

PR 93-02

ANALYSIS OF THE JULY 1992 EXAM

Stephen P. Klein, Ph.D.
Roger Bolus, Ph.D.

GANSK & ASSOCIATES

January 18, 1992

SUMMARY

The July 1992 exam had two sections: the Multistate Bar Examination (MBE), a 200-item multiple choice test; and a Written Examination composed of six essay questions and two Performance Test (PT) problems. There were 7,108 applicants who completed both sections, 26 percent of whom had taken the GBX at least once before.

MBE raw scores (the number of questions answered correctly) were converted to scale scores that controlled for possible differences in average item difficulty from one administration of the exam to the next. Essay and PT answers were graded on a 40 to 100 point scale. Scores on this scale were assigned in 5-point intervals. Each PT score was then multiplied by 2 so that the maximum possible Written raw score was 1,000 points.

Written raw scores were converted to the same scale of measurement as was used on the MBE. This was done to adjust for possible differences over time in the difficulty of the questions asked and the leniency with which the answers to them are graded. An applicant's total score was computed using the formula below:

$$\text{Total Scale Score} = (.35)(\text{MBE Scale}) + (.65)(\text{Essay Scale})$$

A three phased grading process was used to determine an applicant's pass/fail status. In Phase 1, applicants were: failed if their Total score was less than 1391, passed if their Total was greater than 1465, and placed in reread if their Total was in the 1391 to 1465 range.

In Phase 2, applicants in the reread group had all of their essay and PT answers read a second time. At the end of this phase, applicants were: failed if their Total scores were below 1412, passed if their scores were above 1439, and placed in reappraisal if their scores were in the 1412 - 1439 range. Applicants also were placed in reappraisal if their Total score after the first reading was greater than 1439 (i.e., even if their average was less than 1412).

An applicant in the reappraisal group had all of his or her answers and scores reviewed as a set by a member of the Board of Reappraisers who then made a final pass/fail decision. Thus, an applicant's answers were read once, twice, or three times depending upon how close that applicant came to the pass/fail line of 1440.

The major findings with the applicants who had all of their answers read at least once and the group of 1364 applicants who had them read at least twice were as follows:

- o After the first reading of all answers, 29 percent of the applicants failed and 51 percent passed. The total percent passing after the second reading and the reappraisal process were 58 and 60 percent, respectively.
- o California applicants scored higher than the national average on all of the six MBE subjects.
- o The degree of agreement between readers in their evaluations of the relative quality of the written answers was about the same on an essay question as it was on a PT problem. However, because reader assigned PT scores were multiplied by 2, differences in final scores between the first and second reading on an answer were larger on a PT than on an essay answer.
- o The reliability of the Written and Total scores (.71 and .86, respectively) was comparable to that of previous exams.
- o The correlation between MBE and Written scores (.64) was in the normal range.
- o The phased grading cutoff scores led to rereading and reappraising the answers of the applicants who were most likely to benefit from these additional reviews.
- o Males tended to earn higher scale scores on the MBE than on the Written section. The reverse was true for females.

ANALYSIS OF THE JULY 1992 EXAM

TEST SECTIONS

The July 1992 exam had two sections: the Multistate Bar Examination (MBE), a 200-item multiple choice test; and a Written section composed of six essay questions and two Performance Test (PT) problems. The exam was administered over three consecutive days, with two 3-hour sessions per day. Day 1 consisted of essay questions 1 - 3 and PT problem A. Day 2 was devoted to the MBE. Day 3 consisted of essay questions 4 - 6 and PT problem B.

SCORING RULES

MBE raw scores (the number of MBE questions answered correctly) were converted by the American College Testing Program (ACT) to equated ("scaled") scores and then multiplied by 10 using the formula below. This procedure adjusts the raw scores for possible differences in mean question difficulty from one administration of the MBE to the next.

$$\text{MBE Scale} = (8.9130)(\text{MBE Raw}) + 269.8348$$

Each essay answer was graded in 5-point intervals on a 40 to 100-point scale. The same procedure was used to grade each PT answer. PT scores were then multiplied by 2 so that the maximum possible Written raw score was 1,000 points (6 essays at 100 points each plus 2 PT problems at 200 points each). Written raw scores were converted to a score distribution that had the same mean and standard deviation as the applicants' MBE scores (see Appendix F). This scaling was done using the MBE and Written scores of the random sample of about 1,000 applicants who had their Written answers graded first. The formula used to convert written raw scores to scale scores was:

$$\text{Written Scale Score} = 3.5694 (\text{Written Raw Score}) - 946.5505$$

An applicant's Total scale score was a weighted combination of that applicant's MBE and Written scale scores. The formula for computing Total scale scores was:

$$\text{Total Scale Score} = (.35)(\text{MBE Scale}) + (.65)(\text{Written Scale})$$

A three-phased grading process was used to determine an applicant's pass/fail status. In Phase 1, applicants were: failed if their Total score was less than 1391, passed if their Total was greater than 1465, and placed in reread if their Total was in the 1391 to 1465 range.

In Phase 2, applicants in the reread group had all of their essay and PT answers read a second time. The second reader on an answer was a different grader than the first reader and the second reader did not know the score assigned by the first reader.

The scores assigned on the first and second readings were averaged and new Total scale scores were computed using the formulas above. At the end of Phase 2, applicants were: failed if their Total scores were below 1412, passed if their scores were above 1439, and placed in reappraisal if their scores were in the 1412 - 1439 range. Applicants also were placed in reappraisal if their Total score on the first reading was greater than 1439 (i.e., even if their average was less than 1412).

An applicant in the reappraisal group had all of his or her answers and scores reviewed as a set by a member of the Board of Reappraisers who then made a final pass/fail decision. Thus, an applicant's answers were read once, twice, or three times depending upon how close that applicant was to the pass/fail line of 1440.

SAMPLE

The analyses of the July 1992 exam were conducted with the 7108 applicants who had both an MBE and a Written score. This sample contained 5288 applicants who were taking the exam for the first time and 1820 repeaters. The percentage of applicants from ABA approved, California Accredited, and Nonaccredited schools were 69, 15, and 4, respectively. The remaining 12 percent were not assigned to a school.

MULTISTATE BAR EXAMINATION (MBE)

Table 1 shows that California applicants scored higher than the national average on all six MBE subtests. California's average total raw score (the average number of questions answered correctly) was 3.8 points higher than the national average (which included California scores).

WRITTEN SECTION

Table 2 presents the means and standard deviations on each question after all readings of all answers.

Applicants who came close to the examination's pass/fail line after one reading of their answers had all of their answers read again by readers who had not graded them previously.

Table 1

NATIONAL AND CALIFORNIA AVERAGE MBE SCORES AND THE
DIFFERENCE BETWEEN THESE AVERAGE SCORES

Test Score	Number of Items	National Mean	CA Mean	Difference
Constitutional Law	30	22.6	22.7	0.1
Contracts	40	25.2	26.0	0.8
Criminal Law	30	17.8	18.7	0.9
Evidence	30	20.6	20.9	0.3
Real Property	30	17.8	18.4	0.6
Torts	40	26.1	27.2	1.1
Total Raw Score	200	130.1	133.9	3.8
NCBE/ACT Scale	200	143.0	146.4	3.4

Table 2

SUMMARY STATISTICAL DATA AFTER ALL READINGS
OF THE ESSAY AND PT ANSWERS (N = 7108)

Question Number	Subject Matter Area	Mean Score	Standard Deviation
1	Community Property	68.3	8.57
2	Torts/Constitutional Law	68.1	6.36
3	Evidence	65.3	8.12
4	Wills/Trusts	68.4	7.38
5	Criminal Law	66.6	6.26
6	Remedies	66.7	6.94
PT-A	Professional Responsibility	135.9	14.45
PT-B	Real Property	135.0	12.78

Table 3 contrasts the applicants' scores on the first reading with their scores on the second reading. The last row shows that as on past exams, the mean total written score after the first set of readings (664) was higher than it was on the second set (660). There was a .38 correlation between the total written raw scores on the first and second set of readings. And, on the average, the total score on the first set of readings differed by 4 points from the total on the second set.

The upper portion of the next to last column of Table 3 shows that the two readers who graded an answer usually assigned scores to that answer that were within 4 points of each other (before multiplying PT scores by 2). The correlation coefficients in the last column indicate that on a typical question, the relative standings of the applicants on the first reading were generally consistent with their relative standings on the second reading. The average of the eight coefficients was .55.

Table 3's data underestimate the true degree of agreement among readers because: (1) there is a much narrower range of ability in the reread sample than in the total applicant pool and (2) the method used to assign the second reader is designed to offset systematic differences among readers in their tendency to give relatively high or low grades.

Table 3

INDICES OF AGREEMENT BETWEEN THE FIRST AND SECOND READINGS (N = 1364)

Question Number	Mean on First Reading	Mean on Second Reading	Difference Between Means	Mean Absolute Difference	Correlation Between Readings
1	66.7	66.3	0.4	4.1	.62
2	66.9	66.8	0.1	3.3	.55
3	63.7	63.4	0.3	4.1	.58
4	67.5	66.5	1.0	4.4	.58
5	65.8	65.7	0.1	3.9	.44
6	65.7	65.6	0.1	3.6	.59
PT-A	133.7	132.9	0.8	7.7	.55
PT-B	133.8	132.5	1.3	7.5	.50
Total	663.8	659.7	4.1	16.4	.38

Reader assigned PT scores are multiplied by 2.0.

RELATIONSHIPS AMONG SECTIONS

Table 4 presents summary statistical data on each section after all readings. Table 5 shows the correlations among sections and the LSAT.

Appendix E shows that MBE and Written scores correlated about as highly with each other (.64) on the July 1992 exam as they did on previous July exams. About 75 percent of the applicants had the same pass/fail status on the MBE as they had on the Written section (where a scale score of 1440 on a section was considered passing).

Table 4

SUMMARY STATISTICAL DATA AFTER ALL READINGS (N = 7108)

Test Statistic	MBE Scale	Written Raw Score	Total Scale
Average Score	1464.2	674.1	1462.2
Standard Deviation	154.8	42.9	140.4
Internal Consistency	.874*	.711	.864

* Computed by ACT using national data.

Table 5

CORRELATIONS AMONG SECTIONS AND LSAT AFTER ALL READINGS

	LSAT*	MBE	Written
MBE	.60		
Written	.49	.64	
Total	.59	.84	.95

*LSAT (Law School Admissions Test) scores were available for 6005 applicants.

SUBGROUP ANALYSES

Table 6 shows that on the average, women scored 32 points higher on the Written section than on the MBE whereas there was a 33 point average differential in the opposite direction for male applicants. This table also shows that Anglos had higher MBE scores than others. However, because the sample of 1,000 candidates on which the formula for scaling Written scores was not perfectly representative of all 7,108 candidates, Written scale scores tended to run somewhat lower than the MBE scale scores. After adjusting for this difference, only Asians had higher Written scores than MBE scores. Moreover, the size of the differences that remained were far less than they were between gender groups and quite small relative to the variation around these means.

Table 6

MEAN SCALE SCORES WITHIN RACIAL/ETHNIC AND SEX GROUPS AND THE NUMBER OF APPLICANTS AND PERCENTAGE OF MALES WITHIN EACH GROUP*

Test	Racial/Ethnic Group				Gender	
	Anglo	Asian	Black	Hispanic	Female	Male
MBE	1491	1405	1350	1394	1450	1475
Written	1483	1414	1350	1396	1482	1442
Total	1486	1411	1350	1395	1471	1454
Number of Applicants	5239	655	486	477	2995	4005
% Male	57%	60%	48%	60%	- 0 -	100%

* There were 143 applicants who did not fall in one of the four largest racial/ethnic groups and 108 applicants who did not indicate their racial/ethnic or gender group.

PHASED GRADING

As noted previously, a three-phased grading process was used to focus reader time on the applicants who were near the pass/fail line. Table 7 presents the number and percentage of applicants in each of the exam's pass/fail categories. These data indicate that 2090 (60%) passed.

Table 7

NUMBER AND PERCENTAGE OF APPLICANTS WHO PASSED AND FAILED IN EACH PHASE OF THE MULTIPHASED GRADING PROCESS

Pass/Fail Category	Number	Percent
Fail - Phase 1	2090	29.4
Fail - Phase 2	466	6.6
Fail - Phase 3	<u>318</u>	<u>4.5</u>
Total Fail	2874	40.4
Pass - Phase 1	3654	51.4
Pass - Phase 2	433	6.1
Pass - Phase 3	<u>147</u>	<u>2.1</u>
Total Pass	4234	59.6
Phase 1 = 1st reading	5744	80.8
Phase 2 = Reread	899	12.6
Phase 3 = Reappraisal	465	6.5

Table 8 shows the number of Phase 2 applicants who passed and failed after all readings relative to their Phase 1 scores. This table illustrates the strong, but far from perfect relationship between Phase 1 scores and final pass/fail status. These data also indicate that the width of the Phase 2 reread band (1389.5 - 1465.0) was about right in that almost all the applicants with relatively high scores on the first reading passed and almost all with low initial scores failed.

Similarly, Table 9 shows that the higher an applicant's score at the end of Phase 2, the greater the likelihood that applicant passed as a result of reappraisal. The lower limit of the reappraisal band (1412) also appears to be appropriate in that only 2 of the 76 applicants with Phase 2 Total scores below 1415 passed as a result of reappraisal.

Table 8

NUMBER OF REREAD APPLICANTS WHO PASSED AND FAILED RELATIVE TO THEIR TOTAL SCORES AFTER THE FIRST READING

Score After 1st Reading	Number of Applicants			Percent Passing
	Fail	Pass	Total	
1450 - 1465	56	254	310	82
1435 - 1449	108	163	271	60
1420 - 1434	159	91	250	36
1405 - 1419	211	52	263	20
1390 - 1404	250	20	270	7
Total	784	580	1364	43

Table 9

NUMBER OF REAPPRAISED APPLICANTS WHO PASSED AND FAILED RELATIVE TO THEIR TOTAL SCORES AFTER REREAD

Total Score After Reread	Number of Applicants			Percent Passing
	Fail	Pass	Total	
1435 - 1439	8	57	65	88
1430 - 1434	37	53	90	59
1425 - 1429	59	21	80	26
1420 - 1424	64	12	76	16
1415 - 1419	76	2	78	3
<1415	74	2	76	3
Total	318	147	465	32

Appendix A: HISTORICAL EXAM RESULTS

Year	FEBRUARY				JULY			
	Mean MBE	Mean Essay	N	Pass Rate	Mean MBE	Mean Essay	N	Pass Rate
1976	137	68	3088	38	145	69	6709	60
1977	139	68	3399	44	143	69	7191	55
1978	139	NA	4222	38	145	70	6835	55
1979	139	69	4166	45	144	70	7152	55
1980	137	67	3758	34	142	69	7379	49
1981	138	67	3837	33	142	69	7080	50
1982	137	66	4033	31	143	68	7038	49
1983	136	66	4200	28	144	68	7277	50
1984	137	66	3899	30	141	68	7201	42
1985	140	66	4661	33	142	68	7622	46
1986	140	66	4689	28	144	68	7780	45
1987	141	66	4682	43	144	67	7481	51
1988	142	67	4530	46	144	68	7146	53
1989	143	66	4230	50	146	68	6970	60
1990	141	67	3819	44	145	69	6963	58
1991	143	66	3685	51	145	67	7219	55
1992	143	66	3907	51	146	67	7108	60

Mean essay = mean score on a written answer after all readings across all answers, but before PT scores are multiplied by a constant). The PT was introduced in July 1983. Scaling the written scores to the MBE began with the February 1987 exam.

Appendix B: COMPARISON OF NATIONAL AND CALIFORNIA MEAN MBE SCORES

Year	February			July		
	Nation	Calif.	Diff.	Nation	Calif	Diff.
1984	135.1	137.1	2.0	139.2	140.6	1.4
1985	136.4	140.2	3.8	140.5	141.9	1.4
1986	135.6	139.8	4.2	140.3	143.6	3.3
1987	136.4	140.6	4.2	140.3	143.9	3.6
1988	137.2	141.9	4.7	139.8	143.7	3.9
1989	137.8	142.9	5.1	142.0	146.3	4.3
1990	135.9	141.5	5.6*	141.4	145.1	3.6
1991	137.8	143.0	5.2	142.1	145.4	3.3
1992	139.1	143.2	4.1	143.0	146.4	3.4

* California data not included in computation of national mean; California MBE scores imputed using AM scores.

Appendix C

INDICES OF READER AGREEMENT ACROSS EXAMS

Exam	Mean difference between		Correlation between	
	Readers on an answer	1st-2nd set of readings	Readers on an answer	1st&2nd set of readings
2/85	4.7	5.3	.64	.55
2/86	4.7	9.3	.60	.60
2/87	4.5	3.8	.59	.45
2/88	4.4	3.1	.54	.34
2/89	4.7	3.5	.54	.41
2/90	4.0	4.0	.59	.43
2/91	4.1	2.7	.59	.37
2/92	3.8	3.9	.59	.40
7/85	4.4	5.5	.58	.62
7/86	4.5	5.4	.61	.58
7/87	4.3	3.8	.56	.42
7/88	5.6	6.4	.54	.46
7/89	4.3	2.9	.54	.34
7/90	3.7	-1.9	.56	.41
7/91	3.9	3.1	.57	.43
7/92	3.9	4.1	.55	.38

Mean difference between readers on an answer = mean of the absolute differences between the scores assigned by the readers (i.e., before multiplying PT scores by a constant). The second column = difference between the mean total written raw score across all essay and PT questions on the first set of readings minus the mean on the second set of readings. Total written raw scores for columns 2 and 4 were computed for the pre-1987 exams using the current scoring rules.

Appendix D

PERCENTAGE OF ALL APPLICANTS WHO WENT TO REREAD AND REAPPRAISAL;
AND, THE PERCENTAGE OF THOSE IN THESE PHASES WHO PASSED THE EXAM

Exam	Reread		Reappraisal	
	Percent of total	Percent passing	Percent of total	Percent passing
2/85	22	29	8	36
2/86	22	36	7	34
2/87	24	42	8	34
2/88	26	42	9	31
2/89	26	41	9	29
2/90	25	41	9	25
2/91	24	44	9	25
2/92	24	44	9	33
7/85	24	43	9	32
7/86	23	42	8	33
7/87	22	40	8	31
7/88	23	41	9	31
7/89	20	45	8	32
7/90	20	47	8	31
7/91	21	42	7	27
7/92	19	43	7	32
Mean	23	41	8	32

Appendix E: CORRELATION BETWEEN MBE AND WRITTEN SCORES

Year	February	July	Year	February	July
1976	.62	.70	1987	.57	.66
1977	.65	.70	1988	.58	.62
1980	NA	.68	1989	.60	.65
1982	.60	.65	1990	.55	.67
1983	NA	.73	1991	.58	.67
1984	.59	.63	1992	.60	.64
1985	.61	.60			
1986	.58	.66			

Data were not available for 1978, 1979, and 1981. Correlation between MBE and Essay through July 1986; after that, it is between MBE and Written scale scores.

Appendix F

MBE raw scores are adjusted for possible differences in mean item difficulty from one exam to the next. As a result of this step, a given MBE "scale" score indicates the same level of proficiency regardless of the administration of the exam on which it was earned (eg., February 1992 versus July 1992). The adjustment procedure involves inserting into the current version of the test a set of items that have been used previously. The scores of current and previous candidates on the repeated items are used to calibrate one exam to another.

Essay and PT scores cannot be adjusted in this way because test security considerations preclude repeating essay and PT questions across exams. However, scaling the candidates' written scores to a score distribution that has the same mean and standard deviation as their MBE scores indirectly adjusts the written scores for possible differences among exams in the average difficulty of the questions asked (and the leniency with which the answers to them are graded).

The appropriateness of scaling to the MBE is supported by two factors. First, there is a very strong relationship between MBE and essay scores among cohorts of candidates; i.e., groups with relatively high essay scores also tend to have high MBE scores. Figure G-1 illustrates this relationship. Each asterisk (*) is a law school. The x-axis shows the school's mean MBE score and the y-axis shows its mean written scale score. The data for this plot are the 39 law schools that had 30 or more candidates taking the July 1992 exam. The correlation between the means is near perfect ($r = .97$ out of a possible 1.00).

The second reason for scaling to the MBE is that MBE scores reflect variations in the ability level of those taking different exams. Figure G-2 illustrates this sensitivity by contrasting mean LSAT and MBE scores over 11 exams for those candidates who had both an LSAT and an MBE score. The y-axis on the left shows mean MBE score for the solid trend line. The y-axis on the right shows mean LSAT score for the dashed trend line. These data show there is a strong relationship between a cohort's mean LSAT and MBE scores. Moreover, on both tests, the mean score for repeaters was always lower than it was for first timers. In short, the more able the cohort of candidates (as reflected by its mean LSAT score), the more likely it is to have a high mean MBE score.

Table G-1 presents mean LSAT, MBE, and essay raw scores (based on six questions) for those candidates who had scores on all three measures. Tests are listed in descending order of mean LSAT score. These data show that mean MBE scores are more sensitive than mean essay scores to differences in average ability level among cohorts.

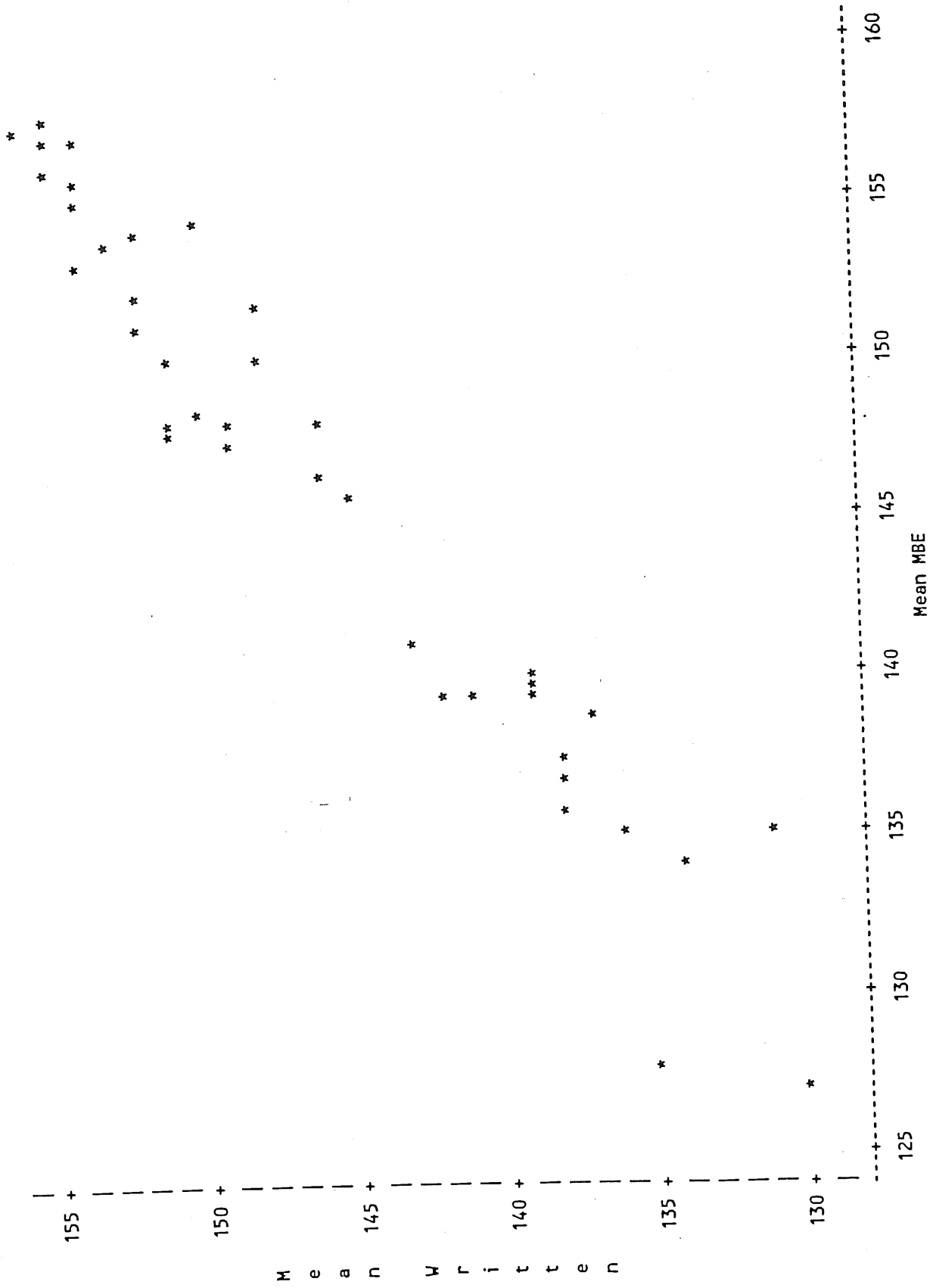


Figure G-1: Relationship between mean MBE and Written scores across law schools.

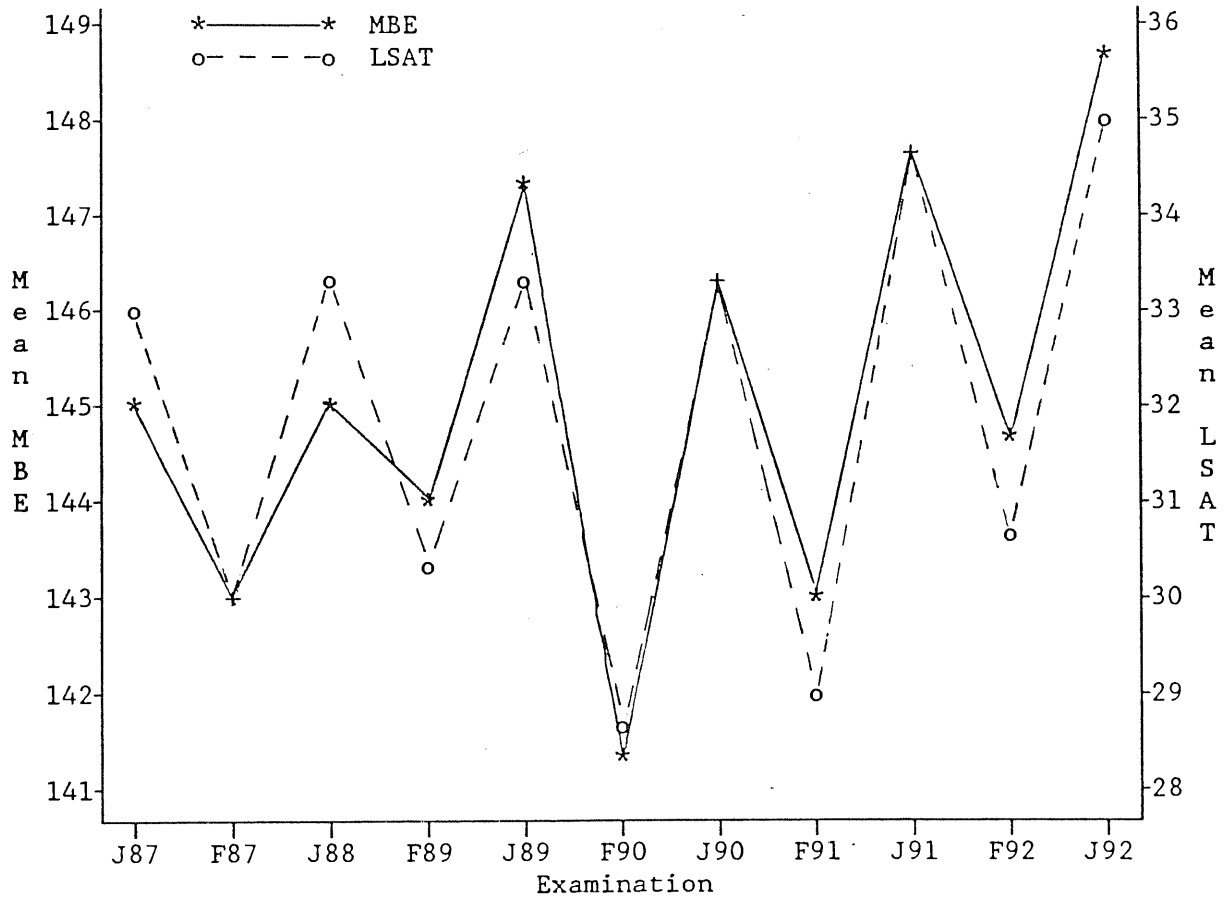


Figure G-2: Relationship between a cohort's mean MBE and LSAT scores.

Table G-1

Exam	N	LSAT	MBE	Essay
2/90	2350	28.78	141.3	401.0
2/91	2207	28.93	143.1	396.3
2/88	3673	30.05	142.9	403.6
2/89	3414	30.32	143.9	395.6
2/92	3073	30.61	144.6	395.8
7/87	6489	33.06	145.1	400.5
7/89	5977	33.17	147.5	405.3
7/88	6237	33.19	144.9	412.1
7/90	5593	33.35	146.3	415.9
7/91	5722	34.65	147.6	407.4
7/92	5909	34.87	148.5	407.4