

A NATIONAL LOOK AT MBE PERFORMANCE DIFFERENCES AMONG ETHNIC GROUPS

by Douglas R. Ripkey and Susan M. Case, Ph.D.

INTRODUCTION

Issues related to diversity are an important concern in bar admissions. Various researchers have examined ethnic performance with regard to admission to law school (1), experiences while in law school (2, 3), and the associated educational outcomes both in law school (4) and on bar examinations (5, 6, 7).

Until now, most of the research related to ethnic differences in performance on bar examinations has focused on individuals from single jurisdictions (6, 7). While these types of studies are generally useful within that given jurisdiction, the results are less generalizable to other jurisdictions. Studies of larger groups of candidates from multiple jurisdictions (5) have their own set of limitations because they are usually focused on bar passage rather than on examination scores. Overall bar passage rates are not directly comparable across jurisdictions since these rates are influenced by the components of the jurisdiction's bar examination (i.e., the mixture of essays, performance tests, and multiple-choice questions), the method for determining overall scores, and the minimum passing score.

The National Conference of Bar Examiners recently began collecting demographic data from examinees at the national level; this data enables us to investigate, among other things, performance

differences on bar examinations among ethnic groups across jurisdictions.

This article reports on our initial investigations into the Multistate Bar Examination (MBE) performance differences across ethnic groups using data from a national population. Since NCBE does not have bar passage information on individual examinees, we cannot undertake a study of overall bar passage rates of ethnic groups. However, we can look at MBE performance, and the commonality of the MBE as a bar examination component in most jurisdictions allows for a clearer interpretation of results across ethnic groups. It avoids problems related to state-specific components and state-specific pass/fail standards. Because we have published previous information about gender differences in MBE performance (8), they are not included as part of this article.

THE STUDY DESIGN

Data from candidates who sat for the MBE in July 2006 were used to make comparisons among ethnic groups. In addition, the relationships between performance on the MBE, prior educational achievement (as indicated by undergraduate grade point average (UGPA)), and scores on the Law School Admissions Test (LSAT) were studied to put the results into context.

Developing the Sample

To aid in the interpretation of the results, we wanted the study to include only those candidates who were generally at the same point in the educational process, i.e., newly graduated from law school. Functionally, this meant that we wanted to limit our study sample to include only those candidates from one MBE administration who were taking the MBE for the first time.

While neither candidates nor jurisdictions directly provide NCBE with information about the number of previous bar examination attempts, we do have some information that allows us to make reasonable estimates about who is taking the MBE for the first time on a given administration. Examinees are asked to provide certain demographic information (e.g., Social Security number (SSN), applicant ID number, jurisdiction code, etc.) on the answer sheets. In jurisdictions that have required candidates to code consistent identifying information (i.e., always an SSN), the initial appearance of a candidate's SSN across all administrations indicates that he or she is a first-time taker.

We selected the July 2006 MBE administration as the basis for our study. Of the 51,155 candidates who sat for the July 2006 MBE in non-territorial U.S. jurisdictions, 5,610 individuals from the seven jurisdictions that had not consistently required the use of an SSN on the MBE answer sheet during each administration in the past five years (Georgia, Hawaii, Michigan, New Hampshire, New Mexico, Texas, and Vermont) were excluded because their first-time taker status could not be reliably determined. An additional 1,221 candidates who coded an invalid SSN (e.g., marking all 0s) were eliminated from the study along

with 10,567 other candidates because the July 2006 administration was not their first MBE administration (their SSNs had appeared on at least one MBE answer sheet from an administration prior to July 2006). The resulting group included 33,757 candidates who were likely July 2006 first-time takers.

Obtaining Ethnicity Data

Because ethnicity is not a demographic code on the MBE answer sheet, we needed to obtain this information separately by using Multistate Professional Responsibility Examination (MPRE) answer sheets. This information is available as part of the data collection program that NCBE has with the Law School Admission Council (LSAC). Beginning in November 2004, candidates who register electronically to take the MPRE have the opportunity to provide consent for NCBE to obtain for research purposes their LSAC records. These records include both demographic information (e.g., birth date, gender, ethnicity, SSN, undergraduate institution, and undergraduate major) and academic performance data (i.e., UGPA and average LSAT score from all attempts). Data pertaining to a large number of MPRE candidates has been obtained as nearly all register electronically and 90 percent consent to the access of their personal information.

The group of 33,757 likely July 2006 first-time takers was then matched with the available LSAC data to determine the final sample. Because LSAC-based information was missing for about 15 percent of these candidates and an additional 5 percent were not matched to any MPRE records, the final group for the analyses included 26,839 candidates, which represented about 80 percent of all known July 2006 MBE first-time takers.

DATA ANALYSIS

Demographic Breakdown

Table 1 provides a breakdown of the racial/ethnic composition within the full group of 26,839 candidates. For purposes of this article, we combined several groups based on the relatively small number of candidates who indicated certain ethnicities. More specifically, candidates who indicated that their ethnicity was “American Indian/Alaskan Native” were grouped together with candidates who chose “Other” as their ethnic indicator. Also, those candidates who indicated either “Chicano/Mexican American,” “Puerto Rican,” or “Hispanic/Latino” were clustered into a broader “Hispanic” group in order to have an appropriate sample size and because the performance of the candidates in these subgroups was similar.

Table 1 Demographic Breakdown by Ethnic Group
(With Some Smaller Groups Combined)

Ethnicity	Count	Percentage
Asian	2,100	7.8%
Black	1,378	5.1%
Hispanic	1,470	5.5%
White	19,479	72.6%
Other	1,213	4.5%
Unavailable	1,199	4.5%
Total	26,839	100.0%

As indicated in Table 1, the most common ethnic group was the White group, which comprised nearly 72.6 percent of the sample. The next largest group was Asian (7.8 percent) followed by Hispanic (5.5 percent), and Black (5.1 percent). The remaining candidates for whom there was data represented 4.5 percent of the July 2006 MBE first-time takers and had an ethnicity classified as “Other.”

MBE Performance

Table 2 provides the July 2006 average MBE score for the likely first-time takers within each ethnic group. It also shows the group’s standard deviation (SD), which is an indicator of the spread or variation of scores in that group; the larger the SD, the more spread out the scores are. (Typically, about 68 percent of individual scores will be within one SD of the average for that group.) The last column in Table 2 shows the effect size, which is a measure of performance differences among groups expressed in terms of SD units. (See The Testing Column in the May 2005 issue of THE BAR EXAMINER for an explanation of SDs.) Typically, the largest or highest-performing group is used as the reference point and comparisons are made between that reference group and the others.

Within the groups with known ethnicity, White candidates had the highest MBE performance (average = 149.3) followed by the Asian (146.1) and the Hispanic (143.3) candidates. The Black group had the lowest average MBE performance (137.9). In terms of the performance difference size, the MBE average for the Black group was about 11.4 points less than the average for the White group. This difference in the average score, when divided by the SD of the reference group (i.e., that of the White group, the highest-scoring group), resulted

Table 2

MBE Performance Breakdown by Ethnic Group

Group	Count	MBE Score		Effect Size in
		Average	SD	SD Units
Asian	2,100	146.1	14.4	-0.23
Black	1,378	137.9	14.4	-0.81
Hispanic	1,470	143.3	14.9	-0.43
White	19,479	149.3	14.0	—
Total *	26,839	148.1	14.5	—

* Includes 2,412 candidates whose ethnic group affiliation was unavailable or listed as “Other.”

in an effect size of -0.81 SD units (i.e., $-11.4/14.0 = -0.81$). For large-scale data, such as in this study, a difference of 0.10 SD units or greater is typically both of statistical and practical significance. As Table 2 shows, the effect size was well above this for the Asian, Black, and Hispanic groups when compared with the White group. The effect size for Black candidates was almost twice as large as for the Hispanic candidates and just less than four times greater than that for the Asian candidates.

Other Academic Performance Measures

Table 3 provides information that compares the MBE ethnic group performance patterns to those on other academic performance measures obtained prior to law school entry (i.e., UGPA and LSAT score). To make the effect sizes comparable to those shown for the MBE in Table 2, the White group is once again used as the reference group for the calculations.

Focusing on UGPA, the rank ordering of the specific ethnic groups (based on average group performance) was about the same as the pattern on the MBE. The White group had the highest average UGPA (3.36), followed by the Asian (3.32), Hispanic (3.25), and Black (3.16) groups. In terms of effect sizes, the differences in average UGPA between the White group and the other groups were smaller than those observed on the MBE. For example, the difference in effect size (i.e., in SD units) between the Black and White groups on the MBE was -0.81, whereas the effect size on the UGPA was only -0.50. Based on the larger effect sizes on the MBE, it appeared that there were greater differences in MBE average scores than would have been expected based on UGPA information. This is a common finding in this kind of research; it reflects the fact that undergraduate institutions and courses taken vary in the rigor of their academic requirements, whereas the MBE is consistent across examinees in that regard.

The performance patterns by ethnic group on the LSAT differed somewhat from those on the MBE. On the LSAT, Asian candidates scored about 1.5 points higher on average than White candidates; the effect size was 0.21 SDs ($1.5 / 7.1 = 0.21$). For the other two main groups, the pattern was similar to the MBE pattern (i.e., White candidates performed better on average), but the magnitude of the difference was slightly higher, especially for the Black group where the LSAT difference was almost a full SD (i.e., about 7 points). Thus, based on LSAT score information, the MBE performance differences (when compared with White candidates)

Table 3 Academic Success Measures by Ethnic Group*

Group	Count	UGPA		LSAT Score		MBE Score	
		Avg.	Effect Size	Avg.	Effect Size	Avg.	Effect Size
Asian	2,100	3.32	-0.10	159.5	+0.21	146.1	-0.23
Black	1,378	3.16	-0.50	151.2	-0.96	137.9	-0.81
Hispanic	1,470	3.25	-0.29	154.7	-0.46	143.3	-0.43
White	19,479	3.36	—	158.0	—	149.3	—
Total **	26,839	3.34	—	157.6	—	148.1	—

* SDs are approximately 0.4 for UGPAs, 7.3 for LSAT scores, and 14.5 for MBE scores.

** Includes 2,412 candidates whose ethnic group affiliation was unavailable or listed as “Other.”

were greater than expected for Asian candidates, less than expected for Black candidates, and about the same as expected for Hispanic candidates.

Relationships Among Performance Measures

While comparing patterns of performance across groups provides some information about the relationship between the performance measures, there are more direct methods to define how one score relates to another. Usually, these relationships are expressed in terms of a correlation coefficient, which describes the linear relationship between the pairs of variables in terms of both the magnitude and the direction. The value indicates the extent to

which performance on one measure is related to performance on another measure. Correlation coefficients always have values between -1.00 and +1.00; in studies like ours these values are typically between 0.00 and +1.00 because candidates who perform well on one test generally perform well on another. Higher correlations, those closer to +1, indicate stronger relationships.

Table 4 provides the correlations among the measures of academic performance in the total group of candidates (n = 23,839). Table 4 shows that the correlation between UGPA and LSAT score was +0.30 for the total group of candidates. This modest, positive correlation indicated that candidates who had high UGPAs also tended to perform well on the LSAT and vice versa. The correlation between the UGPA and the MBE was +0.29, indicating about the same modest, positive relationship between these two measures.

In general, the relationships represented in Table 4 are not surprising. As expected, the relationships among all scores were positive (i.e., those who performed well on one measure performed well on the other). The fact that the correlations were all relatively modest was also expected because the measures assess somewhat different knowledge and skills and they are assessed at different points in time along an educational continuum.

Table 4 Correlations Among Performance Measures for the Total Group

Performance Measures	Correlation Coefficients Among Performance Measures	
	LSAT Score	MBE Score
UGPA	+0.30	+0.29
LSAT Score	—	+0.49

OBSERVATIONS BASED ON THE DATA

We are at the beginning stages of studying ethnic group bar examination performance at the national level, and in many ways, the data that we have begun to analyze have led to more questions than answers.

The academic performance measures available about candidates prior to their entry into law school indicate that performance differentials already occur among the different ethnic groups. The White group in the sample had slightly higher UGPAs than the Asian group (about 1/10 of an SD). White candidates also had higher UGPAs than Hispanic candidates (about 1/4 of an SD) and Black candidates (about 1/2 of an SD). The pattern of the ethnic group performance differences on the LSAT was consistent, with the White group scoring highest on average and the Black group scoring lowest. However, the magnitude of the performance differential for the Hispanic and Black groups was nearly twice as high on the LSAT as it was for UGPAs. Finally, both these observed pre-law school performance differentials are similar in size to those found in several large single jurisdiction studies (6, 7) and are similar to

the ethnic group differences for individuals seeking entry into the medical profession (9).

Unfortunately, information about performance in law school isn't available nationally for the group of candidates who were first-time takers on the MBE in July 2006. Given the similarity in the patterns in pre-law school performance in our group as those encountered in large single jurisdiction studies (6, 7), there is no indication that the law school performance differentials seen in those studies would not have been replicated in the national group as well.

Given observed differences in the average pre-matriculation credentials of candidates from different ethnic groups in this study and suspected differences among groups in law school performance based upon other research (6, 7), it would be reasonable to expect that similar differences would also occur within the bar examination process. While we did not have access to all the relevant components of the bar examination process (e.g., essay, performance, or total test scores), looking at patterns in MBE scores was more useful in some ways because the meaning of these scores is standardized across all candidates (i.e., the tailored nature of the essays, score combination rules, and performance standards would have made pass/fail rate comparisons problematic across jurisdictions).

On the MBE component of the bar examination, the White group performed best followed by the Asian, Hispanic, and Black groups. The magnitudes of the differences between groups on MBE scores were similar to those relating to LSAT scores (with the exception of Asian candidates). Again, our findings were consistent with single-jurisdiction studies that identified similar performance differentials

across ethnic groups on all components of the bar examinations, including the MBE, essays, performance tests, and total score (6, 7).

As indicated when we originally reported on gender differences on the MBE (8), there are several follow-up efforts that we will continue to take to evaluate the legitimacy of the results (i.e., to determine whether some groups are more proficient than other groups on the knowledge and skills that the MBE is designed to assess).

From an item-writing perspective, we will continue to refine the processes used to evaluate the individual items used on the MBE. Following long-standing NCBE practice, our drafting committees (groups of content experts who write and edit the exam questions) will be instructed to eliminate all material that does not directly relate to the competencies tested by the MBE. Similarly, we will continue to encourage these committees to avoid unnecessarily complicating the language used in exam questions and to avoid non-standard English that might affect one ethnic group in a more pronounced way than another. In addition, the composition of these drafting groups will continue the current practice of overrepresentation of women and minorities (relative to the population in general) to help identify and avoid possible subtle substantive biases in the questions.

Research will also continue to be part of the item evaluation process. One form of analysis, called Differential Item Performance (DIF), involves an investigation of individual items to see if some of them are more difficult for one group than another. With continued development of our internal databases to include information about aspects of both the structure and content of items, we will have

the capability to augment the human review of items with the inclusion of statistical evaluations of potential factors that may contribute to differential performance.

Research will also continue at the test level. We are still trying to investigate what factors help determine candidates' overall success on the MBE and which factors may contribute to differential performance. Based on the correlations observed among the performance measures, we have begun to examine some other psychometrically sophisticated models that can be used to predict candidate performance and to better understand how particular factors are related to bar examination performance (7, 10). Unfortunately, we are somewhat limited in this study sample because of the lack of availability of candidates' law school performance data which has been shown to have one of the strongest relationships with bar examination performance.

As our research progresses, we will continue to provide updates. This progress depends, at least in part, on a growing network of collaborative relationships among the many stakeholders in the legal education system. A thorough understanding of the determinants of candidate success will require a combination of data from many information sources. As always, NCBE welcomes opportunities to work with other groups to help facilitate and foster the process of collating and evaluating the data that will help the legal profession ensure that clients are well served by the new lawyers who are entering the practice. 📄

REFERENCES

1. Stilwell, Lisa Anthony, Reese Lynda M., and Pashley, Peter J. (1998). Analysis of Differential Prediction of Law School Performance by Racial/Ethnic Subgroups (Research Report 98-02). Newtown, PA: Law School Admission Council.

2. Pruitt, Lisa R., and Cassman, Celestail S.D. (2005). A Kinder, Gentler Law School? Race, Ethnicity, and Legal Education at King Hall. *U.C. DAVIS LAW REVIEW*, Vol. 38(4), 1209–1297.
3. Dowd, Nancy E., Nunn, Kenneth B., and Pendergast, Jane E. (2003). Diversity Matters: Race, Gender, and Ethnicity in Legal Education. *UNIVERSITY OF FLORIDA JOURNAL OF LAW AND PUBLIC POLICY*, Vol. 15(1), 11–57.
4. Wightman, Linda F. (2000). Beyond FYA: Analysis of the Utility of LSAT Scores and UGPA for Predicting Academic Success in Law School (Research Report 99-05). Newtown, PA: Law School Admission Council.
5. Wightman, Linda F. (1998). LSAC National Longitudinal Bar Passage Study. Newtown, PA: Law School Admission Council.
6. Klein, Stephen P., and Bolus, Roger (2004). Analysis of July 2004 Texas Bar Exam Results by Gender and Racial/Ethnic Group. See http://www.ble.state.tx.us/one/analysis_0704tbe.htm.
7. Kane, Michael T., Mroch, Andrew P., Ripkey, Douglas R., and Case, Susan M. (2006). Impact of the Increase in the Passing Score on the New York Bar Examination: July 2005 Bar Administration. Madison, WI: National Conference of Bar Examiners. See <http://www.nybarexam.org/ncberep.htm>.
8. Case, Susan M. (2006). Men and Women: Differences in Performance on the MBE. *THE BAR EXAMINER*, Vol. 75(2), 44–46.
9. Case, Susan M., Swanson, David B., and Ripkey, Douglas R. (1996). Performance of the Class of 1994 in the New Era of USMLE. *ACADEMIC MEDICINE*, Vol. 71(10), S91–S93.
10. Mroch, Andrew P., and Ripkey, Douglas R. (2007). Structural models relating LSAT, undergraduate GPAs, law school GPAs, and bar examinations. Paper presentation, Annual Meeting of the National Council on Measurement in Education. Chicago, IL, April 2007.



DOUGLAS R. RIPKEY is the Associate Director of Testing for the National Conference of Bar Examiners.



SUSAN M. CASE, PH.D., is the Director of Testing for the National Conference of Bar Examiners.