

RESEARCH ON THE CALIFORNIA BAR EXAMINATION

Stephen P. Klein, Ph.D.
September 9, 1987

This report summarizes the major findings of studies conducted on the California bar examination during the past 12 years. These studies investigated various characteristics of the exam (such as the adequacy of its time limits), essay grading practices, the stability of pass/fail standards, possible racial/ethnic and sex biases, and the relationship between traditional and non-traditional measures of lawyering skills.

A. ESSAY SECTION

- 1) The more questions an applicant answers, the more reliable the total essay score; i.e., the less likely the total essay score will be influenced by chance, such as an applicant's knowledge or ignorance of a particular facet of the law.
- 2) Total essay scores, by themselves, are not sufficiently reliable for making pass/fail decisions on individual applicants. This is true even if the essay section contains 12 questions.
- 3) Essay scores on different questions in the same content area usually correlate no more highly with each other than they do with scores on essay questions in other content areas. This finding suggests that differences in total essay scores among applicants are more a function of general legal knowledge and ability than they are of differences in the mastery of specific subjects.
- 4) Increasing the time applicants have to answer a question from 55 to 90 minutes usually resulted in only a slight improvement in score. It did not systematically change the relative standings of various groups. For instance, applicants who earned low scores under the shorter time limit did not benefit any more or less from the increased time than did other applicants.
- 5) Analytic (scorecard) grading is not as cost effective as holistic grading. The small increase in reliability gained through scorecard grading is more than offset by the substantially longer time it takes readers to grade papers using this approach.
- 6) Applicants who type (versus handwrite) their answers earn about the same essay scores as would be expected on the basis of their MBE scores; i.e., typing does not tend to raise or lower essay scores. We do not know whether applicants would earn the essay same scores if they were required to change from their preferred response mode.

B. PERFORMANCE TEST (PT)

- 1) PT answers are graded just as reliably as essay answers. And, the score on a PT problem is about as reliable as the score on an essay question.
- 2) Two PT problems usually correlate about as highly with each other as each one does with a typical essay question. This finding was considered in the decision to combine essay and PT scores into a single written score starting with the February 1987 exam.
- 3) Applicants who have been admitted to practice in other states generally score no higher or lower on the PT than would be expected on the basis of their scores on the other sections of the exam. This finding holds up even after controlling for length of time in practice and for attorneys who take just the PT and essay as well as for those who take all three sections.
- 4) Applicants believe the PT is a more valid measure of their ability to practice law than is either the Essay or MBE.

C. MULTISTATE BAR EXAMINATION (MBE)

- 1) There is a high correlation between MBE and essay scores. If the percent passing both sections was the same, the MBE and essay would make the same pass/fail decision for about 70 to 80 percent of the applicants. The rank ordering of law schools in terms of average MBE score is essentially identical to their rank ordering in terms of average essay score.
- 2) Similarity of content coverage does not appear to be the source of the close relationship between MBE and essay scores. For instance, scores on the MBE Contracts subtest usually correlate no more highly with scores on an essay question in Contracts than they do with scores on an essay question in Evidence. And, when an essay question is converted to a set of multiple choice questions, applicant scores on these items correlate more highly with MBE scores than they do with essay scores.
- 3) Differences in an applicant's grades among courses in law school do not necessarily coincide with differences in that applicant's scores on MBE subtests; e.g., doing especially well in contracts in law school relative to other subjects may or may not result in an elevated score on the MBE contracts section. The same finding also holds for the essay section.
- 4) MBE scores are very reliable. For example, an applicant's score on the first 100 questions provides a very accurate prediction of that applicant's score on the next 100 questions. Thus, even though applicants are allowed to guess, their scores are not due to chance.

- 5) Lengthening the time applicants are given to answer a set of MBE questions tends to increase scores slightly for all groups. For instance, all of the four most populous racial/ethnic groups who take the exam (Anglo, Asian, Black, and Hispanic) had about a 5 percent increase in average MBE score after being given more than a 50 percent increase in time limits.

D. REREADING AND REAPPRAISAL

- 1) The degree of agreement between readers in the scores they assign to the same essay answer is almost as high as when a single reader grades the same answer twice.
- 2) The average of two readers' grades on the same essay answer produces a more reliable score than does a single reading of that answer.
- 3) It is rare for two readers to assign substantially different grades to the same answer. When large differences do occur, they are almost always a function of chance variation around the applicant's true score rather than one reader making a clear mistake in grading.
- 4) Consistently using the higher (or lower) of the two scores assigned to an answer results in lower test reliability than does using the average of the two scores.
- 5) The scores assigned to an answer on the second reading tend to be slightly lower than those assigned on the first reading.
- 6) The reread and reappraisal bands have been set at about the right place in that they include essentially all of the applicants whose pass/fail status might be affected by reread and reappraisal.
- 7) The degree of agreement between reappraisers is comparable to the degree of agreement between readers (when the reappraisers do not see the grades assigned by the readers).
- 8) The reappraisal process does not increase test reliability, but it does increase the percent passing the exam. It also widens the gap in score between those who do and do not pass; i.e., as a result of reappraisal, only a handful of applicants fail by a few points.

E. TOTAL SCORES AND PASS/FAIL STANDARDS

- 1) The Total bar exam score (after all readings) provides a sufficiently reliable index for making pass/fail decisions about individual applicants.
- 2) The score required for passing in California is similar to that used by Oregon, but higher than the one used by most other states. For instance, California's passing rate for first timers would

increase substantially if it used New York's or Massachusetts' passing score.

- 3) There is a strong correlation between total bar exam scores and performance on legal problems that mirror closely several of the tasks that newly licensed attorneys would be expected to perform. And, the standard for passing these practice oriented problems was identical to the one used on the bar exam even though it was set without knowledge of the bar exam's standard.
- 4) In the past, the actual standard for passing in California was slightly lower in February than in July (as indicated by the MBE score that was required to produce the same passing rate as was actually observed on each exam). And, the standard for passing on both the February and July exams rose slightly during the last few years. These differences should disappear with the use of scaling.
- 5) First time takers are much more likely to pass than are repeaters. And, the higher the initial score, the greater the likelihood of eventually passing and doing so with just one or two more attempts. However, some repeaters eventually pass even though they had very low initial scores and/or made several previous attempts to pass.
- 6) Studies of first time takers on two recent exams reveal that about 70 percent of them eventually pass (i.e., after as many as five subsequent attempts). Almost 85 percent of all first time ABA takers eventually pass and about 63 percent of all minority ABA first timers eventually pass.
- 7) Most of the repeaters who eventually pass do so after their second or third attempt. The likelihood of passing after the third attempt is quite low.

F. CHANGES IN BAR EXAM PROCEDURES AND RULES

During the past 15 years, several procedures and rules were changed in an effort to improve the exam. Some of these changes, such as the use of multiple readers for applicants near the pass/fail line and greater emphasis on reader calibration, have had their anticipated beneficial effects on overall exam quality.

Other procedures did not have their intended impacts and were later discontinued. This latter category includes: (1) the use of extra long essay time limits, (2) giving applicants some choice in which essay questions they answered, and (3) allowing repeaters to pass the exam in sections (the so called "bifurcation" rule). For example, the repeaters' passing rate dropped when this rule was introduced and rose after it could no longer influence their success rate.

In February 1987, scores on the essay and PT were combined into one total score and this score was then converted ("scaled") to the same score distribution as that used on the MBE. The standard for passing

was set at the average of the standards used on the previous 20 exams. The February 1987 passing rate was the highest February rate in the last 10 years, thus "scaling" (and its associated pass/fail rules) did not lower the passing rate.

G. CORRELATES OF SUCCESS ON THE BAR EXAM

- 1) Law School Admission Test (LSAT) scores and law school grade point average (LGPA) are highly correlated with scores on all three parts of the bar exam (MBE, Essay, and PT). However, even when used together, LSAT and LGPA provide less than a perfect prediction of an applicant's pass/fail status on the exam.
- 2) School type (ABA, California Accredited, or Unaccredited) is not correlated with success on the bar exam after factoring out each type's average LSAT score. In other words, the generally higher bar exam passing rates of ABA schools is closely associated with these schools having students with relatively high LSAT scores.
- 3) The rank ordering of law schools in terms of the average LSAT scores of their graduates produces an almost perfect prediction of these schools' bar exam passing rates. It also provides a very good prediction of their average scores on the various sections of the exam. These findings may not apply to the many unaccredited schools that do not provide their graduates' LSAT scores.
- 4) Scores on the first year law student's exam are highly correlated with bar exam scores. An applicant who just passes the first year test is likely to fail the bar exam on the first attempt but pass on a subsequent attempt. Thus, the passing score on this test appears to be set at about the right place in that it tends to pass those students who have a reasonable chance of eventually passing the bar exam.
- 5) Applicants whose LSAT scores are well below the average of their classmates' LSAT scores (such as those who enter law school via special admissions programs) do about as well in law school and on the bar exam as would be expected on the basis of their LSAT scores. In other words, attending a school where most of the other students have substantially higher LGPAs and LSAT scores does not reduce or enhance an applicant's chances of passing the bar exam.
- 6) Applicants whose undergraduate major was English, Journalism, or Philosophy tended to earn slightly higher bar exam scores and especially essay scores than other applicants. However, undergraduate major was only a very weak predictor of bar scores. The pattern of undergraduate majors was quite similar across all racial/ethnic groups; e.g., about one third of the applicants in each group majored in History, Government, or Political Science.

H. ANALYSES BY SEX AND RACIAL/ETHNIC GROUP

- 1) After controlling for differences in overall difficulty across test sections, men tend to earn slightly higher MBE scores than essay or PT scores while the reverse is true for women. This trend even held for women who had an undergraduate major in the hard sciences.
- 2) In the past, women used to have higher average scores than men. However, this difference has been declining rapidly. This decline parallels the increasing percentage of applicants who are women.
- 3) Applicant sex does not predict bar exam scores after controlling on LGPA; i.e., overall, the exam does not tend to favor one group.
- 4) There are large differences in average scores among racial/ethnic groups. However, there also is considerable overlap of score distributions. Because of this overlap, racial/ethnic group is only a very weak predictor of an applicant's pass/fail status.
- 5) The small correlation between bar scores and racial/ethnic group disappears when there is control for the applicant's LGPA. In other words, the differences in passing rates among groups on the bar exam is comparable to the differences among them in law school grades. The bar exam neither widens or narrows this gap.
- 6) The foregoing findings hold for all three types of exam questions (MBE, essay, and PT). Thus, no section tends to favor a given group (after controlling for differences in reliability among sections).
- 7) None of the MBE, essay, or PT subtests tend to widen or narrow the differences in average score among groups. For instance, MBE Torts questions are not especially difficult for a given group relative to this group's performance on other MBE questions.
- 8) The grades assigned to an essay answer are not affected by whether the reader's racial/ethnic group is the same or different than that of the applicant who wrote the answer.
- 9) During each of the past 10 years, one person in 10 who passes the California bar exam belongs to a racial/ethnic minority group. About one applicant in seven who takes the exam for the first time is a member of a racial/ethnic minority group.
- 10) Minority applicants are more likely to come from the most selective law schools in the state (as reflected by their school's average LSAT score) than are Anglo applicants.
- 11) The eventual passing rates among minority group members is much higher than their initial rates. As a result, most minority applicants eventually pass the exam (although they are likely to require more attempts to do so than their Anglo classmates).

I. NOVICE/GRADUATE STUDY

Sections of the July 1985 bar exam were given to students at four ABA schools shortly after they began law school. Their scores were then compared with those of 1985 graduates from these same schools. The major findings of this study were:

- 1) The graduates scored substantially higher than the novices. In fact, none of the novices passed any section. And, the chances are essentially nil that any novice would even come close to passing the general bar exam.
- 2) LSAT scores were not related to the novices' bar scores, but were related to the graduates' bar scores.
- 3) Although there were large differences in average LSAT scores among racial/ethnic groups within the novice sample, there was no difference in their average bar exam scores.
- 4) The more reliable the bar exam test section, the greater the difference in average score between novices and graduates.

Taken together, these findings debunk various criticisms of the bar exam. They show that bar scores are not just a function of guessing, chance, or being adept at taking certain types of tests. Similarly, scores on essays and PT problems are not primarily a function of general writing or reasoning skills as distinct from the abilities developed in law school. And finally, although the PT provides applicants with all of the law, case materials, and facts on which to base their answers, applicants still need a legal education to do well on these problems.

J. ASSESSMENT CENTER AND SPECIAL SESSION STUDIES

Several studies were conducted in conjunction with the July 1980 exam. These studies examined how well the MBE and essay sections correlated with more practice oriented measures of lawyering skills. These measures included tests of trial practice skills, legal research skills, interviewing and counseling clients, examining witnesses, making opening statements and closing arguments, preparing legal memos, etc. Some of these tasks were oral whereas others were written and/or observational. The major findings of this research were that:

- 1) Applicants who had the abilities needed to do well on these tasks also tended to have the skills and knowledge that were required to do well on the bar exam. This conclusion stems from the finding that scores on the special observational, oral, and written tasks correlated highly with regular bar exam scores (and to about the same degree as the MBE and essay correlated with each other).
- 2) Although an applicant's responses on an experimental tasks usually took longer to grade than a typical essay answer, the scores on these tasks were just as reliable as they were on an essay answer.

- 3) After controlling for differences in reliability among measures, scores on the practice oriented tests did not widen or narrow the gap in average performance level among racial/ethnic groups
- 4) Most applicants thought the practice oriented tasks were a better indicator of their ability to function as an attorney than were either the MBE or the essay.
- 5) The pass/fail standards established by an independent panel for the practice oriented tasks coincided with the pass/fail standard for the bar exam even though the panel did not know the applicants' bar exam scores.

REFERENCES

Two reports are prepared for each administration of the California bar exam: Summary Statistics (prepared by Philip Schoner) and Technical Report (by Stephen Klein). The American College Testing Program (ACT) provides a report on each administration of the MBE (the Educational Testing Service prepared similar reports on the MBEs given prior to the July 1984 exam).

Except as noted otherwise, all the references below were prepared by Stephen Klein as reports for the Committee of Bar Examiners.

A. Essay Section

An analysis of grading practices on the California Bar Examination (1977).

Summary of studies conducted for the Committee of Bar Examiners and the statistical rationale underlying proposed changes (1978).

Comparison of an 8, 9, and 12 answer essay test in a multiphased pass/fail decision model (1978).

The effect of time limits, item sequence, and question format on applicant performance on the California Bar Examination (1981). (PR-81-7).

Essay grading: Fictions, facts, and forecasts. The Bar Examiner, 1985, 54, 23-29.

Relationship between MBE and essay scores among applicants who did and did not type their answers (Memo of February 28, 1987).

B. PERFORMANCE TEST

An analysis of the Performance Test on the July 1983 California Bar Examination (1984).

See also the Technical Report on each exam since July 1983.

C. MULTISTATE BAR EXAMINATION

Carlson, A. B. and C. Werts (1976). Relationships among law school predictors, law school performance, and bar examination results. Princeton, NJ: Educational Testing Service.

The effect of time limits, item sequence, and question format on applicant performance on the California Bar Examination (1981).

An evaluation of the Multistate Bar Examination. National Conference of Bar Examiners, Chicago (1982).

Summary of research on the Multistate Bar Examination. The Bar Examiner, 1983, 52, 10-15.

See also the Technical Report on each exam and the reports prepared by ACT and ETS on each administration of the MBE.

D. REREADING AND REAPPRAISAL

Intra- and inter-reader agreement on the essay section of the California Bar Examination (1980).

A comparison of the effectiveness of a single versus multiphased grading system (1980).

Large discrepancies between readers: Their source and policy implications (1985).

Schoner, Philip G. and Alex Millar (1987). The relationship between tentative/calibration grades and general bar examination grades.

See also the Technical Report on each exam.

E. TOTAL SCORES

Summary of studies conducted for the Committee of Bar Examiners and the statistical rationale underlying proposed changes (1978).

An analysis of possible variations in pass/fail standards on the California Bar Examination (1981). (PR-81-1)

An analysis of the relationship between clinical skills and bar examination results (1982).

An analysis of the relationship between initial score and eventual pass/fail status on the California Bar Examination (1982). (PR-81-9)

Results of the July 1982 and July 1984 Tracking Studies (DR-87-5 in preparation).

See also the Summary Statistics and Technical Report on each exam.

F. CHANGES IN BAR EXAM PROCEDURES AND RULES

How the bifurcation rule affected the percent passing California's General Bar Examination (1985). (PR-85-5)

Minority group performance on the California bar examination (in preparation DR-87-2).

See also the Technical Reports on each exam.

G. CORRELATES OF SUCCESS ON THE BAR EXAM

An analysis of possible sex and racial/ethnic biases in the July, 1976 California State Bar Examination (1978).

An analysis of the relationships between bar examination scores and an applicant's law school, admissions test scores, grades, sex, and racial/ethnic group (1979). Summary reprinted in The Bar Examiner, 1980, 49, 14-18.

Preliminary report on the relationship between undergraduate major and bar examination scores (1982).

Minority group performance on the California bar examination (in preparation DR-87-2).

H. ANALYSES BY SEX AND RACIAL/ETHNIC GROUP

An investigation of possible item and grader biases in a state bar examination. Paper presented at the meetings of the Western Psychological Association, Los Angeles, California, April 9, 1976.

Summary of studies conducted for the Committee of Bar Examiners and the statistical rationale underlying proposed changes (1978).

An analysis of the relationships between bar examination scores and an applicant's law school, admissions test scores, grades, sex, and racial/ethnic group (1979). Summary reprinted in The Bar Examiner, 1980, 49, 14-18.

An analysis of the relationship between initial score and eventual pass/fail status on the California Bar Examination (1982).

Minority group performance on the California bar examination (in preparation DR-87-2).

See also the Summary Statistics and Technical Report on each exam.

I. NOVICE/GRADUATE STUDY

The performance of novice law students and law school graduates on the bar exam (1986).

J. ASSESSMENT CENTER AND SPECIAL SESSION STUDIES

Testing research skills on the California Bar Examination (1981).

An analysis of the relationship between clinical skills and bar examination results (1982).

An analysis of the relationship between trial practice skills and bar examination results (1982).

