Information Regarding the Grading of the February 2004 California Bar Examination And Two Incidents That Occurred During the Examination

Two incidents occurred during the February 2004 administration of the California Bar Examination that impacted some of the applicants who took the examination. One incident stemmed from the heavy rains in southern California on Wednesday evening that led to flooding of the Pasadena laptop test center. As a result of this flooding, the start of the Thursday morning session at this center was delayed and the afternoon session (during which Performance Test B (PT-B) was scheduled to be administered) had to be cancelled. This cancellation occurred because several hours were needed to return the Pasadena test center to examination quality status, which necessitated a very late commencement of the morning session, and that delay did not leave enough time to administer the afternoon session.

The second incident also occurred on Thursday morning when it was discovered that the laptop users' computers appeared to stall while saving their answers to essay questions 4, 5, and 6. This stalling occurred at the Pasadena and the other laptop test centers. The stalls became longer toward the end of the test session. The length of the stalls also was a function of the applicants' editing styles and their computers' hardware characteristics. Consequently, some applicants had longer total stall times than did other applicants.

Immediately following the conclusion of the examination, the Committee of Bar Examiners (Committee) consulted with several experts in testing and statistics to investigate what effect the delay and stalls may have had on applicant scores. The experts also were asked to recommend the best way to impute PT-B scores for the applicants who had their Thursday afternoon PT-B session cancelled.

The team of experts analyzed the examination data and found that the laptop users' mean score on an essay question was about one point higher than the non-laptop users' mean. This was true on Tuesday when there was no stalling and on Thursday when there was stalling. In addition, both the laptop and non-laptop users had a three-point higher average essay score on Tuesday than on Thursday. In short, after controlling for the difference in average question difficulty between days, the laptop users did just as well when the stalls occurred as when they did not occur. Thus, the stalls did not appear to affect the laptop users' scores.

Further analyses revealed a slight (and not statistically significant) positive relationship between Thursday morning scores and how much stall time a candidate encountered; i.e., the laptop users' scores tended to *increase* very slightly as their stall times increased. The only exception to this trend was with candidates who averaged more than five minutes of stall time per hour. Their Thursday morning scores decreased slightly as their average hourly stall times increased. To adjust for this decrease, the laptop users had 0.01 points

added to their essay question 6 score for each second of average hourly stall time they experienced in excess of 300 seconds (five minutes). For example, an applicant who experienced an average of 400 seconds of stall time per hour of testing time received a total of (400-300) x 0.01 = 1 additional point on the essay section. This adjustment was based on a regression analysis that predicted an applicant's score on essay questions 4–6 on the basis of that applicant's stall time and (Multistate Bar Examination) MBE, essays 1–3, and PT-A scores.

The analysis team also found that the applicants who switched from using their laptops to hand writing during the Thursday morning session (when the stalling occurred) did just as well in this session as would be expected given their Tuesday morning scores (when the stalling did not occur).

As noted above, the flooding at the Pasadena laptop test center delayed the start of the Thursday morning session at this site. A regression analysis using the same variables as discussed above found that this delay did not adversely affect scores at this test center. However, the delay did force canceling the administration of the PT-B task that was scheduled to be administered on Thursday afternoon.

The analysis team examined two methods (regression and pro-rata) for imputing PT-B scores for the Pasadena applicants who had this session cancelled. Their studies found that these methods yielded identical passing rates. The two methods also agreed almost perfectly in which candidates they would pass. In addition, the pro-rata method can be used with those taking the Attorneys' Examination as well as those taking the General Bar Examination and it is consistent with the procedures the Committee used several years ago to estimate missing scores when an earthquake led to canceling the afternoon portion of the multiple-choice section of the examination at one test center. Taken together, these considerations led the Committee to adopt the pro-rata method.

The pro-rata method consists of multiplying an applicant's grader-assigned PT-A score by 2, adding this product to the sum of an applicant's scores across essay questions 1 through 6, and then dividing the sum of these scores by 7.86. The 7.86 scaling factor in this equation adjusts for differences in mean raw scores among the different parts of the written (essay + PT) portion of the examination.

For example, a Pasadena laptop user with a grader-assigned raw score of 55 on PT-A would have this score multiplied by 2. If this applicant earned 415 raw score points on essays 1–6, then this applicant would have an imputed PT-B score of 67 because [(2 x 55) + 415]/7.86 = 67. For score reporting purposes, all estimated values are rounded to the nearest whole number.

The Committee regrets that these incidents occurred, but it believes that it has taken appropriate steps to adjust the scores of those who were impacted by them.

resultsnotice-1