THE TESTING COLUMN
THE PURPOSE OF THE BAR EXAM:
TO PREDICT FUTURE SUCCESS AS A LAWYER?

by Susan M. Case, Ph.D.

Last week, I had a conversation with someone who was concerned that research on the bar exam had not focused on predictive validity (i.e., the extent to which bar exam scores predict future success as a lawyer). I told her that the central purpose of the bar exam (and other high-stakes exams used for licensure and certification) is to identify examinees who are sufficiently competent to practice in the area covered by the license. In other words, the bar exam strives to assess the extent to which examinees demonstrate their ability to apply their knowledge and skills to fact patterns and cases presented to them at that time; it does not attempt to predict their performance in the future. She was not convinced.

In an ideal world, everyone who passed the bar exam would practice for an entire career without a single mistake due to a lack of knowledge or skills. And everyone who failed the bar exam would have performed below this level if they had been allowed to practice. Obviously, the first scenario is unreasonable, and there is no way to determine if the second scenario would occur. Demonstrating sufficient competency on the bar exam does not guarantee success in practice, but lack of competence would have a deleterious effect on practice, putting clients and the general public at risk.

This position regarding the purpose of licensing exams (more broadly referred to as credentialing exams) is not just my personal belief. This is the view of testing professionals in general. The STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING, a set of testing guidelines developed by the three associations of testing professionals, is the generally accepted definitive authority on professional and technical issues in test development and use. According to the STANDARDS, credentialing tests are designed to determine whether the essential knowledge and skills of a specified domain have been mastered by the candidate.

In contrast, in certain employment tests and in tests used for admission into college or law school, such as the SAT and LSAT, the purpose is to predict subsequent performance in employment or in school. In these instances, there are rather clear criterion measures that serve to quantify the extent to which the examinee is successful. Such a criterion might be first-year law-school grades in the case of the LSAT, or job performance ratings in the case of an employment test. In order to use test scores as predictors of future performance, there must be a method of evaluating the future performance of those who pass the examination.

There are a number of reasons why predicting future success as a lawyer is not a reasonable goal for the bar exam.

1. **In order to evaluate competence in practice, there must be useful, measurable criteria that define and measure competence.**

    Competence must be defined; criteria must be specified; the criteria must be measurable; and the scores must be valid, reliable, and fair.

    The definition of competence and adequate criteria must apply to practitioners in all settings. Some common measures of success such as income levels and awards can be measured but are clearly not appropriate. Other criteria that would be appropriate, reflecting the quality of services provided to clients and the public, are difficult to define and impossible to measure in a standardized fashion.

    One of the complications of specifying criteria of competence for the entire population of newly licensed lawyers is that, in order for these criteria to be widely applicable, they have to be quite general. This is a double-edged sword: the more general the criteria, the harder they are to quantify; and the harder they are to quantify, the more subjective they become.

    General criteria that are widely applicable tend to be evaluated based on ratings; ratings are problematic because it is difficult to find appropriate raters and to ensure that the raters are unbiased and are equally stringent or lenient.

2. **Only passing candidates are in practice; failing candidates cannot be evaluated.**

    Determining whether or not the bar exam differentiates between those who should be admitted and those who should not requires that both be included in the analysis. Predictions about passing candidates could in principle be evaluated if suitable performance criteria were specified, but predictions about failing candidates cannot be examined at all.

3. **The results would be difficult to interpret.**

    Even if good measurable criteria were available, a poor result could reflect a problem with the pass/fail standard rather than with the exam itself. If too many poorly qualified applicants are being licensed, the problem could be that the passing standard is too low. Alternatively, if all newly licensed lawyers greatly exceed that minimum standard, perhaps the passing standard is too high.

4. **Lawyers who fail to perform adequately may have the requisite knowledge and skills, but demonstrate poor performance due to other factors.**

    Factors that affect performance include personality and behavioral traits such as disorganization, inability to meet deadlines, drug or alcohol abuse, and dishonesty.

    The bar exam itself is not expected to weed out examinees who have the requisite knowledge and skills but who have other performance or character flaws. It should also be noted that these characteristics may change over time; behavioral problems that were not present at the time of the bar exam could manifest themselves later as stress levels and circumstances change.

If we do not evaluate the validity of bar exam scores by examining the success of licensed lawyers, how do we evaluate their validity? Evaluating the validity of test scores does not generally involve statistical analysis but rather involves reviewing a collection of evidence that is indicative of what the test scores reflect. One appropriate form of evidence review focuses on the examination content—both the content specifications and the questions themselves. This review might involve collecting judgments from various individuals ranging from experts in the field to newly licensed practitioners about the extent to which the content of the examination and the style of the questions are directed at assessing the knowledge and skills needed by a new

practitioner. A second area of investigation might focus on the extent to which extraneous factors, including a broad range of issues from handwriting to time limits, impact examination scores. Typically, these studies would be designed to ascertain that these extraneous factors are absent from consideration during the scoring process.

Using a variety of techniques, NCBE conducts ongoing evaluations of the quality of the questions on the bar exam. For each test, we evaluate the reliability of the scores; reliability is an essential component of validity. We also focus on the quality of each question. All questions are written by subject-matter experts and reviewed by committees with responsibility for generating the tests; none of our questions for any exam are written by staff. In addition, each MBE question is subjected to an evaluation of the appropriateness of the task posed and the credibility of the scenario in the question. This evaluation is done twice a year by an outside reviewer (a professor not otherwise connected with the exam) and a practitioner member of the MBE Policy Committee. The MPRE, MEE, and MPT exam questions undergo independent review by subject matter experts who were not involved in writing or otherwise reviewing the test questions. The MEE and MPT questions are pretested by newly licensed lawyers who take the questions under proctored test conditions. Their answers are graded, and the grading guidelines are reviewed to ensure that the resultant scores are appropriate in terms of difficulty and discrimination. The lawyers who pretest the questions also complete an evaluation of the questions regarding their clarity and fairness.

NCBE also conducts research regarding the examination. The MPRE includes a questionnaire that collects information from examinees regarding their preparation for the test, the quality of the test administration, and other topics such as the pacing of the test. We analyze scores from all the tests and look at whether they are reasonable. We know, for example, that performance on each test is correlated in reasonable ways not only with LSAT scores but also with law school grades. We also know that performance on each section of the exam is related to performance on other sections of the exam in expected ways (MPRE scores are correlated with MBE, MEE, and MPT scores; MBE scores are correlated with MEE and MPT scores; MEE and MPT scores are also correlated.) We have looked at gender and ethnic performance; these results are consistent with those found in other professions. These and other research results form the basis for many articles and Testing Columns included in THE BAR EXAMINER.

The purpose of the bar exam is to protect the public against risks that would be posed by lawyers who lack the knowledge and skills expected of entry-level practitioners. Validity is an assessment of the extent to which exam scores measure competence in the kinds of issues and problems encountered by newly licensed lawyers in practice. In order to evaluate the validity of exam scores, we undertake evaluation of the scores and amass information about exam performance.

We realize that we do not measure every competency and skill expected of newly licensed lawyers. We evaluate performance on a sample of these competencies and make an inference that the scores would generalize to a broader set of competencies. Clearly, we could add topics to the exam—in particular, to the MBE—that would broaden the sample that we are assessing. As noted in previous BAR EXAMINERs, Civil Procedure and Legal Research Skills have been suggested as additional topics. No plans have been made to begin work on adding any topic, but conversations continue. ▣

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