

# A Proposal to Suspend the GRE as an Admissions Criterion in the C&I Doctoral Program

## The Problem

The Graduate Record Examination (GRE) is a mainstay of contemporary graduate student admissions processes. The GRE is a norm-referenced, standardized assessment of verbal, quantitative, and writing skills that prospective graduate students take in hopes of attaining admission into graduate-level programming, including doctoral programs. The GRE is an established set of assessments that many argue are highly reliable and psychometrically sound. Standardized performance on these tests is compared against norming samples in order to compare individual students' performance on the test, and thus derive percentile rankings that denote higher and lower performance. Universities use the GRE in a variety of ways, ranging from high stakes (e.g., minimum cut scores required to be considered for admission) to holistic (e.g., variably weighted among other admissions criteria), which has varying degrees of alignment with how the Educational Testing Service recommends these assessments be used (see ETS guidance document [here](#)). It has become commonplace to expect the GRE as an admission requirement for graduate study, and indeed many may consider the notion of not using the GRE unthinkable.

Yet the GRE, and other standardized instruments like it, has long been under scrutiny. A dubious psychometric history, alongside long-term and persistent performance differences between U.S.-born Black and Latinx students and their white and Asian counterparts clearly play a role in skewing admissions decisions in favor of more privileged populations (Awad, 2007; Louderback, 2008; Pousselt, 2016; Pruitt & Isaac, 1985; Sternberg & Williams, 1997; Scheuneman, 1987). In addition to persistent race- and ethnicity-based performance differences, recent research notes multiple confounds that obscure any links between a student's GRE scores and their overall performance in doctoral studies (Evans, 2017; Ingram, 1987; Kuncel & Hezlet, 2010; Kuncel, Hezlett & Ones, 2001; Williams, 1987). These latter facts lead one to wonder why universities would continue to require an assessment tool that clearly ranks students as a function of ethnic and racial background but simultaneously fails to significantly predict graduate student performance.

Adding complexity to this issue are national rankings of schools of education. Contemporary algorithms for assessing the quality of a given school require reporting of a school's average GRE scores (verbal and quantitative) for incoming graduate students. Such an algorithm is used by the *U.S. News and World Report* (USNWR) in determining their annual ranking of schools of education in the United States. This average score is included as part of a ranking algorithm that includes a broad array of school-based indicators (e.g., grant dollars per faculty, student-to-teacher ratio, etc.). A school's score on this algorithm is then used to rank order it relative to other schools of education nationwide. Higher performance on this metric is highly desired by many schools of education, and GRE performance, by design, is positively correlated with that score. The rankings thus incentivize schools of education to admit graduate students with higher GRE scores.

A problem arises in this scenario. Higher GRE scores are *positively associated* with USNWR rankings. Higher GRE scores are *inversely associated* with U.S.-born Black and Latinx populations. Thus, higher performing white and Asian students are more likely to be admitted to graduate school, while their Black and Latinx counterparts are less likely to be admitted, which results in a noteworthy lack of diversity at schools of education. The effect is distal, but difficult to dismiss, and highlights the problems of continued use of an assessment with a design history rooted in generating the very results we see today (Kendi, 2019).

Recently, universities have taken note of the arguments outlined above and opted to remove the GRE as an admissions requirement. Notably, Brown University undertook such an endeavor in 2019. Out of the 50 programs offering doctoral degrees, half decided the GREs would no longer be a part of their admissions process. Schools were given the opportunity to alter components of the admissions requirements by, for example, adding questions to the Letter of Recommendation or essay prompts. Additionally, programs were asked to submit the logic of their decisions to the graduate school administration prior to admission season.

### Analysis of GRE Scores & Time to Completion

We conducted an analysis of C&I PhD students admitted between 2010 and 2014 ( $n = 43$ ). Of the 43, one had withdrawn and two have not yet finished the program. A total of 40 have or will have completed the program at the end of this academic year. For the 40 completers, we calculated the time to program completion as an important outcome variable, as some concerns appear to be centered on low GREs being associated with difficulty in completing the program. Among the 40 students, we ran separate regression analyses that yielded contradictory evidence. First, the average time to completion was 5.5 years. Controlling for verbal GRE, quantitative GRE was negatively associated with time to graduation. That is, stronger GREs were associated with quicker time to graduation. See Figure 1.

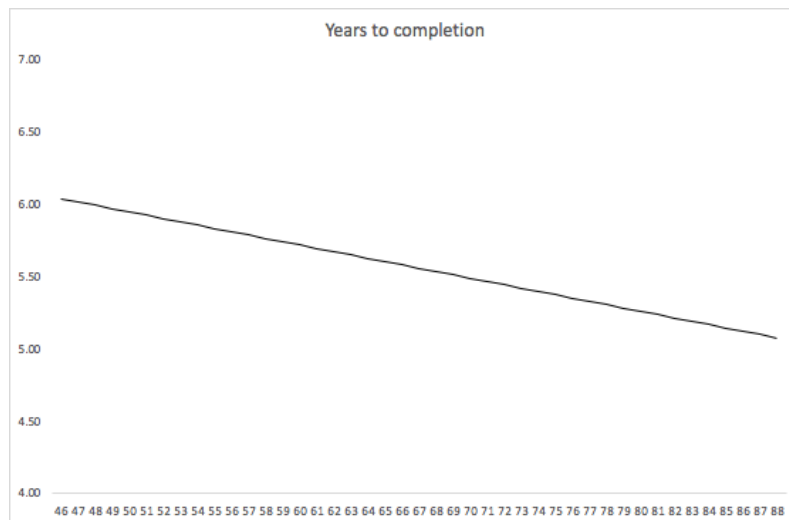


Figure 1. GRE quantitative percentile predicting time to completion

On the other hand, controlling for quantitative GRE score, Verbal GREs were associated with a longer time to graduation. See Figure 2.

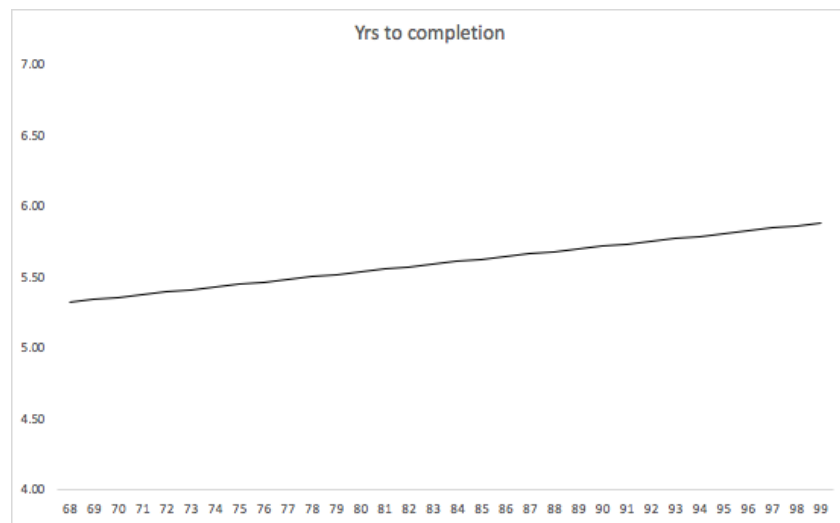


Figure 2. GRE verbal percentile predicting time to completion

In Figure 1, the x-axis represents the range for the sample for GRE quantitative percentile, and is 46 - 88, while GRE verbal percentile is 68 - 99. Both reflect actual sample performance. The effect was stronger for quantitative ( $B = -.023$ ,  $p = .002$ ) than for it was for verbal ( $B = .018$ ,  $p = .055$ ), but in general, these effects cancel each other out, suggesting null evidence for GRE performance on time to completion.

### A Proposal

We propose a small-scale experiment at the Lynch School of Education and Human Development. Specifically, the Curriculum & Instruction doctoral program proposes to conduct a 3-year trial during which time the GRE would be removed as an admissions criterion, and admissions would be conducted with weighting redistributed among remaining admissions criteria (described below).

The C&I Doctoral Advisory Committee met with Jessica Greene, Boston College Director of Assessment and Accreditation, to determine the impact such a decision would have on the Lynch School's USNWR rankings. In response to the committee's questioning, she assembled [a report](#) structured around the USNWR GRE reporting criterion: *The USNWR requires that **at least 50%** of entering doctoral student GRE scores be used to calculate the reporting average. As long as the average GRE scores are derived from >50% of incoming students, there is no scoring penalty attributed to the school.* Given this, the report's main conclusions were as follows:

- 1) For the U.S. News rankings, three sets of GRE scores are submitted -- average for MA students, average for doctoral students, average for the combination of MA and doctoral students (noting that we do not know how they then use these scores in their ranking calculations -- U.S. News does not share its methodology);

- 2) Excluding C&I PhD GRE scores would not put the Lynch School below USNWR's 50% threshold. That is, GRE scores for 50% or more of newly enrolled doctoral students need to be submitted in order to avoid a ranking penalization. The C&I doctoral program only admits 8 of 24 school-wide doctoral students (33%).
- 3) The Curriculum, Teaching, and Society's CAEP accreditation would be unaffected because GRE reporting for CAEP does not involve C&I doctoral or Master's level students;
- 4) The New England Commission of Higher Education's (NECHE) perspective, in terms of admissions, is that there is no GRE requirement or mandate. The NECHE simply needs to be sure there are "appropriate methods of evaluation" and that "BC demonstrates its ability to admit students who can be successful in its academic programs"

### Proposal Specifics

In removing the GRE from C&I doctoral admissions, the criteria for applicant evaluation would be as follows in Table 1.

Table 1. *Admissions criteria for C&I doctoral program*

<b>Undergraduate GPA</b>	1 = 3.67-4.00 2 = 3.34-3.65 3 = 3.00-3.33 4 = <3.00 +special circumstances
<b>Graduate GPA</b>	1 = 3.67-4.00 2 = 3.34-3.65 3 = 3.00-3.33 4 = <3.00 +special circumstances
<b>Experience</b> <ul style="list-style-type: none"> <li>● Consider range and richness of experiences and accomplishments related to working with pupils or others in educational settings OR experience with or study of educational issues in other ways (not just teaching or school-based experiences)</li> <li>● Rate the candidate based on a combination of background experiences (as above) AND how the candidate makes sense of that experience and relates this to interest in pursuing the PhD (e.g., the questions that have emerged from experience that candidate wants to explore through PhD work)</li> <li>● Draw on information from resume, personal statement, recommendations and possibly writing sample</li> <li>● Generally speaking, applicants should have some minimum level of experience in K-12 teaching, school administration, museum education, adult education, college teaching, teacher education, education policy development, education research</li> </ul>	1 = Excellent 2 = Very Good 3 = Good 4 = Poor

<p><b>Research</b></p> <ul style="list-style-type: none"> <li>• The range and richness of the applicant’s research experience or demonstrated potential for doing advanced and original work</li> <li>• Draw on information from personal statement, resume or application form, recommendations</li> </ul>	<p>1 = Excellent 2 = Very Good 3 = Good 4 = Poor</p>
<p><b>Diversity</b></p> <ul style="list-style-type: none"> <li>• The goal here is to establish a high quality and diverse pool of final applicants</li> <li>• By diversity, we agreed that this includes a wide range of characteristics that may differ in importance among faculty members</li> </ul>	<p>1 = Excellent 2 = Very Good 3 = Good 4 = Poor</p>
<p><b>Writing Sample</b></p> <ul style="list-style-type: none"> <li>• Writing should reflect the applicant’s individual thinking and viewpoint</li> <li>• Writing should reflect applicant’s ability to analyze, synthesize ideas</li> <li>• All academic writing conventions (mechanics, form, citations) should be observed</li> <li>• Writing should be well organized, clear and cognizant of audience</li> <li>• Writing sample should show applicant’s promise as a scholar and ability to complete the kind of academic/essay-text genre of writing required in PhD courses, comps and dissertations</li> <li>• Reviewers should keep in mind the challenges faced by non-native users of standard English</li> </ul>	<p>1 = Excellent 2 = Very Good 3 = Good 4 = Poor</p>
<p><b>Personal Statement</b></p> <ul style="list-style-type: none"> <li>• Personal statement should connect experiences to reasons for pursuing PhD—should articulate applicant’s reasoning and his/her pathway in the decision to pursue the PhD</li> <li>• Should address the questions of why BC, why PhD, why now?</li> <li>• Should of course be well written and should present a compelling story that is original, but intellectual</li> <li>• Statement should show that applicant’s interests, experiences, and questions would fit well with our C&amp;I doc program mission and what we can offer, including (in many cases) some connection to the theme of social justice</li> </ul>	<p>1 = Excellent 2 = Very Good 3 = Good 4 = Poor</p>
<p><b>Recommendations</b></p> <ul style="list-style-type: none"> <li>• AT LEAST 2 of the 3 recommendations must be from credible sources (e.g., known scholars, faculty in professorial positions at other research institutions; or reflect the reviewer’s knowledge about doctoral programs and Ph.D. level careers. For example, there might be a recommendation from a museum director as opposed to a university academic.)</li> <li>• Letters should speak to applicant’s ability to complete doctoral level academic work</li> <li>• Letters should be persuasive (often through detail) that applicant has</li> </ul>	<p>1 = Excellent 2 = Very Good 3 = Good 4 = Poor</p>

<p>outstanding academic potential and may speak to applicant's commitments, promise for advanced and original work, maturity, intellectual capacity and initiative, etc.</p> <ul style="list-style-type: none"> <li>Keep cultural and international differences in mind (e.g. some international recommendations cite weaknesses and are not nearly as "glowing" as many American recommendations tend to be)</li> </ul>	
<p><b>Admit?</b></p> <ul style="list-style-type: none"> <li>Faculty member renders a decision on the candidate.</li> </ul>	<p>1 = Recommend Admit 2 = Unsure 3 = Recommend Reject</p>
<p><b>Written Rationale</b></p> <ul style="list-style-type: none"> <li>Faculty free write any comments they have about their rationale for the rankings, whether they have assistantships and interest in the candidate.</li> <li>The section to provide a RATIONALE for why the faculty is choosing to admit, deny, or waitlist an applicant.</li> </ul>	

Using the criteria outlined in Table 1, faculty allocate a numeric score for each admissions construct (GPA, Experience, Research, Diversity, etc.), and finalize their evaluation of an applicant with a written rationale.

This process occurs twice during the C&I admissions process. In the first round of reviews, all doctoral applicants who have complete files receive a first round review from three TCS faculty members. Following this first round of reviews, the Doctoral Advisory Committee meets to summarize faculty reviews and arrives at a "short list" of applicants (typically between 25 and 30). These applications move on to a second round and receive two more faculty reviews, for a total of 5 faculty reviews. Finally, these finalists are individually discussed in an all-faculty admissions meeting at which a finalist pool (between 12 and 15) is selected. These finalists are then personally interviewed by a faculty member who would likely be working with the applicant using a standardized interview protocol. Last, the department chair and doctoral program director meet to make final decisions on the 8 initial acceptance letters that will be sent out. The remaining students move to a wait list. See Table 2 for an overview of the process.

Table 2. *Curriculum & Instruction doctoral application review process.*

Step	Action
1	All qualifying applications assigned to 3 faculty per application
2	Doctoral advisory committee shortlist meeting
3	Second round applications assigned to 2 additional faculty members
4	Full faculty meeting to discuss second round applicants & determine finalist pool
5	Individual interviews with finalist pool

6	Department chair and doctoral program director determine initial 8 acceptances
---	--

### **The Experiment**

In proposing to eliminate the GRE from admissions consideration, we pose no threat to the Lynch School's USNWR viability because the C&I doctoral student numbers are less than 50% of all admitted doctoral students. We have a robust rating system alongside a rigorous review process with three tiers of applicant review. We believe the process is sound and is not improved by including GRE scores, which we contend introduces as much noise as signal into our deliberations.

In proposing a 3-year moratorium, we offer a set of research questions we can ask that might help evaluate the effectiveness of the change in procedure:

1. Does the number of applications increase?
  - a. Baseline: In the 2019-2020 we received 69 applications, up from 48 the previous year.
  - b. Hypothesis: Applications will rise overall to the C&I doctoral program.
2. What percent of applications come from people of minoritized and non-traditional U.S. backgrounds?
  - a. Baseline: In 2019-2020, a small percentage of our overall applicant pool was African-American (5%), which our Latinx applicant pool was even smaller (1%)
  - b. Hypothesis: The percentages of U.S.-born Black and Latinx applicants will rise.
3. What are the demographics of enrolled cohorts of doctoral students under the experimental approach?
  - a. Baseline: In this year's first year cohort of 8 doctoral students, while diversity is present in many respects (e.g., international students, languages spoken, sexual orientation), there are no U.S.-born Black or Latinx students represented
  - b. Hypothesis: Representation of diversity will increase, including greater numbers of U.S.-born Black and Latinx students.

### **Conclusion**

After much deliberation, we believe that the 3-year experiment proposed here would be helpful in understanding how doctoral admissions processes can be made more equitable. We do not ask that this become a school-wide policy, just that we be given a chance to understand more about this GRE phenomenon in doctoral admissions within our program, with the goal of promoting greater equity and justice, and less oppression, in doctoral programming.

Respectfully Submitted,

Department of Teaching, Curriculum, and Society

### References Cited

- Awad, G. (2007). The Role of Racial Identity, Academic Self-Concept, and Self-Esteem in the Prediction of Academic Outcomes for African American Students. *Journal of Black Psychology*, 33(2). 188-207
- Evans, C. (2017). Predictive Validity and Impact of CAEP Standard 3.2: Results From One Master's-Level Teacher Preparation Program. *Journal of Teacher Education*, 68(4). 363-376
- Ingram, R. E. (1981). The GRE in the Graduate Admissions Process: Is How it is Used Justified by the Evidence of Its Validity? *Professional Psychology: Research and Practice*, 14: 711-714.\
- Kendi, I. (2019). *How to be an anti-racist*. New York: Penguin.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127, 162-181
- Kuncel, N. R. and Hezlett, S. A. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current directions in psychological science*, 19(6). 339-345.
- Louderback, P. J. (2008). Comparison of the predictive validity of traditional intellectual measures and sociocultural factors on American Indian student academic achievement. Doctoral dissertation, Oklahoma State University.
- Miller, C.W. (2013). Admissions Criteria and Diversity in Graduate School. *The Back Page*. Retrieved from <https://www.aps.org/publications/apsnews/201302/backpage.cfm>
- Posselt, J. R. (2016). *Inside graduate admissions: Merit, diversity, and faculty gatekeeping*. Cambridge, MA: Harvard University Press.
- Pruitt, A. & Isaac, P. (1985). Discrimination in Recruitment, Admission, and Retention of Minority Graduate Students. *The Journal of Negro Education*, 54(4). 526-536
- Sternberg, R. J. & Williams, W. (1997). Does the Graduate Record Examination Predict Meaningful Success in the Graduate Training of Psychologists? A Case Study. *American Psychologist* 52(6),
- Scheuneman, J. (1987). An Experimental, Exploratory Study of Causes of Bias in Test Items. *Journal of Educational Measurement* 24(2). 97-118
- Williams, W. M. (1997). Reliance on Test Scores Is a Conspiracy of Lethargy. *The Chronicle of Higher Education*. October 10: A60