Case No. _____

IN THE SUPREME COURT OF THE STATE OF CALIFORNIA

**REQUEST THAT THE SUPREME COURT APPROVE
PROPOSED RAW PASSING SCORE AND SCORING
ADJUSTMENTS FOR THE FEBRUARY 2025
CALIFORNIA BAR EXAMINATION**

PREPARED BY
THE STATE BAR OF CALIFORNIA

OFFICE OF GENERAL COUNSEL
Ellin Davtyan (238608), General Counsel
Kirsten Galler (227171), Deputy General Counsel
Jean Krasilnikoff (280450), Assistant General Counsel

180 Howard Street
San Francisco, CA 94105
Telephone: (415) 538-2369

845 South Figueroa Street
Los Angeles, CA 90017
Telephone: (213) 765-1000

Facsimile: (415) 538-2321
Email: OGC@calbar.ca.gov

TABLE OF CONTENTS

IN THE SUPREME COURT OF THE STATE OF CALIFORNIA

REQUEST THAT THE SUPREME COURT APPROVE PROPOSED RAW PASSING SCORE AND SCORING ADJUSTMENTS FOR THE FEBRUARY 2025 CALIFORNIA BAR EXAMINATION

I. INTRODUCTION

In October 2024, the State Bar filed a renewed petition with this Court seeking an order that would approve modifications to the California Bar Examination to allow the State Bar to utilize multiple-choice questions other than the Multistate Bar Examination (MBE) and to allow for the examination to be held remotely and/or in-person test centers. The petition specifically contemplated that, as part of transitioning away from the MBE, the Committee of Bar Examiners (CBE) would need to set a new raw passing score for the February 2025 bar examination and that future bar examinations would be statistically equated to the expectations of the February 2025 bar examination to maintain consistency of interpretation of the passing score across different administrations of the examination.

During and following the February 2025 bar examination, the State Bar received widespread reports of technological and other issues impacting test takers. With the assistance of a psychometric consulting company, the CBE evaluated the impacts of the issues on test takers to determine whether remedial measures were appropriate. During its April 18, 2025, meeting, the CBE considered the reported experiences of test takers, available test-taker performance data, and the challenges associated with fairly and accurately categorizing the level of disruption experienced by individual test takers. At that meeting, the CBE adopted a resolution recommending, as remediation measures, psychometrically imputing scores for eligible test takers and setting a raw passing score of 534 for the February 2025 bar examination. This raw passing score factors in that 171 of the 200 multiple-choice questions would be scored. The 171 scored questions have a reliability of 0.89, well above the desired range of 0.80 or above.

This petition also addresses this Court's request for information relating to the use of artificial intelligence in the development of multiple-choice questions and the performance of

those questions as compared to multiple-choice questions that were not drafted with the use of artificial intelligence, including whether any multiple-choice questions were removed from scoring because they were determined to be unreliable, and the reliability of the remaining multiple-choice questions used for scoring.

This petition requests that this Court adopt the CBE's recommendations on psychometric imputation of scores and on setting the raw score of the February 2025 bar examination, and issue an order as soon as possible, so that the State Bar can release the results of the bar examination on or around May 2, 2025.

II.   BACKGROUND

A.   The State Bar's Petition and the Supreme Court's Order Eliminating References to the MBE and Authorizing a Remote Examination

On October 4, 2024, the State Bar filed a renewed petition with this Court seeking modifications to the California Bar Examination, including allowing the examination to be administered remotely and/or in-person and eliminating reference to the MBE. (Appendix of Exhibits (AE), Ex. 1 [State

Bar's Renewed Request Seeking Modifications to the Bar

Examination Case No. S287231] at pp. 3–35.)[1] These

modifications were required because the Court's prior order on

the administration of the bar examination specifically referred to

the MBE and required administration to be in-person.

The petition explained that the State Bar had entered into

a contract with Kaplan Exam Services, LLC (Kaplan) to develop

multiple-choice questions to be administered starting with the

February 2025 bar examination and attached a copy of the

contract to that petition. (Ex. 2 [Excerpts of the Appendix of

Exhibits in Support of Renewed Request] at pp. 75–108.) The

contract stated that the State Bar would provide First-Year Law

Students' Examination (FYLSX) questions to Kaplan as source

materials. (*Id*. at p. 77.) However, the petition did not

contemplate the use of multiple-choice questions drafted by other

sources on the bar examination. (See Ex. 1.)

---

[1] Hereafter, all references to exhibits refer to exhibits included in
the AE and the page number references are to the consecutively
paginated pages in the AE.

The petition also explained that the State Bar, in consultation with its psychometrician, developed a plan to ensure that the multiple-choice questions would be properly vetted and prepared before the February 2025 bar examination. (Ex. 1 at pp. 22–26.) As described in the petition, following the receipt of questions from Kaplan, the State Bar planned to convene a content validation panel led by psychometricians and comprised of recently licensed attorneys, individuals that supervise recently licensed attorneys, and law school faculty. (*Id.* at pp. 23–24.) The content validation panel would review each question to ensure that the question: (1) tests for minimum competence to practice law; (2) is not biased; (3) is clear; (4) is cohesive in style with other questions; and (5) accurately tests the intended legal issue. (*Ibid.*) The content validation panel would then recommend edits, as needed, to achieve these criteria and return them to Kaplan, which would then finalize the questions and return them to the State Bar. (*Ibid.*)

The petition further explained that, if the Court permitted the State Bar to use non-MBE multiple-choice questions, the CBE would be required to set a raw passing score through a

standard validation study because the examination would no longer be anchored to the MBE. (*Id.* at p. 25.) The petition described that a standard validation panel would be convened after the February 2025 examination administration to develop a recommendation for a raw passing score that is equivalent to current standards and then provide that information to the CBE. (*Id.* at p. 26–26.) The CBE, in turn, would make a policy decision to set the raw passing score. (*Ibid.*) At the time of the petition, the State Bar anticipated that the raw passing score as determined by the CBE following the February 2025 bar examination would carry over to future administrations of the bar examination.

The Court approved modifications to the California Bar Examination as set forth in the Court's Administrative Order 2024-10-21-01, filed on October 22, 2024. (Ex. 3 [Administrative Order 2024-10-21-01] at pp. 110–111.) The Order eliminated reference to the MBE in the prior administrative order and instead specifies that the "second day of the General Bar Examination will constitute the multiple-choice portion of the examination and will consist of 200 multiple-choice questions

administered over four 90-minute sessions." (*Ibid*.) The Order

also maintained the scaled passing score of 1390. (*Ibid*.) Finally,

the Order authorized the bar examination to be administered

"remotely and/or in-person at vendor-run or State-Bar run test

centers." (*Ibid*.)

B. Multiple-Choice Question Development, Selection, and Scoring

The State Bar expected that, after entering into a contract

with Kaplan, the multiple-choice questions for future bar

examinations, including the February 2025 bar examination,

would consist of Kaplan-drafted questions, as well as FYLSX

questions that were provided to Kaplan as source materials.

However, in late October 2024, Office of Admissions' staff

determined that there were not enough multiple-choice questions

for each of the subtopics of the seven subject areas tested. As

such, staff requested that ACS Ventures, LLC (ACS)—the

psychometric and test development consulting company with

which the State Bar contracts to assist with examination

analysis, grading, and related services—draft additional

questions for the February 2025 bar examination.

Regardless of their origins, all multiple-choice questions went through both the content validation process and review by a subject-matter expert before the commencement of the February 2025 examination. Following the examination, the questions that appeared on the examination were subject to a psychometric analysis and reviewed by standard validation panels.

1. Kaplan Question Development

Kaplan employs a team of attorneys who write, edit, and review the multiple-choice questions that are provided to the State Bar. The questions go through multiple rounds of review prior to being submitted to the State Bar, and if any questions are returned to Kaplan for editing following review by the State Bar's content validation panels, Kaplan staff review any feedback and edit the questions, where appropriate. This general process is described in the State Bar's agreement with Kaplan. (Ex. 2 at pp. 75–80.)

As for the FYLSX questions, Kaplan did not make any edits to those questions, and such questions proceeded to content validation and subject-matter expert review as described further below.

In total, 117 Kaplan-drafted questions and 54 FYLSX questions were administered on the February 2025 bar examination.[2] Of those, 100 Kaplan-drafted questions and 48 FYLSX questions were scored for the examination. (See Section II.B.4, *post.*)

2. ACS Question Development

On or around October 30, 2024, State Bar Admissions' staff requested that ACS draft additional questions for the February 2025 bar examination to ensure that there were a sufficient number of questions in all subtopics of the subject areas.[3]

---

[2] Although the agreement between the parties states that the State Bar would provide Kaplan questions that had "previously appeared" on the FYLSX (Ex. 2 at p. 77), Admissions staff has confirmed that all FYLSX questions used in the February 2025 bar examination have not been previously used on the FYLSX.

[3] Prior to this, in late September 2024, Admissions staff requested that ACS draft two questions per subject area for inclusion in the November bar examination study, so that the November study would test a total of 49 questions. ACS drafted 14 questions for the November bar examination study using OpenAI ChatGPT, and the remaining 35 questions were drafted by Kaplan. Of the 14 ACS-drafted questions, 11 were carried over and used on the February 2025 bar examination because they were among the top performing questions from the November study. The decisions by Admissions staff to request that ACS develop questions for the November bar examination study and

Admissions' staff identified both the topic areas and the number of items needed. ACS drafted prompts to yield multiple-choice questions that aligned with the topic areas identified by Admissions' staff and ran the prompts through OpenAI ChatGPT.

ACS performed an initial edit on each draft item to ensure the questions had a standardized structure, but did not review for content accuracy, bias, or to determine whether the question was appropriate for entry-level attorneys. ACS sent the questions to the State Bar and the questions were reviewed by the content validation panels described below. (See Section II.B.3, *post*.)

Ultimately, there were 29 ACS-developed questions on the February 2025 bar examination, of which 23 were included as part of the scored items. (See Section II.B.4, *post*.) All the ACS-developed questions that appeared on the February 2025 bar examination are reflected in the chart below (with the underscore

---

for use on the February 2025 bar examination were not clearly communicated to State Bar leadership. Structural changes within Admissions have been made to address this issue.

indicating the correct answer), and if the question was removed

from scoring, the reason it was removed is also provided.

Subject	Origin	Answer Selection Frequency				Diffi-culty (Target 0.30-0.80)	Discr. (Target >0.10)	Reason Removed from Scoring
		A	B	C	D			
Civ. Proc.	ACS	12%	17%	68%	2%	0.68	0.15	n/a
Civ. Proc.	ACS	29%	6%	26%	39%	0.29	0.09	n/a
Contracts	ACS	6%	23%	55%	17%	0.55	0.20	n/a
Contracts	ACS	28%	50%	18%	4%	0.50	- 0.13	Negative discrimination
Crim. Law and Proc.	ACS	6%	77%	10%	6%	0.77	0.21	n/a
Crim. Law and Proc.	ACS	13%	87%	0%	0%	0.87	0.11	n/a
Crim. Law and Proc.	ACS	3%	77%	8%	12%	0.77	0.31	n/a
Crim. Law and Proc.	ACS	34%	50%	1%	14%	0.34	- 0.15	Negative discrimination
Crim. Law and Proc.	ACS	16%	46%	4%	33%	0.33	- 0.10	Negative discrimination
Crim. Law and Proc.	ACS	11%	55%	2%	31%	0.55	0.03	n/a
Crim. Law and Proc.	ACS	18%	60%	20%	2%	0.60	- 0.11	Negative discrimination
Crim. Law and Proc.	ACS	58%	28%	6%	7%	0.28	- 0.05	Negative discrimination
Crim. Law and Proc.	ACS	1%	96%	1%	2%	0.96	0.13	n/a
Crim. Law and Proc.	ACS	2%	96%	1%	0%	0.96	0.18	n/a
Crim. Law and Proc.	ACS	34%	1%	41%	23%	0.34	0.24	n/a
Crim. Law and Proc.	ACS	56%	13%	18%	11%	0.18	0.14	n/a
Crim. Law and Proc.	ACS	10%	20%	64%	5%	0.64	0.32	n/a
Crim. Law and Proc.	ACS	60%	25%	2%	13%	0.60	0.09	n/a
Evidence	ACS	3%	3%	88%	6%	0.88	0.20	n/a
Evidence	ACS	12%	79%	5%	4%	0.79	0.28	n/a
Real Property	ACS	8%	5%	22%	65%	0.65	0.32	n/a
Real Property	ACS	2%	88%	6%	4%	0.88	0.28	n/a

Subject	Origin	Answer Selection Frequency				Diffi-culty (Target 0.30-0.80)	Discr. (Target >0.10)	Reason Removed from Scoring
		A	B	C	D			
Torts	ACS	5%	83%	2%	10%	0.83	0.10	n/a
Torts	ACS	2%	11%	72%	16%	0.72	0.12	n/a
Torts	ACS	21%	1%	1%	77%	0.77	0.23	n/a
Torts	ACS	2%	12%	85%	0%	0.85	0.14	n/a
Torts	ACS	1%	2%	6%	91%	0.91	0.26	n/a
Torts	ACS	12%	68%	0%	19%	0.68	0.07	n/a
Torts	ACS	1%	34%	1%	64%	0.64	- 0.03	Negative discrimination

According to Dr. Buckendahl, individual questions do not have a reliability value, but item discrimination is a good indicator of question performance. Item discrimination is a measure of how well a question differentiates between individuals who score higher on the test, indicating more knowledge about the subject matter, and those who score lower on the test, indicating less knowledge about the subject matter. The target range for item discrimination is over 0.10; however, questions that are below the target with a low discrimination value (above 0.00 and below 0.10) are not necessarily excluded from the scored items if they contribute to content coverage. (Ex. 4 [February 2025 GBX Item Analysis Summary] at pp. 114–116; see also Section II.B.4, *post*.)

As set forth in the chart above, all ACS questions that had a negative discrimination value were removed from the scored items. A question that has a negative discrimination value (a discrimination value below zero) means that lower-performing test takers are more likely to get the item correct than higher-performing test takers, so including questions with a negative discrimination adversely impacts examination reliability. The average item discrimination for ACS-drafted questions that were selected for scoring was 0.18 as compared to 0.21 for Kaplan-drafted and 0.19 for FYLSX questions. (*Id*. at p. 116.)

The difficulty range of the ACS questions largely fell within the target of between 0.30 and 0.80 (with over 0.90 generally considered as too easy and below 0.25 generally considered as too difficult). As with questions that have low item discrimination, questions that do not meet the difficult target are not necessarily excluded from the scoring if they contribute to content coverage. (See Section II.B.4, *post*.) Of the questions that were selected for scoring, the average item difficulty for ACS-drafted questions was 0.70 as compared to 0.66 for Kaplan-drafted and 0.63 for FYLSX questions. (Ex. 4 at p. 115.)

3. Content Validation and Subject-Matter Expert Review

Prior to the administration of the February 2025 bar examination, based on the ACS's guidance, the State Bar convened content validation panels, consisting of recently licensed attorneys, individuals that supervise recently licensed attorneys, and law school faculty to review the drafted multiple-choice questions.[4] Admissions staff identified and selected potential participants who generally fell within one of the following categories: California attorneys who supervise Provisionally Licensed Lawyers or certified law students; law school deans and faculty from American Bar Association-approved (ABA) law schools, California-accredited law schools, and unaccredited law schools who previously volunteered on

_____

[4] Likewise, the National Conference of Bar Examiners (NCBE) sought panelists with a similar composition for its NextGen Content Scope Committee in 2021. At that time NCBE sought out newly licensed lawyers with one to three years of experience practicing law; experienced lawyers who work closely with newly licensed lawyers; legal educators; members of the judiciary; and lawyer members of the admissions community (bar examiners, graders, and administrators). (See https://nextgenbarexam.ncbex. org/join-ncbe-in-developing-the-next-generation-of-the-bar-exam/.)

other Admissions' projects and initiatives; and law school faculty who specialize in bar examination preparation.[5]

Panelists signed volunteer services agreements in which they agreed to maintain the confidentiality of the examination materials and made representations and warranties regarding their recent access to NCBE intellectual property, any licensing agreements they have with NCBE, and commercial bar examination preparation activities. These representations and warranties were intended to limit the State Bar's exposure to copyright infringement and related claims as well as reduce the appearance of a conflict of interest where the proposed participant has a financial interest in the commercial bar examination preparation business. Panelists were required to

---

[5] Law school deans and faculty who attended the content validation panels represented ABA-approved law schools (four individuals), California-accredited law schools (three individuals) and unaccredited law schools (one individual). Because Admissions' staff were seeking a maximum of seven attorneys from each category, staff chose not to do a broad solicitation and application process out of concern they would need to turn attorneys away. In reviewing the process used, it is apparent to State Bar leadership that there were shortcomings that will be improved upon. (See Section VI, *post.*)

volunteer their time and attend two-day meetings in Los

Angeles.[6]

The content validation panels were facilitated by

assessment specialists and psychometricians from ACS who

provided training for the panelists regarding the content validity

evaluation criteria and compiled the panelists' judgments and

feedback about the questions on rating forms designed for that

purpose.

For the content validation panels, the questions were not

identified by source, and so panelists did not know whether the

questions were drafted by Kaplan, ACS, or had been part of the

FYLSX question bank. The panels received batches of multiple-

choice questions and reviewed the questions together for content

accuracy, bias, minimum competency alignment. In addition, the

panels ensured the questions aligned with the State Bar's content

maps and provided feedback to the State Bar about the questions.

---

[6] Certain individuals that Admissions' staff identified through
their outreach were ultimately unable to participate on the
panels due to their inability to make the representations and
warranties in the volunteer services agreement.

Thereafter, a psychometric facilitator from ACS produced a report for the State Bar with the panel feedback, which noted whether the question was fine as drafted, whether it should be revised, or whether it should be replaced. Kaplan-drafted questions that required revision were sent back to Kaplan for further editing, and all questions were reviewed by a subject-matter expert before the commencement of the February bar examination.

In addition to the content validation process, prior to the commencement of the February bar examination, all multiple-choice questions were reviewed by least one subject-matter expert for legal accuracy.

State Bar staff selected multiple-choice questions for inclusion on the February 2025 bar examination to ensure that the questions met the subject-matter allocations set forth in the State Bar's content maps.

4. **Reducing the Number of Scored Multiple-Choice Questions to Increase Examination Reliability**

When the bar examination included the MBE's 200 multiple-choice questions, only 175 of those questions were

scored, while the remaining 25 were unscored and considered

experimental. Having moved away from the MBE, the State Bar

adopted a similar approach and intended to score only 175 of the

200 questions, with the expectation of having 25 questions scored

in each tested subject.[7]

To identify which questions should be scored and which

should not, following the administration of the February 2025 bar

examination, ACS, led by psychometrician Dr. Buckendahl,

analyzed the performance of the multiple-choice questions. This

analysis reviewed whether each question was within the target

difficulty and discrimination ranges, whether at least five percent

of applicants selected at least three of the four answer choices,

---

[7] The 25 items not included in the scored items on the MBE are
identified in advance of the examination and characterized as
experimental. The concept of having 25 preidentified
experimental questions was inapplicable to the State Bar's
February 2025 bar examination because the 200 questions had
not been tested on a bar examination before (though 35 of the
questions had been tested during the November bar examination
study), and thus the State Bar did not have data regarding how
each of the questions performed on a bar examination in advance
of the February 2025 examination. As such, the State Bar
identified which questions it would not score after the questions
had been subject to psychometric analysis.

and whether there was adequate coverage across subjects. The standard validation panel, discussed in Section II.B.5, *post*, then assessed which of the questions a minimally competent lawyer would be expected to answer correctly. Dr. Buckendahl presented the results of his analysis to the CBE at its April 18 meeting.

In reviewing the statistical performance of the questions, any question that had a negative item discrimination value was not included in the scored items because such questions adversely impact examination reliability. This is because a negative item discrimination means that those who performed poorly on the overall test were more likely to answer that question correctly compared to those who performed well on the overall test, which suggests that the question may not be measuring what was intended or may be flawed in some way. Of the 200 tested questions, 17 had a negative item discrimination and were removed from the scored items, as set forth below: [8]

---

[8] As with the chart in Section II.B.2, *ante*, the underscore in the answer selection column in the charts in this Section indicates the correct answer.

Subject	Origin	Answer Selection Frequency				Difficulty (Target 0.30-0.80)	Discr. (Target >0.10)
		A	B	C	D		
Civ. Proc.	Kaplan	15%	36%	7%	42%	0.36	- 0.16
Con. Law	Kaplan	40%	1%	58%	1%	0.58	- 0.09
Con. Law.[9]	Kaplan	3%	1%	0%	0%	0.03	-0.08
Con. Law	Kaplan	36%	6%	24%	34%	0.34	- 0.02
Contracts	ACS	28%	50%	18%	4%	0.50	- 0.13
Contracts	FYLSX	10%	16%	15%	58%	0.15	- 0.05
Crim. Law and Proc.	ACS	34%	50%	1%	14%	0.34	- 0.15
Crim. Law and Proc.	ACS	16%	46%	4%	33%	0.33	- 0.10
Crim. Law and Proc.	ACS	18%	60%	20%	2%	0.60	- 0.11
Crim. Law and Proc.	ACS	58%	28%	6%	7%	0.28	- 0.05
Evidence	Kaplan	10%	55%	5%	29%	0.29	- 0.05
Evidence	Kaplan	6%	42%	19%	32%	0.32	- 0.02
Real Property	Kaplan	68%	8%	5%	18%	0.08	- 0.12
Real Property	Kaplan	1%	14%	37%	47%	0.37	- 0.03
Torts	ACS	1%	34%	1%	64%	0.64	- 0.03
Torts	Kaplan	38%	16%	42%	3%	0.42	- 0.05
Torts	FYLSX	18%	2%	26%	53%	0.18	- 0.01

In total, nine Kaplan-drafted questions, two FYLSX questions, and six ACS-drafted questions were removed from the scored items due to negative item discrimination.

To reach 175 scored items, Dr. Buckendahl and his team then clustered the remaining 183 questions by subject area to

---

[9] This question also appeared to have a delivery malfunction where only a small number of applicants were able to answer the question.

ensure that 25 questions in each subject area were selected for

scoring. By subject area, Dr. Buckendahl's team reviewed the

other performance statistics, including item discrimination, item

difficulty, and available response options, to identify additional

questions to be removed from the scored questions. The target

ranges for the performance statistics include: item difficulty

between 0.30 to 0.80; item discrimination above 0.10; and

whether at least three response options were selected by five

percent or more of test takers. Significantly, questions that do not

fall within these targets are not automatically removed from

scoring, because including them in the scored items does not

negatively impact examination reliability (the targets are goals,

not brightline disqualifiers) and because it is important to

adequately cover the tested subject areas. Moreover, including

options that are outside of the target difficulty range (and thus

statistically may be considered "too easy" or "too difficult") or

have less than the ideal number of response options, or even

multiple correct response options does not negatively impact

examination reliability. Among other things, the standard

validation panel considers the likelihood that a minimally

24

competent lawyer would answer such questions correctly in developing the raw score, which reflects the performance expectations on the questions. In other words, the standard validation panel would expect a minimally competent lawyer to answer a question that is "too easy" correctly, but not a question that is "too difficult." While questions that do not meet the target ranges may be included in the scored questions, the statistical data on question performance has continuing utility, as the data may be used to inform potential improvements to the questions for future administrations of the examination.[10]

Dr. Buckendahl and his team reviewed the question performance data by subject area and identified eight additional questions to remove from the scored items to reach 175 scored items. Those questions are as follows:

Subject	Origin	Answer Selection Frequency				Difficulty (Target 0.30-0.80)	Discr. (Target >0.10)
		A	B	C	D		
Contracts	FYLSX	4%	22%	7%	66%	0.66	0.23
Contracts	FYLSX	23%	70%	1%	6%	0.70	0.09
Crim. Law and Proc.	FYLSX	33%	34%	3%	30%	0.30	0.04

---

[10] NCBE does not publicly release data on the MBE's question performance, including which of its questions are statistically deemed "too easy" or "too difficult," and, as such, the State Bar is unable to compare the performance its questions with the MBE.

Subject	Origin	Answer Selection Frequency				Difficulty (Target 0.30-0.80)	Discr. (Target >0.10)
		A	B	C	D		
Evidence	Kaplan	57%	1%	2%	40%	0.40	0.00
Real Property	Kaplan	7%	2%	90%	1%	0.90	0.20
Real Property	Kaplan	17%	21%	50%	11%	0.21	0.10
Real Property	Kaplan	4%	3%	90%	3%	0.90	0.30
Torts	FYLSX	56%	20%	3%	20%	0.20	0.05

After Dr. Buckendahl and his team selected the 175 multiple-choice questions for scoring, the State Bar informed Dr. Buckendahl that its subject-matter expert previously identified an additional four questions as potentially having accuracy issues.[11] Dr. Buckendahl, therefore, recommended removing those four additional questions from the scored items, resulting in a total of 171 scored items. Those additional four questions are as follows:

---

[11] Prior to the administration of the bar examination but after the questions had already been uploaded to the Meazure Learning platform (and thus those questions could not be revised further), the State Bar's subject-matter expert identified a total of six questions that had potential accuracy concerns. In identifying the additional eight questions for removal to achieve the target of 25 unscored questions, Dr. Buckendahl and his team had already selected two of the questions with legal accuracy concerns (the Criminal Law and Procedure question and the Torts question in the chart above) for removal from the scored items.

Subject	Origin	Answer Selection Frequency				Diffi-culty (Target 0.30-0.80)	Discr. (Target >0.10)
		A	B	C	D		
Civ. Proc.	Kaplan	1%	34%	61%	4%	0.61	0.05
Con. Law	Kaplan	54%	41%	3%	2%	0.54	0.09
Contracts	Kaplan	22%	30%	40%	7%	0.62	0.11
Crim. Law and Proc.	Kaplan	5%	4%	90%	1%	0.90	0.20

In summary, of the 29 questions that were not included as scored items, 17 were drafted by Kaplan, six were questions from the State Bar's FYLSX question bank, and six were drafted by ACS. (Ex. 4 at pp. 113–114.) Of the questions that were scored, 100 were drafted by Kaplan, 48 were questions from the State Bar's FYSLX question bank, and 23 were drafted by ACS. (*Ibid*.) As a result of removing these questions from the scored items, this left 24 scored questions for civil procedure, 23 scored questions for criminal law and procedure, and 24 scored questions for torts. (*Id*. at p. 114.) The remaining subject matter areas have 25 scored questions. (*Ibid*.)

Removing these 29 questions improved examination reliability. Indeed, by scoring 171 multiple-choice questions instead of 200, the reliability of the multiple-choice section of the examination increased from 0.87 to 0.89 on a scale of 0.00 to 1.00, where values of 0.80 or higher are desired. (See Ex. 4 at p. 113.)

5. Post-Examination Standard Validation Review

In addition to Dr. Buckendahl's psychometric analysis of question performance, the State Bar convened standard validation panels to develop a recommended raw passing score for the February 2025 bar examination that is equivalent to current expectations. This process was required because the scores on the examination were no longer anchored to the MBE. The purpose of these panels was to determine how many of the responses a minimally competent lawyer should be expected to answer correctly.

One panel focused on the multiple-choice questions and the other panel focused on the written portions of the examination. The panels were comprised of recently licensed lawyers, experienced lawyers who supervise recently licensed lawyers, and law school faculty. While the State Bar convened two panels, the

individuals who participated in the panels were largely the same.[12]

The first panel reviewed the 200 multiple-choice questions to evaluate questions against the current passing standard. The purpose of this panel was to determine how many questions one would expect a minimally competent test taker to answer correctly. As such, the panel was responsible for making judgments on expected test-taker performance for each of the 200 multiple-choice questions administered during the February 2025 bar examination by using the performance-level descriptors for a minimally competent lawyer that were developed for the State Bar's 2017 standard setting study. (See July 28, 2017, Final Report: Conducting a Standard Setting Study for the California Bar Exam, and Appendix C available at https://apps.calbar.ca.gov/cbe/docs/agendaItem/Public/agendaitem1000001929.pdf

---

[12] The criteria for participation in the standard validation panels were similar to the criteria for the content validation panels. Law school faculty on the standard validation panels included one participant from an ABA-approved law school, three from California-accredited law schools, and two from unaccredited law schools.

and https://apps.calbar.ca.gov/cbe/docs/agendaItem/Public/ agendaitem1000001932.pdf.) The panel made minimum competence judgments on each of the 200 questions, so that all questions that are included within the scored items have an accompanying judgment by the panel. As guidance, Dr. Buckendahl informed the panel that a range of 110 to 124 correctly answered multiple-choice questions was the typical range of a test taker who passed the 2023 and 2024 California February bar examinations when applying that performance expectation to the February 2025 bar examination. The multiple-choice question standard validation panel results are discussed in Section V.A, *post*.

A second standard validation panel convened to recommend a raw passing score for the written portion of the February 2025 bar examination. The purpose of this panel was to determine the raw passing score for the written component of the examination that reflects minimum competence, using the same performance-level descriptors for a minimally competent lawyer that were used for the multiple-choice questions. The written question

standard validation panel results are discussed in Section V.A,

post.

C. Test Takers Experienced Unacceptable Technological Issues During the February 2025 Administration of the Bar Examination

Following the Supreme Court's approval to administer the

bar examination remotely, the State Bar finalized its contract

and corresponding statements of work with ProctorU d/b/a

Meazure Learning (Meazure Learning) to administer the

February 2025 bar examination both remotely and in-person. A

total of 4,231 applicants sat for the examination;[13] nearly two-

thirds of test takers took the examination remotely and

approximately one-third at in-person test centers. Both remote

and in-person test takers reported experiencing unacceptable

technological issues during the examination, including:

---

[13] Due to a discrepancy in data collection, a presentation to the CBE incorrectly stated the total was 4,232. (Ex. 5 [MAAD Presentation Analytical Work to Support February 2025 Bar Exam Scoring] at p. 125.) An additional discrepancy stated that there were 3,733 test takers who answered all 171 scored multiple-choice questions (Ex. 6 [California Bar Exam February 2025 – Exam Disruption Presentation] at p. 158), and the correct total is 3,503.

significant delays in accessing the examination; frequent disconnections from the examination platform; inability to submit essay responses without technical support; screen freezing; lags in typed responses appearing on the screen; missing examination materials, specifically the performance test file and library; and challenges with functionality—including cut and paste, spell check, and highlighting. In addition to these technological disruptions, test takers reported that proctors and technical support were unable to timely and effectively resolve technological issues, proctors were unprofessional, lacked knowledge of the components of the examination, provided incorrect information, and repeatedly disrupted test takers during the examination. Additionally, a smaller but significant number of test takers and stakeholders expressed concern about the multiple-choice questions, including complaints that they believed questions contained typographical errors, multiple correct answers, and had been drafted by artificial intelligence.

At its April 18, 2025, meeting, the CBE received a presentation from the State Bar's Mission Advancement and Accountability Division (MAAD) that detailed the results of the

State Bar's post-examination survey, which had a 66 percent

response rate. (Ex. 5 at pp. 129–145].)[14] According to that survey,

the majority of respondents experienced technology issues during

the examination, with 96 percent reporting at least one issue

during the essays or performance test and 86 percent during the

multiple-choice question section of the examination. (*Id.* at p.

133.) The majority of respondents also reported experiencing an

issue related to proctors, with 93 percent reporting at least one

issue during the essays/performance test and 84 percent during

the multiple-choice question section of the examination. (*Ibid.*)

Delays or lags between typing and the characters showing up on

the screen were reported by 79 percent of survey respondents (*id.*

at p. 134); 75 percent reported problems with the cut and paste

functionality during the essays and performance test (*ibid.*); and

more than 50 percent reported that the platform crashed or froze

---

[14] A version of this presentation was publicly posted online, including additional slides relating to word-count analyses and other clarifications, including that the information reported was based on then-available data. (See https://www.calbar.ca.gov/Portals/0/documents/admissions/Examinations/MAAD-Feb-2025-Exam-Data.pdf.)

during the multiple-choice portion of the examination. (*Ibid.*) In summary, 86 percent of survey respondents reported that they believed technology issues interfered with their ability to perform their best on the examination either to a significant or moderate extent. (*Id.* at p. 135.) Despite this, nearly 75 percent reported that they completed the examination (*id.* at 132), which is consistent with essay submission reports indicating that 92% of test takers submitted answers to all six essays. (Ex. 6 [California Bar Exam February 2025 – Exam Disruption Presentation] at p. 168.)[15]

D. CBE Considered Standard Validation Panel and Psychometric Recommendations and Available Data to Recommend Imputation and a Raw Passing Score

During its April 18, 2025, meeting, the CBE received the recommendations from ACS and the standard validation panel, as well as available information and data on the significant and unacceptable technological issues experienced by test takers, along with other data. In addition, the CBE received information

---

[15] Exhibit 6 is redacted to protect confidential examination data. (Ex. 6 at p. 171)

on historical pass rates for February bar examinations and the challenges associated with fairly and accurately categorizing the level of disruption experienced by individual test takers during the February 2025 bar examination.

As further discussed in Sections IV, *post*, the CBE deliberated and recommended that, for the February 2025 bar examination, scores be psychometrically imputed where a test taker had missing multiple-choice answers but had answered at least 114 of the 171 scored multiple-choice questions,[16] and where a test taker had missing essay or performance test answers but had answered at least four of the six written sections of the examination. (Ex. 7 [CBE Resolution, dated April 18, 2025]

---

[16] In order to psychometrically impute scores for the multiple-choice portion of the examination, test takers must have answered at least two-thirds of the scored items, which here is 114 of the 171 scored items. As discussed in section II.B.4, *ante*, although the State Bar planned to score only 175 of the 200 multiple-choice questions, in the same way only 175 out of 200 administered questions are scored on the MBE, 171 multiple-choice questions were scored to increase examination reliability.

at pp. 177–179.)[17] The CBE further recommended that the raw passing score for the February 2025 bar examination be set at 534. (See Section V, *post*.)

The CBE's April 18, 2025, resolution directed State Bar staff to immediately seek an order from this Court approving imputation and the recommended raw passing score, and to request a decision by this Court no later than April 28, 2025, to ensure the results of the February 2025 bar examination could be released by May 2, 2025. (Ex. 7 at p. 178.)[18]

---

[17] Similarly, test takers must have completed at least two-thirds of the written section to allow missing scores to be psychometrically imputed. Test takers who were unable to launch the Meazure Learning platform at all, answered fewer than 175 multiple choice questions, or did not successfully submit four or more written responses (essay and/or the performance test) were offered an opportunity to retake the examination in March 2025. (See https://www.calbar.ca.gov/Portals/0/documents/admissions/Examinations/2025-04-18-April-CBE-email-to-applicants.pdf.)

[18] To address the additional information requested on the use of artificial intelligence in the development of multiple-choice questions and multiple-choice development and reliability more generally, the State Bar was unable to file this petition prior to April 29. Notwithstanding this delay, the State Bar respectfully requests that this Court acts promptly so that the bar examination results can be released on or close to May 2.

III. AUTHORITY TO APPROVE A RAW PASSING SCORE AND SCORING ADJUSTMENTS

Under rule 9.6(a) of the California Rules of Court, "[t]he CBE, pursuant to the authority delegated to it by the Board of Trustees, is responsible for determining the bar examination's format, scope, topics, content, questions, and grading process, subject to review and approval by the Supreme Court. The Supreme Court must set the passing score of the examination."

Setting a new raw passing score for the California Bar Examination was contemplated as part of the modifications approved by this Court on October 22, 2024, and as a policy decision that the CBE could undertake. (See Ex. 3 at p. 110; Ex. 1 at pp. 22–26.) Because the CBE determined that it could not recommend individualized scoring adjustments for specific examination disruptions, the CBE addressed the collective effects of disruption through recommending a raw passing score that is lower than they would otherwise have recommended. The CBE's recommended raw passing score is specific to the February 2025 bar examination, and the CBE anticipates that it will set a new raw passing score after the next administration of the bar

37

examination. This Court may consider other remedial measures for February 2025 bar examination test takers under its plenary authority over admission to the bar, and because the recommended raw passing score was conceived as one such remediation measure, the State Bar seeks this Court's approval of the raw passing score.

IV. RECOMMENDED SCORING ADJUSTMENTS FOR THE FEBRUARY 2025 CALIFORNIA BAR EXAMINATION FOR THIS COURT'S APPROVAL

The available data presented to the CBE showed pervasive and significant technology issues impacting test takers. (See Section II.C, *ante.*) After consideration of that data, as well as imputation that was previously applied to address technological issues arising out of the July 2021 bar examination, at its April 18, 2025, meeting, the CBE recommended, as a remediation measure, psychometrically imputing scores for eligible test takers as further explained below.

A. Prior Imputation of Scores on the July 2021 Bar Examination

The State Bar has previously utilized imputation to address technological disruptions to the California Bar

Examination. The State Bar most recently imputed scores for the July 2021 bar examination. (Ex. 8 [Scoring Adjustments for Applicants Negatively Affected During the July 2021 California Bar Exam] at pp. 181–182.) During that examination, technical problems with ExamSoft resulted in test takers' screens going black or blue and required test takers to reboot their computers. (*Ibid*.) The State Bar determined that some test takers were able to reboot and did not lose time on the examination, some applicants were able to reboot but did lose time on the examination, and some applicants were able to reboot but lost content they had typed. (*Ibid*.) Applicants who could not return to where they left off when their screen went black or blue were given the opportunity to retake the examination session (the particular essay, performance test, or multiple-choice session they were taking at the time the problem arose) from the start of that session of the examination. (*Ibid*.)

For those who were negatively impacted by the ExamSoft issues during the written sections of the examination, the State Bar applied a pro rata grading adjustment for each affected question, which utilized data from the unaffected population of

examinees, as well as the affected individual's scores on questions where there were no recorded problems. (*Ibid*.) The adjustment consisted of two components:

- The first component accounted for the relative difficulty of each question. This component was calculated by first obtaining the average score on each written question among examinees who did not experience technical issues on any written question and the overall average of all scores in that group. The difference of the averages on each question and the overall average was used to represent the relative difficulty of each question. For example, if the overall average across all questions was 78 out of 100, and the average on the first question was 76, that question was considered more difficult than the average question by two points. The two-point difference was considered the "adjustment factor" for that question. Each written question was given an "adjustment factor." (*Ibid*.)

- The second component was based on the scores of each individual examinee who experienced the technical issue. To determine this amount, an average was calculated for

40

the examinee's scores for questions where the applicant did not experience a technical issue. That average was considered to be the best estimate of that individual's ability. For each question during which the applicant did experience the technical issues, the average score from the unimpacted sessions was adjusted by the "adjustment factor" to arrive at an "Expected Score." The "Expected Score" was then compared to the examinee's actual score on the question. If the actual score was less than the "Expected Score," an adjustment was made, and the final score on the question was the "Expected Score." If the actual score was greater than the "Expected Score," the score was not changed. (*Ibid.*)

For any examinees who experienced the technical issue during the MBE portion of the examination, the NCBE provided an adjusted score that also used the pro rata method described above. (*Ibid.*)

The approach utilized in 2021 for both the written portions of the examination and the MBE was based on estimating different scores when comparing unaffected to affected test

41

takers. Given the variance in the types and severity of issues test

takers encountered during the February 2025 bar examination,

and the resulting difficulty in categorizing those experiences, the

psychometric model described below and recommended for the

February 2025 bar examination provides better information for

estimating test taker performance than what was used following

the July 2021 bar examination.

B. Proposed Scoring Adjustment: Imputation of Missing Scores for the February 2025 California Bar Examination

Concluding that technology, proctor, or administration

issues likely contributed to some test takers having blank

responses for some multiple-choice questions, essays, or the

performance test, Dr. Buckendahl recommended imputing scores

for portions of the examination that had missing data (i.e., where

a response was blank). Dr. Buckendahl recommended utilizing

Item Response Theory (IRT), specifically the Rasch model, for

multiple-choice questions, and a variation of the Rasch model for

the essays and performance test, to impute missing scores. (Ex. 6

at p. 157.)

The Rasch model provides a probabilistic framework for understanding the interaction between a person's ability and an item's difficulty. Specifically, it models the probability that a person will answer an item correctly based on these two parameters. By estimating these parameters from the available data, the model allows psychometricians to make informed predictions about how likely an individual test taker would be to respond correctly to items they did not answer.

For the multiple-choice questions, as long as a test taker responded to at least 66 percent of the 171 scored questions, scores could be imputed for each question that did not have a recorded response. (*Ibid*.) In those circumstances, Dr. Buckendahl is able to estimate the likelihood that the test taker would answer the question correctly using the Rasch model. By leveraging the relationship between item difficulty and individual ability, this approach allows for defensible interpretations of scores and improved use of partially completed tests.

Dr. Buckendahl did note that one limitation on the modeling is that for test takers who were missing just one or a few responses to multiple-choice questions, the reasons for the

43

blank answer were unknown.[19] While the answer could have

been blank because of technical disruptions, the test taker

running out of time, the test taker skipping the question, or other

reasons, the model imputes a score regardless of the reason for

the missing answer. In light of the breadth of issues faced by test

takers, Dr. Buckendahl defined this as a minor limitation of this

approach.

For the written components of the examination,

Dr. Buckendahl recommended utilizing the one-parameter

Partial Credit Model (PCM), which estimates the likely essay

---

[19] Of the 3,886 February 2025 examinees who took the General
Bar Examination, 3,503 test takers answered all 171 scored
multiple-choice questions. Based on the parameters set forth
above, 382 test takers would qualify to have values for multiple-
choice questions imputed under this approach; 205 test takers
had only one missing multiple-choice response; 125 had between
two and four missing responses, and 52 had five or more missing
responses. There was one test taker who did not answer at least
114 questions and therefore would not qualify to have scores
imputed. For the written component, 4,231 test takers took either
the General Bar Examination or the Attorneys' Examination of
which 4,186 had scores for all six written responses. The
remaining 45 test takers would qualify to have values for missing
written responses imputed; 33 test takers had one written
response missing and 12 test takers had two responses missing.
This data has been updated since it was presented to the CBE at
its April 18 meeting. (See Ex. 6 at p. 158.)

score a test taker would earn for a constructed response question based on the difficulty of the question and the test taker's performance on the questions they did answer. (Ex. 6 at p. 157.) These estimates could be made for test takers who answered at least four of the six written components of the examination. (*Ibid*.) The PCM is a member of the Rasch family of IRT models and is specifically designed for polytomous items (i.e., items that are not scored right or wrong but have multiple possible scoring values), such as essays and performance tests, that represent degrees of performance. The model provides a probability distribution across all score categories for each item-person pair. For example, for a missing response on a 100-point item, like an essay, the model might indicate that a particular test taker has a 10 percent chance of scoring a 40, a 30 percent chance of scoring a 50, a 40 percent chance of scoring a 60, and a 20 percent chance of scoring a 70. From here, missing responses were imputed using model imputation which assigns the score category with the highest probability. This model was selected by Dr. Buckendahl because, among other reasons, it ties the imputed values to item

properties and individual test-taker ability, resulting in imputations that are statistically grounded and interpretable.

At its April 18 meeting, the CBE agreed that the technological and other issues experienced by test takers for the February 2025 bar examination required that missing responses have a score imputed using the IRT models described by the psychometrician and recommends that the Court adopt imputation as set forth herein to score the February 2025 bar examination.

Finally, while not specifically considered by the CBE, a small number of test takers had content in the response field for the performance test, but did not have access to the file and library, which are critical to responding to the question. State Bar Admissions' staff interpret the CBE's action that "missing responses" for purposes of imputation to include these test takers, who otherwise qualify for imputation (i.e., they had completed at least four other sections of the written portion of the examination), and have shared that interpretation with CBE

leadership.[20] Pursuant to that interpretation, these test takers would also receive imputation for "missing responses."

C. Proposed Scoring Adjustment: Accounting for Specific Issues Encountered

Following the administration of the February 2025 bar examination, Dr. Buckendahl reviewed multiple sources of data to determine whether there was any valid way to classify test takers based on the type and severity of disruptions they experienced during the examination. In addition to test-taker reported experiences through calls and emails to the State Bar, and the post-examination survey responses, Dr. Buckendahl also reviewed available data comparing the average word count on the essays and performance tests for the February 2024 and February 2025 bar examinations, the number of remote test takers who were escalated to Meazure Learning's technical support and the number of escalations for each test taker, the amount of time test takers spent in each question, and the total amount of time spent in the examination. During its April 18,

---

[20] There were 18 test takers who experienced this issue.

2025, meeting, the CBE also received reports based on the available data showing that most test takers were able to submit answers to all six questions on the written portion of the examination, and that the word count for those answers was in alignment with the word count for the written portion of the examination in 2024. (Ex. 6 at p. 168.) Additionally, the average first-read total raw written score for 2025 test takers was 423, which was approximately ten points higher than in 2023 (413) or 2024 (411). (*Id.* at p. 167.)

At the April 18 meeting, Dr. Buckendahl reported that classification of test takers into unaffected and affected categories proved elusive due the variances in the types and numbers of disruptions encountered. Indeed, MAAD attempted to synthesize the available data on examination disruptions and in some instances the various pieces of information examined validated one another and, in other cases, there was a lack of convergence. (See Ex. 5.) As a result, Dr. Buckendahl concluded that the available data were not sufficiently reliable to classify test takers into different tiers of disruption impacts or to recommend differential adjustments. (Ex. 6 at p. 163.)

The CBE agreed that instead of recommending scoring adjustments for specific disruptions, the CBE would address the collective effects of disruption through the recommended raw passing score. In other words, the CBE recommends a raw passing score for February 2025 that is lower than would otherwise have been recommended if there had been no or minimal technological issues, as a remediation measure for the significant disruptions faced by test takers during the examination.

V. PROPOSED RAW PASSING SCORE FOR THE FEBRUARY 2025 CALIFORNIA BAR EXAMINATION FOR THIS COURT'S APPROVAL

As part of the move to new multiple-choice questions, the State Bar expected and advised this Court that the CBE would be setting a raw passing score for the bar examination, since it was no longer anchored to the MBE. The Court's Administrative Order 2024-10-21-01 described the number of raw points available on the written portion of the examination (100 raw points for each of the five essay questions and 200 raw points for the performance test question). (Ex. 3 at p. 110.)

To set a raw passing score, the State Bar, with the
assistance of Dr. Buckendahl, conducted a psychometric analysis
of question performance and also convened standard validation
panels to evaluate minimum threshold of performance, informed
by historical data, to establish minimum competence. (See
Sections II.B, *ante*.) The CBE considered this and additional
information in adopting its recommendation that this Court
should adopt a score of 534 points for the February 2025 bar
examination.

A. CBE Approved a Raw Score of 534 Following Consideration of the Recommendations by the Standard Validation Panels and the Psychometrician

As discussed in Section II.B.5, *ante*, the State Bar convened
standard validation panels to review the results of the multiple-
choice questions and the written portions of the examination. The
recommendations of the standard validation panel as well as of
the psychometrician were considered by the CBE at its April 18,
2025, meeting. At that meeting, the CBE adopted the
recommendation that this Court set the raw passing score for the
February 2025 bar examination at 534. (Ex. 7 at pp. 177–179.) As

described below, this figure comprises 114 for the multiple-choice portion and 420 for the written portion.

After reviewing the February 2025 bar examination multiple-choice questions, the standard validation panel concluded that a minimally competent lawyer should be expected to correctly respond to 133 of the 175 multiple-choice questions that were intended to be scored. Following the standard validation meeting, the number of scored multiple-choice questions was further reduced to 171 as a result of additional questions having been identified as potentially having accuracy issues (see Section II.B.4, *ante*). The number of questions a minimally competent lawyer was expected to answer correctly likewise was reduced to 129. That number remains above the recommended range of 110 to 124 provided to the panel as guidance about historical performance. (See Section II.B.5, *ante.*)

At the April 18, 2025, meeting, Dr. Buckendahl presented the results of the standard validation study. Dr. Buckendahl recommended that the CBE adopt a raw score of 120. (Ex. 6 at p. 160.) This raw score is a value that is two standard errors of

the mean below the standard validation panel's recommendation of 129.

Dr. Buckendahl's recommendation was informed by historical performance. As he explained at the CBE meeting, if the test takers who took the last three February bar examinations (2023, 2024, 2025) possessed equivalent abilities, then, for the February 2025 bar examination, first-time test takers would be expected to correctly answer 122 or more multiple-choice questions to pass the examination and repeat test takers would be expected to correctly answer 120 or more multiple-choice questions to pass the examination. (Ex. 6 at p. 160.) That baseline for the multiple-choice section would replicate the pass rate from the two previous years if the pass rate were based on the multiple-choice questions only (45 percent pass rate for first-time test takers and 29 percent pass rate for repeat test takers) and correspond to the lower bound of the recommendation of the standard validation panel for the multiple-choice questions. Thus, based on the distribution of first-time and repeat test takers for the examination, the baseline for

the calculation of the raw passing score would be 120 correct

responses for the 171 scored items.

The CBE considered the recommendations by

Dr. Buckendahl and the standard validation panel. However, due

to the pervasive and significant reported difficulties faced by test

takers, and considering all data available to them about test-

taker performance, the CBE recommended further reducing

Dr. Buckendahl's recommended passing score of 120 to 114

correct responses out of the 171 scored questions, which is two

standard errors of measurement below Dr. Buckendahl's

recommendation. Dr. Buckendahl explained that, approximately

95 percent of the time, test takers score within two standard

errors of measurement of their observed score.[21]

When combined with the raw passing score for the written

portion of the examination, as further explained below, the CBE's

---

[21] A standard error of measurement reflects that probability that
a test taker's true score falls within a given range of scores. This
means that if the same test taker took the same examination
multiple times, they might not obtain the same score each time.
A standard error of measurement reflects how much the scores
are likely to vary around the test taker's true score.

recommendation reflects a reduction of two standard errors of measurement lower than what Dr. Buckendahl recommended in total.[22] (See Ex. 6.)

In addition to receiving the results of the standard validation panel's review of the multiple-choice questions, the CBE also received information on the standard validation panel's review of the written portion of the examination. The standard validation panel determined that the raw passing score for the written portion of the examination should be 518 out of 700 points. Were such a raw score utilized, it would result in only a 1.8 percent pass rate if only the written component were utilized in grading.

At the April 18 CBE meeting, given that the standard validation panel's recommendation would have resulted in an

---

[22] Because the multiple-choice questions on one examination may have a different average level of difficulty than the questions tested on another examination, a raw score may not signify the same level of proficiency across different administrations of the examination. This is addressed by equating, which adjusts for differences in the difficulty of the multiple-choice questions across administrations. (See https://www.calbar.ca.gov /Admissions/Examinations/California-Bar-Examination/Scaling.)

exceptionally low pass rate, Dr. Buckendahl instead recommended that the CBE link the multiple-choice question expectations to the written section. This approach is more consistent with past practice of using the MBE to scale the written section of the examination. Specifically, Dr. Buckendahl recommended setting the baseline raw score for the written portion of the examination at 420 by linking it to the pass rate that would result if the raw passing score for the multiple-choice component was set at 114—the raw score the CBE recommended—to establish a comparable expected performance. In other words, to correlate the pass rate for the written component to the multiple-choice component, Dr. Buckendahl recommended that the CBE select 420 raw points as the raw passing score for the written component.

The 420 figure was derived as follows: (1) to align with the pass rate of 45 percent for first-time examinees and 21 percent for repeat examinees, Dr. Buckendahl initially recommended a raw score on the written component of 445 and 440, respectively, or an overall baseline of 440 points; and (2) reducing that recommendation by two standard errors of measurement, in

alignment with the CBE's recommendation for the multiple-choice questions, yields a minimum raw passing score of 420. (See Ex. 6 at p. 161.)

The CBE adopted the recommendation that the raw score for the written component be 420. (Ex. 7.)

*   *   *

Accordingly, following the CBE's action at its April 18 meeting, the State Bar requests that the Supreme Court adopt 534 as the minimum raw passing score for the February 2025 bar examination.[23]

---

[23] If the determinations of the standard validation panels were adopted by this Court, the minimum raw passing score would be 647 (129 for the multiple-choice questions and 518 for the written component). If Dr. Buckendahl's initial recommendation—which sought to replicate historical performance—were adopted, the minimum raw passing score would be 560 (120 for the multiple-choice questions and 440 for the written component). If Dr. Buckendahl's secondary recommendation, one standard error of measurement below 560, the minimum raw passing score would be 547 (117 for the multiple-choice and 430 for the written component).

B. The CBE's Proposed Raw Passing Score is Recommended for the February 2025 Bar Examination Only

When the State Bar sought modifications to the California Bar Examination in October 2024, it anticipated that the CBE would set a new raw passing score for the February 2025 bar examination and that future bar examinations would be statistically equated to the expectations of the February 2025 bar examination to maintain consistency of interpretation. However, due to the myriad reported issues that occurred during the February 2025 bar examination, the CBE's recommendation for the raw passing score is specific to the February 2025 bar examination. Dr. Buckendahl has recommended convening standard validation panels following the July 2025 bar examination so that the CBE may adopt a recommended raw passing score for that examination and all future examinations instead of using the February 2025 raw score as a baseline.

The recommended raw passing score for February 2025, while lower than prior administrations, is psychometrically sound for measuring minimum competence. Prior to and during the April 18, 2025, meeting, CBE members were also provided

57

with data regarding multiple-choice question performance, including performance broken down by question source (Ex. 4 at pp. 114–116), historical test-taker performance on the essay and performance test portion of the examination (Ex. 6 at pp. 167–169), estimated comparative data regarding the multiple-choice portion of the examination, and available data related to reported test-taker experiences. (Ex. 5 at pp. 129–145).

The multiple-choice question performance data indicated that the internal consistency reliability estimate for scored items on the multiple-choice section of the examination yielded a coefficient alpha of 0.89 on a scale of 0.00 to 1.00 where values of higher than 0.80 are desired. (Ex. 4 at p. 113.) In addition to overall reliability, the multiple-choice questions performed well on other metrics including item discrimination (with an average of 0.18 when the target is 0.10 or greater). As noted above, questions that adversely impacted examination reliability were not scored, increasing overall reliability. (See Section II.B.4, *ante.*) The historical test-taker data on the essay and performance test portion of the examination reflected that the average first-read total raw scores as determined by the graders

the written portion of the February 2025 bar examination was approximately ten points *higher* than either the average first-read total raw scores for the written portions of the February 2023 and February 2024 bar examinations. (Ex. 6 at p. 167.) The data suggest that the examination appropriately measured minimum competence and that February 2025 test takers *performed in alignment with or better* than prior February bar examination administrations.

VI.   IMPROVEMENTS PLANNED FOR BAR EXAMINATION EXPERIENCE FOR JULY 2025

As described in Section II.B, *ante*, in developing the multiple-choice questions for the February 2025 bar examination, the State Bar subjected all multiple-choice questions to content validation and subject-matter expert review. However, the State Bar has identified process improvements that could improve examination integrity and question performance for future administrations. As part of process improvements for the July 2025 bar examination and beyond, the State Bar, with direction and assistance from the CBE, intends to introduce another layer of subject-matter expert review for the multiple-choice questions.

Specifically, all questions would be reviewed by two subject-matter experts to weigh in on the accuracy of the question, and if those two experts have diverging views on a question, a third subject-matter expert would provide a tie-breaker review. Admissions staff sent a solicitation to the deans of all law schools in California, requesting that they forward the solicitation to all faculty. In addition, at the request of Admissions' staff, the solicitation was posted to a listserv for the Association of American Law Schools. In just over one week, more than 110 law school faculty applied to serve as a subject-matter expert. The State Bar also plans to send the solicitation to retired judges and justices to further expand the pool of possible subject-matter experts.

As part of this effort, the CBE's agenda for its meeting on May 5, 2025, calls for the CBE to review the eligibility criteria for subject-matter experts and take action on establishing a policy for their recruitment and selection. After adopting criteria, the CBE leadership recommend that the CBE be involved in the selection of the experts. The State Bar is working closely with the CBE leadership on an aggressive timeline to institute these

changes, but following the adoption of the new policy, ample time is required to do the following:

- Select applicants with the input of the CBE;

- Confirm eligibility of selected applicants based on the CBE's adopted criteria;

- Perform attorney complaint and disciplinary history checks on selected applicants (which is anticipated to be part of the CBE's selection policy);

- Execute contracts;

- Securely transmit the questions to the subject-matter experts; and

- Provide ample opportunity for subject-matter experts to review the questions for legal accuracy.

The State Bar has also committed to ensuring that all multiple-choice questions are subjected to professional copyediting as part of the final step in the review process.[24]

---

[24] The CBE will also be asked to adopt a policy for recruitment and selection of content validation and standard validation panelists going forward. It is anticipated that this policy would result in a broader solicitation and applications would be

Up until this point, neither the State Bar nor the CBE have considered returning to the MBE. But as the State Bar continues to work with the CBE to improve the multiple-choice question review process, this Court may conclude that, pursuant to its plenary authority over admissions to the bar in this State, the State Bar should be directed to utilize the MBE for the July 2025 General Bar Examination so that there is not a risk that the process improvements are not effectively implemented before the next administration of bar examination.[25] The Court's Administrative Order 2024-10-21-01, filed on October 22, 2024, which refers only to 200 multiple-choice questions, does not require amendment for the Court to make this directive.

---

reviewed in accordance with that policy in advance of the content validation panels scheduled to convene in August.

[25] If directed to utilize the MBE in July 2025, the State Bar would be required to purchase the MBE, which has a per applicant price of $72. The State Bar currently estimates approximately 10,000 applicants will sit for the the July 2025 examination, and thus estimates the cost of utilizing the MBE as $720,000. The State Bar anticipates contracting with Examsoft to administer the multiple-choice and written portions of the July 2025 bar examination. If the State Bar is directed to utilize the MBE, which is not administered via Examsoft, the per-applicant Examsoft fee may be reduced by $10 per applicant.

VII. CONCLUSION

For the foregoing reasons, in accordance with the CBE's recommendations at its April 18, 2025, meeting, the State Bar respectfully requests that the Court issue an administrative order approving a raw passing score for the February 2025 California Bar Examination of 534 and imputation of scores as described herein, and as set forth in the attached proposed order.

Dated: April 29, 2025   Respectfully submitted,

ELLIN DAVTYAN
KIRSTEN GALLER
JEAN KRASILNIKOFF

By: */s/ ELLIN DAVTYAN*
ELLIN DAVTYAN

General Counsel
Office of General Counsel
The State Bar of California

IN THE SUPREME COURT OF CALIFORNIA

EN BANC

ORDER APPROVING RAW PASSING SCORE AND SCORING ADJUSTMENT FOR THE FEBRUARY 2025 CALIFORNIA BAR EXAMINATION

The Court is in receipt of the State Bar of California's Request to Approve a Proposed Raw Passing Score and Scoring Adjustments for the February 2025 California Bar Examination, filed on April 29, 2025. The Court, having considered the State Bar's request, sets the raw passing score for the February 2025 California Bar Examination only as 534. The Court also approves psychometric imputation of scores as follows: for missing multiple-choice answers, where the test taker answered at least 114 of the 171 scored multiple-choice question and for missing essay or performance test answers (including 18 test takers who had content in the performance test response field, but did not

have access to the file and library), where the test taker

answered at least four of six written sections of the examination.

_____

Chief Justice