

Update (26 July): see statement from a member of the “pro” team, pages 7-9

Update (2 Aug): white paper from the “pro”-GRE team, pages 10-11

A White Paper Presenting Arguments Against Using Graduate Record Exam (GRE) Scores for Graduate Admission

*Prepared by: Beth Moore, Pat Schloss, Kate Barald, Leilani Marty-Santos, Claire Ofiara
with help from Lori Isom, Kathy Spindler, Jennifer Linderman*

Evidence suggests that GRE scores don't predict future success in graduate programs (1-9). Importantly, data from these studies indicate the test is biased against underrepresented minority (URM), economically disadvantaged, female or rural students (**Figure 1**). Standardized tests like the SAT or GRE gained popularity as screening metrics for admissions as baby boomers sought higher education in numbers that dwarfed available slots (10). Because the 4.0-GPA scale left little room for differentiation, SAT/GRE tests gained popularity because greater variability among test-takers was considered a more quantitative “measure of merit.” However, this premise is based on assumptions that **1)** the GRE measures relevant metrics needed for higher education and graduate school success, **2)** the GRE doesn't introduce bias and **3)** all test-takers have equal access to information and resources needed to prepare for the test. As highlighted in *Science* and *Nature* editorials (1, 9) and in the preponderance of current literature cited below, **all these assumptions are false**. The following are critical discussion points.

The GRE is an inconsistent and modest predictor of student success. Burton and Wang (2) reported that **1)** students who withdrew from programs had mean quantitative GRE averages actually higher than those completing degrees or remaining in training, and **2)** the GRE is at best a predictor of first-year graduate school grades, but undergraduate GPA does this as well. A larger study (5) using data from Vanderbilt's PIBS-like graduate program (2003-2011) reached similar conclusions and highlighted a lack of correlation in predicting who would graduate with a PhD, time to defense, number of conference presentations, number of first author papers, or ability to receive an individual grant/fellowship. Additionally, study of the Tetrad graduate program at UCSF concluded that GRE scores varied by fewer than 6 points among students ranked by faculty from lowest to highest using multiple criteria; *most importantly, some of the highest-ranked students had scores below the 30th percentile on verbal and analytical portions of the GRE* (6). A study of UM-PIBS graduate students (1999-2009) confirms these previous results: undergraduate GPA was a predictor of success as defined by their programs ($p < 0.01$), whereas verbal and quantitative GRE scores were not (7). Additionally, GREs may not be reliable for screening foreign students given evidence of widespread cheating in test-taking, especially in China (11-13). Furthermore, the GRE doesn't measure quantitative skills beyond the high school level, covering only arithmetic, algebra, geometry, and elementary statistics (14, 15). Therefore, it is unconvincing that GRE scores would have greater predictive power than undergraduate math grades. Ultimately, the predictor emerging as most powerful in analyses of UNC (3; **Figure 2**) and UCSF (6) students was recommendation letters from applicants' undergraduate research advisors. However, we caution that letters of recommendation can also introduce bias (16). **Given the inability of the GRE to accurately assess/predict success in graduate school, it should not be included as part of the rubric that evaluators use when considering prospective students.**

Cost and format of the GRE disadvantages rural and low-income students: Beyond lack of any positive predictive value for the GRE, there is ample evidence that the GRE is a biased tool. The test itself costs \$205 (USA)-\$220.70 (China). This includes sending scores to 4 institutions, but it costs \$27 more for each additional report. The test is generally offered in larger metropolitan areas, meaning students from rural areas must incur travel expenses to take the test. For ETS to report only their highest cumulative score, it costs students an additional \$150. GRE prep courses cost \$699 (self-paced) to \$1299 (in-person class) to \$2499 (private tutoring), based on web searches. These factors disproportionately impact low-income (often URM) and rural students. As an example of these effects, a Puerto Rican applicant flew to Florida and took a 24-hour bus ride to attempt a retest at the only possible location prior to UM's application deadline at great personal cost with little change in score. That student has since won several research awards and published well at the University of Michigan. Two other URM students with verbal GRE scores in single digits have graduated, done postdocs, and now hold prestigious jobs in industry. **Considering that other application components can be used to predict success in graduate school, it is unnecessary to use such a biased tool with no predictive power.**

The GRE has been abandoned by many: Based on lack of positive predictive value of the test and its bias against certain socioeconomic groups, many USA entities are working to remove the GRE as a tool for admissions decisions. Since 2001 Rackham has not required GRE scores nor used them to assess program quality. The UM Astronomy Department has stopped accepting GRE scores, as have similar departments at other institutions based on a resolution of the *American Astronomical Society* stipulating that programs no longer require the GRE: <https://aas.org/governance/council-resolutions#GRE>. Other institutions including UT Southwestern (17), USC (18), Harvard (19) and UNC Chapel Hill (20) are strongly considering not accepting GRE scores as well. As an indicator of where policies stand at the national level, the prestigious **NSF Graduate Research Fellowship stopped accepting GRE scores over 5 years ago. NIH T32 progress reports and new submissions no longer require GRE information.**

The University of Michigan could be a national leader in eliminating GRE requirements, which could improve recruiting: Ample literature documents that diverse teams make better decisions (21-24). We are categorically in favor of pursuing a diverse graduate student body. This is important for the excellence of our science and is a mandate for institutional training grants. Based on the above, *we believe the GRE disadvantages students we are trying to recruit*. Elimination of the GRE requirement could improve our competitive advantage with these students. **We encourage PIBS to be the first major US biomedical sciences program to eliminate this requirement!**

Figure 1: GRE quantitative reasoning scores by gender and ethnicity (from ref. 9).

THE GREAT DIVIDE

The data represent the scores typically achieved in the quantitative reasoning test of the graduate record examinations (GRE) by US students from different ethnic groups applying for graduate school. In the physical sciences, a minimum score of 700 is required by many PhD programmes.



Can Applications Predict Biomedical Graduate Student Productivity?

Joshua D. Hall, Anna B. O'Connell, Jeanette G. Cook

University of North Carolina at Chapel Hill School of Medicine, Office of Graduate Education, Chapel Hill NC, 27599

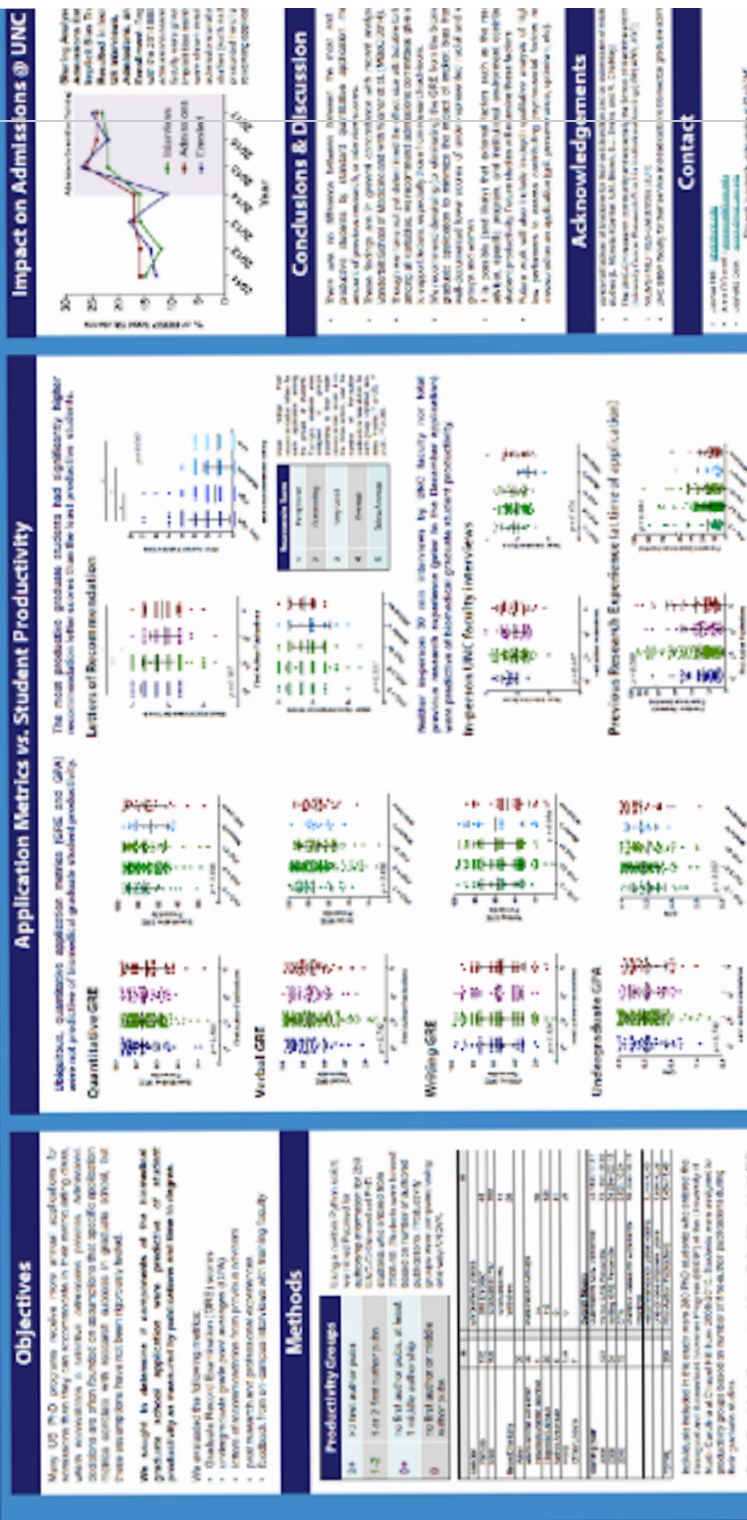


Figure 2: Updated study of application components that predict success (courtesy of Josh Hall, UNC). Extension of work by Hall et al., 2017 (refs. 3, 19, 20).

[Link to higher-res poster file](#)

1. Benderly, B.L. GREs don't predict grad school success. What does? *Science*, Jun. 7, 2017. <http://www.sciencemag.org/careers/2017/06/gres-dont-predict-grad-school-success-what-does>

2. Burton, N. W., and M.-m. Wang. 2005. Predicting Long-term Success In Graduate School: A Collaborative Validity Study. *ETS Research Report Series* 2005: i-61. <https://www.ets.org/Media/Research/pdf/RR-05-03.pdf>
3. Hall, J. D., A. B. O'Connell, and J. G. Cook. 2017. Predictors of Student Productivity in Biomedical Graduate School Applications. *PLoS One* 12: e0169121. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0169121>
4. Kuncel, NR., Ones, DA and Hezlett, SA. 2001. A comprehensive meta-analysis of the predictive ability of the GRE: Implications for graduate student selection and performance. *Psychological Bulletin* 127: 162-181. <https://www.ncbi.nlm.nih.gov/pubmed/11271753>
5. Moneta-Koehler, L., Brown, AM, Petrie, KA, Evans, BJ and R. Chalkley. 2017. The Limitations of the GRE in Predicting Success in Biomedical Graduate School. *PLoS One* 12: e0166742. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0166742>
6. Weiner, O. 2013. How should we be selecting our graduate students? *Mol Biol Cell* 25: 429-430. <http://www.molbiolcell.org/content/25/4/429.long>
7. Curtin, N., Meader E, and Stewart A. 2011. Predictors of Career Success among Rackham Doctoral Students. *Rackham Report* (unpublished).
8. Posselt, J. R. 2014. Toward Inclusive Excellence in Graduate Education: Constructing Merit and Diversity in PhD Admissions. *American Journal of Education* 120: 481-514. <https://eric.ed.gov/?id=EJ1034780>
9. Miller, C., and K. Stassun. 2014. A test that fails. *Nature* 510: 303-304. <https://www.nature.com/naturejobs/2014/140612/pdf/nj7504-303a.pdf>
10. Alon, S., and M. Tienda. 2007. Diversity, Opportunity, and the Shifting Meritocracy in Higher Education. *American Sociological Review* 72: 487-511. <http://journals.sagepub.com/doi/abs/10.1177/000312240707200401>
11. Tyre, P. How Sophisticated Test Scams From China Are Making Their Way Into the U.S. *The Atlantic*, March 21, 2016. <https://www.theatlantic.com/education/archive/2016/03/how-sophisticated-test-scams-from-china-are-making-their-way-into-the-us/474474/>.
12. Redden, E. and Jaschik, S. Indicted for Cheating. *Inside Higher Ed*, May 29, 2015. <https://www.insidehighered.com/news/2015/05/29/chinese-nationals-indicted-elaborate-cheating-scheme-standardized-admissions-tests>.
13. Mandak, J. The College Board has canceled college entrance exams on Saturday in China and Macau because of cheating concerns. *US News & World Report*, Jan. 21, 2016. <https://www.usnews.com/news/us/articles/2016-01-21/sat-tests-canceled-in-china-macau-over-cheating-concerns>.
14. ETS: Overview of the Quantitative Reasoning Measure. https://www.ets.org/gre/revised_general/prepare/quantitative_reasoning/
15. ETS: Math Review. https://www.ets.org/s/gre/pdf/gre_math_review.pdf
16. Trix, F., and C. Psenka. 2003. Exploring the Color of Glass: Letters of Recommendation for Female and Male Medical Faculty. *Discourse & Society* 14 191-220. <http://advance.uci.edu/images/Color%20of%20Glass.pdf>
17. Nancy Street, Assoc. Dean of Admissions, U.T. Southwestern, personal communication.
18. Julie Posselt, Asst. Prof., USC Rossier School of Education, personal communication.
19. Josh Hall, PREP Director, UNC Chapel Hill, personal communication regarding Harvard presentation at TWD meeting of PREP directors.
20. Jean Cook, Dean of Graduate and Postdoctoral Studies, UNC Chapel Hill, personal communication.

21. Ellison, S.F., and W.P. Mullin. 2014. Diversity, Social Goods Provision, and Performance in the Firm. *Economics and Management Strategy* 23: 465-481. <http://economics.mit.edu/files/8851>

22. Sommers. 2006. On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations. *J of Personality and Social Psychology* 90: 597-612. <http://www.apa.org/pubs/journals/releases/psp-904597.pdf>

23. Temm, T. 2008. If You Meet the Expectations of Women, You Exceed the Expectations of Men: How Volvo Designed a Car for Women Customers and Made World Headlines. In Schiebinger, L. (Ed.), *Gendered Innovations in Science and Engineering*, pp. 131-149. Stanford: Stanford University Press.

24. Page, S. E. 2007. *The Difference: How the power of diversity creates better groups, firms, schools and societies*. Princeton University Press.

UPDATE: Statement from Steven Ragsdale of the “pro” team (updated 27 July)

I have headed a cross-campus graduate program, chaired a departmental admissions committee for over 10 years (Biochemistry at Univ. of Nebraska), and served as a member of the Biological Chemistry admissions committee for over 5 years. I support the “Pro” side. I feel that the GRE does provide valuable information, both for admission and for decisions by potential mentors about which laboratory would be best suited for a student. I recognize the drawbacks of the GRE if it is used as the singular or major determinant, but I feel that it provides some data points that indicate where a student is in the academic environment (which has its own/similar drawbacks). Below, I have considered the various factors that are available to the admission committee, summarized the information that the GRE provides and offered my thoughts about how the information can be (and should be) used in assessment.

1. What discriminating indicators are available for the admissions committee to decide who they should recommend for admission to the PIBS program. This information is also useful for a faculty member to decide who to admit into their lab?

- Academic history: Undergraduate school(s) attended, Transcript, GPA
- Quantitative measures of student’s academic mastery:
 - GRE - Quantitative indicators (raw scores and percentiles) in these areas: Writing, Verbal, Quantitative; Subject test.
 - Grade Point Average (GPA)
- Prior Research Experience: List of Publications, Years of research experience, Areas of research interests as well as information usually provided in the Statement of Purpose, Personal Statement, CV and Recommendation Letters.
- Work Experience: Curriculum vitae
- Academic statement of purpose
- Personal statement
- Recommendation Letters
- Student’s choice of program

2. What information does the GRE provide? Because many of us may not be familiar with what information is covered by the GRE, I have provided a brief summary. There are three General sections and a Subject test.

- Analytical writing section. Consists of two different essays, scored by at least two readers on a six- point holist scale. The first (issue task) essay is about a topic selected from a list (available

to the student before the test); the second (argument task) essay is to critique and offer suggestions to improve the logic of argument given to the students.

- Verbal section. Three sub-sections (reading comprehension, critical reasoning, and vocabulary usage), each of which contains 20 questions to be completed in 30 minutes. The student is asked questions involving text completion, sentence equivalence, and critical reading.

- Quantitative section. Each quantitative section consists of 20 questions to be completed in 35 minutes and consists of questions involving quantitative comparisons, problem solving, and data interpretation.

- Subject Tests in 6 areas: Biology, Chemistry, English Literature, Mathematics, Physics and Psychology. The length of each exam is 170 minutes. Note that the Biochemistry, Cell and Molecular Biology exam was discontinued in December 2016.

3. Intention of the GRE

The test is intended to measure how well the student has developed their verbal and quantitative skills in abstract thinking during their undergraduate education.

4. How important is the information provided by the GRE relative to the other admissions criteria?

In my personal experience, the GRE has been a valuable indicator in assessing and understanding the skill sets of a student. Students with very poor scores have to work harder to develop the communication, analytical and quantitative skills tested by the GRE. I also feel optimistic that students can overcome whatever limitations they enter graduate school as well as further develop those strong skill areas.

Students with very high scores on the writing exam are typically excellent at written communication skills and have good skills at assessing and developing the logic of an argument. These students are better prepared for exercises like writing grant proposals and academic papers. Similarly, a student with outstanding verbal GRE scores will typically be rather advanced in their use of grammar, delivering seminars, and writing papers and proposals. Students with verbal GRE scores in the sub-40 percentile will typically experience problems in an academic environment where verbal and written communication skills are important. High or Low verbal GRE scores are usually accompanied by similarly excellent or poor grades in subjects that require verbal or written communication.

Students with high quantitative scores generally have good skills in math and science and are generally strong at analyzing data and remembering and communicating factual information.

5. How should the information be used?

- I do not feel that we should set a cut-off value and exclude students who drop below that mark. Some graduate schools will not consider applicants that score below a particular score (say 310).

- I do not feel that we should rank students according to the GRE scores.

- I feel that we should be holistic in our admissions decisions and weight the GRE as a valuable criterion along with those others listed in the first point.

I can understand the claims that the GRE alone is not a good predictor of success. However, none of these available criteria can be used as THE admissions criterion. I argue that the GRE scores should be valued with a spirit of curiosity as one useful indicator in the admissions package.

Overvaluing the GPA might lead to rejecting students with outstanding promise, especially those students who experienced a difficult transition from high school into living on their own in the new college environment. Overweighting the research experience might lead us to reject a student who attended an undergraduate school that did not offer strong research opportunities or a student who had to work to pay their way through school (or help support their family). Each of the individual

criteria can be similarly criticized. In both of these cases, the GRE may help identify a promising student whose GPA or research experience is below that of their cohorts.

6. Impact of omitting the GRE scores from students' records.

Omitting GRE results would have a negative impact both in the admissions decision and in the decision about which lab the student might choose to pursue their PhD or MS degree. Ignoring a poor GRE during the admissions process will potentially lead to acceptance of a student who is not sufficiently prepared for graduate school or for a particular program. Ignoring an outstanding GRE score might lead to premature rejection of an outstanding candidate.

Regarding the choice of laboratory, different faculty members tend to find portions of the GRE to be useful. For example, research in some laboratories require a high degree of quantitative sophistication; thus a student with very poor scores on the quantitative portion of the GRE will have a difficult time succeeding in that environment. The faculty member should be provided access to that information.

Update: Position statement in support of the use of GRE as a component of holistic PIBS graduate admissions

Prepared by:

Jorge Iñiguez-Lluhí, Miriam Meisler, Margit Burmeister, Aaron King, Brian Athey, Christin Carter-Su, Sally A. Camper, Stephen Ragsdale, Philip Andrews, Alon Kahana, Julie Douglas. With input from additional faculty members.

The University of Michigan considers diversity in its academic mission to be essential to its core goals and values. As a community responsible for the training of future scientists, we are committed to both diversity and academic excellence, especially given the national and international stature of our institution. This diversity is also reflected in the different competencies required for successful careers in the broad set of disciplines encompassed by PIBS. We believe that to recruit and train a diverse and talented cohort, it behooves us to use a broad range of criteria in the evaluation of applicants to our diverse programs.

There are many faculty members at the University of Michigan with experience in graduate education who believe that GRE scores make a contribution to admissions decisions, in combination with the other components of 'holistic evaluation'.

This is supported by the largest primary study that we are aware of (Kuncel and Hezlett, *Science*, 2007), a meta-analysis which has a sample size of 259,640, which was not cited by the "con" GRE white paper. This study concludes, "Results from a large body of literature indicate that standardized tests are useful predictors of subsequent performance in graduate school, predict more accurately than college GPA, do not demonstrate bias, and are not damaged by test coaching. Despite differences across disciplines in grading standards, content, and pedagogy, standardized admissions tests such as the GRE have positive and useful relationships with subsequent student accomplishments."

In this context, it is relevant that the claim of low predictive value of GRE scores is derived from analysis only of data from students with high enough scores to be admitted to graduate programs. This is a form of statistical selection bias. In many cases, a high GRE score has contributed to admission of successful students who have weaker performance on undergraduate GPA or lacked research experience. Similarly, students that were admitted despite low GRE scores may be equally successful because they possessed other positive attributes. Thus, the lack of correlation between GRE scores and success in the admitted subset is an indication that the holistic approach is working effectively, not that the GRE scores are without value.

Our goal should be to ensure the success of all our students, regardless of background. GRE scores can therefore be very useful to identify areas where specific students could benefit from targeted support. This can help mentors tailor course selection and training to the student according to their discipline. Omitting the GRE scores will unnecessarily decrease the diversity of metrics that

inform the holistic review process, and will undercut our current process of identifying a diverse and talented pool of students. The University of Michigan can be a leader by removing barriers to GRE testing by offering support to those students that consider its cost a barrier to application.

The experience of many of us was well stated by Stephen Ragsdale in his material appended to the white paper:

"In my personal experience, the GRE has been a valuable indicator in assessing and understanding the skill sets of a student. I feel we should be holistic in our admissions decisions and weigh the GRE as a valuable criterion along with the others available."

Ignoring an outstanding GRE score might lead to premature rejection of an outstanding candidate. Finally, regarding the choice of laboratory, different scores may be important for the work in different laboratories, and faculty members should have access to that information, e.g. for work that require a high level of quantitative sophistication.

Our concluding position is summarized by Alon Kahana: "We see the GRE as data and, as scientists, we should embrace, not disregard data." and Aaron King: "We believe that unless one can unequivocally demonstrate that GRE scores contain zero information, there is no sensible rationale for denying faculty the information they do contain."