

Examinees Retaking the Graduate Management Admission Test[®]

Lawrence M. Rudner

GMAC[®] Research Reports • RR-05-01 • March 17, 2005

The decision to retake an admissions test is not an easy one. The test taker must weigh the required time and money against the likelihood of significant gains. Decisions then need to be made about how to prepare for a second or third testing. The difficulty of the decision is compounded by the fact that there is little objective and almost no current information regarding the gains a serial examinee can expect.

This paper provides information on the Graduate Management Admission Test[®] (GMAT[®]), based on analysis of approximately 28,000 examinees first taking the GMAT[®] in 2003 and retaking the test at least once between February 2003 and August 2004. Differences between test takers who only take the GMAT[®] once and those who take the test multiple times are explored. Average gains for different groups of retakers are identified.

Background

The Graduate Management Admission Council[®], publisher of the GMAT[®], urges examinees to adequately prepare for the test. As Powers and Alderman (1983) so aptly point out, the concerns of test publishers are twofold: “a) to prevent sophisticated examinees from ‘beating the test,’ and b) to keep less experienced examinees from ‘being beaten’ by it.” Good test-construction principles help minimize the first threat. To prevent examinees from being “beaten,” a variety of test-familiarization materials, including official guides, sample tests, and free software, are provided. Because the GMAT[®] is computer adaptive and employs unique item types and content, a test taker who is unfamiliar with the format and general structure of the test may not obtain a score that reflects their true likelihood of success in business school.

In the years following the protracted Federal Trade Commission coaching study (1979), there have been many investigations concerning test preparation, especially for admissions tests. Messick and Jungeblut (1981) and Slack and Porter (1980) are perhaps the most frequently cited studies of that era. Messick and Jungeblut found that the effectiveness of test preparation is primarily a function of the amount of time invested by the test taker. Averaged across all investigated conditions, Slack and Porter found a score increase of about 33 points (a third of a standard deviation), on the 200- to 800-point scale for the SAT[®].

More recently, Becker (1990) provided a review of SAT[®] coaching studies and found that reported gains were mostly a function of the research design. Confirming DerSimonian and Laird (1983), Becker found that studies that only examined gains for students who had attended formal test-preparation programs had mean gains of about 40 points on the SAT[®] Verbal scale and about 50 points on the SAT[®] Math scale. Studies with a more rigorous design that examined gains for a group of students enrolled in a formal program compared to gains of a control group not enrolled in a formal program showed much smaller increases of only 10 to 15 Math or Verbal SAT[®] score points.

Test preparation can range from simply reviewing a few sample items to taking a lengthy, formal test-preparation course. In an extremely large study involving more than 650,000 Law School Admission Test[®] (LSAT[®]) test takers from 1991 to 1997, Thornton (1998) examined eight different forms of test preparation. Every form of preparation, including review of the LSAT[®] Information Booklet, led to higher examinee performance than no preparation. Multiple forms of preparation were better than single forms. Preparation closest to the actual test administration, such as studying items from the official guide, tended to be the single most effective method of

preparation. As with Slack and Porter, the average gain appeared to be about one-third of a standard deviation. A logical extension of the Thornton study is that, for those candidates retaking an admissions test, prior actual test experience is the best form of test preparation. The candidate obtains experience with the look, feel, and content of the test under actual conditions.

Studies of coaching, however, typically lack adequate control groups. Some examine score differences between coached and uncoached students on one test administration. Others examine gains across test administrations for coached students, but lack a comparison group. As Powers (1993, and Rock 1999) repeatedly points out, many of the investigations into coaching do not provide credible scientific evidence in support of their stated claims.

Unfortunately, there is a lack of published literature using examinees as their own control group. Use of such a repeated measures design provides a direct examination of the effects of test practice. Early studies (Lane, Penn, and Fischer 1966; Moore 1966) and a more recent study (Kaufman 1990) have found that simply retaking some intelligence tests can result in increased scores. Because intelligence tests purport to measure enduring traits, practice effects can be problematic. A logical inference from these studies is that other tests are susceptible, to some degree, to practice effects.

Nathan and Camara (1998) provide an extensive analysis of gains for high school students retaking the SAT® I. Very modest mean gains of 7 to 13 points on the Verbal scale and 8 to 16 points on the Math scale, about a 10th of a standard deviation, were observed. Nathan and Camara noted that the gains may be due to a variety of factors, including regression toward the mean, academic growth, and practice effects.

Methodology

A database of 463,630 unique candidates taking the GMAT® between January 1, 2002, and August 31, 2004, built from the daily item-response data feed formed the initial dataset for this study. The database was sorted by last name, first name, and date of birth. Examinees with the same name and date of birth taking the test multiple times were considered test retakers. The test dates for each candidate were then sorted to identify the test-taking

order. This report analyzes the 161,221 unique candidates first taking the GMAT® in the 2003 calendar year.

Candidates taking the GMAT® are asked to complete an optional Background Information Questionnaire (BIQ). Several variables were constructed from the examination and BIQ data:

- Number of tests: A binary variable indicating whether an examinee took the GMAT® once during the time period or multiple times.
- Gender: A binary variable indicating male or female.
- Ethnicity: A binary variable indicating white/non-Hispanic or other. Ethnicity data was only available for United States citizens.
- Native Language: A binary variable indicating English or a language other than English.
- Finished Quant: A binary variable indicating whether the candidate finished the Quantitative portion of the GMAT®. Of those who did not finish the Quantitative section, the majority (54%) had only one or two unanswered items remaining at the end of the testing period.
- Finished Verbal: A binary variable indicating whether the candidate finished the Verbal portion of the GMAT®. The majority of those who did not finish the Verbal section (55%) had one to three unanswered items remaining at the end of the testing period.
- Age: A grouping variable with three equally sized groups. The youngest group was 25 or younger on the day of testing, the middle group was 26 to 29 years old, and the oldest was 30 years old or older.
- Undergraduate Grade Point Average: A continuous variable containing self-report data. This variable was recoded to a grouping variable with three equally sized groups representing the lowest to highest UGPA.
- Discrepancy: A continuous variable constructed by subtracting the normalized first administration GMAT® Total score from the normalized continuous UGPA. Examinees with positive discrepancy values have UGPAs that are high relative to their GMAT® Total scores. A grouping variable with three equally sized groups was formed from the discrepancy variable.

Results

One-time and multiple-time test takers

Of the 161,221 unique test takers during the time period, 27,763 (about 17.5%) took the GMAT® more than once. As shown in Table I, repeat test takers have lower GMAT® Total and GMAT® Verbal scores and slightly

lower GMAT® Quantitative scores. There is no difference between the groups in their Undergraduate Grade Point Averages.

Table I GMAT® scores and UGPA for one-time and multiple-time test takers

	One time			Multiple times		
	N	Mean	sd	N	Mean	sd
GMAT® Quant.	133,448	34.2	10.4	27,763	33.7	10.2
GMAT® Verbal	133,448	28.4	9.0	27,763	24.6	8.2
GMAT® Total	133,448	529.2	119.9	27,763	496.2	102.7
Undergraduate GPA	117,668	3.2	0.5	23,059	3.2	0.4

Note: Ns are not equal because UGPA is self-reported and has missing data.

There are meaningful differences in a variety of background characteristics between one-time GMAT® test takers and candidates who take the test multiple times. As shown in Table 2, higher percentages of non-White and non-native English speakers retake the GMAT® compared to one-time-only test takers. Repeat test takers are also

less likely to have finished either the Quantitative or Verbal sections and more likely to have high self-reported UGPAs relative to their GMAT® Total scaled scores. There are no meaningful differences between genders for one-time test takers and repeat test takers.

Table 2 Characteristics of GMAT® examinees taking the test one time and multiple times			
		One time	Multiple times
Gender	N	131,148	27,072
	Male	61.0%	60.2%
	Female	39.0%	39.8%
Ethnicity	N	77,896	12,577
	White non-Hispanic	75.1%	64.8%
	Other	24.9%	35.2%
Native Language	N	129,378	26,702
	English	63.7%	48.2%
	Non-English	36.3%	51.8%
Finished Quant.	N	113,730	28,106
	Finished	84.9%	80.3%
	Did not finish	15.1%	19.7%
Finished Verbal	N	133,569	27,919
	Finished	90.7%	84.2%
	Did not finish	9.3%	15.8%
Discrepancy	N	117,654	23,044
	UGPA>GMAT	32.5%	39.6%
	UGPA<GMAT	34.3%	25.2%
Note: Ns are not equal because some variables are self-reported and have missing data.			

Time between Testing

Examinees are permitted to take the GMAT® only once in a calendar month and only five times in a 12-month period. Thus, a candidate can take the GMAT® at the end of one month and then again at the beginning of the next month. Table 3 shows the distribution of the time between the first and second

test sitting for candidates retaking the GMAT® between February 2003 and August 2004. Almost a majority of candidates who retake the GMAT® do so within 60 days of the first administration. The distribution is quite skewed with a median of 56 and an interquartile range of 32 to 119 days.

Table 3: Number of days between first and second testing

	I–30	31–60	61–90	91–120	121–150	151–180	> 180
%	20.4	27.4	13.6	8.7	6.3	4.6	15.0
cumulative %	20.4	47.8	61.4	70.1	81.0	85.0	100.0

Overall Gains

GMAT® scores are reported in terms of—

1. scaled scores, which range from 200 to 800 and 11 to 51 for the GMAT® Total and the GMAT® Quantitative and Verbal tests, respectively; and
2. percentiles, which represent the percent of examinees scoring below a given scaled score during the prior three years.

Repeat test takers were able to obtain modest gains in their GMAT® scaled and percentile scores. As shown in Table 4, the mean increase in the GMAT® Total score was about 31 scaled-score points and 8 percentile points. This is less than one-third of a standard deviation and, on average, not much of an improvement. The standard deviation of the gain, however, indicates that there is substantial variability in the observed gains. Quantitative scores tended to show greater improvement than Verbal scores, but the average gains for both sections were only about 2 scaled-score points, which represents only 6 or 7 percentile-score points.

Table 4 Overall gains in GMAT® scores from first to second administration

	Mean gain	Standard deviation
GMAT® Total score	31.1	56.6
GMAT® Total percentile	8.2	15.0
GMAT® Quantitative scaled score	2.4	5.3
GMAT® Quantitative percentile	6.8	14.7
GMAT® Verbal scaled score	1.8	5.3
GMAT® Verbal percentile	6.2	17.1

Note: Based on 27,763 repeat test takers.

The variability in gains is shown by the cumulative frequency distribution of the gains in GMAT® Total scaled scores shown in Table 5. Nearly one-quarter of all retakers actually score worse the second time. About half

observe a gain between 0 and 60 points on their second testing. Only slightly more than 10% of all retakers witness a gain of 100 points or more.

Table 5 Cumulative frequency distribution of GMAT® Total score gains			
Score	Cumulative frequency	Score	Cumulative frequency
<-100	1.0	40.00	60.9
-100.00	1.4	50.00	67.8
-90.00	2.0	60.00	74.1
-80.00	2.8	70.00	79.6
-70.00	3.9	80.00	84.3
-60.00	5.4	90.00	88.2
-50.00	7.7	100.00	91.5
-40.00	10.6	110.00	93.8
-30.00	14.5	120.00	95.6
-20.00	19.2	130.00	96.9
-10.00	24.5	140.00	97.8
0.00	31.3	150.00	98.4
10.00	38.0	160.00	98.9
20.00	45.7	>160	100.0
30.00	53.5	—	—

Note: Based on 27,763 repeat test takers.

Group Differences in Gains

Table 6 shows the gains in GMAT® Total, Quantitative, and Verbal scaled scores by gender, native/first language, and age group. Just as there are no meaningful gender differences between those examinees who retake the GMAT® and those who do not, there are no meaningful gender differences in score gains. The gains for male and female examinees retaking the GMAT® were virtually the same on all scaled scores.

Slightly more than half of the examinees who retake the GMAT® do not have English as their first or native language. Their gains were about the same as those of test takers whose first language is English. The differences in the average GMAT® Total and Quantitative scaled score gains were only .9 and .6 points, respectively.

Larger Total scaled-score gain differences were observed for older versus younger examinees. Test takers who were 25 years old or younger on the day of testing scored, on average, 2.8 points higher than their 30-year-old or older counterparts. Those in the middle age group, ages 26 to 30, showed similarly minor differences.

The largest gains were observed for candidates who tended to have Undergraduate Grade Point Averages that were relatively high compared to their first-time GMAT® Total scaled scores. These test takers gained an average of 39 points, compared to an average gain of only 21 points for candidates whose first GMAT® test scores were high relative to their UGPAs.

Table 6 Gains in GMAT® Total, Quantitative, and Verbal scaled scores by gender, native/first language, and age group						
Gender						
	Male			Female		
	Mean	sd	N	Mean	sd	N
Total gain	31.2	56.7	16,127	30.8	55.2	10,684
Quant. gain	2.4	5.3	16,127	2.5	5.4	10,684
Verbal gain	1.8	5.3	16,127	1.7	5.1	10,684
Native/First Language						
	English			Other		
	Mean	sd	N	Mean	sd	N
Total gain	31.4	55.5	12,728	30.5	56.7	13,714
Quant. gain	2.7	5.4	12,728	2.1	5.2	13,714
Verbal gain	1.6	5.2	12,728	1.9	5.3	13,714
Age Group						
	Youngest (≤ 25.00)			Oldest (30.00+)		
	Mean	sd	N	Mean	sd	N
Total gain	32.9	54.2	10,988	29.1	58.4	7,260
Quant. gain	2.5	5.3	10,988	2.3	5.5	7,260
Verbal gain	1.9	5.1	10,988	1.6	5.3	7,260
Discrepancy Score Group						
	UGPA < GMAT			UGPA > GMAT		
	Mean	sd	N	Mean	sd	N
Total gain	20.9	54.3	5,730	39.4	57.1	9,072
Quant. gain	1.8	4.8	5,730	3.2	5.7	9,072
Verbal gain	1.0	5.3	5,730	2.2	5.1	9,072

As previously noted, examinees who retake the GMAT® are less likely to have finished either the Quantitative or Verbal sections than those who take the exam only once. As shown in Table 7, test takers who did not finish the examination on their first sitting tend to have much larger gains than examinees who finished all items. Examinees

who did not finish the Quantitative section gained, on average, 10 more Total scaled-score points and 1.3 more Quantitative scaled-score points the second time than examinees who did finish. Examinees who did not finish the Verbal section the first time experienced only modest gains.

Table 7 Gains in GMAT® Total, Quantitative, and Verbal scaled scores for examinees who do and do not finish the Quantitative or Verbal sections on their first administration						
Quantitative	Finished			Did not finish		
	Mean	sd	N	Mean	sd	N
Total gain	29.2	55.2	22,244	39.3	61.5	5,272
Quant. gain	2.2	5.2	22,244	3.5	5.8	5,272
Verbal gain	1.8	5.2	22,244	1.8	5.5	5,272
Verbal	Finished			Did not finish		
	Mean	sd	N	Mean	sd	N
Total gain	30.4	55.7	23,283	34.8	61.1	4,233
Quant. gain	2.4	5.3	23,283	2.3	5.5	4,233
Verbal gain	1.7	5.2	23,283	2.3	5.6	4,233

Large gains were realized for examinees with below-average first-time GMAT® Total scaled scores. As shown in Table 8, these examinees had an average gain of 15

scaled-score points higher than the examinees with above-average first-time scores.

Table 8 Gains in GMAT® Total, Quantitative, and Verbal scaled scores for examinees with below-average and above-average GMAT® Total scores						
	Below average (<= 530)			Above average (>530)		
	Mean	sd	N	Mean	sd	N
Total gain	36.3	58.0	17,347	22.3	52.9	10,169
Quant. gain	2.9	5.8	17,347	1.6	4.4	10,169
Verbal gain	2.0	5.1	17,347	1.4	5.5	10,169

Summary

The 18% of the GMAT® test-taking population who retake the examination are a special, self-selected group. Compared with one-time-only test takers, they tend to have lower Total and Verbal scaled scores, are less likely to have finished their first examination, and are more likely to have GMAT® scores that are not aligned with their undergraduate grade point averages.

The largest gains for repeat examinees were found for—

- the youngest examinees,
- examinees whose native language is English,
- examinees who did not finish the Quantitative section on their first sitting,
- examinees with below average first-time scores, and
- examinees whose self-reported UGPAs were relatively higher than their first-time GMAT® Total scaled scores.

There were no differences noted by gender.

On average, examinees who retake the GMAT® obtain only a modest increase in scores—about 31 points or one-third of a standard deviation. These gains are consistent with the gains found by Slack and Porter (1980) for the SAT® and Thornton (1998) for the LSAT®.

The observed gains for repeat test-takers are greater than the gains one would expect solely based on measurement error. Due to measurement error, one would expect only 17% of a random sample of examinees to gain 28 Total scaled-score points or more, yet more than 50% of those who retake the GMAT® witness such gains.

For most examinees, the gains should have little practical meaning. In workshops for school admissions personnel, the Graduate Management Admission Council® emphasizes that GMAT® scores, like the scores for all standardized examinations, are not perfect and contain error. For the GMAT®, the standard error of the difference between two scores is about 40 points. In other words, when comparing two scores, differences of up to 40 points are to be expected by chance fluctuations and schools are encouraged to treat those scores as if they were equivalent. For individual scores, the standard error of measurement is about 28 scaled-score points. With no learning between testing sessions, one should expect an

individual's score to fluctuate up or down by as much as 28 points. About 40% of those who retake the GMAT® see gains within that one standard error of measurement.

Nevertheless, a small portion of examinees who retake the GMAT®, about 10%, obtain a significant gain of 100 scaled-score points or more. An analysis contrasting examinees with gains of 100 or more points with candidates who gained 30 or fewer points, however, did not find any consistent patterns. Test takers with different characteristics tend to obtain higher gains, but the effects do not appear to be cumulative and, based on this data, it is not possible to predict which examinees will be in this high-gain group.

Test takers should carefully consider whether retesting will significantly raise their GMAT® scores. On average, retest gains are modest and can be expected to have a negligible effect on admissions decisions. Candidates who did not finish all the items on their first GMAT® administration and candidates whose scores are well below what they would expect based on their undergraduate grade point averages are more likely to have better than average gains on a GMAT® retest.

The author is the executive director of research and development at the Graduate Management Admission Council®.

References

- Becker, B. J. (1990). Coaching for the Scholastic Aptitude Test: Further synthesis and appraisal. *Review of Educational Research, 60*, 373–417.
- DerSimonian, R. & Laird, N. M. (1983). Evaluating the effect of coaching on SAT scores: A meta-analysis. *Harvard Educational Review, 53*, 1–15.
- Levine, A. E. (1979). *Effects of coaching on standardized admission examinations: Revised statistical analysis of data gathered by the Boston Regional Office of the Federal Trade Commission*. Washington, DC: U.S. Government Printing Office.
- Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. 98–124. Boston: Allyn & Bacon.
- Lane, R. G., Penn, N. E. & Fischer, R. F. (1966). Miller Analogies Test: A note on permissive retesting. *Journal of Applied Psychology, 50*, 409–411.
- Messick, S. M. & Jungeblut, A. (1981). Time and method in coaching for the SAT. *Psychological Bulletin, 89*, 191–216.
- Moore, J. C. (1966). The relationship between time taken to complete the example questions and raw score on the Miller Analogies Test. *Journal of Educational Measurement, 3*(2), 175–177.

- Nathan, J. S. & Camara, W. J. (1998). *Score Change When Retaking the SAT I: Reasoning Test* (Research Notes RN-05). New York, NY: College Board. Retrieved from http://www.collegeboard.com/research/pdf/rn05_10756.pdf.
- Powers, D. E. (1993). Coaching for the SAT: A summary of the summaries and an update. *Educational Measurement: Issues and Practice*, 12(2), 24–30, 39. Powers, D. E. & Alderman, D. L. (1983). Effects of test familiarization on SAT performance. *Journal of Educational Measurement*, 20(1), 71–79.
- Powers, D. E. & Rock D. A. (1999). Effects of coaching on SAT I: Reasoning Test scores. *Journal of Educational Measurement*, 36(2), 93–118.
- Slack, N. V. & Porter, D. (1980). The Scholastic Aptitude Test: A critical appraisal. *Harvard Educational Review*, 50, 154–175.
- Spielbercer, C. D. (1959). Evidence of a practice effect on the Miller Analogies Test. *Journal of Applied Psychology*, 43, 259–263.
- Thornton, A. E. (1998). *Summary of Self-Reported Methods of Test Preparation by LSAT Takers for Testing Years 1991–1992 Through 1996–1997*. Law School Admission Council Research Report Series. Retrieved from <http://www.lsacnet.org/ljac/research-reports/TR-97-02.pdf>.

© 2005 Graduate Management Admission Council® (GMAC®). All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, distributed or transmitted in any form by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of GMAC®. For permission contact the GMAC® legal department at legal@gmac.com.

Creating Access to Graduate Business EducationSM is a service mark of the Graduate Management Admission Council®. GMAC®, GMAT®, Graduate Management Admission Council®, and Graduate Management Admission Test® are registered trademarks of the Graduate Management Admission Council® (GMAC®). SAT® is a registered trademark of the College Entrance Examination Board. LSAT® is a registered trademark of the Law School Admission Council, Inc.