

THE EFFECT OF TIME LIMITS, ITEM SEQUENCE, AND QUESTION FORMAT
ON APPLICANT PERFORMANCE ON THE CALIFORNIA BAR EXAMINATION

A Report Submitted to the
Committee of Bar Examiners of the State Bar of California
and the National Conference of Bar Examiners

Prepared by

Stephen P. Klein
GANSK & ASSOCIATES

October 15, 1981

PREFACE

This report describes a series of studies that was funded by the Committee of Bar Examiners of the State Bar of California and the National Conference of Bar Examiners. These studies, which were conducted in conjunction with the July 1980 administration of the California State Bar Examination, investigated the effects on applicant performance of certain bar examination procedures so as to assess the probable impacts of various changes that might be made in these procedures.

Two other sets of studies also were conducted in conjunction with the July 1980 administration of the California examination. These studies focused on measuring the degree to which scores on the essay and multiple choice portions of the bar examination corresponded with scores on tests of an applicant's ability to perform various practice related tasks, such as conducting legal research. A description of these other studies has been provided by O'Hara and Klein (1981).

SUMMARY

The July 1980 version of the California Bar Examination was divided into two parts. One part, called the General Bar Examination (GBX), had two sections; the Multistate Bar Examination (MBE) and the Essay Examination. The MBE had 200 multiple choice questions (items) and all the applicants participating in the research described in this report had MBE scores. The Essay section had nine questions, however, because of California's multi-phased grading process, about 30 percent of the applicants had only three of their essay answers graded.

The other part of the July 1980 examination was called the Special Session (SS). The SS was optional; i.e., an applicant could choose not to participate in it without jeopardizing his/her chances of passing the GBX. If an applicant participated and did well in the SS, it could increase that applicant's chances of passing. About 99 percent of the applicants who sat for the complete GBX also took the SS.

Applicants were assigned randomly to one of the SS's four sections. Since these sections were administered concurrently, a given applicant could take only one of them.

The results obtained with two of the SS's four sections are described in this report. These two sections were used to investigate the effects on applicant performance of certain bar examination procedures so as to assess the probable impacts of various changes that might be made in these procedures. The major questions addressed by this research were:

- o Is performance on multiple choice questions affected by the order in which the questions are asked and/or the amount of time applicants are given to answer them?
- o Is performance on essay questions affected by the amount of time applicants are given to answer them?
- o When content coverage is held constant, is the relative standings of applicants affected by whether their legal skills and knowledge are measured by essay or multiple choice questions?

MBE TIME LIMITS AND ITEM ORDER STUDIES

The MBE Time Limits and Item Order Studies involved four groups of applicants (with 705 to 750 applicants per group). Each group took two tests and each test consisted of 30 multiple choice questions (items). These items were drawn from previous, but still secure versions of the MBE. The applicants had 55 minutes to complete the first test (i.e., the regular time per item) and 90 minutes to complete the second test. There were two versions of each test (designated as 1 and 2). These versions differed solely in terms of the sequence in which blocks of items were asked. Test forms were assigned to groups so that each version of each set of 30 items was answered by one group under the 55 minute time limit and by another group under the 90 minute time limit.

Results with the MBE Time Limits portion of this research indicated that:

- o The average score under the 90 minute time limit was 0.87 points higher than it was under the 55 minute time limit. This finding suggests that if applicants were given essentially unlimited time to answer the MBE's 200 items, their average raw score would increase by about 6 points.
- o Applicants did not benefit equally from the extra time (only 58 percent had higher scores under the 90 minute time limit than they did under the 55 minute time limit). Some of the increase and decrease in scores was due to chance (e.g., guessing, the degree to which an applicant was familiar with the specific issues addressed by the items on each test, etc.); i.e., the scores were not perfectly reliable.
- o When scores on the tests were corrected for unreliability, there was a near perfect correlation between them. Corrected correlations between scores on the full MBE and the 30 item tests were the same and near perfect under both time limits. These findings suggest that the differences in the relative standings of the applicants under the different time limits were due to their taking different MBE questions in each time limit rather than to the extra time per se.
- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's background characteristics (such as age, sex, race, type of law school attended, and repeater status).
- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's performance level on the full examination or that applicant's attitudes about the experimental and full examination (e.g., such as the adequacy of time limits and the efficacy the tests as measures of legal ability).

Raw scores on the MBE are adjusted through a scaling process in order to control for possible differences in average question difficulty across administrations of this test. In California, pass/fail decision on the MBE are based on the scale scores (rather than the raw scores). Thus, while allowing applicants more time on the MBE would generally result in slightly higher raw scores, it would not affect the scale scores. The percent passing the MBE in California (or any other jurisdiction that uses scale scores) would therefore not be increased by an increase in the time applicants had to answer. The analyses conducted in this study also indicated that giving applicants more time to answer MBE questions would have no significant impact on their relative standings; i.e., who passes versus fails this test.

Results with the Item Order portion of this research indicated that the ordinal position of an item in a test did not significantly affect the percentage of applicants answering it correctly or its correlation with other items. Variations in item order also did not affect the reliability of the test (there was only a .01 difference in reliability coefficients between versions under the regular time limit).

It may be inferred from these results that varying the sequence in which MBE items are presented will not affect an applicant's chances of passing this test. Thus, multiple forms of the MBE could be used on a given administration of the examination so as to eliminate any benefit an applicant might derive from observing another applicant's answers.

ESSAY TIME LIMITS STUDY

The Essay Time Limits Study involved eight groups of applicants (with 198 to 229 applicants per group). Four essay questions were used; two from California's regular pool of essay questions and two that were specially constructed from MBE fact patterns and items. Applicants were given one question from each source. One of these questions dealt with Contracts or Real Property while the other question dealt with Torts or Evidence.

The applicants had 55 minutes to answer the first question and 90 minutes to answer the second question. Questions were assigned to groups so that some groups answered a given question under the 55 minute time limit while other groups answered this same question under the 90 minute time limit. All the answers to a question were graded in a random order so that the readers did not know under which time limit a given answer was written.

The major findings of the Essay Time Limits Study were as follows:

- o The average score on a question under the 90 minute time limit was higher than it was under the 55 minute time limit on all four of the experimental questions. There was an average increase of 4 points per question, but performance on some questions was more affected by the extra time than it was on other questions.
- o Attorneys with extensive experience in essay test construction and grading were not able to predict (either individually or collectively) on which questions scores would be most affected by the different time limits.
- o Applicants did not benefit equally from the extra time (only 67 percent had higher scores under the 90 minute time limit than under the 55 minute time limit). Some of the increase and decrease in scores was due to chance (e.g., grader unreliability and the degree to which an applicant was familiar with the issues addressed by each question). It also is possible that the extra time gave some applicants sufficient opportunity to demonstrate their true performance level (which in turn resulted in their not being given the benefit of any doubt that might have been created by a shorter answer).
- o When essay scores were corrected for unreliability, there was a near perfect correlation between scores on an experimental question under the 55 minute time limit and Phase I essay scores (i.e., the total score on three essay questions given under the normal time limit). This correlation dipped only slightly when applicants were given 90 minutes to answer an experimental question. These findings suggest that the differences in the relative standings of the applicants under the different time limits were due to their answering different questions in each time limit rather than to the extra time per se.

- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's background characteristics (such as age, sex, race, type of law school attended, and repeater status).
- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's performance level on the full examination or that applicant's attitudes about the experimental and full examination (e.g., such as the adequacy of time limits and the efficacy of the tests as measures of legal skill and knowledge).

It is evident from the foregoing results that giving applicants 64 percent more time to answer essay questions would not affect their relative standings any more than would just giving them a different set of essay questions to answer under the regular time limits. However, giving all applicants extra time might result in generally higher total essay scores for most (but not necessarily all) applicants. If this increase was as large as the one obtained in this research, it would substantially increase the percent passing the California examination. Thus, in terms of policy implications, the 4 point per question increase with the extra time is potentially the most significant finding presented in this report.

There are several reasons why it is not certain that the 4 point per question increase observed in this study would be maintained if all applicants were given additional time. Some of the reasons for not being able to predict what would occur include:

- o If more time were allowed, the nature of the questions asked might change, such as their containing more issues and/or more difficult issues.
- o Scores on some questions were much more affected by the extra time than were scores on other questions. Thus, the size of the increase would be affected by the nature of the question asked and there did not appear to be any way of determining in advance on which questions scores would be most affected by the time allowed to answer them.
- o Grading criteria and standards may shift when readers score answers written under an expanded time limit. They may consciously or unconsciously place more weight on organization as well as expect a higher quality of answer for passing even if they are given instructions not to make such changes in their grading practices.
- o It is not known whether the average increase in score was due to applicants identifying more issues, to discussing the issues better, or to their just writing more. Specifically, the answers written under the 90 minute time limit may have been longer (but not necessarily better) than those written under the 55 minute time limit. Previous research has indicated that longer answers tend to receive higher grades than shorter ones. If all the answers are written under the same time limit, any affect of answer length on score probably would be reduced.

The foregoing issues are not raised to minimize the potential significance of the 4 point per question increase with the extra time. Rather, they indicate the need for additional studies to determine the source of the increase so that appropriate policy decisions can be made.

Giving applicants more time to answer essay questions could affect costs, test reliability, and validity. For example, if the total time devoted to the essay section was maintained (i.e., 9 hours) but the amount of time allocated per question increased, then there would be a reduction in the number of questions that could be asked which in turn would reduce scoring costs and probably reliability. It also would reduce the breadth of content that could be covered in any one administration of the examination. If the number of questions remained the same but the amount of time per question increased, readers might have to be paid more (because the answers would take longer to score) and there would be the additional costs associated with longer test administrations. Applicant fatigue also might become a factor affecting scores.

Finally, giving applicants more time to answer essay questions would not affect the percent passing in jurisdictions that scale their essay scores to the MBE. The reason for this is that such scaling removes the effect of any variations in essay test difficulty across administrations (whether this variation stems from differences in question difficulty, grader leniency, or the amount of time applicants had to answer the questions).

MBE/ESSAY CONGRUENCE STUDY

A 35 minute, 20 item, multiple choice test was administered to all the applicants who participated in the Essay Time Limits study. This test had one set of five items for each of the four essay questions used in that study. Each set of five items was designed to measure the same issues as were addressed in its corresponding essay question. Ten of the 20 items were constructed by the staff of the Committee of Bar Examiners to coincide with the two essay questions developed by the Committee. The other ten items, which were drawn from previous administrations of the MBE, served as the basis for the staff's construction of the other two experimental essay questions. Since a given applicant answered only two of the four essay questions used in the Essay Time Limits study, 10 of the multiple choice items dealt with issues an applicant had encountered in the essay study and 10 dealt with new issues. Four forms of the 20 item test were used (two groups per form). These forms differed in terms of the sequence in which the item sets appeared in the test booklet.

The major findings of the MBE/Essay Congruence Study were as follows:

- o The sequence in which the 20 multiple choice items were asked did not affect performance on these items or test reliability. These findings are consistent with those obtained in the Item Order Study.
- o Applicants did only very slightly better on the items sets corresponding to essay questions they answered previously than they did on item sets that dealt with essay questions they had not answered.

- o Whether an applicant answered an essay question under the 55 or 90 minute time limit had no affect on that applicant's score on item set corresponding to that essay question.
- o Total scores on an item set were generally no more highly correlated with scores on that item set's corresponding essay question than they were with any other essay question.
- o Scores on essay questions in the same general area (e.g., contracts) correlated no more highly with each other than did scores on essay questions in different areas. This finding could result from applicants balancing their preparation for the examination and/or the essay test measuring skills that are not highly content specific.

The foregoing findings suggest that MBE and California essay questions will not rank order applicants in exactly the same way even when the content on these two types of questions is held constant. Nevertheless, scores on the the two types of questions are highly related. This means that applicants who have the skills and knowledge required to do well on MBE items also tend to have the abilities needed to do well on essay questions.

ACKNOWLEDGMENTS

The Committee of Bar Examiners of the State Bar of California provided the advice and cooperation that were necessary for carrying out the studies described in this report. Two members of this committee, Armando M. Menocal, III and Martin R. Glick, made especially important contributions in the design, implementation, analysis, and report preparation phases of this research.

The Committee's staff constructed, printed, and administered the tests and questionnaires that were developed for the project. This staff, along with the Committee's Board of Reappraisers, also selected, trained, and calibrated the cadre of lawyers who graded the essay questions. Kenneth D. McCloskey, the Committee's former Director of Testing, was in charge of and participated actively in all of these activities. The efficiency with which these tasks were completed, the quality of the measures produced and the data obtained from them, and the fact these activities were carried out concurrently with testing and processing over 7,500 applicants on the regular bar examination testify to the skills, energies, and unstinting dedication that Ken provided. Ken was ably assisted in these activities by James B. Tippin, Jr., Suzanne M. Obermeier, and Philip Schoner. Phil also provided invaluable help in coordinating readers, data cleaning, and records management.

John Bianchini and Andrew York of the Educational Testing Service developed and implemented the computer systems that were used in assigning applicants to experiments and in cleaning and linking diverse data files. Randy Onishi and Roger Bolus of GANSK & Associates also participated in the computer data cleaning and file management activities. Roger was further responsible for conducting the statistical analyses presented in this report.

Ralph Hoepfner of the System Development Corporation provided a thorough and extremely helpful technical review of a draft of this report.

Finally, the project could not have been carried out without the cooperation of the applicants who participated in it. Their tolerance and generally good nature in putting up with the many unique demands of the research at a time when they were under great pressure were most appreciated by all who were involved in this project.

CONTENTS

PREFACE..... iii

SUMMARY..... v

ACKNOWLEDGMENTS..... xi

LIST OF TABLES..... xiii

Chapter

1. INTRODUCTION..... 1

2. GENERAL PROCEDURES..... 2

 Overview..... 2

 General Bar Examination..... 2

 Special Session..... 5

 Background Characteristics..... 5

3. MBE TIME LIMITS AND ITEM ORDER STUDIES..... 7

 Purposes..... 7

 Procedures..... 7

 Results..... 8

 Summary and Conclusions..... 13

4. ESSAY TIME LIMITS STUDY..... 15

 Purpose..... 15

 Procedures..... 15

 Preliminary Analyses..... 16

 Results..... 17

 Summary and Conclusions..... 21

5. MBE/ESSAY CONGRUENCE STUDY..... 24

 Purpose..... 24

 Procedures..... 24

 Results..... 25

 Summary and Conclusions..... 27

REFERENCES..... 29

Appendix

A. Glossary of Abbreviations and Technical Terms..... 30

B. Policy Regarding the Assignment of Grades..... 31

C. Description of the Multiphased Grading Process..... 32

D. Excerpts from Notice to Applicants Regarding Special Session..... 34

E. Pre-examination Questionnaire..... 35

F. Post-examination Questionnaire..... 37

G. Directions and Questions for Essay Time Limits and Essay/MBE..... 39

 Congruence Studies

H. Test Administration Instructions for the Special Sessions..... 44

LIST OF TABLES

2.2	Area(s) Covered by Each Essay Question.....	3
2.2	Summary Statistical Data on the July 1980 GBX.....	4
3.1	Sequence of Items Within Forms.....	8
3.2	Group Means and Standard Deviations on the MBE.....	9
3.3	Means and Standard Deviations by Group and Time Limit.....	9
3.4	Percentage of Applicants Changing Score Relative to the Percentage that Would have Changed if the MBE's Time Limits had no Affect on Performance	9
3.5	Correlations with and Between Item Set Scores.....	10
3.6	Differences Between Version Means.....	11
3.7	Differences Between Version Standard Deviations.....	11
3.8	Differences Between Version Reliabilities.....	11
3.9	Correlation of Item Statistics Between Versions and Time Limits...	12
4.1	Assignment of Questions to Groups.....	15
4.2	Average Essay Question Score on the GBX.....	16
4.3	Reader Reliability Statistics.....	17
4.4	Average Score in Each Time Period by Group.....	17
4.5	Average Score on Each Question Under Each Time Limit.....	18
4.6	Percentage of Applicants Changing Score Relative to the Percentage that Would have Changed if the Essay's Time Limits had no Affect on Performance	19
4.7	Correlations with Phase I Essay Scores Under Each Time Limit.....	20
5.1	Sequence of Item Sets Within Forms.....	25
5.2	Group Means on Item Sets and Total Test.....	25
5.3	Correlations Between Item Sets.....	26
5.4	Correlations Between Item Sets and Essay Questions.....	27
C.1	Distribution of July 1980 Applicants Across Pass/Fail Categories..	33

Chapter 1
INTRODUCTION

The studies described in this report were conducted to determine the effects on applicant performance of certain bar examination procedures so as to assess the probable impacts of various changes that might be made in these procedures. The major questions addressed by this research were as follows:

- o Is performance on multiple choice questions affected by the order in which the questions are asked and/or the amount of time applicants are given to answer them?
- o Is performance on essay questions affected by the amount of time applicants are given to answer them?
- o When content coverage is held constant, is the relative standings of applicants affected by whether their legal skills and knowledge are measured by essay or multiple choice questions?

In order to answer these questions, several experiments were conducted in conjunction with the July 1980 administration of the California bar examination. Chapter 2 of this report describes the major features of this examination and the background characteristics of the applicants who took it. The next three chapters correspond to the three questions above; i.e., for each question, there is a chapter that describes the procedures that were used to answer it and the results obtained. Throughout the report, special attention is given to determining whether general trends hold for applicants who differ with respect to such variables as racial /ethnic group, sex, type of law school from which they graduated, and whether they have taken (and failed) the examination previously.

A glossary has been provided in Appendix A for those readers who may not be familiar with some of the technical terms used in this report.

Chapter 2 GENERAL PROCEDURES

OVERVIEW

The July 1980 version of the California Bar Examination was divided into two parts. One part, called the General Bar Examination (GBX), had two sections; the Multistate Bar Examination (MBE) and the Essay examination. Most of the applicants who took one section also took the other section.

The other part of the July 1980 examination was called the Special Session (SS). The SS was optional; i.e., applicants could choose not to take it without jeopardizing their chances of passing the GBX. However, if an applicant participated and did well in the SS, it could increase that applicant's chances of passing. Thus, most applicants who sat for the GBX also took the SS.

The SS had four sections that were administered concurrently. A given applicant could, therefore, take only one of these sections. The results obtained with two of the SS's four sections are the focus of the remaining chapters of this report. Information about the background characteristics of the applicants who participated in the SS was obtained from state bar records and questionnaires.

GENERAL BAR EXAMINATION (GBX)

The July 1980 General Bar Examination (GBX) was administered over a three day period. The first day and the afternoon of the second day were devoted to the Essay. The MBE was given on the third day.

Multistate Bar Examination

The Multistate Bar Examination (MBE) is developed by the National Conference of Bar Examiners and is published and scored by the Educational Testing Service. About 45 states and the District of Columbia use the MBE as part of their bar examinations.

The MBE consists of 200 multiple choice questions (or "items"). The six areas covered by the MBE and the number of items in each area are as follows: Constitutional Law (30), Contracts (40), Criminal Law (30), Evidence (30), Real Property (30), and Torts (40). Some of the questions relate to a common set of facts while other items stand by themselves. Items are not grouped by area within the test.

The MBE is administered in two test sessions. Each session has 100 items (drawn proportionately from the six areas). Applicants are given three hours to complete each session and a break between sessions for lunch.

Each MBE item has four choices. An applicant's raw score is the number of questions answered correctly; i.e., there is no correction for guessing. The raw scores are converted to "scale scores" by the Educational Testing Service so that an applicant's final score is not affected by possible differences in overall test difficulty across administrations of the test. California multiplies these scale scores by 3.0 so that the theoretical maximum score on the MBE in California is 600 points.

Essay Examination

The Essay portion of the July 1980 GBX contained nine questions. These questions were developed by law professors teaching at non-California schools and by the California Committee of Bar Examiners and its staff. Some of the questions focused on only one subject matter area while others, called "cross-over questions," covered two areas. Table 2.1 lists the area(s) covered by each question.

Table 2.1
AREA(S) COVERED BY EACH ESSAY QUESTION

Question	Area(s) Covered
1	Torts/Civil Procedure
2	Community Property
3	Constitutional Law
4	Remedies/Contracts
5	Wills & Succession/Trusts
6	Evidence
7	Remedies/Contracts
8	Corporations
9	Criminal Law/Constitutional Law

The Essay section was divided into three test sessions of three hours each with three questions per session. Applicants were allowed to allocate the three hours in any way they wished across questions within a session.

The grading of the essay answers was done by a team of 12 lawyers per question (i.e., there were a total of 108 readers). Each reader was trained to use the same criteria and standards in evaluating the answers to his/her question. The answers were graded in five point increments on a scale of 0 to 100 points. A score of 70 was assigned to an answer if the reader was in doubt as to whether the answer was a minimally adequate response to the question (see Appendix B). The general criteria for grading are presented in The Bar Examiner's Handbook (Sprecher, 1968).

Pass/Fail Decisions

There are two ways an applicant can pass the GBX. Method 1 consists of receiving a combined Total score (i.e., MBE + Essay) of 1050 or greater. Method 2 involves receiving an MBE score of 420 or greater on one administration of the GBX and an Essay score of 630 or greater on another administration; i.e., in order to pass by the second method, an applicant must pass both sections of the examination. If an applicant fails the GBX because the applicant's score on one section was not high enough to compensate for a failing score on the other section, the applicant can retake the entire GBX (and thereby have the opportunity to pass under either method) without jeopardizing the passing status on the section passed previously.

Multiphased Grading

Research on California bar examinations (Klein, 1979 and 1980a) has indicated that final pass/fail decisions can be predicted with 99.5% accuracy without evaluating all the essay answers written by many applicants. This finding along with other considerations led the Committee of Bar Examiners to adopt a multiphased grading process. This process involves using the scores on three of the nine essay questions plus the MBE to identify those applicants who are very likely to pass and then focusing reader time on the applicants whose pass/fail status is most in doubt. A more complete description of this process appears in Appendix C.

One consequence of the multiphased system is that total Essay scores were available for only 70 percent of the applicants taking this section. The group with complete data were not representative of the total applicant pool. Thus, for the purposes of this research, a total Essay score was created for an applicant with only three answers graded by multiplying by 3.0 that applicant's total score on the three questions that were graded. If an applicant did not have at least three essay questions graded, then no Essay score was created for that applicant.

Summary Statistical Data

Table 2.2 presents summary statistical data on the July 1980 GBX for the 7379 applicants who took both the MBE and Essay portions of this test. The reliability of the MBE was computed by Dorans and Wright (1980). The reliability of the Essay portion was computed by using the Spearman-Brown formula to step up the average inter-question correlation within sessions ($r = .32$). Since the multiphased grading process resulted in only some of the applicants having all of their essay answers read and since these applicants were not representative of all the applicants who took the examination, it was not possible to compute the average correlation across all questions. The reliability of the total examination score was estimated using the procedures described by Gulliksen (1950) for a linear composite. There was a .68 correlation between the MBE and Essay sections.

Table 2.2

SUMMARY STATISTICAL DATA ON THE JULY 1980 GBX

<u>Statistic</u>	<u>MBE</u>	<u>Essay</u>	<u>Total</u>
Average Score	424.5	617.5	1042.0
Standard Deviation	46.0	54.4	92.2
Reliability	.88	.81	.91

The average standard deviation of an essay question in Phase I was 9.5.

Some of the other characteristics of the applicant pool were: 49 percent passed the GBX, 69 percent took the GBX for the first time, and 62 percent were graduates of American Bar Association approved law schools. Within this latter group, 81 percent took the GBX for the first time. The passing rate among these ABA first timers was 73 percent. These results are very consistent with those obtained with past administrations of the examination (Klein, 1981a).

SPECIAL SESSION (SS)

Prior to the administration of the examination, applicants were advised that there would be a Special Session (SS). They were further advised that (1) their scores in the SS would not be counted if they passed the regular examination (i.e., based on the total of their Essay and MBE scores) and (2) their scores in the SS would be counted as one-sixth of their total grade if they failed the regular examination (see Appendix D). Thus, participating in the SS could increase but not decrease their chances of passing. Applicants also were informed that they had to take the MBE and Essay portions of the GBX in order to derive any benefit from taking the SS and that they would be assigned randomly to one of its four sections.

An applicant was assigned to one of SS's four concurrently administered sections. A stratified random sampling plan was used for this purpose. The stratification variables were race, school type, and repeater status. The number of applicants assigned to a section was based on how many would be needed for statistical analyses of the data and the cost of scoring and administering each section.

The SS was administered on the morning of the second day of examination. Of the 7379 applicants who took the complete GBX, 98.5 percent elected to participate in the SS.

The data used in this report came from two of the four SS sections. One of these sections involved giving applicants 55 minutes to answer one set of 30 multiple choice question and then 90 minutes to answer a different set of 30 multiple choice questions. Both sets of questions were drawn from previous administrations of the MBE. The other SS section involved giving applicants 55 minutes to answer one essay question, 90 minutes to answer a different essay question, and then 35 minutes to answer a set of 20 multiple choice questions. Ten of these multiple choice questions covered the same content areas and issues as were addressed by the essay questions the applicants answered in the first two parts of the section. A more complete description of each section is presented in the subsequent chapters of this report.

BACKGROUND CHARACTERISTICS

Law School Type and Repeater Status

All persons applying to take California's bar examination are required to provide a transcript of their law school grades. These transcripts made it possible to determine the type of law school from which an applicant graduated; i.e., American Bar Association (ABA) approved, California Accredited (but not ABA approved), and Unaccredited. State bar records also were used to determine whether an applicant had taken and failed the GBX previously; i.e., whether or not the applicant was a repeater.

Sex and Racial/Ethnic (R/E) Group

All applicants taking the bar examination in California are requested to complete a form on which they indicate their sex and racial/ethnic (R/E) group affiliations. This information, which is used solely for research purposes, was provided by 99.6 percent of those taking the July 1980 examination. The distribution of sex and racial/ethnic groups on this examination was as follows: males (71%) and females (29%); Anglos (82%), Asians (4%), Blacks (6%), and Hispanics (6%). The Asian group is composed mainly of Chinese and Japanese Americans, but it also includes those with Philippino and Pacific Islander backgrounds. The Hispanic group is largely Mexican-American, although it contains small percentages of applicants with Central American, South American, and Puerto Rican backgrounds.

Questionnaire Data

Applicants were advised that in order to derive any benefit from their participation in the Special Session (i.e., in terms of passing the GBX), they had to complete and return the questionnaire that appears in Appendix E. This questionnaire requested information about certain potentially relevant background characteristics, such as the extent to which English was spoken in the home, and data about an applicant's legal training and experience.

All applicants who applied to take the July 1980 examination were mailed a copy of the questionnaire prior to the examination. Applicants who did not return the questionnaire by early July were sent another copy of it. They also were advised again that it had to be completed in order for them to derive any benefit from the Special Session and that they could turn it in at the examination site. Of the 7379 applicants taking the complete GBX (and thereby eligible to benefit from the Special Session), 96 percent returned questionnaires. And, almost all of those who completed both the full GBX and the Special Session returned questionnaires.

A questionnaire also was administered between the morning and afternoon sessions on the third day of the examination. This questionnaire inquired about the applicants' preparation for the examination; assessment of the adequacy of the time limits for the MBE, Essay, and Special Session; and opinions about how well each of these sections measured their legal skills and knowledge. A copy of this questionnaire appears in Appendix F. About 95 percent of the applicants who participated in the SS and took the complete GBX turned in this questionnaire.

Chapter 3
MBE TIME LIMITS AND ITEM ORDER STUDIES

PURPOSES

Although the MBE Time Limits and Item Order studies were conducted with the same groups of applicants and tests, their purposes were quite different.

The MBE Time Limits study investigated the validity of a frequent applicant complaint about this test, namely; "there is not enough time to answer the questions." The fact that 99.5 percent of the applicants answer every question has not quelled this concern since the 'no penalty for guessing' rule encourages applicants to mark one choice for every question. The specific issues addressed in the MBE Time Limits Study were as follows:

- o Would giving applicants substantially more time to answer MBE questions improve their level of performance on this test?
- o Would all applicants benefit equally from the extra time; i.e., would their relative standings be affected?
- o What are the characteristics of applicants who tend to benefit most versus least from the extra time?

The Item Order Study was conducted to determine whether performance on the MBE is affected by the order in which the questions are asked. If the sequence of questions does not affect scores, then different forms of the same test could be used during a given administration. And, the use of multiple forms would facilitate the task of proctoring the examination since there would be less temptation for an applicant to look at some other applicant's answer sheet.

PROCEDURES

This study used 60 multiple choice questions (items) that were drawn from previous but still secure versions of the MBE. The four content areas covered by these items were: Contracts, Criminal Law, Real Property, and Torts. The 60 items were divided into two 30 item sets, designated as Set A and Set B. Two versions of each set (designated as Versions 1 and 2) were constructed. These versions differed solely in terms of the sequence in which three blocks of 10 items were asked. Thus, there were four test forms; designated as Forms A-1, A-2, B-1, and B-2 (see Table 3.1).

Table 3.1

SEQUENCE OF ITEMS WITHIN TEST FORMS

TEST FORM		SEQUENCE		
Set	Version	1st	2nd	3rd
A	1	1-10	11-20	21-30
A	2	21-30	11-20	1-10
B	1	31-40	41-50	51-60
B	2	51-60	41-50	31-40

The 2940 applicants that participated in this research were assigned randomly to four groups. These groups had comparable means and standard deviations on the MBE portion of the GBX (see Table 3.2).

Each group was given 55 minutes, the standard time per MBE item, to answer item set A or B and then 90 minutes to answer the other set. Table 3.3 shows the assignment of forms to groups relative to the time given to answer these forms.

RESULTS

Time Limit Effect on Level of Performance

Table 3.3 presents the mean and standard deviation in each group under each time limit. An inspection of these data indicates that the average score under the 55 and 90 minute time limits differed by slightly less than one point (i.e., 21.19 versus 20.32 = .87). Moreover, almost all of the applicants who participated in this experiment indicated they had more than enough time to answer the items when they were given 90 minutes to do so. These findings suggest that if applicants were given essentially unlimited time to complete the MBE's 200 items, the mean raw score on this test would increase by about six points.

The larger difference between means in Groups 1 and 3 than in Groups 2 and 4 was apparently due to item Set A being slightly more difficult than item Set B. For example, the average Set A and B scores under the 55 minute time limit were 20.00 and 20.64, respectively.

An analysis also was conducted to determine whether the slight average increase in scores with the extra time was due to all applicants receiving a small advantage or whether some profited more than others from the extra time. This analysis involved subtracting an applicant's score on the 55 minute session from that applicant's score in the 90 minute session. The distribution of these difference scores is presented in Table 3.4 along with the percent who would have changed scores in each direction if the time limits had no affect on performance; i.e., if all the observed differences were due to answering a different set of items in each session rather than to the time limits per se.*

The data in Table 3.4 indicate that while 58 percent of the applicants benefited from the extra time, 42 percent had lower scores when they had more time to answer the questions.

*The procedures used to construct the chance distribution were as follows: Two sets of 30 items, X and Y, were selected from the full MBE. The items in each set covered the same four subject areas as the two experimental sets used in this study. The two sets had about the same means (19.83), standard deviations (3.90), and reliabilities (.65) as the experimental sets. The scores on Form X correlated with the scores on Form Y to the same degree (.63) as the scores in the 55 minute session correlated with the scores in the 90 minute session. A random sample of 50 percent of all the applicants who took the MBE had their score on Form X subtracted from their score on Form Y while the remaining 50 percent had their score on Form Y subtracted from their score on Form X. The difference scores in the two samples were combined to form the chance distribution.

Table 3.2

GROUP MEANS AND STANDARD DEVIATIONS ON THE MBE

	Group Number			
	1	2	3	4
Number of Applicants	705	755	730	750
Mean Scale Score	430.4	428.8	430.4	428.8
Standard Deviation	44.7	49.6	47.3	45.9

Table 3.3

MEANS AND STANDARD DEVIATIONS BY GROUP AND TIME LIMIT

Group	55 MINUTE			90 MINUTE			Difference Between Means
	Form	Mean	SD	Form	Mean	SD	
1	A-1	20.01	3.79	B-1	21.71	3.88	1.70
2	B-1	20.70	4.14	A-1	20.83	3.61	.13
3	A-2	20.00	3.74	B-2	21.56	4.14	1.56
4	B-2	20.58	4.18	A-2	20.67	3.89	.09
Mean		20.32	3.96		21.19	3.88	.87

Table 3.4

PERCENTAGE OF APPLICANTS CHANGING SCORE RELATIVE TO THE PERCENTAGE THAT WOULD HAVE CHANGED IF THE MBE'S TIME LIMITS HAD NO AFFECT ON PERFORMANCE

	Difference in Scores Between Time Limits					
	<-4	-4 to -3	-2 to 0	0 to 2	3 to 4	>4
Observed %	7.3	11.3	23.0	26.1	17.7	14.6
Chance %	9.0	14.0	27.0	27.0	14.0	9.0
Difference	-1.7	-2.7	-4.0	-0.9	3.7	5.6

Effect of Background Characteristics

An analysis was conducted to determine whether an applicant's background characteristics were related to how much that applicant benefited from the extra time. This was done by computing the correlation between the difference scores between sessions (see Table 3.4) and all of the background information on the applicants. The results of this analysis (which used dummy variables for categorical measures in a series of regression equations) indicated that there was no systematic relationship between the difference score and an applicant's sex, race, law school type, repeater status, or GBX scores (all the correlations were less than .04).

The same result was obtained when the difference scores were correlated with all the background characteristics and attitudes that were assessed through the questionnaires completed before and after the examination, including attitudes about the appropriateness of the time limits on the full MBE and the experimental tests.

Time Limit Effect on Relative Standings

Table 3.5 contains the correlations between the MBE and item set scores under each time limit. When these correlations were corrected for the less than perfect reliability of the measures, the average correlation of an item set with the MBE under both time limits and the average correlation between item sets increased to .94. These data indicate that the observed differences in the relative standings of the applicants were due to their answering a different set of questions in each time period rather than to differences in the amount of time they had to answer these questions.

Table 3.5

CORRELATIONS WITH AND BETWEEN ITEM SET SCORES

Group	MBE PORTION OF GBX WITH		Between Item Sets	Corrected
	55 Minute Item Set	90 Minute Item Set		Correlation Between Item Sets
1	.71	.74	.61	.96
2	.67	.62	.58	.92
3	.69	.74	.62	.96
4	.74	.72	.62	.93
Mean	.70	.71	.61	.94

Effect of Item Sequence on Test Statistics

The data in Tables 3.6, 3.7, and 3.8 indicate that the mean, standard deviation, and reliability of an item set were not appreciably affected by the sequence in which the questions were asked. For example, the average difference between version means, standard deviations, and reliabilities across item sets and time limits were .11, .16, and .03 respectively.

Table 3.6

DIFFERENCES BETWEEN VERSION MEANS

	55 MINUTES			90 MINUTES			Mean
	Ver 1	Ver 2	Diff.	Ver 1	Ver 2	Diff.	Diff.
Set A	20.01	20.00	.01	20.70	20.58	.12	.06
Set B	20.83	20.67	.16	21.71	21.56	.15	.16

Table 3.7

DIFFERENCES BETWEEN VERSION STANDARD DEVIATIONS

	55 MINUTES			90 MINUTES			Mean
	Ver 1	Ver 2	Diff.	Ver 1	Ver 2	Diff.	Diff.
Set A	3.79	3.74	.05	4.14	4.18	.04	.05
Set B	3.61	3.89	.28	3.88	4.14	.26	.27

Table 3.8

DIFFERENCES BETWEEN VERSION RELIABILITIES

	55 MINUTES			90 MINUTES			Mean
	Ver 1	Ver 2	Diff.	Ver 1	Ver 2	Diff.	Diff.
Set A	.61	.60	.01	.59	.65	.06	.04
Set B	.68	.69	.01	.66	.70	.04	.03

Effect of Sequence on Item Statistics

Each applicant's data file was reorganized so that the applicant's response to a particular item within a set appeared in the same ordinal position in the file as all the other applicants who answered this question; i.e., regardless of whether they took version 1 or 2 of the set. An item analysis was then run on each test form under each time limit. This analysis yielded eight sets of item means and "corrected item-total score" correlations (i.e., the point biserial correlation between an item and the total score on the other 29 items). These statistics (and the ordinal position of the item in its own test form) were entered as variables in an intercorrelation matrix. A separate matrix was computed for each item set; i.e., there were 30 observations per variable.

Table 3.9 contains a summary of the correlations of interest from the two matrices. An inspection of these data indicates that the relative difficulty of an item was not affected by its ordinal position in the test or by the amount of time applicants had to answer it.

Table 3.9

CORRELATION OF ITEM STATISTICS BETWEEN VERSIONS AND TIME LIMITS

	55 MINUTE		90 MINUTE		55 vs 90 MINUTE			
	Set A	Set B	Set A	Set B	A-1	A-2	B-1	B-2
Item Means	.98	.98	.99	.96	.98	.98	.96	.97
Point Biseri-als	.69	.87	.64	.81	.52	.71	.89	.89

The correlations among the point biseri-als were higher in item Set B than in Set A. This difference may have stemmed from the the average standard deviation of the Set B biseri-als (.098) being almost twice as large as the Set A biseri-als (.055). The standard deviations of the biseri-als were very homogeneous within sets across versions and time limits. It also was observed that Set B was slightly more reliable than Set A (see Table 3.8).

The average value of the biseri-als under each time limit was .173 and .178, respectively. And, the average correlation between the ordinal position of the items and their biseri-als were .10 and .02 for the 55 and 90 minute time limits, respectively. These findings strongly suggest that the time that is now allocated per item on the MBE (as represented by the 55 minute time limit) is not affecting an applicant's score. The reason for this conclusion is that if time were a factor, then the applicants would tend to be guessing more towards the end of the test (especially under the 55 minute time limit) which in turn would result in a large difference between the means of the biseri-als and a negative correlation between the biseri-als and item order.

SUMMARY AND CONCLUSIONS

Results with the MBE Time Limits portion of this research indicated that:

- o The average score under the 90 minute time limit was 0.87 points higher than it was under the 55 minute time limit. This finding suggests that if applicants were given essentially unlimited time to answer the MBE's 200 items, their average raw score would increase by about 6 points.
- o Applicants did not benefit equally from the extra time (only 58 percent had higher scores under the 90 minute time limit than they did under the 55 minute time limit). Some of the increase and decrease in scores was due to chance (e.g., guessing, the degree to which an applicant was familiar with the specific issues addressed by the items on each test, etc.); i.e., the scores were not perfectly reliable.
- o When scores on the tests were corrected for unreliability, there was a near perfect correlation between them. Corrected correlations between scores on the full MBE and the 30 item tests were the same and near perfect under both time limits. These findings suggest that the differences in the relative standings of the applicants under the different time limits were due to their taking different MBE questions in each time limit rather than to the extra time per se.
- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's background characteristics (such as age, sex, race, type of law school attended, and repeater status).
- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's performance level on the full examination or that applicant's attitudes about the experimental and full examination (e.g., such as the adequacy of time limits and efficacy as a test of legal skills and knowledge).

Raw scores on the MBE are adjusted through a scaling process in order to control for possible differences in average question difficulty across administrations of this test. Pass/fail decision on the MBE are therefore based on the scale scores (rather than the raw scores). While allowing applicants more time on the MBE would generally result in slightly higher raw scores, it would not affect the scale scores. Thus, the percent passing the MBE would not be increased by an increase in the time applicants had to answer. The analyses conducted in this study also indicated that giving applicants more time to answer MBE questions would have no significant impact on their relative standings; i.e., who passes versus fails this test.

Results with the Item Order portion of this research indicated that the ordinal position of an item in a test did not significantly affect the percentage of applicants answering it correctly or its correlation with other items. Variations in item order also did not affect the reliability of the test (there was only a .01 difference in reliability coefficients between versions under the regular time limit).

It may be inferred from these results that varying the sequence in which MBE items are presented will not affect an applicant's chances of passing this test. Thus, multiple forms of the MBE could be used on a given administration of the examination so as to substantially reduce any benefit an applicant might derive from observing another applicant's answers.

Chapter 4
ESSAY TIME LIMITS STUDY

PURPOSE

The essay time limits study was conducted to determine whether performance on the bar examination's essay questions would be affected if applicants were given substantially more than the usual time to answer these questions. The major questions addressed by this study were as follows:

- o Would the extra time improve the general quality of the answers?
- o Would all applicants benefit equally from the extra time; i.e., would their relative standings be affected?
- o Would the effect of the extra time be the same across questions?
- o What are the characteristics of applicants who tend to benefit most versus least from the extra time?

PROCEDURES

Four essay questions, numbered 21 through 24, were used in this study (Appendix G contains a copy of each question). Question #21 (Contracts) and #23 (Evidence) were constructed by the staff of the Committee of Bar Examiners. These questions were based on sets of multiple choice questions that were used in pre-1980, but still secure, versions of the MBE. Question #22 (Real Property) and #24 (Torts) were drawn from the pool of questions that were developed for the Essay portion of the GBX; i.e., they went through all the regular developmental steps and the Committee of Bar Examiners judged them to be eligible for inclusion in the GBX.

An applicant participating in this study was assigned randomly to one of eight groups. Each group had 55 minutes to answer one essay question and then 90 minutes to answer a second question. The test administration instructions appear in Appendix H.

Within a group, an applicant was assigned one of the "business law" type questions (i.e., contracts or real property) and one of the "trial law" type questions (i.e., evidence or torts). The assignment of questions to groups and the time the applicants had to answer these questions are presented in Table 4.1.

Table 4.1

ASSIGNMENT OF QUESTIONS TO GROUPS

Time Limit	Group Number							
	1	2	3	4	5	6	7	8
55 minutes	21	22	23	24	22	21	23	24
90 minutes	23	24	21	22	23	24	22	21

All the answers to a question were mixed together (i.e., regardless of group number and the time applicants had to answer the question). The answers were then assigned randomly to the three readers who graded that question. These readers were drawn from among those who graded the Essay portion of the GBX.

The readers assigned to a question participated in a one day calibration session that was similar to the type of session used for standardizing reader teams on the GBX. They also were told to use the same grading criteria as was used on the GBX and to "give more credit to answers that deal with the issues in a concise and well-organized fashion than...to those that are merely longer." Applicants were informed of these criteria and advised that when they had extra time, they should use it "to write a better answer, not just a lengthier one."

The readers assigned grades to answers on the same 100 point scale as is used with the GBX (i.e., 70 = passing). However, unlike the GBX which requires readers to assign grades in five point intervals (e.g., 50, 55, etc.), they assigned them in one point intervals for this study. This was done to determine whether the one point interval scale would lead to more or less reliable grades than the five point interval system.

PRELIMINARY ANALYSES

The number of applicants in each group and each group's mean and standard deviation on a question in the Essay portion of the GBX are presented in Table 4.2. An inspection of these data (and a one-way analysis of variance) indicated that the groups had comparable Essay scores; i.e., the random assignment process was successful in producing equivalent groups.

Table 4.2

AVERAGE ESSAY QUESTION SCORE ON GBX

<u>Group</u>	<u>Number of Applicants</u>	<u>Mean</u>	<u>Standard Deviation</u>
1	228	68.0	6.0
2	211	67.8	6.4
3	220	67.5	6.9
4	229	67.8	5.8
5	215	67.7	6.7
6	213	67.8	5.7
7	215	68.2	6.1
8	198	67.7	5.7
Mean	216	67.8	6.2

A check was made on the degree to which Questions #21 - #24 were scored reliably. This check involved embedding in the group of answers that were graded by a reader a set of answers that also were graded by the other two readers assigned to the question. Table 4.3 contains the results of this

check. These data indicate that on the average, there was a five point difference between two readers in the grades they assigned to an answer and that the average correlation between two readers over all the applicants they graded in common was .73. These levels of inter-reader agreement in evaluating the absolute and relative quality of the answers are adequate for the purposes of this research and about the same as that obtained in the grading of the Essay portion of GBX (Klein, 1980b). These findings also suggest that there would be no major increase in inter-reader agreement on the Essay portion of the GBX if grades were assigned in one rather than five point intervals on the 100 point scale.

Table 4.3

READER RELIABILITY STATISTICS

Question Number	Number of Applicants	Average Deviation	Average Correlation
21	15	5.5	.64
22	20	6.6	.77
23	30	2.7	.88
24	32	7.0	.63
Average	24	5.5	.73

The average deviation is the average difference in score between two readers of the same answer. The average correlation is the average of the correlations between readers.

RESULTS

Time Limit's Effect on Level of Performance

Table 4.4 contains each group's average score on each question. An inspection of this table indicates that overall, there was a four point increase in average score with the expanded time limit.

Table 4.4

AVERAGE SCORE IN EACH TIME PERIOD BY GROUP

Time Limit	Group Number								Average
	1	2	3	4	5	6	7	8	
55 minutes	66	63	63	68	63	65	63	69	65
90 minutes	70	73	68	67	68	74	64	68	69

Table 4.5 presents the average score on each question in each time period. An inspection of this table indicates that Question #24 was easier (or graded more leniently) than the other questions. This difference was the apparent source of Groups #4 and #8 having higher average scores on their 55 than on their 90 minute question.

Table 4.5

AVERAGE SCORE ON EACH QUESTION UNDER EACH TIME LIMIT

Time Limit	#21 Contracts	#22 Real Prop.	#23 Evidence	#24 Torts	Average
55 minutes	65	63	63	69	65
90 minutes	68	66	69	74	69
Difference*	3	3	6	5	4

*All the differences in average score between time limits were statistically significant (alpha = .01). The average standard deviation under the 55 and 90 minute time limits were 8.6 and 8.3, respectively.

The data in Table 4.5 also indicate that the average increase in score attributable to the extra time was about twice as great on the Evidence and Torts questions as it was on the Contracts and Real Property questions. This pattern was unrelated to the source of the questions (i.e., MBE versus GBX), but it was related to the distinction between "business" and "trial" law. The possible source of this pattern was investigated further by having a group of experts (the Committee of Bar Examiner's Board of Reappraisers) estimate the probable impact of increased time on each question. While many interesting rationales were suggested, none of them matched the observed pattern nor did the experts generally agree as to which questions would be more or less influenced by the extra time.

An analysis was conducted to determine whether all applicants benefited equally from having 64 percent more time to answer an essay question. In this analysis, an applicant's essay score in the 55 minutes session was subtracted from that applicant's score in the 90 minute session. The distribution of these difference scores is presented in Table 4.6.

An inspection of the data in Table 4.6 indicates that 67 percent of the applicants received a higher score as a result of the added time. In other words, 33 percent received lower scores when they had more time to answer! The percent of applicants whose scores would have changed in each direction solely as a result of their answering a different question (as distinct from having more time to answer) is presented in the "Chance %" row of Table 4.6.

Table 4.6

PERCENTAGE OF APPLICANTS CHANGING SCORE RELATIVE TO THE PERCENTAGE THAT WOULD HAVE CHANGED IF THE ESSAY'S TIME LIMITS HAD NO AFFECT ON PERFORMANCE

	Difference in Scores Between Time Limits					
	<-10	-10 to -6	-5 to 0	0 to 5	6 to 10	>10
Observed %	9.4	9.4	14.1	18.6	20.8	27.7
Chance %	13.5	11.5	25.0	25.0	11.5	13.5
Difference	-4.1	-2.1	-10.9	-6.4	9.3	14.2

The procedures used to construct the chance distribution in Table 4.6 were as follows: All the applicants taking the essay portion of the GBX were assigned randomly to two groups, X and Y. An applicant's score on each essay question that was used to compute that applicant's Phase I score was tabulated. The following difference scores were computed among the three questions in this session for all the applicants in Group X: 1st - 2nd, 1st - 3rd, and 2nd - 3rd. For example, if the second session was used to compute an applicant's Phase I score, then the difference scores between questions would be computed as follows: #4 - #5, #4 - #6, and #5 - #6. The difference scores in Group Y were computed as follows: 2nd - 1st, 3rd - 1st, and 3rd - 2nd. Thus, there were three difference scores for each applicant. The distribution of these difference scores across all applicants constituted the chance distribution of difference scores.

Time Limit's Effect on Relative Standing

The correlation between scores on the 55 and 90 minute questions in Groups 1 through 8 were: .26, .30, .24, .22, .21, .08, .30, and .13, respectively. The two unusually low correlations, .08 and .13, were between Questions #21 and #24 (see Table 4.1 for the pairings of questions across groups). The average of the eight correlations, .22, was lower than the .32 average correlation among essay questions on the GBX. This finding suggests that the relative standings of the applicants on a question may have been affected by the time they had to answer that question.

This matter was investigated further by computing the correlation between the applicants' Phase I Essay scores on the GBX and their scores on Questions #21 - #24 under both time limits. These data are presented in Table 4.7 and an inspection of them indicates that the correlations were only slightly higher under the normal than under the expanded time limits (average r's were .41 and .38, respectively). When these correlations were corrected for the less than perfect reliability of the measures, a strong relationship emerged under both time limits between an applicant's relative standing on an experimental question and that applicant's relative standing on the Essay portion of the GBX (r's of .95 and .88 for the 55 and 90 minute time limits, respectively...based on a reliability of .32 for a question under either time limit and .59 for the Phase I Essay score).

Table 4.7

CORRELATIONS WITH PHASE I ESSAY SCORES UNDER EACH TIME LIMIT

Time Limit	#21 Contracts	#22 Real Prop.	#23 Evidence	#24 Torts	Average	Corrected Average
55 minutes	.29	.48	.48	.39	.41	.95
90 minutes	.37	.40	.42	.32	.38	.88
Difference	-.08	.08	.06	.07	.03	.07

These findings suggest that the observed differences in the relative standings of the applicants on the 55 and 90 minute questions were typical of the differences that occur on different questions under the same time limit. They were not due to differences in the amount of time the applicants had to answer the questions. The difference between the correlations would have been greater were it not for Question #21's correlations deviating from the general pattern of relationships (its correlation with the Phase I essay score was higher under the expanded than under the normal time limit). It also was observed that Question #21 had unusually low correlations with Question #24 (.08 and .13 in groups 6 and 8, respectively).

Effect of Background Characteristics

An analysis was conducted to determine whether an applicant's background characteristics were related to how much that applicant benefited from the extra time. The first step in this analysis involved scaling the scores on each of the experimental questions to the mean and standard deviation on the Essay portion of the GBX of all the applicants who answered that question. This scaling permitted combining applicants across groups into one sample since it controlled for differences in question difficulty.

The second step in the analysis involved determining the extent to which an applicant's background characteristics were correlated with the difference between that applicant's scale score on the 55 and 90 minute questions. The results of this analysis (which used dummy variables for categorical measures in a series of regression equations) indicated that there was no systematic relationship between the difference score and an applicant's sex, R/E group, law school type, repeater status, or GBX scores (all the correlations were less than .04).

The same result was obtained when the difference scores were correlated with all the background characteristics and attitudes that were assessed through the questionnaires completed before and during the examination, including attitudes about the appropriateness of the time limits on the essay questions.

SUMMARY AND CONCLUSIONS

The major findings of the Essay Time Limits Study were as follows:

- o The average score on a question under the 90 minute time limit was higher than it was under the 55 minute time limit on all four of the experimental questions.
- o There was an average increase of 4 points per question, but performance on some questions was more affected by the extra time than it was on other questions.
- o A panel of attorneys with extensive experience in essay test construction and grading was not able to predict (either individually or collectively) on which questions scores would be most affected by the different time limits.
- o Applicants did not benefit equally from the extra time (only 67 percent had higher scores under the 90 minute time limit than under the 55 minute time limit). Some of the increase and decrease in scores was due to chance (e.g., grader unreliability and the degree to which an applicant was familiar with the issues addressed by each question). It also is possible that the extra time gave some applicants sufficient opportunity to demonstrate their true performance level (which in turn resulted in their not being given the benefit of any doubt that might have been created by a shorter answer).
- o When essay scores were corrected for unreliability, there was a near perfect correlation between scores on an experimental question under the 55 minute time limit and Phase I essay scores (i.e., the total score on three essay questions given under the normal time limit). This correlation dipped only slightly when applicants were given 90 minutes to answer an experimental question. These findings suggest that the differences in the relative standings of the applicants under the different time limits were due to their answering different questions in each time limit rather than to the extra time per se.
- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's background characteristics (such as age, sex, race, type of law school attended, and repeater status).
- o There were no statistically significant correlations between how much an applicant's score changed as a result of the extra time and that applicant's performance level on the full examination or that applicant's attitudes about the experimental and full examination (e.g., such as the adequacy of time limits and the efficacy of the tests measures of legal skills and knowledge).

It is evident from the foregoing results that giving applicants 64 percent more time to answer essay questions would not affect their relative standings any more than would just giving them a different set of essay questions to answer under the regular time limits. However, giving all applicants extra time might result in generally higher total essay scores for most applicants.

If the average increase in essay scores was as large as the one obtained in this research, it would substantially increase the percent passing the California examination. Thus, in terms of policy implications, the 4 point per question increase with the extra time is potentially the most significant finding presented in this report.

There are several reasons why it is not certain that the 4 point per question increase observed in this study would be maintained if all applicants were given additional time. Some of the reasons for not being able to predict what would occur include:

- o If more time were allowed, the nature of the questions asked might change, such as their containing more issues and/or more difficult issues.
- o Scores on some questions were much more affected by the extra time than were scores on other questions. Thus, the size of the increase would be affected by the nature of the question asked and there did not appear to be any way of determining in advance on which questions scores would be most affected by the time allowed to answer them.
- o Grading criteria and standards may shift when readers score answers written under an expanded time limit. They may consciously or unconsciously place more weight on organization as well as expect a higher quality of answer for passing even if they are given instructions not to make such changes in their grading practices.
- o It is not known whether the average increase in score was due to applicants identifying more issues, to discussing the issues better, or to their just writing more. Specifically, the answers written under the 90 minute time limit may have been longer (but not necessarily better) than those written under the 55 minute time limit. Previous research has indicated that longer answers tend to receive higher grades than shorter ones. If all the answers are written under the same time limit, any affect of answer length on score probably would be reduced.

The foregoing issues are not raised to minimize the potential significance of the 4 point per question increase with the extra time. Rather, they indicate the need for additional studies to determine the source of the increase so that appropriate policy decisions can be made.

Giving applicants more time to answer essay questions could affect costs, test reliability, and validity. For example, if the total time devoted to the essay section was maintained (i.e., 9 hours) but the amount of time allocated per question increased, then there would be a reduction in the number of questions that could be asked which in turn would reduce scoring costs and probably reliability. It also would reduce the breadth of content that could be covered in any one administration of the examination. If the number of questions remained the same but the amount of time per question increased, readers might have to be paid more (because the answers would take longer to score) and there would be the additional costs associated with longer test administrations. Applicant fatigue also might become a factor affecting scores.

Finally, giving applicants more time to answer essay questions would not affect the percent passing in jurisdictions that scale their essay scores to the MBE. The reason for this is that such scaling removes the effect of any variations in essay test difficulty across administrations (whether this variation stems from differences in question difficulty, grader leniency, or the amount of time applicants had to answer the questions). Thus, while the 4 point per question increase that was observed in this study could have major policy ramifications for California, it would not affect pass/fail decisions in many other jurisdictions because the increase in time had no systematic impact on the relative standings of the applicants.

PURPOSE

Several studies at the elementary school, high school, and college level have sought to determine whether essay and multiple choice questions can be used to test the same skills and knowledge. The results of these studies suggest that the two types of questions can be used interchangeably for some subject areas, but not for others (Traub and Fisher, 1977).

Previous studies of California's bar examination (Klein, 1979) and the bar examinations given in other states (Carlson and Werts, 1976; Klein, 1981b) have generally found a strong correlation between MBE and Essay scores. However, this correlation is still not perfect even after it has been corrected for the unreliability of the measures (the average corrected r is about .80). One explanation for this finding is that there is only a moderate degree of overlap between the content coverage of the MBE and the Essay sections of the examination. Another explanation is that the two sections measure highly related but not identical skills and knowledge.

The MBE/Essay Congruence study was conducted to determine the extent to which the less than perfect relationship between MBE and Essay scores is due to differences in their subject matter coverage (e.g. the specific issues addressed by a question) versus the types of analytic skills measured by the two kinds of tests.

PROCEDURES

A 35 minute, 20 item, multiple choice test was administered to all the applicants who participated in the Essay Time Limits study (see Chapter 4). This test had one set of five items for each of the four essay questions used in that study. Each set of items was designed to measure the same issues as were addressed in its corresponding essay question. Ten of the 20 items were constructed by the staff of the Committee of Bar Examiners to coincide with the two essay questions, #22 and #24, that developed by the Committee. The other 10 items, which were drawn from previous administrations of the MBE, served as the basis for the staff's construction of essay Questions #21 and #23.

The 20 item test was administered immediately after the completion of the Essay Time Limits study. Since an applicant answered only two of the four essay questions in that study, 10 of the 20 items corresponded to essay questions the applicant had seen previously while 10 were completely novel. For example, by the time they took the 20 item test, applicants in Group 1 were exposed to the facts and issues that served as the basis for the multiple choice items corresponding to Questions #21 and #23, but they were not exposed to the facts and issues involved in Questions #22 and #24 (see Table 4.1). Group 4 had just the opposite exposure pattern.

Four forms of the 20 item test were used to help insure the independence of responses across applicants. These forms differed in terms of the sequence in which the item sets appeared in the test booklet. The sequence used in each form is presented in Table 5.1.

Table 5.1

SEQUENCE OF ITEM SETS WITHIN FORMS

Form	1st	2nd	3rd	4th
A	22	24	21	23
B	21	23	22	24
C	22	23	21	24
D	21	24	22	23

RESULTS

Table 5.2 contains each group's mean score on each item set and on the total test. An inspection of these data (and a one way analysis of variance) indicated that the eight groups had essentially equivalent total scores.

The item set corresponding to Question #21 had a lower average score than did the other three item sets. The standard deviations of the item sets ranged from 1.15 to 1.33.

The average reliability of the four forms was .53 (which when stepped up to a 200 item test yields a reliability of .92). There was no apparent reason for test Form D having a slightly lower internal consistency than the other three forms.

Table 5.2

GROUP MEANS ON ITEM SETS AND TOTAL TEST

Test Form	Group	MEANS ON ITEM SETS				TOTAL SCORE		
		#21	#22	#23	#24	Mean	SD	Alpha
A	1	3.1	3.2	3.3	3.1	12.8	2.9	.53
A	3	3.1	3.3	3.4	3.2	12.9	2.9	.55
B	2	2.7	3.3	3.4	3.2	12.6	3.0	.56
B	4	2.8	3.5	3.4	3.2	13.0	2.9	.52
C	5	2.8	3.5	3.2	3.1	12.7	3.1	.56
C	7	2.8	3.4	3.3	3.1	12.6	3.1	.59
D	6	3.1	3.3	3.3	3.2	12.9	2.7	.45
D	8	3.2	3.2	3.3	3.1	12.8	2.8	.47
Total Sample		2.9	3.3	3.3	3.2	12.8	2.9	.53

The correlations between the MBE portion of the GBX and the four item sets, regardless of test form, are presented in Table 5.3. An inspection of these data indicates that sets #21 and #24 had unusually low correlations with each other. This finding is consistent with the unusually low correlations between essay Questions #21 and #24 noted in Chapter 4. Item set #22, which was developed by the staff of the Committee of Bar Examiners, had an unusually high correlation with the MBE.

Table 5.3

CORRELATIONS BETWEEN ITEM SETS

Item Set	#21	#22	#23	#24	MBE
#21	-	.21	.18	.07	.35
#22	.21	-	.19	.18	.49
#23	.18	.19	-	.14	.33
#24	.07	.18	.14	-	.29

After adjusting the item set scores for differences in overall item difficulty, it was observed that the applicants did slightly better on the item sets corresponding to the essay questions they answered than they did on the item sets corresponding to the essays questions they did not answer. The mean scores on the two groups of items (10 items per group) were 6.48 and 6.22, respectively. This effect, while still very small, was most apparent with the applicants who answered essay Question #21.

There was no difference in average score on an item set corresponding to an essay question taken under the regular versus expanded time limit. Both sets had an average of 3.26 out of a possible five points. While the interaction effect described above was still present, it was not affected further by differences in time limits; i.e., applicants benefited more from having taken Question #21 (regardless of how much time they had to answer it) than they did from having taken other questions. It is important to emphasize, however, that the size of Question #21's unique effect on an applicant's score was quite small. For example, the average overall scores of the applicants who answered versus did not answer Question #21 during the essay portion of the session were 12.84 and 12.68, respectively.

Table 5.4 presents the correlations between item set and essay scores under the regular (55 minute) and expanded (90 minute) time limits. The average number of applicants used to compute each correlation coefficient was 432 (because all applicants did not answer all essay questions). If the relationship between scores on an item set and its corresponding essay question was stronger than it was between the item set and the other essay questions, then the underlined correlations in Table 5.4 would be higher than the off diagonal (not underlined) coefficients.

The data in Table 5.4 indicate that there were only two atypically high diagonal coefficients; namely, the .49 and .50 correlations associated with item set and essay question #22. Thus, the correlation between essay and multiple choice questions generally did not improve when the facts and issues addressed by the two types of questions were held constant.

Table 5.4

CORRELATIONS BETWEEN ITEM SETS AND ESSAY QUESTIONS

Multiple Choice Item Set	55 MINUTE ESSAY				90 MINUTE ESSAY			
	#21	#22	#23	#24	#21	#22	#23	#24
#21	<u>.15</u>	.22	.21	.14	<u>.12</u>	.23	.13	.14
#22	.08	<u>.49</u>	.20	.13	.19	<u>.50</u>	.13	.17
#23	.14	.22	<u>.22</u>	.21	.16	.11	<u>.18</u>	.16
#24	.11	.12	.17	<u>.13</u>	.13	.18	.09	<u>.10</u>

Correlations Among Essay Questions in the Same Area

Some of the experimental questions were in the same content area (such as Evidence) as essay questions on the GBX (see Tables 2.1 and 4.5). This situation provided an opportunity to explore whether essay questions in the same subject area correlated higher with each other than did questions in general. The results of this investigation did not reveal any systematic relationship due to content when the experimental questions were answered under the 55 or 90 minute time limits.

SUMMARY AND CONCLUSIONS

The major findings of the MBE/Essay Congruence Study were as follows:

- o The sequence in which the 20 multiple choice items were asked did not affect performance on these items or test reliability. These findings are consistent with those obtained in the Item Order Study discussed in Chapter 3.
- o Applicants did only very slightly better on the items sets corresponding to essay questions they answered previously than they did on item sets that dealt with essay questions they had not answered.
- o Whether an applicant answered an essay question under the 55 or 90 minute time limit had no affect on that applicant's score on the item set corresponding to that essay question.
- o Total scores on an item set were generally no more highly correlated with scores on that item set's corresponding essay question than they were with any other essay question.

- o Scores on essay questions in the same general area (e.g., Evidence) correlated no more highly with each other than did scores on essay questions in different areas. This finding could result from applicants balancing their preparation for the examination across areas and/or the essay test measuring skills that are not highly content specific.

The foregoing findings suggest that MBE and California essay questions will not rank order applicants in exactly the same way even when the content on these two types of questions is held constant. Nevertheless, scores on the the two types of questions are highly related. This means that applicants who have the skills and knowledge required to do well on MBE items also tend to have the abilities needed to do well on essay questions.

REFERENCES

- Carlson, A. B. and Werts, C. E. Relationship among law school predictors, law school performance, and bar examination results. Princeton, N.J.: Educational Testing Service, 1976.
- Dorans, N. and Wright, R. Test Analysis: Multistate Bar Examination, Forms 3CEB2 and S-3CEB2. Unpublished Statistical Report, SR-80-117. Princeton, N.J.: Educational Testing Service, October, 1980.
- Gulliksen, Harold Theory of Mental Tests. New York: Wiley, 1950.
- Klein, S. P. An analysis of the relationships between bar examination scores and an applicant's law school, admissions test scores, grades, sex, and racial/ethnic group. Paper presented at the American Bar Association meetings, Dallas, Texas; August 14, 1979. Reprinted in The Bar Examiner, 1980, 49, 14-18.
- Klein, S. P. A comparison of the effectiveness of a single versus multiphased grading system. Unpublished report submitted to the Committee of Bar Examiners of the State Bar of California, May, 1980a.
- Klein, S. P. Intra- and inter-reader agreement on the essay section of the California State Bar Examination. Unpublished report submitted to the Committee of Bar Examiners of the State Bar of California, June, 1980b.
- Klein, S. P. An analysis of possible variations in pass/fail standards on the California State Bar Examination. Unpublished report submitted to the Committee of Bar Examiners of the State Bar of California, January, 1981a.
- Klein, S. P. Factors associated with the difference in passing rate between Anglo and Hispanic applicants on the New Mexico Bar Examination. Unpublished report submitted to the New Mexico Board of Bar Examiners, January, 1981b.
- O'Hara, J. F. and Klein, S. P. Is the bar examination an adequate measure of lawyer competence? The Bar Examiner, 1981, 50, #3, 28-35.
- Sprecher, R. A. (ed.). The Bar Examiner's Handbook. Chicago, IL: National Conference of Bar Examiners, 1968.
- Traub, R. E. and Fisher, C. W. On the equivalence of constructed-response and multiple-choice tests. Applied Psychological Measurement, 1977, 1, 355-369.

APPENDIX A

GLOSSARY OF ABBREVIATIONS AND TECHNICAL TERMS

Corrected Correlation Coefficient	This is the correlation coefficient that would be obtained between two variables if scores on both variables were perfectly reliable; i.e., it indicates the strength of the true underlying relationship between two variables. A formula, called the "correction for attenuation," is used to estimate the corrected correlation.
Correlation Coefficient (r)	The correlation coefficient is an index of the degree to which the relative performance of the applicants on one variable corresponds to their relative performance on another variable. A positive correlation indicates that applicants with high scores on one variable tend to have high scores on the other variable. A negative correlation indicates that applicants with high scores on one variable tend to have low scores on the other variable. Correlation coefficients range between ± 1.00 ; the larger the r, the stronger the relationship (regardless of algebraic sign). A 0.00 correlation means that there is no relationship between the variables.
GBX	General Bar Examination. This examination consists of the MBE and 9 essay questions.
Mean or Average	The mean is the arithmetic average. It is computed by summing all the scores on a variable and then dividing that sum by the number of scores that entered it.
MBE	Multistate Bar Examination
Phase I Essay Score	The sum of the scores assigned to three essay questions on the Essay portion of the GBX. The Phase I essay score is used in combination with the MBE to make the Phase I screening decision (see Appendix C).
Reliability	Reliability refers to the consistency with which a variable is measured; e.g., the degree to which two readers agree on the relative quality of different essay answers or how much an applicant's relative standing on the MBE would be affected if a different set of MBE questions were asked.
Scaling	In this report, scaling refers to converting the scores on a variable to a different unit of measurement (in terms of the mean and standard deviation of the scores).
Standard Deviation (SD)	The standard deviation is an index of the degree to which scores on a variable spread out on either side of the mean. The larger the standard deviation, the greater the spread. If the shape of the score distribution is normal, then 34% of the scores fall between the mean and one SD above it.
Variable	Any characteristic of persons on which they differ. Score level on the MBE is an example of a variable.

APPENDIX B

COMMITTEE OF BAR EXAMINERS'
POLICY REGARDING ASSIGNMENT OF GRADES

The passing grade on the bar examination is 70% of the highest possible grade. The grade range is from 0-100. Consequently a very bad answer can hurt the applicant more than two very good answers can help. The applicant could fall 70 points below passing on the one bad answer and could gain no more than 60 points above the required average on two outstanding answers.

It is the Committee's policy with regard to assigning grades that the readers' attention should first be focused upon the question whether, considering the overall quality of the answer, the applicant has exhibited sufficient judgment, analytical ability and knowledge of the subject matter involved in the particular question to merit a pass or a fail on that question. If, after careful consideration, the reader is unable to make up his mind whether the applicant deserves a pass or fail on the question, the reader shall assign a grade of 70.

If the reader is satisfied that the applicant has written a passing answer the reader shall enter a grade ranging from 75 to 100, depending on the quality of the answer:

A passing answer	-	75
A good passing answer	-	80
A very good to outstanding answer	-	85-100

The score of 100 is not reserved for a perfect answer or for the single best answer which the reader may encounter on a particular examination. A grade of 100 shall be assigned when the reader believes the applicant has done as well as can be reasonably be expected of any applicant on that question. Therefore a grade of 100 may be assigned even if the applicant failed to develop fully the analysis of the issues discussed.

If the reader is satisfied that the applicant has failed the question, the grade must be between 0 and 65.

If the applicant has shown any understanding of the question and has made any serious attempt to grapple with it, the applicant should not receive a grade below 40. If the applicant has written an answer that is wholly, or almost wholly without redeeming quality the grade range is from 0 to 40.

If an answer contains matter that is irrelevant, the irrelevant matter must be disregarded and no credit may be deducted or added because of the inclusion of such matter.

Revised August 1977

APPENDIX C

DESCRIPTION OF THE MULTIPHASED GRADING PROCESS

The major steps in the multiphased grading process are listed below:

- 1) In Phase I, all three answers in one of the three essay sessions are graded (different applicants have different sessions read in this phase so that all reader teams can begin simultaneously). If the sum of an applicant's MBE and three essay scores is greater than 665, the applicant passes; i.e., there is no further reading of that applicant's answers. If the applicant's Phase I score is 665 or less, the applicant goes on to Phase II.
- 2) All the applicants in Phase II have their remaining six essay answers read once. Applicants with total scores (i.e., MBE + Essay) below 1010 are failed, those with scores above 1065 are passed, and those with scores between 1009 and 1066 are placed in Phase III.
- 3) All the applicants in Phase III have all nine of their essay answers read again. The second reader assigned to an answer is a different person than the first reader and the second reader does not know the grade given by the first reader. The Phase III applicants have their Phase II and III essay scores averaged. If the total of an applicant's MBE and average essay score is less than 1030, the applicant fails; if it is more than 1049, the applicant passes; and if it is between 1029 and 1050, the applicant is placed in Phase IV.
- 4) Phase IV is called "reappraisal." In this phase, all nine of an applicant's essay answers are evaluated as a set by a member of the Board of Reappraisers (i.e., a member of a panel of consultants who have extensive experience in developing and grading the essay portion of California's examination). The reappraiser is advised of all the grades assigned to the applicant (i.e., MBE and essay scores from Phases I - III) and is asked to review the essay answers again to see if enough points can be found to pass the applicant. If enough points cannot be found, the applicant fails.

A parallel procedure (but with different cutoff scores) is used to determine whether an applicant passed the essay section by itself.

Of the 7379 applicants who took the complete July 1980 GBX, 29% passed in Phase I. The percent passing as a result of Phases II, III, and IV were 11, 5, and 3, respectively. Another 1% passed as a result of having passed one portion of the GBX previously and having received in July 1980 a passing grade on the portion they had failed previously. Table C.1 shows how the 7379 applicants were distributed across pass/fail categories.

Table C.1

DISTRIBUTION OF JULY 1980 APPLICANTS ACROSS PASS/FAIL CATEGORIES

Final Status	Basis for Pass/Fail Status	Number of Applicants	Percent of Applicants
Pass	Phase I Pass	2141	29.0
Pass	Phase II Pass	783	10.6
Pass	Phase III Pass	380	5.1
Pass	Phase IV (Reappraisal) Pass	234	3.2
Pass	Previous MBE Pass + Phase II Essay Pass	34	0.5
Pass	Previous MBE Pass + Phase III Essay Pass	29	0.4
Pass	Previous MBE Pass + Phase IV Essay Pass	12	0.2
Pass	Previous Essay Pass + Current MBE Pass	10	0.1
Fail	Phase II Fail	1932	26.2
Fail	Phase III Fail	1476	20.0
Fail	Phase IV Fail	348	4.7
Total		7379	100.0

APPENDIX D

EXCERPTS FROM NOTICE TO APPLICANTS REGARDING THE SPECIAL SESSION

In addition to the usual examination sessions, each applicant for the Fall 1980 Examination will be expected to participate in a one-half day Special Session to be administered either on Monday, July 28, 1980, or Tuesday, July 29, 1980.

The Special Session is designed to give applicants an opportunity to improve their scores by demonstrating their capabilities in other than the standard methods or procedures and to test alternative means of examining for the future.

To derive any benefit from the Special Session, applicants must complete and return designated questionnaires by specified deadlines and complete all other portions of the examination which they are taking.

There will be four types of problems given during the Special Session. For convenience, they will be referred to as... (c) Special Essay Questions, and (d) Special Multiple-Choice... Individual applicants will be assigned to the various types of problems by the Committee based on a statistical analysis of the composition of the entire examinee population. General descriptions of the four types of problems are:

(c) Special Essay Questions

The time for this session will be divided into three segments of 55, 90, and 35 minutes. Each applicant will receive two essay questions to answer - one during the 90-minute segment and the other during the 55-minute segment. The answers to these questions will be graded in the same manner as answers written during the regular essay portion of the examination. The readers will be instructed to give more credit to answers that deal with the issues in a concise and well-organized fashion than they give to those that are merely longer. In short, applicants should use the extra time to attempt to write a better answer, not just a lengthier one. During the 35-minute segment, each applicant will be given 20 multiple-choice questions to answer.

(d) Special Multiple-Choice

In the 90 and 55 minute segments (of this session), applicants will answer MBE type questions under conditions that might improve performance on this type of test (e.g., having more time to answer each question). (The 35-minute segment was used for another study.)

The scoring for the Special Session for the various groups of applicants will be scaled so that the Special Session would be equivalent to one session of a six-session examination. An applicant's score on the Special Session will then be utilized if and only if (a) the applicant did not pass the regular part of the examination based upon the usual scoring procedures and (b) the applicant's score on the Special Session is above passing. Thus, an applicant's performance on the Special Session can increase but cannot decrease the likelihood that the applicant will pass the examination. Since the score on the Special Session can count for one-sixth of an applicant's total grade, it can have a substantial beneficial impact for an applicant who scores well on it.

7. Circle the box below that best describes your undergraduate major:

- [1] Economics, Business, Accounting
- [2] Physical Sciences, Engineering, Mathematics, Biology
- [3] Social Sciences (Anthropology, Psychology, Sociology)
- [4] History, Government, Political Science
- [5] English, Journalism, Classical Studies, Philosophy
- [6] Fine Arts, Theater Arts, Music
- [7] Education
- [8] Other (specify)

8. For each of the activities below, indicate the number of hours, if any, you have spent doing them for: SIMULATED cases, ACTUAL cases associated with supervised law school programs, and ACTUAL cases as part of paid and/or volunteer employment. Insert the number 99 in a box if you spent 100 or more hours doing the activity.

	ACTUAL CASES		
	SIMULATED CASES	Supervised Law School Programs	Paid or Volunteer Employment
	-----	-----	-----
a. Conduct legal research.	[]	[]	[]
b. Prepare briefs, petitions, or motions.	[]	[]	[]
c. Conduct direct examinations.	[]	[]	[]
d. Conduct cross examinations.	[]	[]	[]
e. Interview a client or a witness for a hearing.	[]	[]	[]
f. Interview a client on general legal matters; e.g. landlord- tenant dispute.	[]	[]	[]
g. Present an oral argument in a legal proceeding.	[]	[]	[]

9. In the box next to each choice below, indicate the number of courses, if any, you have taken in:

- [] Evidence
- [] Clinical or Trial Practice involving your participation in simulated or actual hearings
- [] Trial Practice NOT involving your participation in simulated or actual hearings; i.e., lecture only.

10. If you have secured employment in California in a law related job commencing by September 15, please circle the box corresponding to how you will be employed.

- [1] Attorney General's Office
- [2] Public office, criminal prosecution (e.g., District Attorney)
- [3] Public office, criminal defense (e.g., Public Defender)
- [4] Legal aid office (e.g., neighborhood legal assistance)
- [5] Public interest law firm
- [6] Law clerk for a judge
- [7] Private law firm, criminal defense
- [8] Private law firm, general practice
- [9] Law department of a corporation
- [0] Other (specify)

11. Do you hereby apply to participate in the two day Assessment Center? (Circle your answer below.)
 (1) No (2) Yes

If selected as a participant in the Assessment Center, I will be available to participate in the location circled below on any date between August 4 and August 24 (inclusive) except for the dates circled below.

Circle one	Mon	Tue	Wed	Thu	Fri	Sat	Sun	
in (1) Los Angeles only	4	5	6	7	8	9	10	
in (2) San Francisco only	11	12	13	14	15	16	17	_____
in (3) Either Los Angeles or San Francisco.	18	19	20	21	22	23	24	_____

(Note: The likelihood of being selected may be improved by increased availability.)

Date: _____, 1980

Signature: _____

APPENDIX F

POST EXAMINATION QUESTIONNAIRE

APPLICANTS MUST COMPLETE AND FILE THIS QUESTIONNAIRE BEFORE THE FINAL SESSION OF THE GENERAL BAR EXAMINATION ON WEDNESDAY, JULY 30, 1980 IN ORDER TO DERIVE ANY BENEFIT FROM THE SPECIAL SESSION OF THE JULY 1980 EXAMINATION OR TO PARTICIPATE IN THE AUGUST 1980 ASSESSMENT CENTER.

Your answers to the questions below will assist the Committee of Bar Examiners in its efforts to improve the bar examination. Your responses will be kept strictly confidential and used solely for statistical purposes. A completed questionnaire is required of all applicants who participated in any one of the special sessions and/or will be participating in the Assessment Center. We are most appreciative of your cooperation in answering the questions below.

Application # [] Center Number [] Birthdate [/ /]

1. Which of the following methods, if any, did you use to prepare for the examination? Check all that apply:

- Commercial bar review course of at least 5 weeks duration which met 4 or more times per week; e.g., BAR or BRC.
- Commercial bar review course of less than 5 weeks duration.
- Intensified commercial writing course; e.g., Beverly Rubens.
- Law school sponsored or administered bar review course.

2. About how many hours per week were you engaged in paid employment during the last five weeks? Please put your answer in the box below.

[] Hours per week Employed

3. About how many hours per week did you spend studying for the examination during the last five weeks? Please put your answer in the box below.

[] Hours per week Studied

4. Circle the number corresponding to the special session you took:

- 1 Videotape of Arbitration and Court Room cases
- 2 Case of Barelvas v. Brown (Research Task)
- 3 Case of State v. Dolan (Research Task)
- 4 60 multiple choice questions and script with questions
- 5 Two essay questions and 20 multiple choice questions
- 9 Did not participate in a special session

5. In your opinion, how good a measure of your LEGAL KNOWLEDGE was each part of the examination? Please circle one number for each part.

	Very Poor	Poor	Fair	Good	Very Good	Did Not Take
a. Special Session	1	2	3	4	5	9
b. Essay Examination	1	2	3	4	5	9
c. Multistate (MBE)	1	2	3	4	5	9

6. In your opinion, how good a measure of your ABILITY TO PERFORM AS AN ATTORNEY was each part of the examination? Please circle one number for each part.

	Very Poor	Poor	Fair	Good	Very Good	Did Not Take
a. Special Session	1	2	3	4	5	9
b. Essay Examination	1	2	3	4	5	9
c. Multistate (MBE)	1	2	3	4	5	9

7. In your opinion, was the time allowed for each part the examination appropriate? Please circle one number for each part.

	Less Than Enough	About Right	More Than Enough	Did Not Take
a. Special Session	1	2	3	9
b. Essay Examination	1	2	3	9
c. Multistate (MBE)	1	2	3	9

8. Please circle one number below for each part of the examination to indicate how much you agree or disagree with the statement: "The case situations were realistic."

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Did Not Take
a. Special Session	1	2	3	4	5	9
b. Essay Examination	1	2	3	4	5	9
c. Multistate (MBE)	1	2	3	4	5	9

9. In general, were the directions and questions in your special sessions clear or ambiguous? Please circle one choice below.

1-Clear 2-Mixed 3-Ambiguous 9-Did Not Take

10. What were your general reactions to the special session? How could it be improved? Should it be included in future examinations?

APPENDIX G

DIRECTIONS AND QUESTIONS FOR ESSAY TIME LIMITS STUDY AND DIRECTIONS FOR ESSAY/MBE CONGRUENCE STUDY

SPECIFIC DIRECTIONS FOR PARTS I AND II - ESSAY QUESTIONS

An answer should demonstrate your ability to analyze the facts presented by the question, to select the material from the immaterial facts, and to discern the points upon which the case turns. It should show your knowledge and understanding of the pertinent principles and theories of law, their relationship to each other, and their qualifications and limitations. It should evidence your ability to apply the law to the facts given, and to reason logically in a lawyer-like manner to a sound conclusion from the premises adopted. Try to demonstrate your proficiency in using and applying legal principles rather than mere memory of them.

An answer containing only a statement of your conclusions will receive little credit. State fully the reasons that support them. All points should be thoroughly discussed. Although your answer should be complete, you should not volunteer information or discuss legal doctrines that are not necessary or pertinent to the solution of the problem.

Unless a question expressly asks for California law, it should be answered according to legal theories and principles of general application.

SPECIFIC DIRECTIONS FOR PART III - MULTIPLE CHOICE QUESTIONS

Each of the questions or incomplete statements below is followed by four suggested answers or completions. You are to choose the best of the stated alternatives. Answer all questions according to the generally accepted view, except where otherwise noted.

For the purposes of this test, you are to assume that the Uniform Commercial Code Articles 1, 2, and the provisions of 9 relating to the assignment of accounts receivable and contract rights have been adopted. The Federal Rules of Evidence are controlling. You are also to assume that there is no relevant comparative negligence rule, no relevant No-Fault Insurance Act, and no applicable community property law unless the specific question informs you to the contrary.

QUESTION NO. 21

In early 1977 the XU Boosters Club, an incorporated group of State X University alumni, offered in a letter to Mr. Chops, a karate teacher at the largest high school in State X, to pay Chops a "commission" of \$500 for each "quality" high school football player Chops might induce to enroll at State X University during a period of two years, commencing in September 1977. The letter concluded: "We emphasize that you must agree to solicit for us only footballers who are large, fast, agile, and mean."

Chops promptly replied by mail: "I accept your proposal and promise to abide by the terms thereof." Before making the offer, the Club's officers knew that for several years Chops had directed dozens of his school's star football players to out-of-state universities, and mentioned their awareness of this fact in their letter to Chops. On receipt of Chops's reply, the Club pledged \$300,000 toward a stadium addition at State X University and induced the hiring of three additional assistant football coaches in anticipation of an upsurge in the University's football success.

Chops did not recruit or try to recruit any players for State X University during 1977, 1978, or 1979 but did recruit a football player named Runner for State Y University in 1978. In a lawsuit against Chops based on breach of contract and promissory estoppel, the XU Boosters Club alleged the foregoing facts plus damages allegedly sustained.

Chops has sought advice from the law firm for which you work. Chops says he believed that the letter from the Boosters Club sought only a non-binding "gentleman's agreement," and that his reply letter meant only that he would give State X University reasonable consideration in any recruiting efforts he might make.

Chops provides you with these additional facts: (a) between June 1977 and September 1979 no player who was large, fast, agile and mean had graduated from the high school where Chops was employed; (b) during that time, Runner was the only player Chops had solicited for any university; (c) Runner is a sensationally fast and agile kick return specialist and the smallest person to have played on his high school's football team since 1962; (d) Chops had received a "commission" of \$5,000 from a group of State Y University alumni when Runner enrolled there; and (e) Chops has received no money from Boosters.

1. Is the court likely to agree with Chops' interpretation of his exchange of correspondence with Boosters and, if it does not, what is the most likely interpretation? Discuss.
2. Will the additional facts which Chops has provided support any good defenses to the suit if the court finds that there was a binding agreement? Discuss.
3. Do the alleged facts support any theory upon which Boosters might recover judgment against Chops for the \$300,000 Boosters had pledged for the stadium addition? Discuss.

QUESTION NO. 22

Adam was the owner of record of Wildacre, a ten-acre tract of vacant, unimproved land located just beyond the city limits of an expanding industrial city. On June 1, 1980, he entered into a valid executory contract with Baxter for the sale of Wildacre to Baxter. June 30, 1980, was the date specified in the contract for Adam to deliver to Baxter a warranty deed to Wildacre and for Baxter to pay Adam the purchase price of \$45,000.

On June 15, 1980, Baxter, in consideration of the sum of \$50,000 paid to him by Cramer, executed and delivered a warranty deed purporting to transfer Wildacre to Cramer. Cramer made no search of the records to determine who appeared as owner of record. However, Cramer immediately recorded his deed to Wildacre.

On June 30, 1980, Baxter paid Adam the agreed upon purchase price of \$45,000, and Adam executed and delivered to Baxter his warranty deed to Wildacre. Baxter had it recorded immediately.

On July 15, 1980, Baxter, in consideration of \$40,000 paid to him by Dawson, executed and delivered to Dawson a warranty deed to Wildacre. Dawson made no search of the records and had no knowledge of any of the prior transactions concerning the tract. Furthermore, Dawson misplaced the deed and never had it recorded.

In the jurisdiction where Wildacre is situated, the recording statute provides that unrecorded deeds are void as to subsequent purchasers who (1) are innocent and (2) have paid a valuable consideration.

The jurisdiction has and maintains both an alphabetical grantor-grantee index and a tract index but there has never been a decision as to which, if either, should take precedence over the other.

1. What, if any, were the legal and equitable interests of Adam, Baxter and Cramer in Wildacre as of June 2, 1980, June 16, 1980, and July 1, 1980? Discuss.
2. In a contest between Cramer and Dawson as to the title to Wildacre, what arguments are each likely to assert and who should prevail? Discuss.

QUESTION NO. 23

James, a wealthy stockbroker, stabbed Allen, his customer, after Allen had slapped James in the face during a violent argument over money at Allen's apartment. James, his clothes spotted with blood from Allen's wound, ran from the apartment screaming "Good Lord, what will the financial world think?" His departure and exclamation were witnessed by Sylvia, Allen's friend, who arrived at the front door of the apartment as James fled.

Upon entering Allen's apartment and discovering Allen wounded and unconscious, Sylvia called the police. When the police arrived at the apartment, Sylvia described to them the man she had seen flee the apartment.

Using the description provided by Sylvia, the police stopped James as he arrived home. James was placed under arrest and immediately taken to Allen's hospital room. Allen, who was on the critical list and not expected to live out the night, identified James as his assailant. Thereafter Allen showed surprising improvement. However, he died three weeks later from a virus contracted while in the hospital.

After the hospital visit, James was taken to police headquarters and placed in a lineup. At the demand of the police and over his objection, James repeated the words, "Good Lord, what will the financial world think?" At the lineup, Sylvia identified James both by sight and by voice as the person she had seen leaving the apartment. Later, over James' objection, a police officer took his clothing. Subsequent laboratory analysis matched Allen's blood type with the blood on James' clothing. James was charged with murder in the first degree.

James has moved to prevent the introduction into evidence at his trial of:

- a. Any testimony as to his identification by Allen at Allen's hospital room;
 - b. Any testimony as to his visual identification by Sylvia at the lineup;
 - c. Any testimony as to the identification of his voice by Sylvia at the lineup; and
 - d. His clothing and the results of the laboratory analysis of the blood found thereon.
1. How should the court rule regarding the admissibility of the above items of testimony and evidence? Discuss.
 2. Has James committed murder in the first degree in the death of Allen and, if not, has James committed any lesser included criminal homicide? Discuss.

QUESTION NO. 24

Al, a professor of biology, was conducting an experiment in the basement of his home in an attempt to mate a docile variety of bee with a more aggressive variety of bee that had begun to move north toward the United States from South America. Al had a hive of both varieties of bees. One Saturday morning while Al was away, Debbie, a six year old girl who lived next door to Al, was standing in her driveway carelessly throwing rocks, one of which hit and broke a basement window of Al's house. Several bees of the aggressive variety escaped and stung Debbie, who immediately became severely ill and was rushed to the hospital. A consulting physician told Debbie's father, Dan, that he needed to test some bees of the type that had stung Debbie in order to treat her.

Dan immediately went to Al's house and, finding no one there, broke down the back door and entered the basement. He caught several bees in a tin can and, in the process of leaving with the bees, knocked over a jar of corrosive chemicals onto Al's antique rug. The chemicals severely damaged the rug.

Debbie subsequently recovered from the bee stings. The shock of the episode, however, released previously undetected latent psychotic tendencies and she has remained in a mental institution.

Al was thereafter prosecuted for not having obtained, as required by federal statute, a license from the Department of Health to conduct experiments with live specimens of animals or insects "not normally found in North America." Al pleaded guilty and paid a \$500 fine.

1. What are Debbie's rights against Al? Discuss.
2. What are Al's rights against Dan? Discuss.