examples. Each section was scored 1 to 4, using the same criteria as the Elbow and honors interest essays.

The on-site essay (OSE) had three evaluation areas: answers the prompt, shows attention to content of lecture, and shows reflection on content of lecture. Each section received a score of 1 to 4, using the same criteria as the previous essays.

Small-group interaction (SMGR) was graded on eight areas: engagement—student is engaged in the conversation, makes eye contact, and asks questions; interest—student demonstrates interest in ideas; enthusiasm—student shows enthusiasm for thinking; resourcefulness—student conveys a sense of resourcefulness and considers a wide variety of possible resources in answering questions; adventurousness—student is willing to take risks, perhaps accompanied by a lack of obsession with the "right answers"; communication—student

tries to talk about something unusual, reaching for new ways to conceptualize or verbalize his/her thoughts; collaboration—student discusses directly with other students; and professionalism—student demonstrates decorum and shows respect to other students. Each section was scored 1 (unacceptable or below expectations), 2 (acceptable or marginal), 3 (commendable or very good) or 4 (exceptional or outstanding). Points from each section were compiled to create a total rubric score (160 possible points).

The faculty member doing the qualitative evaluation placed the applicant in one of five categories (coding for analysis in parentheses): absolutely, the student is extremely well qualified and an exceptional candidate, put in my small group tomorrow (9); yes, the student is commendable with good potential and could become a great honors student (7); some doubts or reservations, the student shows some potential but something is missing (5); not preferable, the student is marginal and would require a lot of work to be successful in honors (3); and absolutely not, the student is unacceptable and below expectations (1). This score was kept separate from the total rubric score and compared to it for consistency of faculty evaluations.

Exploratory data analysis on each variable revealed completeness and no coding errors. Recursive partitioning, using the first- and second-semester college GPA response outcomes with the remaining variables except retention as explanatory variables, yielded two cross-validated regression trees. Next, a cross-validated classification tree was constructed using retention as the response and the rest of the variables except first- and second-semester college GPA as independent variables.

RESULTS

A full regression analysis was conducted. The results depicted a regression tree dividing subjects into more homogeneous subgroups (R Development). Next, the cross-validated mean error rate for the sub-trees was run on one hundred partitions of the data. The regression tree was pruned based on the cross validation of means. The key finding was that freshmen with a high school GPA below 3.80 earned on average a 3.40 GPA in their freshman year in college. Freshmen with a high school GPA of 3.80 and higher on average earned a GPA of 3.77. Next, a cross-validated classification tree with an error rate of retention was used to identify predictors of retention in the admissions model. No single variable meaningfully predicted retention. In other words, no variable predicted retention better than random guessing.

Finally, freshman fall-semester college GPA was compared with high school GPA using a scatter plot (Figure 1). Results showed that 19.4% of students with less than a 3.80 GPA in high school scored a 3.80 or better at college in their freshman year whereas 56.1% of students with a 3.80 GPA or

Figure 1: High School GPA vs. Freshman Fall GPA
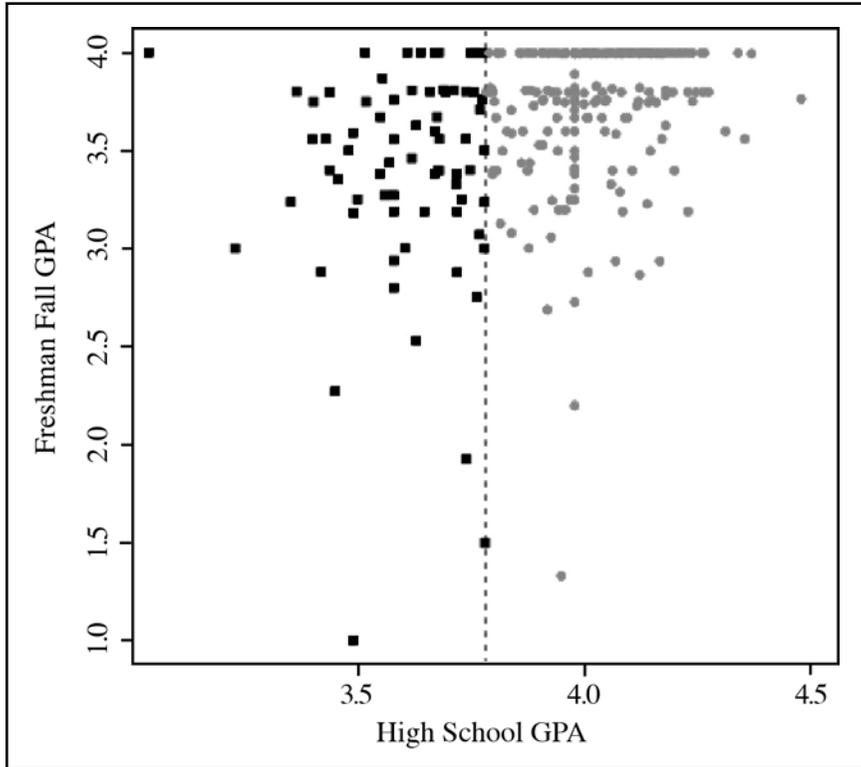
better in high school scored a 3.80 or better at college in their freshman year. This finding helped explain the cross-validation result from the full regression analysis.

DISCUSSION

In response to our first research question—"Is there inter-rater reliability between the qualitative impression and the quantitative scores?"—we found that, although there was some difference in faculty ratings (intra-rater reliability), the qualitative impression of each faculty matched the quantitative scores (inter-rater reliability). In other words, the scoring rubric accounted for criteria that honors faculty defined as qualities of ideal honors students or "fit."

The results of the second research question—"Do items being evaluated predict retention (the operational definition of "fit")?—yielded the simple finding that no single predictor of retention existed within this admissions model. The main issue with retention in honors other than "fit," however, tended to be low college GPA, and the scatterplot allowed us to see if any

predictive relationship existed between admissions variables and freshman fall and spring GPA. The strongest relationship with freshman year college GPA was high school GPA. Not surprisingly, students with a higher high school GPA have a higher freshman year college GPA. The significance of this relationship could not be accurately measured because the sample was limited in range (with most students earning an above-average GPA), possibly limiting the ability of GPA to serve as a good predictor of outcome. Equally important, however, was that there was no predictive relationship between freshman-year college GPA and ACT score.

CHANGES TO ADMISSIONS MODEL

Results from the scatter plot indicated that setting a high GPA floor for applicants would improve their chances of earning a higher college GPA at the end of the freshman year, thereby increasing persistence rates into the sophomore year. Adjustments to the admissions model re-weighted assessment points collected in the application, including assignment of equal value to the two parts of the application, namely, assessment of previous academic achievement (ACT, GPA, class rank, and letter of recommendation) and assessment of data collected to measure fit with the values and practices of the Schedler Honors College (evaluations of Peter Elbow essay response, public service essay, and small group discussion).

Research results supported increasing the importance of hsGPA and decreasing it for ACT. The administration addressed these results in two ways. First, they increased the minimum high school GPA from 3.25 to 3.50. Second, they assigned a greater weight to hsGPA and a lesser weight to ACT in the overall applicant assessment (Figure 2). A standardized recommendation format was also introduced so that specific qualities (scholarship, service, and leadership) could be assessed more consistently, and it was given greater weight.

This new version of the admissions model has yielded greater freshman-to-sophomore and freshman-to-junior retention. Over the last three years, the program has averaged freshman-to-sophomore retention rates greater than 97% (Figure 3). This climb in retention rates appears to be based on adjustments to the holistic, multi-criterion rubric being used to assess applicants for admission as well as on the Inform and Interview day process that allows applicants to better understanding the program before deciding to attend and allows faculty to better identify ideal honors students. Some students who are offered admission decline the offer because it is not the education they want—also a sign of success in the admissions process. Prospects are making better-informed choices about whether their values and goals align with the Schedler Honors College at UCA, and faculty are recommending students that have a

**Figure 2: Components of Revised Rubric Score**
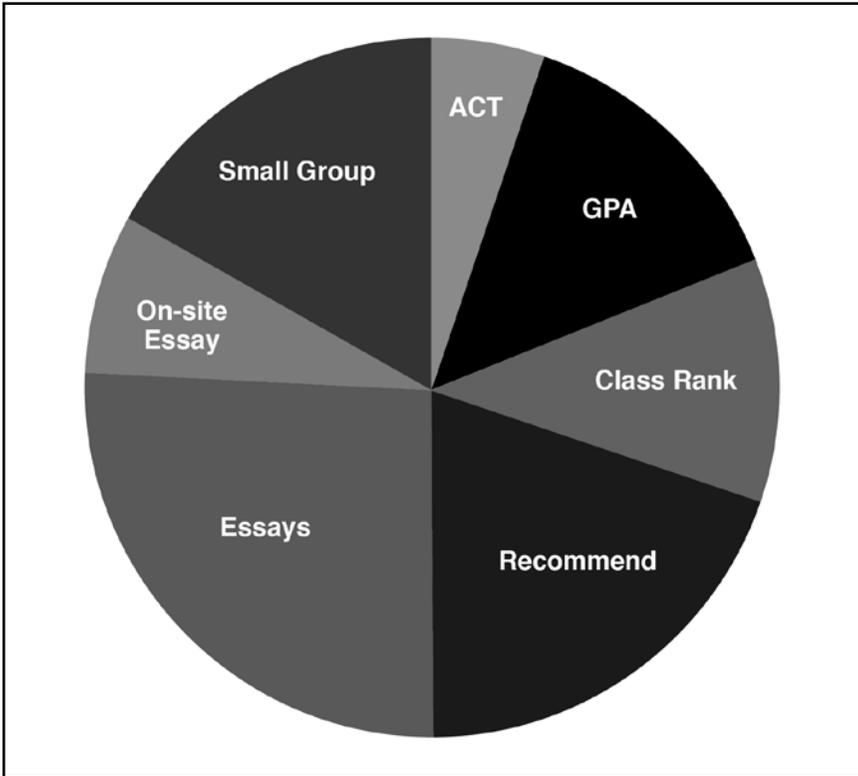
higher likelihood of success as indicated in this figure illustrating the freshman to sophomore and freshman to junior retention rates of entering UCA Honors students.

Racial and ethnic diversity of the student population is important to the Schedler Honors College administration for both inclusive access and enrichment of the living and learning experience. In the first year that the holistic process was revised to place greater emphasis on GPA and less on ACT, the freshman non-white student population increased from a prior average of 12.3% to 16%. Even with this change, continuous and improved outreach efforts are merited to sustain annual admission of diverse entering classes.

The selection process, though labor intensive, has been used to admit entering classes ranging in size from 57 to 150, with applicant pools ranging from two to four times the size of the incoming classes. Using a password-protected, web-accessible admissions technology with user-friendly interfaces for applicants, recommenders, faculty, and administrators has made the process more efficient and accurate than reliance on paper applications and has eased the labor for all participants.

Figure 3: Retention Rates
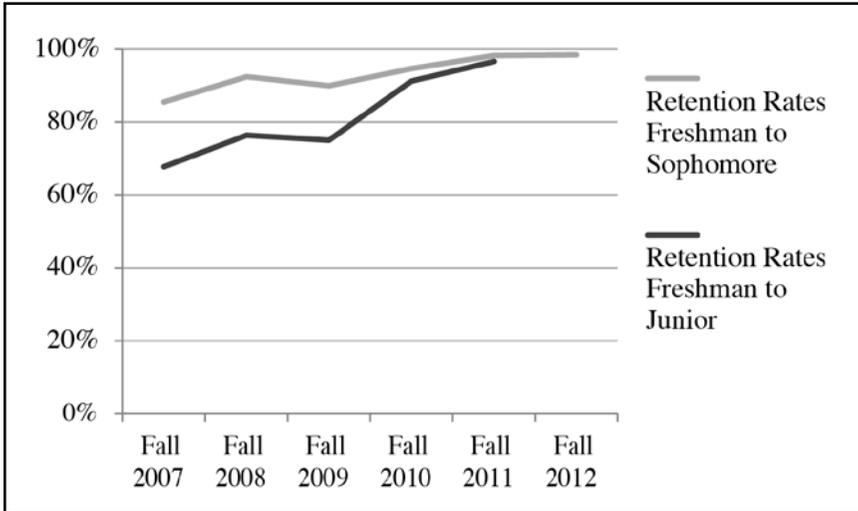
The Schedler Honors College at UCA has accomplished the two main goals of the analysis: (1) to create a process of admission that better predicts student fit and success and therefore retention; and (2) to improve enrollment of a more racially diverse population of students. Having faculty evaluate applicants' values, reasoning, writing, past civic engagement, and interpersonal skills in the classroom has proven critical to the first goal. Emphasizing high school GPA rather than standardized test scores has proven to be the single most important factor influencing achievement of both goals. This finding reveals an important take-away for honors units that do not have a sufficient number of faculty to carry out intensive applicant screening. Simply shifting from ACT/SAT minimums to high school GPA minimums for selection could lead to greater diversity and better retention rates.

CONCLUSION

Tierney et al. describe the socioeconomic and cultural biases in standardized testing that result from unequal resources among students and their parents, indicating that standardized testing cannot predict success for all groups of applicants. Two other issues are that the predictive power of testing is relatively weak and that the ACT and SAT are not interchangeable. Banerji further suggests that not using standardized tests is just as effective as using them in predicting success. This study presents additional supporting evidence since, for the Schedler Honors College at UCA, no statistical relationship exists between freshman-year college GPA and ACT score.

Honors administrators would benefit from further research, including within their own programs and colleges, on the extent of cultural and socio-economic bias introduced through standardized testing as well as the actual predictive power of standardized testing for higher-achieving students. Logically, the reliability of the relationship of any interval-level variable to another interval-level decreases when its distribution is truncated. Applicants to honors programs and colleges typically exhibit a small upper range of scores if standardized tests are being used. Empirical results and the logic of truncated distributions argue that honors programs and colleges have an obligation to examine their admissions practices and determine whether selection criteria offer accurate measures of success or are standing in the way of greater diversity and retention.

REFERENCES

Banerji, S. (2006). George Mason's SAT-Optional Admissions Policy Could Boost Diversity. *Diverse Issues Higher Education*, 23 (9), 12.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification and Regression Trees*. Boca Raton, FL: CRC Press.

Chenowith, K. (1997). A Measurement of What? *Black Issues in Higher Education*, 14 (14), 18–22.

Elbow, P. (1973). *Writing Without Teachers*. New York: Oxford UP.

FairTest. (2013). FairTest National Center for Fair and Open Testing. Web. Retrieved August 1, 2013 from <http://fairtest.org>.

Gilroy, M. (2007). Colleges Making SAT Optional as Admissions Requirement. *Education Digest*, 73 (4), 35–39.

Madden, J. (1994). What is an Honors College? *The National Honors Report*, XV (2), 35–40. Retrieved from <http://www.tntech.edu/honors/nhr>

National Collegiate Honors Council (2012). Membership News, National Collegiate Honors Council Newsletter. Retrieved February 2, 2013, <http://nchchonors.org/newsletters/2012-newsletters/may-2012/#membership>.

R Development Core Team (2005). R: A language and environment for statistical computing, reference index version 2.x.x. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org>.

Rinn, A. (2006). Major Forerunners to Honors Education at the Collegiate Level. *Journal of the National Collegiate Honors Council*, 7 (2), 63–69. Retrieved from <http://digitalcommons.unl.edu/nchcjournal/17>.

Sadler, P., & Tai, R. (2007). Weighting for Recognition: Accounting for Advanced Placement and Honors Courses When Calculating High School Grade Point Average. *NASSP Bulletin*, 91 (1), 5–32.

Sederberg, P. (2005). Characteristics of the Contemporary Honors College: A Descriptive Analysis of a Survey of NCHC Member Colleges. *Journal of the National Collegiate Honors Council*, 6 (2), 121–136. Retrieved from <http://digitalcommons.unl.edu/nchcjournal/180>.

Syverson, S. (2007). The Role of Standardized Tests in College Admissions: Test-Optional Admissions. *New Directions in Student Services*, 118, 55–70.

Tierney, W., & Venegas, K. (2009). Finding Money on the Table: Information, Financial Aid, and Access to College. *The Journal of Higher Education*, 80 (4), 363–88.

****

The authors may be contacted at

psmith@uca.edu.

APPENDIX A

ASSESSMENT OF STUDENT WRITING

Writing Assessment	Exceptional/ Outstanding (4)	Commendable/ Very Good (3)	Acceptable/ Marginal (2)	Unacceptable/ Below Expectations (1)
Elbow Essay				
Shows understanding of purpose				
Shows understanding of main parts of argument				
Delineates the primary strengths of the essay				
Raises questions and issues related to the essay				
Total				
Honors College Essay				
Shows understanding of Honors purpose				
Distinguishes self creatively				
Distinguishes self academically				
Distinguishes self in leadership/ service/collaborative work				
Total				

Mechanics & Organization of Elbow & HC Essays						
Coherent sentences and appropriate word choices						
Strong paragraphs and sentence variety						
Sufficiently addresses the question						
Use of specific detail and examples						
Total						
On-site Essay						
Answers prompt completely						
Shows attention to content of lecture						
Shows reflection on content of lecture						
Total						

APPENDIX B

ASSESSMENT OF STUDENT SMALL GROUP INTERACTION

Small Group Interaction	Exceptional/ Outstanding (4)	Commendable/ Very Good (3)	Acceptable/ Marginal (2)	Unacceptable/Below Expectations (1)
Engagement: The student is engaged in the conversation, makes eye contact, and asks questions				
Interest: The student demonstrates interest in ideas				
Enthusiasm: The student shows enthusiasm for thinking				
Resourcefulness: The student conveys a sense of resourcefulness and considers a wide variety of possible resources in answering questions				
Adventurous: The student is willing to take risks, perhaps accompanied by a lack of obsession with the "right answers"				

Communication: The student tries to talk about something unusual, reaching for new ways to conceptualize or verbalize their thoughts				
Collaboration: The student discusses directly with other students				
Professionalism: The student demonstrates decorum and shows respect to other students				
Overall Assessment:	25–32 points	17–24 points	9–16 points	8 points