

Scholarly Commons @ UNLV Boyd Law

Scholarly Works

Faculty Scholarship

9-2018

New York Leads from the Middle: Crowdsourcing the Bar Exam Cut Score

Joan W. Howarth

University of Nevada, Las Vegas – William S. Boyd School of Law

Follow this and additional works at: <https://scholars.law.unlv.edu/facpub>



Part of the [Legal Education Commons](#), and the [Legal Profession Commons](#)

Recommended Citation

Howarth, Joan W., "New York Leads from the Middle: Crowdsourcing the Bar Exam Cut Score" (2018).
Scholarly Works. 1215.

<https://scholars.law.unlv.edu/facpub/1215>

This Article is brought to you by the Scholarly Commons @ UNLV Boyd Law, an institutional repository administered by the Wiener-Rogers Law Library at the William S. Boyd School of Law. For more information, please contact youngwoo.ban@unlv.edu.

New York Leads Crowdsourcing the Bar

By Joan W. Howarth

Most lawyers prefer to forget about bar exams as soon as they have cleared the hurdle, but attorney licensing deserves our attention. New York recently adopted the Uniform Bar Exam (UBE), now approved by 34 jurisdictions. The time is ripe to consider the fact that even UBE bar exams are much harder to pass in some jurisdictions and easier to pass in others. Despite its name, the Uniform Bar Exam does not address this longstanding bar exam incongruity. Oddly, the main reason that bar exams are harder or easier from state to state is that states choose different passing scores on the one part of the test that is identical across the country, the multiple choice questions, also known as the Multistate Bar Exam (MBE). Candidates across the country are answering the same MBE multiple-choice questions to establish the same proposition – minimal competence to practice law – but are measured by different passing scores on that single test. Differences in difficulty based on different numbers of subjects tested or intricacies of local law could be easier to justify.

The range of MBE passing scores is dramatic. Cut scores extend from 129 in Wisconsin to 145 in Delaware.¹ The states with the most licensed attorneys, New York and California, use MBE cut scores of 133 and 144 respectively. Not even Californians pretend that these disparities are justified on the grounds that practicing law as a new lawyer is more difficult in California than in New York. It isn't. Our wide state-to-state spread in MBE cut scores is an artifact of an earlier era of jurisdictional isolation and less reliable bar exams.

Defensible attorney licensing today requires moving to a uniform cut score for multiple reasons. First, our current MBE cut score disparities constitute bad logic because

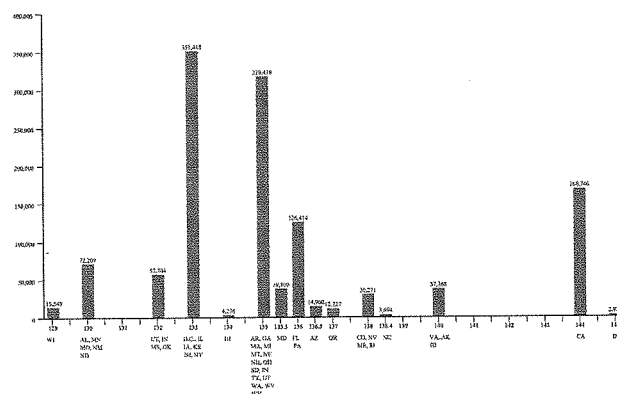
every state is attempting to predict exactly the same thing: minimum competence to practice law, by using the same test but with different passing scores. Second, the cut score disparities are bad science because precision in setting a cut score is critical in assuring the validity of the use of the exam. Third, MBE cut score disparities are also bad policy, which explains why other professions have moved to uniform cut scores. Law, too, needs to arrive at a consensus, a uniform cut score. But, how?

Crowdsourcing follows the wisdom of the multitudes. We now routinely use the crowd to determine our driving routes and to support charitable causes. Following the crowd also offers the best way to eliminate bar exam cut score disparities. Based on attorney population, the leading MBE cut score is 133, now used by New York, New Jersey, Illinois, Kansas, Iowa, and the District of Columbia.² Attorney licensing will be more valid across the country when other jurisdictions follow the crowd of licensed attorneys now figuratively gathered at an MBE cut score of 133.

To truly appreciate the problem and embrace this solution, lawyers need to set aside their natural reluctance to think back to bar exams, and, probably for the first time, learn how bar exams are scored.

FIGURE 1

MBE Cut Scores and Lawyer Populations



Joan W. Howarth is a Distinguished Visiting Professor at Boyd School of Law, University of Nevada, Las Vegas, and Dean Emerita, Michigan State University College of Law. This article is based on Joan W. Howarth, *The Case for a Uniform Cut Score*, 42 J. of Leg. Prof. 69 (2017).



from the Middle: Exam Cut Score

PSYCHOMETRIC BASICS FOR LAWYERS

Many of us distrust multiple-choice tests, but multiple-choice tests like the MBE are ubiquitous throughout professional licensing because of the science of testing. Two touchstones for standardized tests are validity and reliability. Validity means that the test measures what it purports to measure, here minimal competence to practice law. Reliability is the extent to which a score on the test means the same thing even when the test is given at different times.³ Every profession uses multiple-choice tests in licensing to enhance reliability, such that a score from February means the same thing as a score from July, for example. Essay questions are too memorable to be used more than once, but multiple choice questions can be repeated. High-stakes multiple-choice tests typically include some repeat questions, whose degree of difficulty is already known. Psychometricians compare how test takers do on the repeat and the new questions, using the scores on the repeat questions to determine the degree of difficulty of the new questions, and then of the entire test. These complex statistical processes convert *raw* multiple-choice scores to *equated* scores that can be compared from February to July and from year to year. This equating process is the first of two big psychometric steps focused on reliability.

The second big step, scaling, makes the MBE multiple-choice scores crucial for determining how many candidates pass the *entire* bar exam, including the essays. Scaling essentially anchors the essay and performance test scores to the multiple-choice scores. Essay grades are notoriously unreliable because the questions change and grading is more subjective. Scaling uses the greater reliability of the equated multiple-choice score to improve the reliability of scores from less objective parts of the test, such as essays and performance tests. The subjectivity in essay grading is less problematic in law school, because law professors generally grade essays by comparing each student's essay to his or her classmates.'

Reliability on a licensing test, by contrast, means that the score would be the same no matter who else is taking the test. To counter the potential inconsistencies and subjectivities in essay scores, psychometricians use

statistical scaling processes to match, in a way, the raw essay scores to the equated multiple-choice scores.⁴

In other words, the MBE scores anchor the scores for all components of the exam. Scaling means that the number of candidates who pass the MBE may also determine how many pass the essays. Scaling makes the MBE cut score a crucial decision concerning the degree of difficulty of the *entire* exam, a much bigger impact than what one would expect.

Other professions use the same equating and scaling practices for the same reasons. Multiple-choice test scores anchor state-specific test components, such as essays or performance tests, for doctors, engineers, nurses, and others. Nurses and engineers first adopted uniform cut scores for their multiple-choice licensing tests in the 1980s. Today, doctors, nurses, dentists, veterinarians, physical therapists, engineers, surveyors, architects, certified public accountants, mortgage loan originators, psychologists, emergency medical technicians, social workers, and real estate appraisers include a national multiple-choice test *with a uniform cut score* as a requirement for state licensure.

FIGURE 2



Longstanding habits of state control are not easily set aside, but architects, social workers, dentists, and other professions have overcome these impediments. We should too.

MBE CUT SCORE DISPARITIES ARE BAD LOGIC, BAD SCIENCE, AND BAD POLICY

Other professions moved to a uniform cut score in part because of the flawed logic of attempting to use the same pass-fail test to measure minimum competence between jurisdictions but setting the passing score at different levels. Nurses, doctors, and social workers do not gain or lose minimum competence by crossing state lines any more than lawyers. The difference is that the nurses, doctors, social workers, engineers, veterinarians, dentists, accountants, and other professions have given up the illogical pretense that minimal competence – *as measured by the same multiple-choice test* – changes from state to state.

The science of testing, psychometrics, supports finding a uniform cut score. The cut score is aimed at the dividing line that separates test takers with minimal competence from test takers who are *barely* below minimal competence. Therefore, not surprisingly, psychometric standards for any high stakes pass-fail test require great care related to the passing score. By adopting uniform cut scores, other professions have taken seriously these fundamental psychometric principles meant to ensure validity of the pass-fail test, meaning that the test actually does what it purports to do.

A uniform cut score is also better policy. The current practice of each state setting its own MBE cut score prevails only because most states do not approach the task in the way that testing standards require. Bar examiners in many states (other than New York) have no idea how their state's MBE cut score was established. This longstanding mystery is directly contrary to professional norms for transparency in licensing tests. Transparency is crucial to counter potential, perceived, or actual conflicts of interest, or anti-competitive behavior, when a profession is setting the bar for new entrants to the profession.

State licensing decision makers in other professions relegate standard-setting to national entities because the process is burdensome and too difficult for states to do well. For example, cut scores of licensing tests must be reviewed periodically. Nurses review their multiple-choice cut score every three years; engineers and physical therapists every five years. With all the other fiscal and operational pressures on state courts and bar examiners, routinely reevaluating cut scores is not a priority.

THE PATH FORWARD

The MBE cut score disparity problem will be addressed by moving to a consensus middle ground. Protection of the public is the touchstone: errors in either direction hurt the public. Setting the bar too low risks licensing attorneys who lack minimal competence. Setting the bar too high risks depriving the public of competent attorneys. Access to justice is implicated if competent attorneys are prevented from practicing, in part because fewer attorneys

may mean increased costs for legal services. Setting the bar too high also has a disproportionate impact on competent attorneys of diverse racial and ethnic backgrounds.⁵

Licensing cut scores usually are difficult to evaluate in part because the professional performance of candidates with scores below the cut score – who do not receive the license – cannot be assessed. But our current MBE cut score variation creates a massive natural experiment. Do states with lower cut scores suffer from less competent attorneys? No evidence supports that conclusion. In the absence of data identifying harms to the public in jurisdictions with lower cut scores,⁶ states should follow the crowd. Arriving at a middle-ground consensus will require states with very low cut scores to move up, and states with very high cut scores to move down. New York's cut score, 133, is a prime candidate for consensus because it is the score currently being used by jurisdictions with the largest total attorney population.

States should cherish their authority over attorney licensing, including their opportunity to provide meaningful public protection in innovative ways. Attorney licensing is, indeed, ripe for innovation.⁷ States should be asking, what is minimum competence to practice law? How do we best protect the public? New York recently added pro bono and experiential experience requirements. But resting the case for state autonomy on setting a different cut score on the common, national portion of the exam is illogical, unfair, unambitious, and harmful. The public deserves valid licensing tests; eliminating MBE cut score disparities would be an important step in that direction. Crowdsourcing suggests that New York has set its cut score in the right place. New Jersey, Illinois, Iowa, Kansas, and the District of Columbia are already on board. Attorney licensing across the country will be more rational when others states follow.

1. Nat'l Conf. of B. Examiners, Comprehensive Guide to Bar Admissions 2018, at 33-34, Chart 9, available at <http://www.ncbex.org/pubs/bar-admissions-guide/2018/mobile/index.html>.

2. Figure 1 is derived from ABA National Lawyer Population Survey 2017, https://www.americanbar.org/content/dam/aba/administrative/market_research/National%20Lawyer%20Population%20by%20State%202017.authcheckdam.pdf.

3. Susan M. Case, *Back to Basic Principles: Validity and Reliability*, B. Examiner 23, Aug. 2006 at 23.

4. See Susan M. Case, *Frequently Asked Questions About Scaling Written Scores to the MBE*, B. Examiner 41 (Nov. 2006); see also Susan M. Case, *Demystifying Scaling to the MBE: How'd You Do That?*, B. Examiner (May 2005), at 45-46).

5. See Stephen P. Klein & Roger Bolus, *The Size and Source of Differences in Bar Exam Passing Rates Among Racial and Ethnic Groups*, 6 B. Examiner 8, 8 (Nov. 1997); Deborah Jones Merritt, *Validity, Competence, and the Bar Exam*, AALS News, Spring 2017, <http://www.aals.org/about/publications/newsletters/aals-news-spring-2017/faculty-perspectives/>; Tracy A. Montez, *Observations of the Standard Setting Study for the California Bar Examination 10* (Calif. Dept. of Consumer Affairs, July 2017), <http://www.calbar.ca.gov/Portals/0/documents/admissions/Examinations/Tracy-Montez-ReviewBarExamstudy.pdf>. Recent California modeling shows that the higher the studied cut score, the larger the racial disparities in pass rates. *Final Report on the 2017 California Bar Exam Standard Setting Study* (State Bar of Cal. Sept. 12, 2017) [hereinafter *Calif. Report*] at Appendix I, available at <https://www.calbar.ca.gov/Portals/0/documents/communications/CA-State-Bar-Bar-Exam09122017.pdf>.

6. "There is no empirical evidence that indicates that California lawyers are more competent than those in other states [with lower cut scores]." *Calif. Report*, *supra* note 5, at 17-18.

7. See, e.g., Eileen Kaufman, Andi Curcio, & Carol Chomsky, *A Better Bar Exam – Look to Upper Canada?*, Law School Café (July 25, 2017), <https://www.lawschoolcafe.org/2017/07/25/a-better-bar-exam-look-to-upper-canada/>.

