Intelligence 46 (2014) 25–34

Contents lists available at ScienceDirect

Intelligence

Trends in GRE scores and graduate enrollments by gender and ethnicity

April Bleske-Rechek [a,*], Kingsley Browne [b]

[a] Psychology Department, University of Wisconsin-Eau Claire, Eau Claire, WI 54702, USA
[b] Wayne State University Law School, 471 West Palmer, Detroit, MI 48202, USA

ARTICLE INFO

Article history:
Received 6 December 2013
Received in revised form 6 May 2014
Accepted 8 May 2014
Available online 27 May 2014

Keywords:
Graduate Record Exam
GRE
Graduate enrollment
Graduate study
High stakes testing

ABSTRACT

The Graduate Record Examination (GRE) is a cognitive abilities test that predicts success in graduate training (Kuncel & Hezlett, 2007; Kuncel, Hezlett, & Ones, 2001; Kuncel, Wee, Serafin, & Hezlett, 2010). Because of its reliability, validity, and predictive utility, it is used by many graduate schools to inform admissions decisions. However, some critics describe the GRE as a gatekeeper that limits equitable access across groups to higher education (Dutka, 1999; Pruitt, 1998; Toyama, 1999). We explored how scores on the GRE have fared over time as a function of test-taker gender and ethnicity, and we investigated whether enrollment patterns over time implicate the GRE as obstructing efforts toward increasing parity in higher education. First, we found that the gap between men's and women's GRE quantitative reasoning scores has changed little since the 1980s, although female representation in science, technology, engineering, and math (STEM) graduate programs has increased substantially. Second, ethnic gaps on the GRE persist, especially in quantitative reasoning, although representation of historically disadvantaged ethnic groups in graduate programs has increased. Enrollment gaps have narrowed despite ethnic and gender GRE gaps persisting, so it appears that continued use of the GRE for admissions decisions has not blocked efforts toward equalizing representation in higher education.

1. Introduction

Institutions of higher education face at least two major challenges: first, to select and enroll those most likely to benefit from and succeed in advanced training; and second, to attract and enroll a diverse student body (Cole, 1998). The Graduate Record Examination (GRE) General Test (and other tests like it, such as the LSAT and MCAT) is an assessment of cognitive abilities that addresses the first challenge of selecting students who are most likely to succeed in graduate training (Kuncel & Hezlett, 2007; Kuncel, Hezlett, & Ones, 2001; Kuncel, Wee, Serafin, & Hezlett, 2010). However, the GRE has also been described as a gatekeeper that preserves inequities by impeding enrollment of a diverse student body (Dutka, 1999; Pruitt, 1998; Toyama, 1999). We analyzed GRE scores and graduate enrollments to investigate two primary questions: First, under the proposition that GRE scores reflect intellectual capital, how does that intellectual capital vary by gender and ethnicity? Second, do trends in graduate enrollment coincide with the notion that use of the GRE impedes enrollment of a diverse student body? If graduate enrollments have become increasingly diverse despite consistent gender differences and ethnic differences in GRE scores, we can conclude that use of the GRE has not been a clear hindrance to diversity efforts in higher education. Such a finding would raise a subsidiary issue, however, which is whether and to what extent efforts to achieve diversity by de-emphasizing GRE scores have impaired the first goal of selecting those applicants who are most likely to benefit from advanced training.

* Corresponding author. Tel.: +1 715 836 4641; fax: +1 715 836 2214.
E-mail addresses: bleskeal@uwec.edu (A. Bleske-Rechek), kingsley.browne@wayne.edu (K. Browne).

http://dx.doi.org/10.1016/j.intell.2014.05.005

1.1. Background

The GRE General Test is a standardized test of developed cognitive abilities. It is not a test of interests, preferences, or motivation; instead, it functions as an intelligence test. Individuals' scores on the verbal, quantitative, and analytic reasoning sections of the GRE general examination provide both overlapping and unique variance (Stricker & Rock, 1987), such that each component draws on a general factor *g*. One lucid simplification of Carroll's (1993) hierarchical organization of cognitive abilities is the radex organization (Snow & Lohman, 1989), which represents verbal ability, mathematical ability, and spatial ability as three content domains encircling *g*, their statistical commonality. In the current analysis, we focus on quantitative reasoning and verbal reasoning, as they represent two familiar components of the radex model (Wai, Lubinski, & Benbow, 2009) and also represent the two component scores of the GRE general examination that test-takers typically use to describe their performance.

Several lines of research have established that the GRE functions as a test of general cognitive ability, theoretically because verbal and quantitative reasoning abilities represent the speed and efficiency with which individuals acquire both declarative and procedural knowledge in their intellectual pursuits (Kuncel et al., 2001). Statistically, the GRE displays many characteristics displayed by other intelligence tests (Bridgeman, Burton, & Cline, 2008; Rock, Werts, & Grandy, 1982; Stricker & Rock, 1987; Wilson, 1988). First, performance on the GRE is nearly redundant with performance on other well-known measures of general cognitive ability, such as the SAT (Angoff & Johnson, 1990; Hsu & Schombert, 2010), which itself has been tied to life outcomes such as earning a doctorate, earning tenure, and receiving patents (Lubinski, Benbow, Webb, & Bleske-Rechek, 2006). Scores on the GRE also correlate strongly with IQ, even in highly restricted samples (Carvajal & Pauls, 1995). Second, slight but systematic variations in the factor structure of GRE scores in young and middle adulthood coincide with hypotheses about the differentiation of intelligence in young and middle adulthood (Stricker & Rock, 1987). Third, individuals display substantial stability in their GRE scores (mean test–retest coefficients of .86), even over 10-year intervals (Wilson, 1988), which is what would be expected from a measure of enduring individual differences. Fourth, as would be expected for a measure of general cognitive ability as opposed to a measure of exposure to test-relevant material, performance on the verbal and quantitative sections of the GRE is not altered significantly by coaching (Powers, 1985).

Graduate training is an intellectual pursuit with demands on both declarative and procedural knowledge (Kuncel et al., 2001). As such, student performance on the GRE is a strong predictor of success in both master's and doctoral graduate programs (Kuncel et al., 2010), as indexed by first-year GPA, graduate GPA, faculty ratings (Kuncel & Hezlett, 2007), citation counts (Kuncel et al., 2001), and even quality of academic hiring placements (Krueger & Wu, 2000). Given the predictive utility of the GRE, most graduate programs require or recommend submission of GRE scores (Hartnett & Oltman, 1984; Norcross, Haynch, & Terranova, 1996).

The GRE is not the only predictor of exceptional achievement. Emotional stability and conscientiousness forecast success in graduate school (Rigdon & Kuncel, 2010) and a variety of other positive life outcomes (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). However, there is substantial rationale for use of the GRE in higher-education admissions decisions: Large-scale databases and meta-analytic reports suggest that cognitively loaded tests in general do not exhibit predictive bias; are not substantially affected by motivation in high-stakes testing; and demonstrate strong predictive utility even after controlling for socioeconomic status (Sackett, Borneman, & Connelly, 2008; Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Performance on the GRE, in particular, is also immune to aspects of the test-center environment, such as exam proctor ethnicity and gender, which might evoke stereotype threat (Walters, Lee, & Trapani, 2004; see also Stoet & Geary, 2012).

Widespread use of the GRE began after World War II, when the pool of individuals seeking graduate education expanded. The GRE served as a common, objective assessment of applicants who came from undergraduate programs that differed widely in prestige and curricula; thus, the GRE was described by the Educational Testing Service (ETS) as a method for enhancing equity, fairness, and access to graduate school (Educational Testing Service, 2004). However, others have expressed concern about use of the GRE in admissions decisions. In particular, educators have suggested that the GRE impedes efforts toward decreasing gender and ethnic disparities in graduate enrollments, particularly in science, technology, engineering, and math (STEM). For example, Toyama (1999) argued, "For many people of color, the GRE has been used as a gatekeeper; one of the barriers to getting an education," (although, of course, the same could be said for people of all demographic descriptions). Similarly, Pruitt (1998) noted, "Those who argue that [admissions] procedures are unfair hold that qualified minority students are being denied admission to graduate study primarily because we rely on standardized test scores and undergraduate grade point averages to predict academic achievement." Given these arguments, the focus of our study is to determine whether there is evidence that disparities in GRE scores among different groups have inhibited efforts to diversify the graduate population.

1.2. Study objectives

In summary, students who take the GRE represent the pool of intellectual talent aspiring to study at the graduate level. Thus, we compiled the current set of data to determine how scores on the GRE have changed over time by gender and ethnicity, and we compared GRE scores with graduate enrollment patterns to explore the concern that use of the GRE for admission decisions has inhibited efforts toward equitable access to higher education.

2. Method

The Educational Testing Service (ETS) has published psychometric and policy reports related to the GRE for several decades, and many reports are publicly available through its website http://www.ets.org/gre/research. GRE scores for 1982 to 1996 were reported by Grandy (1999); ETS supplied us with all available reports on factors related to performance on the GRE General Test for subsequent years (J. Livingood, personal communication, September 23, 2010): 2002–03, 2003–04, 2004–05, 2005–06, and 2006–07 (Educational Testing Service, 2004, 2005, 2007a, 2007b, 2008). Repeated searches on the ETS website revealed no additional reports on factors related to performance on the GRE General Test. The tables to follow report all available scores for these two spans of years: 1982 to 1996 and 2003 to 2007.

The GRE Board and the Council of Graduate Schools (CGS) launched the CGS/GRE Enrollment and Degrees Survey in 1986. We obtained initial enrollment and degree trends from CGS reports between 1986 and 2009 (Bell, 2008, 2009, 2010, 2011; Brown, 2006; Redd, 2007). Nathan Bell, Director of Research and Policy Analysis at CGS, provided us with annual records of graduate enrollments, by sex and ethnicity, from 1986 to 2009 (N. Bell, personal communication, November 15, 2010).

3. Results

3.1. GRE scores and graduate enrollments, by gender

Table 1 lists men's and women's mean GRE-Verbal Reasoning scores and mean GRE-Quantitative Reasoning scores from 1982–1996 and 2003–2007. During the periods studied, scores on the Verbal and Quantitative sections of the GRE could be as low as 200 and as high as 800. As displayed in Table 1, mean GRE-Verbal scores hovered above 500 for men and below 500 for women through the 1980s. In the 1990s, verbal scores declined gradually for both sexes and then plateaued, with a consistent 20–30 point gap between the sexes across the time period (see Fig. 1, upper panel). The gender gap in quantitative reasoning scores is larger, at approximately one-half of a standard deviation. The lower panel of Fig. 1 shows that although quantitative reasoning scores increased slightly for both sexes over the time period, men consistently scored at least 75 points higher, on average, than women. In 1982, men outscored women by 79 points, and in 2007 by 78 points.

Despite the gender gap in quantitative reasoning scores, female representation in STEM disciplines began to grow in the 1970s (National Research Council, 2001) and continued to grow over the years under investigation here. As shown in Table 2, women comprised an increasing proportion of graduate students across all disciplines between 1986 and 2009, and their representation in the physical sciences and engineering, in particular, increased at a greater rate than did their representation in graduate training as a whole. Fig. 2 displays women's increasing representation in primary STEM disciplines from 1986 to 2009.

3.2. GRE scores and graduate enrollments, by ethnicity

Table 3 shows GRE-Verbal scores by major ethnic group for the years 1982–1996 and 2003–2007; the pattern is displayed graphically in the upper panel of Fig. 3. Overall, there was little change in verbal reasoning scores in any ethnic group. White test takers tended to have the highest average GRE-Verbal scores, while Black test takers obtained the lowest scores. The gap between White and Asian test-takers on the GRE-Verbal test diminished by the early 1990s, and the GRE-Verbal gap between White test-takers and historically underrepresented groups decreased somewhat over time as well. For example, in 1982, the GRE-Verbal gap between White and Black test-takers was 145 points; in 2007, it was 98 points.

Table 3 also shows GRE-Quantitative scores by ethnic group for the years 1982–1996 and 2003–2007; the pattern is displayed graphically in the lower panel of Fig. 3. Although GRE-Quantitative scores increased slightly over time for all ethnic groups, and greater increases occurred among

Table 1
Female representation among test-takers, and GRE-Verbal reasoning and GRE-Quantitative reasoning scores by sex, from 1982–1996 and 2003–2007*.

			GRE-Verbal			GRE-Quantitative		
Year	Test-taker *N*	% Female	Male	Female	Male–female difference	Male	Female	Male–female difference
1982	167,080	57	509	489	20	578	499	79
1983	156,336	56	516	492	24	587	506	81
1984	157,016	56	517	492	25	587	505	82
1985	177,755	57	513	490	23	589	506	83
1986	184,853	56	517	494	23	594	513	81
1987	196,209	56	516	492	24	594	512	82
1988	221,592	57	518	495	23	601	520	81
1989	235,506	58	522	495	27	603	519	84
1990	254,257	59	521	494	27	601	519	82
1991	277,314	59	520	492	28	601	517	84
1992	291,080	60	519	490	29	594	515	81
1993	270,214	61	513	481	32	592	511	81
1994	311,771	62	510	477	33	589	509	80
1995	302,751	63	505	474	31	588	511	77
1996	280,972	64	502	472	30	588	513	75
1997–02	**	**	**	**	**	**	**	**
2003	312,004	64	503 (110)	473 (106)	30	603 (137)	526 (133)	77
2004	298,735	64	504 (110)	472 (106)	32	603 (137)	527 (134)	76
2005	312, 526	65	506 (111)	472 (107)	34	600 (138)	523 (135)	77
2006	321,726	65	504 (112)	470 (109)	34	598 (141)	521 (137)	77
2007	333,200	66	502 (113)	470 (109)	32	599 (141)	521 (138)	78

Note. GRE-Verbal and GRE-Quantitative scores range from 200 to 800, U.S. citizens only.
* Standard deviations, when available, are included in parentheses.
** Information/reports not available.

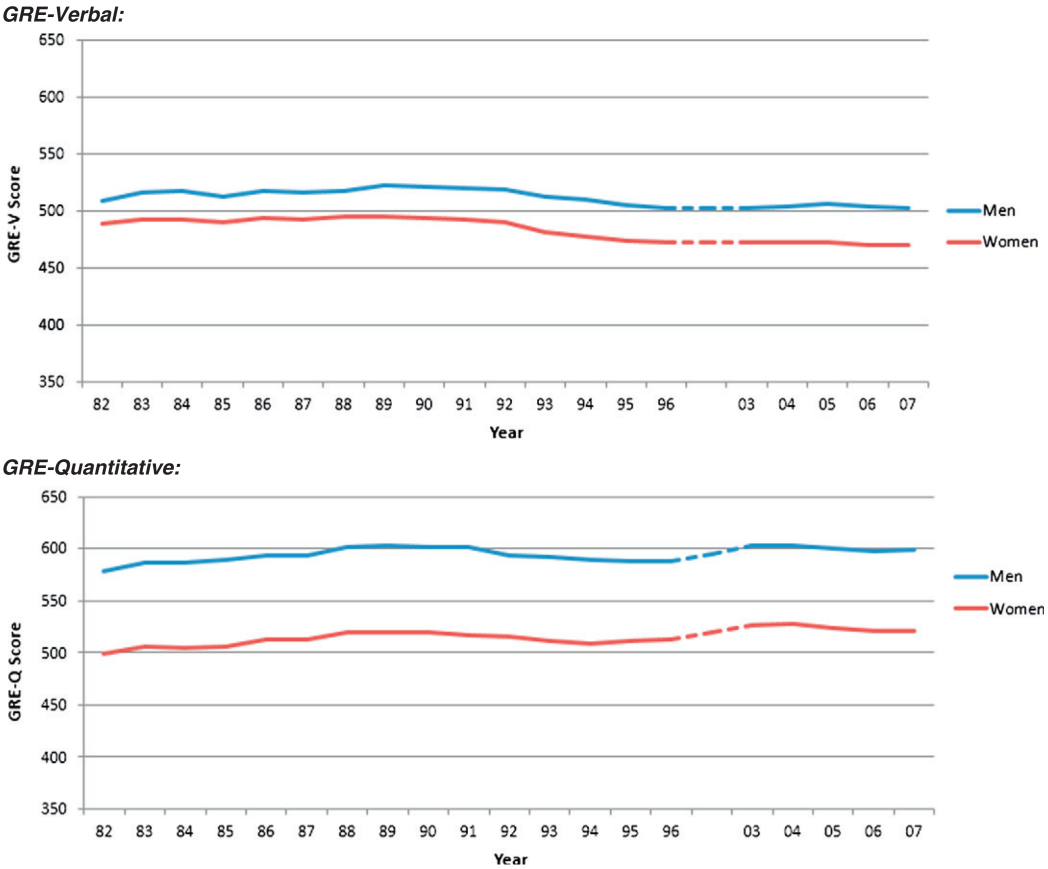

Fig. 1. Men's and women's mean GRE-Verbal scores (upper panel) and mean GRE-Quantitative scores (lower panel), by year. GRE-Verbal scores decreased slightly for both sexes in the 1990s and then plateaued, with a consistent 20–30 point difference favoring men across time. GRE-Quantitative scores increased slightly over time, with a consistent 75–80 point difference favoring men across time.

Table 2
Female representation in Master's and Doctoral graduate programs, 1986 to 2009.

	Master's programs				Doctoral programs			
Year	Engrg.	Math & CS	Physical & earth sciences	Across all disciplines	Engrg.	Math & CS	Physical & earth sciences	Across all disciplines
1986	12	26	21	44	10	17	18	33
1987	13	29	24	46	10	19	18	32
1988	14	28	25	45	11	16	19	32
1989	15	28	26	48	13	19	21	36
1990	14	29	25	46	11	18	21	34
1991	15	29	27	49	12	20	21	36
1992	15	30	28	50	12	19	23	36
1993	16	30	30	51	12	22	23	38
1994	17	28	31	51	14	21	23	39
1995	18	29	31	53	14	24	25	39
1996	18	28	34	53	15	20	24	40
1997	19	30	33	54	14	22	24	40
1998	20	29	35	52	15	23	26	40
1999	20	31	35	54	17	23	25	41
2000	21	33	35	55	15	23	26	42
2001	22	33	37	55	18	25	28	42
2002	22	33	38	56	20	27	29	44
2003	21	33	38	55	20	25	30	45
2004	22	32	41	57	19	25	30	46
2005	22	31	41	54	20	29	29	46
2006	23	31	42	59	20	26	32	47
2007	23	30	41	59	23	25	31	49
2008	23	32	41	59	22	27	31	49
2009	23	31	42	60	22	27	33	50

Note. Engrg. = Engineering, CS = Computer Science. Values reflect the percent of enrolled students identified as female.

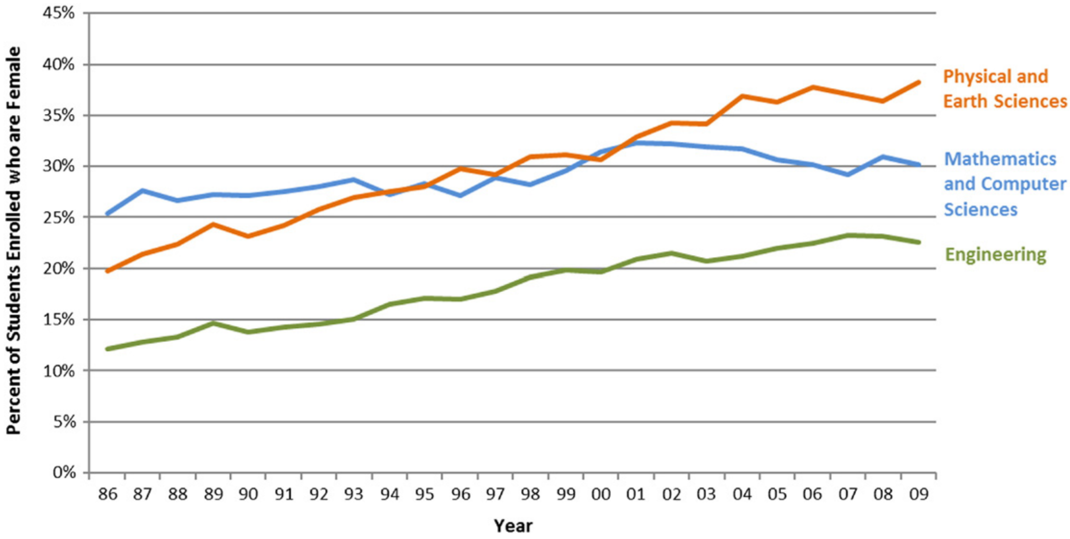

Fig. 2. Percent women in engineering, math and computer science, and physical sciences graduate programs (masters and doctoral programs combined), by year. Female representation in graduate STEM programs increased from 1986 to 2009.

historically disadvantaged groups, group discrepancies in GRE-Quantitative scores were large throughout these periods. Test takers who self-identified as Asian consistently earned higher GRE-Quantitative scores than did test-takers from any other ethnic group. In 1982, Asian test-takers scored 49 points higher on average than did White test-takers, who scored 171 points higher than Black test-takers; 25 years later, Asian test-takers scored 55 points higher than White test-takers, who scored 143 points higher than Black test-takers.

Although the GRE scores revealed persistent gaps between racial/ethnic groups, with deficits primarily among historically disadvantaged minority examinees, the graduate enrollments

Table 3
Representation of major ethnic groups among GRE test-takers, and their Verbal and Quantitative reasoning scores, 1982–1996 and 2003–2007*.

Year	White			Black			Mexican American			Puerto Rican			American Indian			Asian/Pacific		
	%	V	Q	%	V	Q	%	V	Q	%	V	Q	%	V	Q	%	V	Q
1982	84	512	535	6	367	364	1	416	422	1	387	421	.8	465	470	2	478	584
1983	83	516	542	6	371	369	1	426	441	1	389	430	.7	473	479	2	482	598
1984	83	517	540	6	378	375	1	429	442	1	391	430	.7	478	475	2	486	601
1985	84	514	538	6	380	378	1	427	434	1	387	429	.6	474	485	2	481	603
1986	84	518	543	5	386	387	1	437	451	1	398	443	.6	475	480	3	480	609
1987	84	518	542	6	388	392	1	441	456	1	392	444	.6	475	478	3	477	606
1988	78	520	547	6	391	394	1	444	461	1	397	445	.4	464	471	3	480	612
1989	78	521	544	6	393	397	1	442	458	1	408	452	.5	469	473	3	484	616
1990	80	520	543	6	396	399	1	453	470	1	410	454	.5	475	473	3	483	617
1991	81	519	543	6	395	404	2	450	469	1	408	455	**	**	**	3	486	612
1992	82	515	541	7	398	410	2	450	469	1	412	459	**	**	**	3	489	608
1993	82	507	539	7	393	408	2	443	470	1	413	454	.5	468	479	3	489	601
1994	80	503	537	8	390	405	2	442	470	1	409	452	.6	467	483	4	489	599
1995	79	498	537	8	390	407	2	439	466	1	413	461	.6	465	478	4	491	599
1996	75	496	538	8	389	409	2	438	469	1	407	465	.6	461	490	4	489	597
1997–02	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**
2003	76	496	566	9	396	425	2	432	492	1	407	487	1	454	507	5	490	637
		(103)	(130)		(93)	(133)		(99)	(137)		(101)	(137)		(106)	(140)		(120)	(128)
2004	75	496	568	9	394	426	2	430	489	1	409	488	1	454	499	5	493	631
		(103)	(130)		(93)	(134)		(99)	(136)		(103)	(141)		(102)	(141)		(120)	(130)
2005	75	497	565	9	396	422	3	436	491	1	409	476	1	459	500	5	493	627
		(104)	(131)		(93)	(137)		(102)	(142)		(105)	(143)		(102)	(136)		(120)	(133)
2006	75	495	563	9	394	418	3	430	484	1	414	476	1	450	498	5	489	621
		(106)	(133)		(94)	(137)		(101)	(141)		(107)	(140)		(106)	(141)		(120)	(136)
2007	75	493	562	9	395	419	3	431	485	1	411	469	1	446	497	6	485	617
		(106)	(135)		(95)	(139)		(102)	(143)		(106)	(144)		(109)	(141)		(121)	(138)

Note. Grandy (1999) did not report GRE scores for American Indian examinees in 1991 and 1992. GRE-Verbal and GRE-Quantitative scores range from 200 to 800. White = White, non-Hispanic, U.S. citizens only.
* Standard deviations, when available, are included in parentheses.
** Information/reports not available.

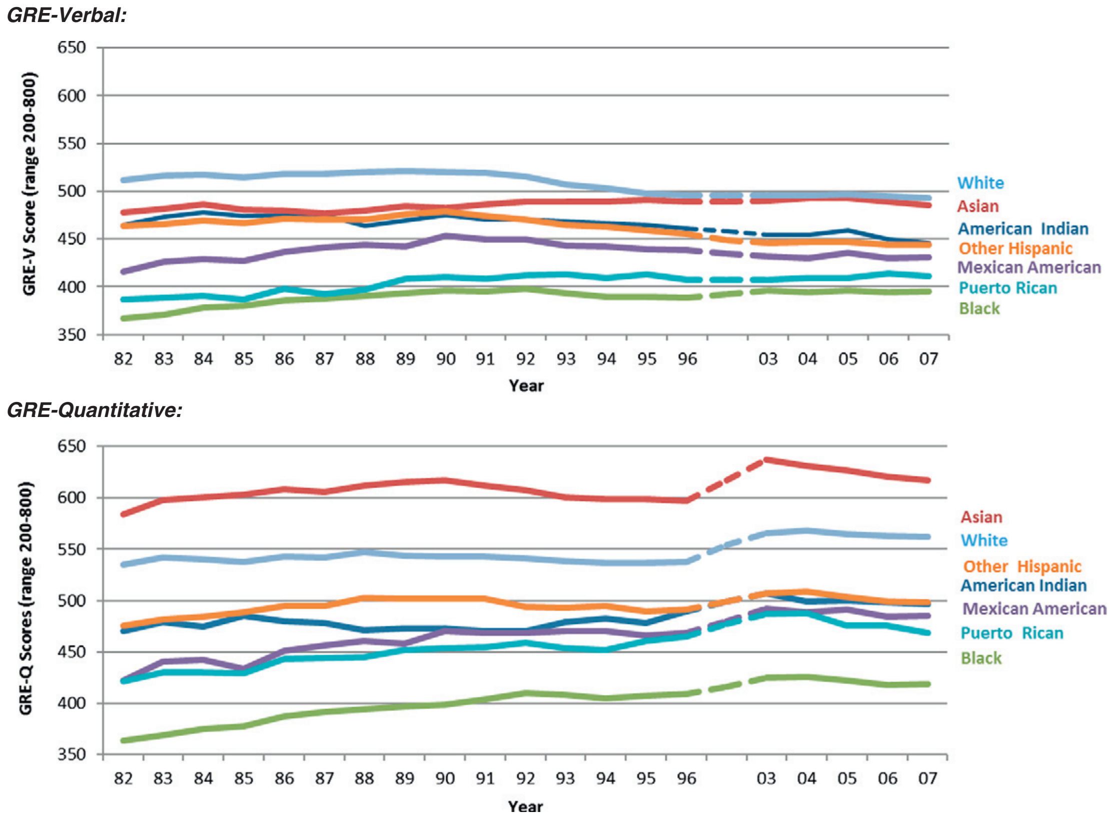

Fig. 3. Primary ethnic groups' mean GRE-Verbal scores (upper panel) and mean GRE-Quantitative scores (lower panel), by year.

from 1986 to 2009 tell a story of increasing diversity. Table 4 lists the percent of enrolled graduate students from each racial/ethnic group in successive years. Since 1986, Blacks, Hispanics, American Indians, and Asian/Pacific Islanders have made up increasing percentages of the graduate student population. Indeed, Table 3 shows that the GRE test-taking population has also become increasingly diverse. These findings coincide with other reports that have examined GRE scores and enrollment patterns. For example, one study that followed GRE examinees into the subsequent year (Grandy, 1990) reported gender and ethnicity differences in examinees' test scores but not in examinees' rate of actual enrollment in graduate school. Over the decade of 2000 to 2010, total graduate enrollment increased at a faster rate for all racial/ethnic minority groups than for Whites (Bell, 2011). Fig. 4 shows the representation of each major racial/ethnic group in graduate training, set against their representation in the resident U.S. population (ages 18 to 44). Because the large majority (over 90%) of GRE test-takers are between 18 and 44 (Grandy, 1999); and because most test-takers go on to pursue graduate training within a few years of taking the test, the numbers suggest that Hispanics are the only minority ethnic group that is currently clearly under-represented in graduate training programs as a whole. Although Fig. 4 suggests that some ethnic groups are currently not underrepresented in graduate programs overall, examinees from historically underrepresented groups are less likely to enroll in STEM programs and more likely to enroll in Business and Education (Grandy, 1999).

4. Discussion

4.1. Gender differences in GRE scores and representation in STEM

The GRE data compiled in this study show a persistent mean gap between the sexes of approximately half a standard deviation in quantitative reasoning. These data coincide with other data on differences between males and females at the upper tail of quantitative reasoning. On various tests of quantitative reasoning, including the math section of the National Assessment of Educational Progress and the quantitative sections of the SAT, ACT, and Cognitive Abilities Test, males outnumber females at the upper end of the distribution, and they outnumber females by larger magnitudes farther along the upper tail (Benbow & Stanley, 1980, 1983; Hedges & Nowell, 1995; Lakin, 2013; Wai, Cacchi, Putallaz, & Makel, 2010). Even among graduate students enrolled in top STEM programs around the country, men are more likely than women to have scored at ceiling (Lubinski, Benbow, Shea, Eftekhari-Sanjani, & Halvorson, 2001). Quantitative reasoning ability is related to the pursuit of advanced education in STEM disciplines (Lubinski & Benbow, 2006), as is spatial ability (Wai et al., 2009); and the predictive utility of quantitative reasoning ability for success in STEM disciplines is largest at the extreme right of the tail (Robertson, Smeets, Lubinski, & Benbow, 2010; see also Arneson, Sackett, & Beatty, 2011, in the context of general cognitive ability).

Table 4
Representation of primary races/ethnic groups in graduate programs (Master's and Doctoral combined), 1986–2009.

Year	White	Black	Hispanic	American Indian	Asian/Pacific
1986	88.2	5.0	3.2	0.5	3.2
1987	87.4	5.1	3.6	0.4	3.5
1988	86.8	5.4	3.7	0.4	3.6
1989	86.6	5.2	3.9	0.4	3.8
1990	86.2	5.4	4.0	0.4	3.9
1991	85.5	5.6	4.2	0.5	4.2
1992	85.0	5.9	4.2	0.5	4.5
1993	83.6	6.6	4.5	0.5	4.8
1994	82.9	6.7	4.7	0.5	5.1
1995	81.9	7.2	5.0	0.6	5.3
1996	81.1	7.8	5.1	0.6	5.5
1997	79.8	8.3	5.8	0.6	5.5
1998	78.6	8.4	6.4	0.6	5.9
1999	78.3	8.5	6.6	0.7	5.9
2000	77.2	9.3	6.9	0.7	5.9
2001	76.5	9.6	7.2	0.7	6.0
2002	75.9	10.1	7.1	0.7	6.3
2003	75.5	10.4	7.0	0.7	6.4
2004	74.1	11.0	7.7	0.7	6.4
2005	74.0	11.7	7.4	0.7	6.1
2006	73.2	12.8	7.9	0.8	6.3
2007	71.7	13.4	7.8	0.8	6.3
2008	71.4	13.3	8.1	0.8	6.4
2009	70.8	13.5	8.4	0.8	6.5

Note. Values are computed out of U.S. citizens and permanent residents with specified race/ethnicity (temporary residents and unknown race/ethnicity not included). Numbers may not add to exactly 100% due to rounding.

Indeed, validity analyses of the GRE have demonstrated associations between quantitative reasoning ability and success in STEM graduate training (Ayers & Quattlebaum, 1992; Holt, Bleckmann, & Zitzmann, 2006; Wang, 2013). Thus, our data suggest that the persistent gender difference in quantitative reasoning ability, at least at the upper tail of the distribution, needs to be a key part of discussions pertaining to women's underrepresentation in STEM disciplines (Templer & Tomeo, 2002).

In addition to outnumbering females at the extreme right tail of quantitative reasoning, males with strong cognitive abilities are more likely than their female counterparts to show quantitative *tilt*, such that their quantitative reasoning scores are higher than their verbal reasoning scores; females with strong cognitive abilities are more likely to show a balanced ability profile (Lubinski et al., 2001; Park, Lubinski, & Benbow, 2007). Further, males more often than females display vocational interests and preferences that are conducive to training in math and science, whereas females more often than males express multiple competing interests, including a strong interest in people-oriented professions (Lippa, 1998). Hence, women's different ability profiles and vocational preferences relative to men's also need to be considered in discussions of women's underrepresentation in STEM disciplines (Ceci & Williams, 2011; Halpern et al., 2007).

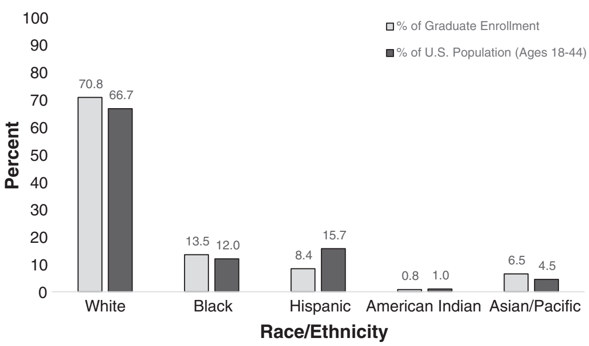

Fig. 4. Primary ethnic group representation in graduate training (masters and doctoral programs combined) in 2009, set against their representation in the U.S. resident population (ages 18 to 44) in 2009. Population representation values taken from the U.S. Census Bureau, Vintage 2009 National Tables at https://www.census.gov/popest/data/historical/2000s/vintage_2009/. Percentages were calculated from the following tables: NC-EST2009-04-WA, NC-EST2009-04-BA, NC-EST2009-04-IA, NC-EST2009-04-AA, and NC-EST2009-04-HISP.

There is widespread concern about female representation in the inorganic sciences, technology, engineering, and math (Fox, 2001; Hill, Corbett, & St. Rose, 2010). That concern notwithstanding, the CGS/GRE data on graduate enrollments suggest that women's representation in STEM disciplines has steadily increased since at least the 1980s. Other records on doctoral training, in particular, indicate that the number of women earning doctorates in physical sciences increased 70% from 1999 to 2009, and the number of female engineering doctoral recipients more than doubled over that decade (National Science Foundation & Division of Science Resources Statistics, 2010). These growth rates were three times larger than the growth in male doctorate recipients during the same period (National Science Foundation & Division of Science Resources Statistics, 2010). By 1995, women were earning over 50% of the doctorates in the social and behavioral sciences (National Research Council, 2001). In fact, in 2009 women were near parity or had actually achieved a majority in every broad field except the physical (inorganic) sciences, math, and engineering (National Science Foundation & Division of Science Resources Statistics, 2010).

4.2. Ethnic differences in GRE scores and representation in STEM

The GRE data compiled in this study show persistent mean gaps between ethnic groups on both verbal and quantitative reasoning. The small gap between White and Asian examinees in GRE-Verbal scores diminished by the 1990s, but the gap between White and Asian examinees in GRE-Quantitative scores has persisted. In Verbal reasoning, White and Asian examinees currently score a full standard deviation above Black examinees and a half standard deviation higher than examinees from other underrepresented groups; in Quantitative reasoning, Asian examinees score about a one-third standard deviation higher than White examinees and well over one standard deviation higher than Black examinees. The Asian–White–Black pattern of ethnic score gaps in GRE scores runs parallel to achievement score gaps reported by Humphreys (1988), and the White–Black gap in GRE scores runs parallel to the persistent White–Black gap in general cognitive ability scores reviewed by Rushton and Jensen (2005). In a review of achievement and ability data from seven different probability samples, Hedges and Nowell (1999) concluded that White–Black disparities in representation at the upper-tails of cognitive ability have not diminished since the 1960s. The ethnic differences in GRE scores reported in the current analysis reinforce that conclusion.

Over the past several decades, numerous programs have worked toward increasing minority representation in higher education, and especially in STEM disciplines (U.S. Government Accountability Office, 2012). The CGS/GRE enrollment data indicate that efforts have been successful: Historically underrepresented groups are increasingly represented in graduate training programs overall. Between 1986 and 2009, American Indian representation in graduate training approached equivalence to its representation in the U.S. resident population (1%); Hispanic representation went from 3% to 8%; and Black representation in graduate training programs went from 5% to 13.5%. Thus, although Blacks are underrepresented in STEM programs (Garrison, 2013), Fig. 4 shows that they are not underrepresented in higher education as a whole. This finding is somewhat surprising given that Blacks consistently scored far below Whites and Asians on both Quantitative and Verbal sections of the GRE, and modestly below other underrepresented groups on the Quantitative section.

4.3. A paradox?

The persistence of gender and ethnic differences in GRE performance coupled with the decline, if not elimination, of gender and ethnic "gaps" in graduate-school representation raises an obvious question: How can these two phenomena co-exist? If the predictors of academic success have not changed but the profile of admitted students has, does this mean that students with less promise of academic success are being admitted to further diversity goals? The GRE is, of course, just one commonly employed predictor of academic success. Undergraduate GPA is another. There are yet other predictors, such as personality (O'Connor & Paunonen, 2007; Rigdon & Kuncel, 2010), that are not generally relied upon, at least in a direct way. Thus, it is possible that relative weakness on the GRE is offset by relative strength in other predictors (e.g., conscientiousness; see Moutafi, Furnham, & Crump, 2006), such that gender and ethnic parity reflect an actual parity in academic promise.

It does not appear, however, that parity in admissions is a simple reflection of parity in qualifications. Attiyeh and Attiyeh (1997) analyzed graduate-school applications at 48 institutions and found that in three of the five fields studied, women were given "modest" preferences, and in all five fields, minorities were given "substantial" ones. The study looked not only at GRE scores (both General and Subject Matter), but also undergraduate grade-point average and undergraduate institution. Given the magnitude of these preferences, it is not surprising that after a number of states outlawed affirmative-action preferences, they saw substantial declines in minority representation in graduate programs, especially in STEM fields (Garces, 2013).

One might expect that if the GRE is a valid predictor of graduate school success, and if universities accept applicants with lower GRE scores for demographic reasons, then those students may not have the same success as students with stronger records even if they are admitted in equal proportions. This is especially a concern in STEM disciplines, because gender and ethnic gaps in quantitative ability increase in magnitude farther along the right tail (Hedges & Nowell, 1999; Lakin, 2013; Wai et al., 2010), and exceptional levels of achievement in STEM are closely linked to exceptional levels of quantitative reasoning ability (Park, Lubinski, & Benbow, 2008). If women in STEM careers are highly able but still discrepant, on average, from their male counterparts, one would expect average gender differences in rates of tenure, publication, citation counts, and funding (Ceci & Williams, 2011). The same logic applies to ethnic differences, although perhaps to a larger degree given the magnitude of the test-score discrepancies. If not understood, these discrepancies could lead to unwarranted perceptions of discrimination. As noted by Ceci and Williams (2011), gender differences in rates of tenure, publication, citation counts, and grant funding in STEM disciplines are tied to women's abilities and preferences; and interventions that focus on discrimination as the primary cause are unlikely to be successful.

Ability distribution differences are also key for understanding concerns about attrition of minorities from the academic pipeline (Griffith, 2010). The foundation for this concern is buttressed by Garrison's (2013) finding that racial disparities in STEM fields, including in graduate education, are more a product of differences in graduation rate than matriculation rate. Suggesting that attrition rates are influenced by qualifications, Baker (1998) found that substantial race differences in Ph.D. completion rates disappear once measures of "ability" (GRE scores, GPA, and NSF Graduate Fellowship panel evaluations) are controlled for. Degree completion is not the only outcome measure that varies by ethnic group. Price and Price (2006) found, for example, that minority graduate students in humanities and social sciences are less likely than non-minorities to publish as graduate students or within three years of finishing graduate school.

Whether the benefits of diversity are worth the costs potentially incurred by less academically promising students is a policy question, not a scientific one. However, policies pursued in the name of diversity may actually produce negative effects on diversity itself. Elliott and colleagues (1996) concluded that the most important reason for Black attrition from science programs was their low level of preparation — not in absolute terms but relative to their colleagues. Black students with SAT scores that would put them near the top of Black performers, but perhaps below average among White performers, are often admitted to the most elite schools, where they are in competition with students who have shown substantially more promise of academic success in the sciences. Consequently, large numbers of Black students in this situation find themselves falling behind and either switch to a less-competitive course of study or withdraw from the school altogether. Elliot et al. point out that elite schools, which are often in a position to attract the most talented Black students, have not been a major source of Black science Ph.D.s. Instead, those have come primarily either from historically Black colleges and universities or from other non-elite schools, where the Black students' qualifications are more competitive with those of their colleagues. Consistent with the view that relative qualifications are important determinants of success, Lott, Gardner, and Powers (2009) found in a study of a major public university that an individual's GRE score relative to others in the program was a much stronger predictor of persistence than was an individual's absolute score. Sander and Taylor (2012) report a similar phenomenon in California after the passage of Proposition 209, which banned affirmative-action preferences. The number of minorities admitted to the University of California went down, but graduation rates went up, so much so that the *absolute number* of Blacks and Hispanics earning bachelor's degrees actually increased, as did the number of such students earning degrees in STEM fields. Similarly, among graduate students, the UC system experienced a decline in Black and Hispanic enrollments but an increase in the number of graduate degrees awarded to them. Given these findings, it seems important for those making admissions decisions with an eye toward diversity to understand that individuals' scores on assessments like the GRE reflect their likelihood of thriving in high-intensity intellectual pursuits (Kuncel & Hezlett, 2007; Kuncel et al., 2001).

5. Conclusion

A recent report on the future of graduate education, prepared jointly by the Educational Testing Service and the Council of Graduate Schools, offered these introductory words: "Our competitiveness in the global economy hinges on our ability to produce sufficient numbers of graduate-degree holders — people with the advanced knowledge and critical-thinking abilities to devise solutions to grand challenges such as energy independence, affordable health care, climate change and others. One of our greatest resources is our human talent, and as a nation we must invest in educating more of our population at the graduate level to ensure our capacity to innovate and to secure our intellectual leadership into the future" (Wendler et al., 2010, p. 4). Indeed, graduate enrollments have grown consistently over recent decades, and since the year 2000, in particular, total enrollment growth has been stronger for minorities than for Whites, and stronger for women than for men (Bell, 2011). This would not occur were the GRE selectively suppressing the admission of women and minorities. In fact, the enrollment numbers from 1986 to 2009 show an ever more ethnically and gender diverse graduate student body, despite persisting GRE test score gaps.

That said, the GRE forecasts success in graduate training and degree completion. It therefore forecasts ethnic and gender differences in grades or degrees obtained, particularly in STEM disciplines, because group gaps in quantitative reasoning are large and success is tied to quantitative reasoning. We suggest that future efforts for obtaining both a diverse and talented student body should take into account the persistent gender differences and ethnic differences in GRE scores, so that those who enroll in graduate training profit from it as intended and subsequently provide the intellectual leadership that such training promises.

References

Angoff, W. H., & Johnson, E. G. (1990). The differential impact of curriculum on aptitude test scores. *Journal of Educational Measurement*, *27*, 291–305.

Arneson, J. J., Sackett, P. R., & Beatty, A. S. (2011). Ability–performance relationships in education and employment settings: Critical tests of the more-is-better and the good-enough hypotheses. *Psychological Science*, *22*, 1336–1342.

Attiyeh, G., & Attiyeh, R. (1997). Testing for bias in graduate school admissions. *Journal of Human Resources*, *32*, 524–548.

Ayers, J. B., & Quattlebaum, R. F. (1992). TOEFL performance and success in a masters program in engineering. *Educational and Psychological Measurement*, *52*, 973–975.

Baker, J. G. (1998). Gender, race and Ph.D. completion in natural science and engineering. *Economics of Education Review*, *17*, 179–188.

Bell, N. E. (2008). *Graduate enrollment and degrees: 1997–2007.* Washington, DC: Council of Graduate Schools.

Bell, N. E. (2009). *Graduate enrollment and degrees: 1998–2008.* Washington, DC: Council of Graduate Schools.

Bell, N. E. (2010). *Graduate enrollment and degrees: 1999–2009.* Washington, DC: Council of Graduate Schools.

Bell, N. E. (2011). *Graduate enrollment and degrees: 2000–2010.* Washington, DC: Council of Graduate Schools.

Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, *210*, 1262–1264.

Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science*, *222*, 1029–1031.

Bridgeman, B., Burton, N., & Cline, F. (2008). *Understanding what the numbers mean: A straightforward approach to GRE predictive validity.* ETS GRE Board Research Report No. 04-03, ETS RR-08-46. Princeton, NJ: Educational Testing Service.

Brown, H. A. (2006). *Graduate enrollment and degrees: 1986–2005.* Washington, DC: Council of Graduate Schools.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, England: Cambridge University Press.

Carvajal, H., & Pauls, K. K. (1995). Relationships among Graduate Record Examination scores, Wechsler Adult Intelligence Scale — Revised IQs, and undergraduate grade point average. *College Student Journal*, *29*, 414–416.

Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*, *108*, 3157–3162.

Cole, N. S. (1998). Merit and opportunity: Testing and higher education at the vortex. *The GRE, FAME Report Series*, *Vol. 1*, Princeton, NJ: Educational Testing Service.

Dutka, J. T. (1999). Beyond inclusion: Multicultural perspectives in the service of transformation. *The GRE, FAME Report Series*, *Vol. 3*, Princeton, NJ: Educational Testing Service.

Educational Testing Service (2004). *Factors that can influence performance on the GRE General Test: 2002–2003.* Princeton, NJ: Educational Testing Service.

Educational Testing Service (2005). *Factors that can influence performance on the GRE General Test: 2003–2004.* Princeton, NJ: Educational Testing Service.

Educational Testing Service (2007a). *Factors that can influence performance on the GRE General Test: 2004–2005.* Princeton, NJ: Educational Testing Service.

Educational Testing Service (2007b). *Factors that can influence performance on the GRE General Test: 2005–2006.* Princeton, NJ: Educational Testing Service.

Educational Testing Service (2008). *Factors that can influence performance on the GRE General Test: 2006–2007.* Princeton, NJ: Educational Testing Service.

Elliott, A. R., Strenta, A. C., Adair, R., Matier, M., & Scott, J. (1996). The role of ethnicity in choosing and leaving science in highly selective institutions. *Research in Higher Education*, *37*, 681–709.

Fox, M. F. (2001). Women, science, and academia: Graduate education and careers. *Gender & Society*, *15*, 654–666.

Garces, L. M. (2013). Understanding the impact of affirmative action bans in different graduate fields of study. *American Educational Research Journal*, *50*, 251–284.

Garrison, H. (2013). Underrepresentation by race-ethnicity across stages of U.S. science and engineering education. *CBE-Life Sciences Education*, *12*, 357–363.

Grandy, J. (1990). *Comparison of expected with actual field of graduate study: An analysis of GRE survey data.* Princeton, NJ: Educational Testing Service (GRE No: 87-02P, RR-90-17).

Grandy, J. (1999). *Trends and profiles: Statistics about GRE general test examinees by gender, age, and ethnicity* (2nd ed.). Princeton, NJ: Educational Testing Service (GRE No: 96-07, RR-99-16).

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, *29*, 911–922.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, *8*, 1–51.

Hartnett, R. T., & Oltman, P. K. (1984). *The role of GRE general and subject test scores in graduate program admission.* Princeton, NJ: Educational Testing Service (GREB-81-08R, RR-84-14).

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*, 41–45.

Hedges, L. V., & Nowell, A. (1999). Changes in the Black–White gap in achievement test scores. *Sociology of Education*, *72*, 111–135.

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics.* Washington, DC: American Association of University Women.

Holt, D. T., Bleckmann, C. A., & Zitzmann, C. C. (2006). The Graduate Record Examination and success in an engineering management program: A case study. *Engineering Management Journal*, *18*, 10–16.

Hsu, S. D. H., & Schombert, J. (2010, April). Data mining the university: College GPA predictions from SAT scores. Available at SSRN: http://ssrn.com/abstract=1589792

Humphreys, L. G. (1988). Trends in levels of academic achievement of Blacks and other minorities. *Intelligence*, *12*, 231–260.

Krueger, A. B., & Wu, S. (2000). Forecasting job placements of economics graduate students. *Journal of Economic Education*, *31*, 81–94.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science*, *315*, 1080–1081.

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, http://dx.doi.org/10.1037/0033-2909.127.1.162.

Kuncel, N. R., Wee, S., Serafin, L., & Hezlett, S. A. (2010). The validity of the graduate record examination for master's and doctoral programs: A meta-analytic investigation. *Educational and Psychological Measurement*, *70*, 340–352, http://dx.doi.org/10.1177/0013164409344508.

Lakin, J. M. (2013). Sex differences in reasoning abilities: Surprising evidence that male–female ratios in the tails of the quantitative reasoning distribution have increased. *Intelligence*, *41*, 263–274, http://dx.doi.org/10.1016/j.intell.2013.04.004.

Lippa, R. A. (1998). Gender-related individual differences and the structure of vocational interests on the people–things dimension. *Journal of Personality and Social Psychology*, *74*, 996–1009.

Lott, J. L., Gardner, S., & Powers, D. A. (2009). Doctoral student attrition in the STEM fields: An exploratory event history analysis. *Journal of College Student Retention*, *11*, 247–266.

Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math–science expertise. *Perspectives on Psychological Science*, *1*, 316–345.

Lubinski, D., Benbow, C. P., Shea, D. L., Eftekhari-Sanjani, H., & Halvorson, M. B. J. (2001). Men and women at promise for scientific excellence: Similarity not dissimilarity. *Psychological Science*, *12*, 309–317.

Lubinski, D., Benbow, C. P., Webb, R. M., & Bleske-Rechek, A. (2006). Tracking exceptional human capital over two decades. *Psychological Science*, *17*, 194–199.

Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score? *Learning and Individual Differences*, *16*, 31–42, http://dx.doi.org/10.1016/j.lindif.2005.06.003.

National Research Council (2001). *From scarcity to visibility: Gender differences in the careers of doctoral scientists and engineers.* Washington, DC: National Academy Press.

National Science Foundation, & Division of Science Resources Statistics (2010). *Doctorate recipients from U.S. universities: 2009.* Special report NSF 11-306. Arlington, VA: National Science Foundation (Available at http: www.nsf.gov/statistics/nsf11306/).

Norcross, J. C., Haynch, J. M., & Terranova, R. D. (1996). Graduate study in psychology: 1992–1993. *American Psychologist*, *51*, 631–643.

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, *43*, 971–990.

Park, G., Lubinski, D., & Benbow, C. P. (2007). Contrasting intellectual patterns for creativity in the arts and sciences: Tracking intellectually precocious youth over 25 years. *Psychological Science*, *18*, 948–952.

Park, G., Lubinski, D., & Benbow, C. P. (2008). Ability differences among people who have commensurate degrees matter for scientific creativity. *Psychological Science*, *19*, 957–961.

Powers, D. E. (1985). Effects of coaching on GRE Aptitude Test scores. *Journal of Educational Measurement*, *22*, 121–136.

Price, J., & Price, J. (2006). Citizenship, gender, and racial differences in the publishing success of graduate students and young academics. *Cornell University School of Industrial and Labor Relations Working Papers*.

Pruitt, A. S. (1998). Human diversity, graduate education, and challenges for assessment. *The GRE, FAME Report Series*, *Vol. 1*, Princeton, NJ: Educational Testing Service.

Redd, K. E. (2007). *Graduate enrollment and degrees: 1996–2006.* Washington, DC: Council of Graduate Schools.

Rigdon, J. L., & Kuncel, N. R. (2010, Augustt). *Using personality to predict graduate student performance: A meta-analysis.* San Diego, CA: Paper presented at the 118th Annual Convention of the American Psychological Association.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*, 313–345.

Robertson, K. F., Smeets, S., Lubinski, D., & Benbow, C. (2010). Beyond the threshold hypothesis: Even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Current Directions in Psychological Science*, *19*, 346–351.

Rock, D. A., Werts, C., & Grandy, J. (1982). *Construct validity of the GRE aptitude test across populations: An empirical confirmatory study.* GRE Board Professional Report No. 78-1P, ETS Research Report 81-57. Princeton, NJ: Educational Testing Service.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235–294.

Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist*, *63*, 215–227.

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary performance? *Psychological Bulletin*, *135*, 1–22, http://dx.doi.org/10.1037/a0013978.

Sander, R. H., & Taylor, S. (2012). *Mismatch: How affirmative action hurts students it's intended to help, and why universities won't admit it.* New York: Basic Books.

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational measurement* (pp. 263–332) (3rd ed.). New York, NY: Macmillan.

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology*, *16*, 93–102, http://dx.doi.org/10.1037/a0026617.

Stricker, L. J., & Rock, D. A. (1987). Factor structure of the GRE general test in young and middle adulthood. *Developmental Psychology*, *23*, 526–536.

Templer, D. I., & Tomeo, M. E. (2002). Mean Graduate Record Examination (GRE) score and gender distribution as a function of academic discipline. *Personality and Individual Differences*, *32*, 175–179.

Toyama, J. S. (1999). What will be the role of the GRE in graduate admissions decisions in the 21st century? *The GRE, FAME Report Series*, *Vol. 3*, Princeton, NJ: Educational Testing Service.

U.S. Government Accountability Office (2012). *Science, technology, engineering, and mathematics education: Strategic planning needed to better manage overlapping programs across multiple agencies.* GAO-12-108, Washington, DC: United States Government Accountability Office (http://gao.gov/products/GAO-12-108 (accessed 23 March 2014)).

Wai, J., Cacchi, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence*, *38*, 412–423, http://dx.doi.org/10.1016/j.intell.2010.04.006.

Wai, J., Lubinski, D., & Benbow, C. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*, 817–835, http://dx.doi.org/10.1037/a0016127.

Walters, A. M., Lee, S., & Trapani, C. (2004). *Stereotype threat, the test-center environment, and the performance on the GRE General Test, GRE Board Research Report No. 01-03R.* Princeton, NJ: Educational Testing Service.

Wang, W. (2013). *Testing the validity of GRE scores on predicting graduate performance for engineering students.* (Unpublished master's thesis). University of Nebraska.

Wendler, C., Bridgeman, B., Cline, F., Millett, C., Rock, J., Bell, N., et al. (2010). *The path forward: The future of graduate education in the United States.* Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1988). *A study of the long-term stability of GRE general test scores.* GRE Board Research Report No. 86-18R, ETS Research Report 88-39. Princeton, NJ: Educational Testing Service.