

A Response to: “NOT SO FAST! A Second Opinion on a University of California Proposal to Endorse the New SAT,” CSHE.16.08, by S. Geiser

Mark M. Rashid
Professor, Civil Engineering
University of California, Davis
Chair, BOARS 2006-08

Sylvia Hurtado
Professor, Education and Director, Higher Education Research Institute
University of California, Los Angeles
Chair, BOARS 2008-10

Michael T. Brown
Professor, Counseling, Clinical, and School Psychology
University of California, Santa Barbara
Chair, BOARS 2004-06
Chair, University of California Academic Senate, 2007-08

Samuel J. Agronow
Coordinator of Admissions Research and Evaluation
University of California Office of the President

The paper, “Not so Fast,” which discusses an Academic Senate proposal that is now being considered by the University of California Board of Regents, is full of errors of fact and misrepresentations. As faculty and staff who were centrally involved in the analyses and deliberations surrounding the Senate’s proposal, we write to correct these errors and misrepresentations.

Earlier this year, the Academic Senate of the University of California recommended to the Board of Regents, via the president of the University, that UC’s freshman eligibility policy be substantially revised. The central element of the Academic Senate’s proposal is creation of a new category called “entitled to review,” whereby high-school graduates who satisfy certain basic criteria of college-readiness would be guaranteed an admissions review at each UC campus to which they apply. A second element of the proposal is elimination of UC’s current requirement that, in order to be eligible for freshman admission, all applicants must submit scores on two SAT Subject tests in different subject domains. The current requirement that applicants take either the ACT with its optional Writing component or the SAT Reasoning test, would be retained. The Senate’s Board of Admissions and Relations with Schools (BOARS) put the proposal forward following more than two years of work, including a lengthy and thorough process of UC-systemwide review, comment, and revision.

The Geiser paper incorrectly refashions the Senate’s proposal as an “endorsement” of the “New SAT,” and further claims that this supposed endorsement “reverses a decade of UC research and

policy development.” To support the latter claim, the author misrepresents a 2002 BOARS policy statement (<http://www.universityofcalifornia.edu/senate/committees/boars/admissionstests.pdf>) as a sweeping endorsement of SAT II (now called SAT Subject) tests over the SAT I (now SAT Reasoning) exam. The paper further implies that the findings of previous UC research relating to the SAT I and II, or at least those findings that are favorable to the paper’s claims, remain unalterably applicable now, even after major changes (in 2006) to both the SAT and UC’s testing requirements. Inexplicably, the author mischaracterizes his and others’ work as indicating that achievement tests, generally, are more predictive of college student performance, with generally “less adverse impact on low-income and minority applicants.”

Any rational view of admissions testing begins with acknowledging that a testing requirement is but one element of a larger policy governing admissions. The value, efficacy, and potentially negative consequences of the tests, and ultimately the advisability of their use, depend heavily not only on the characteristics of the tests themselves, but also on the manner in which they are used to effect admissions decisions, the circumstances of the institution, and the types of students the institution aims to serve. Also, and importantly, a clear distinction must be maintained between the functioning of the tests themselves, and the *requirement that the tests be taken*. This is especially important in the case of SAT Subject tests, which represent an unusual requirement (in fact unique to UC among public university systems in the US), and for which the test-taking population is both much smaller than, and significantly different than, that of the SAT-R.

As BOARS worked to formulate its proposal to reform the eligibility construct, it was well aware of the above-cited landmark 2002 position paper on admissions testing and, in fact, studied the document and its supporting evidence closely. BOARS’ recommendation to eliminate Subject tests from among UC’s freshman admissions requirements represents a reasoned synthesis of the following points: 1) With inclusion of the old SAT II Writing exam in the new SAT Reasoning test (SAT-R), essentially in its entirety, the SAT Subject tests are left without an instrument for the assessment of writing; it no longer exists as a stand-alone Subject test. Indeed, it is not now possible to fashion a “general skills” testing instrument from SAT Subject tests that would serve UC’s purposes. 2) A broad consensus appears to exist among UC faculty that standardized assessments of writing and math should be required of all applicants. 3) Changes to the SAT-R and elimination of the SAT II Writing exam by the College Board leave UC, as well as other institutions, without any specific Subject test(s) that can reasonably and profitably be required of applicants on a national basis. 4) UC’s Subject test requirement is unusual, and in fact unique among public University systems in the country. About twice as many California college-bound students take the SAT-R than take one or more Subject tests; among African Americans, the ratio is over three-to-one. 5) Each additional required test incurs an additional expense and burden on students and their families, and an additional chance that deserving and high-achieving students will find themselves excluded from UC altogether, because of failure to take the required test for whatever reason. 6) UC’s Subject-test requirement so constrains the pool of UC-eligible California high school graduates that the statewide GPA/test-score index must be set at a surprisingly modest level in order to identify nominally 12.5% of all CA graduates. The result is that there are large numbers of *ineligible* students with stronger academic credentials than some eligible students. *Thousands of these high-achieving, but ineligible, students actually apply to UC each year, and the vast majority of them are denied while their lower performing but eligible*

peers are guaranteed admission. 7) In the context of UC's testing requirement and UC's applicants, the two Subject test scores have been found to contribute negligible increments of predictive validity in predictions of freshman-year success at UC, once other factors like GPA and ACT/SAT-R scores are taken into account.

The requirement that particular tests be taken by all applicants should meet an appropriate standard of educational justification, in relation to the tests' contribution to high-quality admissions decisions. In the case of UC's current Subject-test requirement, in which the specific tests are left to the discretion of the student with no reasonable alternative to this format, the predictive validity of the scores is so small that no compelling educational argument can be made for their requirement. The conclusion is inescapable that UC's Subject test requirement is not only unnecessary, but also harmful to the ideals of fairness and equity in access to the institution.

In this context, the Geiser paper asserts, without evidence, that "...the primary obstacle is not the tests, *per se*, but the aspiration to attend UC and the perception that it is a realistic possibility," attributing blame to students for failure to gain UC access instead of to unnecessary testing. The paper declares that student aspirations and college objectives pose the actual barriers in the form for student application behavior. But data show that thousands of students lacking the complete test pattern apply to UC every year, thereby demonstrating their UC aspirations and objectives. A large subset of these students each year have academic records superior to those deemed eligible, with the exception that, for reasons that remain unclear, they did not complete the entirety of UC's unusual testing requirements.

The "Not so fast" paper is silent to all of the aforementioned realities. Instead, the paper ardently advocates for Subject tests as if they represent some kind of monolithic entity (there are 20 of them, of which 19 are accepted by UC), while disparaging the SAT Reasoning test that has arguably changed from the time he conducted his analyses (it now contains more advanced mathematics and the writing component). It is a matter of fact that the subject tests vary greatly in terms of their ability to predict later college performance, as well as in their racial/ethnic impact.¹ BOARS, as the Academic Senate agency responsible for recommending admissions policy to the Regents, is obliged to take a more reasoned and objective approach to its policy work.

In developing the eligibility reform proposal, for more than two years BOARS discussed and reviewed data that implicated elements of the eventual proposal. For example, data considered by BOARS at its November 2006 meeting bears on the impact of the SAT II requirement on various racial/ethnic groups; this data suggests that the SAT II requirement constitutes a major barrier to eligibility. In June of 2007, after extensive engagement with the campuses, BOARS issued its recommendation regarding eligibility reform that included the recommendation to eliminate the superfluous Subject-test requirement. Because of the lateness in the academic year, UC's Academic Council released BOARS' eligibility-reform proposal for systemwide review in the Fall of 2007, at the beginning of the 07-08 academic year. During the 2007-08 academic year, the proposal, ultimately in revised form, was subjected to two exhaustively thorough systemwide reviews in which all UC divisions and relevant systemwide Senate committees

¹ For a review of relevant research, see Zwick, Rebecca (2004). *Rethinking the SAT: The Future of Standardized Testing in University Admissions*. Routledge: New York and London.

vigorously participated. Few issues that have come before the Academic Senate in recent years have prompted the kind of rigorous and detailed debate that attended the review of the eligibility reform proposal. Representatives of BOARS visited every undergraduate UC campus (some campuses more than once) to discuss the proposal with Senate and administrative leaders on these campuses, and to solicit feedback. The chair of BOARS met with legislative officials in Sacramento to explain the proposal and its rationale. The Senate review process culminated in endorsement of the proposal by the Academic Assembly, the final parliamentary authority of the Academic Senate, in June of this year on a vote of 38 in favor, 12 opposed.

It is worth noting that, throughout the entire Senate review process, the one element of the proposal that was consistently supported by an overwhelming majority of Senate faculty – including those who expressed concerns and reservations about various aspects of the proposal – was elimination of the Subject test requirement.

The “Not so fast” paper attributes to BOARS the finding that the SAT Reasoning test is more predictive of UC freshman performance than high-school grades. The author supports this attribution by citing a reference (Agronow and Rashid, 2007; Table 1, Models 1 and 3). The paper further suggests that this finding forms the basis for BOARS’ proposal to eliminate UC’s Subject-test requirement. There is no basis for this major claim in the paper. The eligibility-reform proposal in no way rests on, or even references, the incorrectly-attributed assertion that the SAT-R is “. . . a better predictor of student performance than the old SAT and is now superior to either high-school grades or achievement tests in predicting how students are likely to perform at UC.” In fact, the text of the reference cited by Dr. Geiser in connection with this assertion, which appears as Appendix IV in the much larger proposal document itself (<http://universityofcalifornia.edu/senate/underreview/sw.rev.eligibility.02.08.pdf>), *makes no mention whatever* of any finding of greater predictive validity of the SAT-R test. For that matter, no such finding is mentioned in the larger proposal document, nor in any other document authored by BOARS. We are unaware of any source that might form a basis for this central claim of Dr. Geiser’s paper -- that the Senate’s eligibility reform proposal constitutes an “endorsement” of the SAT-R test, and that this supposed endorsement is, in turn, driven by a comparison of its predictive validity to that of high school GPA.

For the record, the empirical evidence about comparisons with the old SAT I and/or high-school grades is found in a 2007 conference paper by Agronow and Studley (<http://www.cair.org/conferences/CAIR2007/pres/Agronow.pdf>) that takes up the issue of the predictive validity of the “new” SAT (i.e. the SAT Reasoning test) in the context of UC freshmen. That work found that the predictive validity of the SAT-R for the Fall 2006 freshman class – the first class subject to the new test – was somewhat greater than the predictive validity of the old SAT I exam for the Fall 2004 freshman class. This can be seen in Figure 1, Model 4 of the Agronow and Studley paper, where adjusted R-squared values of 0.172 and 0.200 in 2004 and 2006 respectively are reported. Further, comparing Models 1 and 4 in the same figure shows a slightly greater predictive validity of the SAT-R (0.200) in 2006 over the HS GPA (0.196) that same year. This relationship is also seen in a true multivariate context in Table 3, where the adjusted R-squared for SAT-R combined with High school GPA in predicting UC GPA is 0.282 in Model 13. The *increase* in adjusted R-squared from Model 1 (HS GPA alone) to Model 13 is 0.086 when SAT-R is added compared with the *increase* from Model 4 (SAT-R alone) to Model

13 of 0.082 when HS GPA is added. Similar results are observed when comparing similar Models in Table 7 that presents the analysis for engineering enrollees.

Dr. Geiser's paper also mistakenly claims that BOARS/UCOP made a "common error of statistical inference" by basing conclusions on "simple bivariate correlations." In fact, the above observations are based on multivariate linear regression modeling, which, in the words of the Geiser paper, are "the preferred methodology" (we agree). (Ironically, according to Table 2 of the 2007 Agronow and Rashid study, which provides bivariate correlations *in addition* to multivariate-regression results, the correlations of UC GPA with HS GPA on the one hand, and of UC GPA with SAT-R composite on the other, are *identical* to the two decimal places reported: 0.44.) In any case, the observations contained in Agronow and Studley's 2007 paper are obviously subject to a number of important limitations and caveats inherent in the models. For this reason, such findings did not form the basis for any generalizations by BOARS or the Academic Senate about the educational value of various measures of achievement.

Of greatest significance is the paper's misrepresentation of the rationale for the eligibility-reform proposal. The paper repeatedly – and falsely – asserts that one element of the eligibility reform proposal – elimination of the Subject-test requirement – is based on the supposed superiority of the SAT-R in comparison with other measures of achievement. In this connection, the paper completely ignores the main and explicitly-stated conclusion of the 2007 Agronow and Rashid study: Once other factors are taken into account (most notably HSGPA and SAT-R or ACT composite), the two Subject-test scores add only very small increments of predictive validity with respect to UC freshman GPA. This conclusion is inescapable based on the results presented in Table 1 of Agronow and Rashid's study. While BOARS was well aware of the modest increase in predictive validity exhibited by the SAT-R over the old SAT I test, the reasoning behind the broad eligibility reform proposal in general, and the testing provision in particular, makes no reference whatever to this observation. That reasoning is carefully explained in the proposal document itself.

We further note that table 1 of the Agronow and Rashid study does not support a major claim of the Geiser paper, namely that Subject tests and the SAT-R are comparable in their predictive validity in the context of recent UC matriculants. Comparing models that have SAT-R instead of Subject tests and vice-versa, but are otherwise identical (Models 3 vs. 4, 8 vs. 9, 12 vs. 13, 16 vs. 17, and 20 vs. 21), it is seen that, in all cases, the model with the SAT-R included gives a better prediction of UC GPA than the one with the Subject-test scores. The paper notes the comparison between the nearly-fully-saturated Models 20 and 21, which contain many other factors that are themselves positively correlated with test scores. In this particular comparison, the SAT-R exhibits only a 1.1% advantage over the Subject tests, out of about 29% variance explained. What the Geiser paper fails to mention, however, is that in all the other pairwise comparisons listed above, the predictive superiority of the SAT-R over the Subject tests was found to be significantly greater. In Model 3 vs. 4, for example, the SAT-R exhibited a 4.7% advantage over the Subject tests, out of about 20% variance explained. In 8 vs. 9, the SAT-R accounts for an increment of 8.6% out of 28%, whereas the Subject tests account for only 5.1%. In light of these results, it would seem strained to claim, as Dr. Geiser does, that the tests are "largely interchangeable as predictors."

We emphasize that the above comparisons are based on empirical findings; like all empirical findings their validity is limited by the circumstances of the observations: in this case, 2006 UC freshman matriculants, subject to UC admissions policies and procedures. No sweeping claims about the intrinsic value of one testing instrument over another, valid in all circumstances and for all institutions, can be – or were – made based on these results. In any case, these results do not represent an “endorsement” of one set of tests over another. They do, however, strongly suggest that there is little value in UC continuing to require both the SAT-R (or ACT) and two Subject tests.

The “Not so fast” paper goes on to make a curious reference to supposed “methodological glitches” in the regression models used by BOARS. Geiser apparently refers to the suggestion that, in regression modeling of freshman-GPA outcomes, scores on the 17 out of 19 Subject tests *not* reported by a given applicant should be treated statistically as “missing data.” Statistical techniques can then be used to impute these “missing” scores from the reported scores, along with the other quantitative information in the application. It would seem highly debatable that, e.g., imputing a score on the Chinese With Listening Subject test for a student who presents scores on English Literature and Chemistry would lead to substantially different conclusions about the predictive efficiencies of the reported scores in general. In any case, use of applicant data in regression models should align, as closely as possible, with its use in actual admissions practice: currently, UC uses the two highest-reported Subject-test scores as generic scores in its admissions processes. Further, we note that the “missing data” approach was suggested (unpersuasively) to BOARS based on the hypothesis that the range of the Subject-test scores among UC applicants is significantly restricted, because the applicants themselves choose which tests to take. A test-by-test examination of the relevant data reveals that, in fact, no such range-restriction problem exists for Subject-test scores reported by UC applicants, in comparison to test-takers nationally as well as in comparison to SAT Reasoning-test scores reported by UC and national applicant pools. Finally, and perhaps most importantly, there is the obvious point that estimation of “missing” test scores in statistical modeling only makes sense if we would contemplate doing the same thing in actual admissions practice. Such a practice would be highly problematic.

A main claim of Dr. Geiser’s paper – that HSGPA is “by far the best predictor of UC freshman GPA” while the “New SAT. . . continues to be a relatively weak indicator of student performance at UC” – is based on a puzzling rationale: the beta weights observed in the fully-saturated regression model (Model 22) in the 2007 Agronow and Rashid reference. Yes, clearly the best-fit linear function of the predictor variables is a strong function of HSGPA, as one might expect. It is also the case that HSPGA is strongly correlated with other predictor variables in the model (SAT-R scores, SAT Subject scores), which are themselves strongly correlated with each other. As is well known, the individual weights emerging from least-squares optimization are of little significance when the predictor variables are highly correlated with each other. We are unwilling to draw conclusions about what is stronger (HSGPA or SAT-R) based on normalized regression weights under these circumstances, particularly when the incremental variance in the models suggests rather different conclusions. In any case, BOARS remains a strong advocate of conferring a position of primary significance on evidence of classroom performance in assessments of academic achievement. BOARS does not support primary reliance on

standardized tests in campus-based admissions decisions, and this position is reflected in, and is clearly supported by, the current reform proposal.

In sum, the Geiser paper attempts to prevent the elimination of UC's SAT Subject test requirement, on the eve of an impending Regent's decision, by mischaracterizing the carefully articulated rationale behind the Academic Senate's eligibility reform proposal. The paper claims misrepresentations on the part of others, even while being highly selective with the facts, and stating conclusions that are unsupported by available evidence. It attributes claims to cited references and statements that do not appear in those references. This approach to policy research is uncharacteristic of the scholarship we should expect on the discourse of major UC policies.

REFERENCES

Agronow, S. and M. Rashid. (2007). "Examining the predictive value of the SAT subject exams in the prediction of first year UC GPA – A report to BOARS."

<http://www.universityofcalifornia.edu/senate/underreview/sw.rev.eligibility.02.08.pdf>.

Agronow, S. and R. Studley. (2007). "Prediction of college GPA from New SAT test scores – a first look." Paper presented at the annual meeting of the California Association for Institutional Research, Nov. 16, 2007. <http://www.cair.org/conferences/CAIR2007/pres/Agronow.pdf>

Assembly of the Academic Senate. (2008). Proposal to Reform UC's Freshman Eligibility Policy.

<http://www.universityofcalifornia.edu/senate/assembly/jun2008/as.eligibility%200608.pdf>

Board of Admissions and Relations with Schools. (2002). The Use of Admissions Tests by the University of California.

<http://www.universityofcalifornia.edu/senate/committees/boars/admissionstests.pdf>

Board of Admissions and Relations with Schools. (2008). "BOARS' revised proposal to reform UC's freshman eligibility policy." Oakland, CA: UC Academic Senate.

<http://www.universityofcalifornia.edu/senate/underreview/sw.rev.elgiblty.reform.02.08.pdf>.

Geiser, Saul (2008) NOT SO FAST! A Second Opinion on a University of California Proposal to Endorse the New SAT. CSHE.16.08.

<http://cshe.berkeley.edu/publications/publications.php?id=319>

Zwick, Rebecca (2004). Rethinking the SAT: The Future of Standardized Testing in University Admissions. Routledge: New York and London.