

The GRE and Its Predictive Validity for Graduate Student Success

Annotated Bibliography

Compiled by the Graduate Diversity Program, UC Berkeley

Table of Contents

I. Predictability page 2

- a. Educational Testing Service (1991). Guide to use of the Graduate Record Examination Program, Princeton, NJ.
- b. Fairtest (2001). *Examining the GRE: Myths, Misuses, and Alternatives*.
- c. Grove, W and Wu, S. 'The Search for Talent: Doctoral Completion and Research Productivity of Economists,' July 2005.
- d. Milner, M., McNeil, J. & King, S.W. (1984). 'The GRE: A Question of Validity in Predicting Performance in Professional Schools of Social Work.' *Educational and Psychological Measurement, Vol 44*: 945-950.
- e. Phillips, D.R. and McAuliffe, 'The Validity of the GRE in Predicting School Psychology Graduate Program Outcomes: A Conditional Probability Analysis.' *The School Psychologist, Vol 58, no 2*: 56—61.
- f. Sternberg & Williams. 'Does the Graduate Record Examination Predict Meaningful Success in the Graduate Training of Psychologists?' *American Psychologist 52* (1997): 630-641.

II. Performance & Meta-Analysis page 3

- a. Hunter, Schmidt & Le. 'Implications of Direct and Indirect Range Restriction for Meta-Analysis Methods and Findings.' *Journal of Applied Psychology 91* (2006): 594-612.
- b. John Orlando. 'The reliability of GRE Scores in Predicting Graduate School Success: A Meta-Analytic, Cross-Functional, Regressive, Unilateral, Post-Kantian, Hyper-Empirical, Quadruple Blind, Verbiage-Intensive and Hemorrhoid-Inducing Study.' *Magazine Ubiquity, Vol 2005* (June).
- c. Kuncel, N., Wee, S., Serafin, L., & Hezlett S. 'The Validity of the Graduate Record Examination for Master's and Doctoral Programs: A Meta-Analytic Investigation.' *Psychological Science in Public Interest* (2010).
- d. Morrison, T. & Morrison, M. (1995). 'A Meta-Analytic Assessment of the Predictive Validity of the Quantitative and Verbal Components of the Graduate Record Examination with Graduate Grade Point Averages Representing the Criterion of the Graduate Success.' *Educational and Psychological Measurement, Vol 55* (no.2): 309-316.
- e. Scott, R.R. & Shaw, M.E. (1985). 'Black and White Performance in Graduate School and Policy Implications for Using GRE Scores in Admission.' *Journal of Negro Education, Vol 54* (no. 1): 14-23.
- f. Steele, Claude M., and Joshua Aronson. 'Stereotype threat and the intellectual test performance of African Americans.' *Journal of Personality and Social Psychology 69*.n5 (Nov 1995): 797(15).

III. Background & Standardized Test Supplements

page 5

- a. Jennifer Hedlund, Jeanne M. Wilt, Kristina L. Nebel, Susan J. Ashford, Robert J. Sternberg. ‘Assessing practical intelligence in business school admissions: A supplement to the graduate management admissions test.’ *Learning and Individual Differences* (2006).
- b. Pencock-Roman, M. (1994) ‘Background Characteristics and Future Plans of High-Scoring GRE and General Test Examinees.’ *Research Report ETS-RR9412*. Submitted to EXXON Education Foundation, Princeton, NJ: Educational Testing Service.
- c. William E. Sedlacek, Research Report # 5-01 ‘Why We Should Use Non-cognitive Variables with Graduate Students’ (2004).

I. Predictability

- a. Educational Testing Service (1991) Guide to use of the Graduate Record Examination Program, Princeton, NJ.
 - i. Updated (2014) online at: http://www.ets.org/s/gre/pdf/gre_guide.pdf
 - ii. ‘Mich Kabay notes that studies often compare GREs to the GPAs of those who were never admitted in the first place, as well as those who drop out. This is an important oversight. If the issue were just distinguishing between those who will be just successful in a graduate program and those who will be highly successful, then the tested group would be sufficient. But because the test is being used to make admissions decisions in the first place, lack of information about applicants not admitted means that we cannot tell if the test is being used to deny admissions to those who would have been successful. Because of this weakness, the Education Testing Service itself flatly states that “a cutoff score based on only GRE scores should never be used as sole criterion for denial of admission.”’
- b. Fairtest (2001). *Examining the GRE: Myths, Misuses, and Alternatives*.
 - i. Online at: <http://www.fairtest.org/examining-gre-myths-misuses-and-alternatives>
 - ii. The article explores the fairness of the GRE and whether it accurately predicts success in graduate school. The findings of many studies suggest that despite its primary purpose of predicting success in graduate school, a GRE score adds little useful information to a student’s application. The test’s own developer admits that undergraduate college grades do a better job at predicting graduate achievement. In addition, ETS admits that scores for students who carefully answer questions at the beginning of the exam but guess quickly at the end may be inaccurate.
- c. Grove, W and Wu, S. ‘The Search for Talent: Doctoral Completion and Research Productivity of Economists,’ July 2005.
 - i. Online at: http://www.google.com/url?sa=t&ret=j&q=&esrc=s&source=web&cd=1&ved=0CB4QFjAA&url=http%3A%2F%2Fwww.researchgate.net%2Fpublication%2F4721205_The_Search_for_Economics_Talent_Doctoral_Completion_and_Research_Productivity%2Flinks%2F00b7d517e73d97d33d000000&ei=bit-VLT8FIjhoAT754D4BA&usq=AFQjCNE3BhBJINvaWK2rihF04_JJzTOTw&bvm=bv.80642063,d.cGU&cad=rja

- ii. Little is known about the predictive validity of graduate school criteria for long run career accomplishments. Using information contained in applications from Princeton University's Economics Ph.D. program in 1989, we predict the determinants of doctoral degree completion and research productivity 15 years later.
- d. Milner, M., McNeil, J. & King, S.W. (1984). 'The GRE: A Question of Validity in Predicting Performance in Professional Schools of Social Work.' *Educational and Psychological Measurement, Vol 44*: 945-950.
 - i. Online at: <http://epm.sagepub.com/content/44/4/945.abstract>
 - ii. "At the University of Texas, one study showed that the GRE is not a valid predictor of success for the Graduate School of Social Work as measured by grades, degree completion, and fieldwork quality. Researchers concluded that eliminating the GRE from the selection process would not lead to lower quality but would create "opportunities for many other individuals of professional quality."
- e. Phillips, D.R. and McAuliffe, 'The Validity of the GRE in Predicting School Psychology Graduate Program Outcomes: A Conditional Probability Analysis.' *The School Psychologist, Vol 58, no 2*: 56—61.
 - i. Online at: <http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCAQFjAA&url=http%3A%2F%2Fwww.apadivisions.org%2Fdivision-16%2Fpublications%2Fnewsletters%2Fschool-psychologist%2F2004%2F04-issue.pdf&ei=2Ct-VLqHD8vUoASY4oHIBg&usg=AFQjCNF6SEwXn1cLrxPbgsdpaAapUCRXuA&bvm=bv.80642063,d.cGU>
 - ii. (A comprehensive analysis of the originations of the GRE, its predictive validity and limitations).
- f. Sternberg & Williams. 'Does the Graduate Record Examination Predict Meaningful Success in the Graduate Training of Psychologists?' *American Psychologist 52* (1997): 630-641.
 - i. "The authors consider the empirical validity of the Graduate Record Examination (GRE) as a predictor of various kinds of performance in a graduate psychology program, including first and second year grades: professor's ratings of students and dissertations; and professors' ratings of students' analytical and creative, practical, research, and teaching abilities. On the basis of triarchic theory of intelligence, the GRE was predicted to be of some use in predicting graduate grades but of limited or no use in predicting other aspects of performance. In fact, the test was found to be useful in predicting 1st year grades but not other kinds of performance, with one exception- performance on the GRE Analytical test as predictive, but only for men. The authors conclude there is a need to develop better theory-based tests."

II. Performance & Meta-Analysis

- a. Hunter, Schmidt & Le. 'Implications of Direct and Indirect Range Restriction for Meta-Analysis Methods and Findings.' *Journal of Applied Psychology 91* (2006): 594-612.
 - i. Online at: <http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCUQF>

[jAA&url=http%3A%2F%2Fwww.biz.uiowa.edu%2Ffaculty%2Ffschmidt%2Fmeta-analysis%2FHunter_Schmidt_Le_2006.pdf&ei=wvN8VPHCC4fqoATr2ILwAQ&usg=AFQjCNEUFC4g5_rBAguy5b9ejg-tD935cQ&bvm=bv.80642063,d.cGU&cad=rja](http://www.biz.uiowa.edu/faculty/fschmidt/meta-analysis/Hunter_Schmidt_Le_2006.pdf&ei=wvN8VPHCC4fqoATr2ILwAQ&usg=AFQjCNEUFC4g5_rBAguy5b9ejg-tD935cQ&bvm=bv.80642063,d.cGU&cad=rja)

- ii. The authors explored the validity of the Graduate Record Examination (GRE) for predicting performance in graduate school alleging that it can only be estimated with samples of students admitted to graduate programs, which poses a restricted sample that does not account for other students that were denied admission due to low GRE scores.
- b. John Orlando. ‘The reliability of GRE Scores in Predicting Graduate School Success: A Meta-Analytic, Cross-Functional, Regressive, Unilateral, Post-Kantian, Hyper-Empirical, Quadruple Blind, Verbiage-Intensive and Hemorrhoid-Inducing Study.’ *Magazine Ubiquity*, vol. 2005 (June).
- i. Online at: <http://ubiquity.acm.org/article.cfm?id=1071921>
 - ii. Research analysis: the validity of GRE Score in predicting student’s success in graduate school. At times, research that supports GRE use utilizes questionable measurement. For example, some data strictly measures the correlation between GRE scores & students who complete their graduate program. Graduate students who do not complete their programs because of non-academic reasons (such as family) are not accounted for. Despite the data that shows relation between highly successful, moderately successful and respective GRE scores, data that presents the scores of students who were not accepted into graduate programs or decided not to attend, is not present. Contradicting studies: in many disciplines (Applied Sciences, Education, Humanities, Life Sciences, and Social Sciences), those who scored better in the GRE were more likely to fail than those who scored worse. It’s a possibility that students in these fields have demanding jobs that interfere with their studies. Some studies also lack age data. Students “who are returning to school after an extended absence” tend to have lower GREs and GPAs, “in fact earned considerably higher grades than all other groups.” Ultimately, this article suggests to avoid the GRE composite score, ignore small score differences and to not use the GRE as the deciding factor.
- c. Kuncel, N., Wee, S., Serafin, L., & Hezlett S. ‘The Validity of the Graduate Record Examination for Master’s and Doctoral Programs: A Meta-Analytic Investigation.’ *Psychological Science in Public Interest* (2010).
- i. Online at: <http://epm.sagepub.com/content/70/2/340.short>
 - ii. An investigation of over 100 studies that seeks correlations between GRE scores and GPAs for both master’s and PhD degrees. Findings: The GRE strongly predicts the *first* year of graduate academic grades. Variations in score comparison and GPA predictability are found with certain student populations (possibly uncontrolled factors). Essentially, the GRE is a good tool, however to enhance discussions about GPA predictability for master’s and doctoral programs, more GPA data needs to be collected for every graduate year level.
- d. Morrison, T. & Morrison, M. (1995). ‘A Meta-Analytic Assessment of the Predictive Validity of the Quantitative and Verbal Components of the Graduate Record Examination with Graduate Grade Point Averages Representing the Criterion of the Graduate Success.’ *Educational and Psychological Measurement*, Vol 55 (no.2): 309-316.
- i. Online at: <http://epm.sagepub.com/content/55/2/309.refs?patientinform-links=yes&legid=spepm;55/2/309>

- ii. “A meta-analytic review was conducted on published studies examining the relationship between performance on the quantitative and verbal components of the Graduate Record Examination (GRE-Q and GRE-V, respectively) and graduate grade point average (GGPA). The weighted average effect sizes (d) for GRE-Q and GRE-V were .445 and .591, respectively, and their resultant correlation coefficients were .22 (GRE-Q and GGPA) and .28 (GRE-V and GGPA), $p = n.s.$ Thus an average of 6.3% of the variance in graduate level academic achievement, as represented by GGPA, was accounted for by performance on the GRE-Q and GRE-V.”
- e. Scott, R.R. & Shaw, M.E. (1985). ‘Black and White Performance in Graduate School and Policy Implications for Using GRE Scores in Admission.’ *Journal of Negro Education*, vol. 54 (no. 1): 14-23.
 - i. “One study in the *Journal of Negro Education* revealed that at the University of Florida Black graduate students with low GRE scores were compared to White peers with high test scores. The Black students attained higher GPAs during graduate study than did the Whites, leading researchers to conclude that the GRE is insensitive to how Black students' skills and resources translate into performance.”
- f. Steele, Claude M., and Joshua Aronson. ‘Stereotype threat and the intellectual test performance of African Americans.’ *Journal of Personality and Social Psychology* 69.n5 (Nov 1995): 797(15).
 - i. The article discusses how the threat of confirming race stereotypes may place additional evaluative pressures on minority test-takers. “Stereotype threat is being at risk of confirming, as self-characteristic, a negative stereotype about one's group. Studies 1 and 2 varied the stereotype vulnerability of Black participants taking a difficult verbal test by varying whether or not their performance was ostensibly diagnostic of ability, and thus, whether or not they were at risk of fulfilling the racial stereotype about their intellectual ability. Reflecting the pressure of this vulnerability, Blacks underperformed in relation to Whites in the ability-diagnostic condition but not in the non-diagnostic condition (with Scholastic Aptitude Tests controlled). Study 3 validated that ability-diagnosticity cognitively activated the racial stereotype in these participants and motivated them not to conform to it, or to be judged by it. Study 4 showed that mere salience of the stereotype could impair Blacks' performance even when the test was not ability diagnostic. The role of stereotype vulnerability in the standardized test performance of ability-stigmatized groups is discussed.”

III. Background & Standardized Test Supplements

- a. Jennifer Hedlund, Jeanne M. Wilt, Kristina L. Nebel, Susan J. Ashford, Robert J. Sternberg. ‘Assessing practical intelligence in business school admissions: A supplement to the graduate management admissions test.’ *Learning and Individual Differences* (2006).
 - i. Online at: <http://www.sciencedirect.com/science/article/pii/S1041608005000397>
 - ii. The GMAT is recognized as a generally used tool to assess student managerial and academic potential. The GMAT has limits concerning student demographics (race, gender, socio-economic status). Although preliminary, this study suggests more

extensive assessments of interpersonal intelligence, practical intelligence, and problem solving skills (in addition to the academic intelligence). Compared to GPAs and GMAT scores, this study showed similar results. Students who scored high in this new study also had strong GMAT scores and GPAs. The study sought to extend/create more thorough and unbiased assessments of student talent and potential.

- b. Pencoek-Roman, M. (1994) ‘Background Characteristics and Future Plans of High-Scoring GRE and General Test Examinees.’ *Research Report ETS-RR9412*. Submitted to EXXON Education Foundation, Princeton, NJ: Educational Testing Service.
 - i. Online at: <http://eric.ed.gov/?id=ED382665>
 - ii. The GRE is particularly susceptible to the influence of socioeconomic class. ETS own research has shown a strong relationship between family background and test scores. One study of applicants who scored between 750 and 800 on the exam found that only 4% of these high-scoring test-takers had fathers who had not completed high school; around half had fathers with bachelor's degrees or more, and of these, a whopping 90% had fathers with graduate or professional degrees. When family income was held constant, most of the test score differences between races disappeared or shrank dramatically.

- c. William E. Sedlacek, Research Report # 5-01 ‘Why We Should Use Non-cognitive Variables with Graduate Students’ (2004).
 - i. Online at:
http://www.google.com/url?sa=t&ret=j&q=&esrc=s&source=web&cd=1&ved=0CCAQFjAA&url=http%3A%2F%2Fwilliamsedlacek.info%2Fpublications%2Farticles%2Fwhy1.pdf&ei=cSx-VJvHNc3SoASOs4C4Ag&usg=AFQjCNGG7PdC3A_sfvLAjlRODdZ1IKzMRw&bvm=bv.80642063,d.cGU
 - ii. “Problems with current predictors of success for graduate students (Graduate Record Examinations and grades) include restriction of range artifacts, grade inflation, and the increasing diversity of examinees. A case is made as to why non-cognitive variables can add to the validity of selecting graduate students. Legal, moral, ethical and statistical arguments are presented and discussed. Key points made in the report include the illogic of reflecting diversity with a single measure, and the necessity of including a range of measures to be fair to all applicants.”