

AccessLex

AccessLex Resource Collections

Grantee Research

AccessLex-Funded Research

4-2023

Putting the Bar Exam to the Test: An Examination of the Predictive Validity of Bar Exam Outcomes on Lawyering Effectiveness

Jason M. Scott

Stephen N. Goggin

Rick Trachok

Jenny S. Kwon

Sara Gordon

See next page for additional authors

Follow this and additional works at: <https://arc.accesslex.org/grantee>



Part of the [Legal Profession Commons](#)

Authors

Jason M. Scott, Stephen N. Goggin, Rick Trachok, Jenny S. Kwon, Sara Gordon, Dean Gould, Fletcher S. Hiigel, Leah Chan Grinvald, and David Faigman

Putting The Bar Exam to the Test: An Examination of the Predictive Validity of Bar Exam Outcomes on Lawyering Effectiveness

WORKING PAPER
(Last Updated March 1, 2023)

Jason M. Scott¹, Stephen N. Goggin², Rick Trachok³,
Jenny S. Kwon⁴, Sara Gordon⁵, Dean Gould⁶,
Fletcher S. Hiigel⁷, Leah Chan Grinvald⁸, David Faigman⁹

ABSTRACT

How well does bar exam performance, on the whole, predict lawyering effectiveness? Is performance on some components of the bar exam more predictive? The current study, the first of its kind to measure the relationship between bar exam scores and a new lawyer's effectiveness, evaluates these questions by combining three unique datasets—bar results from the State Bar of Nevada, a survey of recently admitted lawyers, and a survey of supervisors, peers, and judges who were asked to evaluate the effectiveness of recently-admitted

¹ CORRESPONDING AUTHOR: AccessLex Institute, jscott@accesslex.org

² Department of Political Science, San Diego State University

³ Nevada Board of Bar Examiners

⁴ University of California College of the Law, San Francisco

⁵ Peter A. Allard School of Law, University of British Columbia

⁶ State Bar of Nevada (retired)

⁷ AccessLex Institute

⁸ William S. Boyd School of Law, University of Nevada, Las Vegas

⁹ University of California College of the Law, San Francisco

lawyers. We find that performance on both the Multistate Bar Examination (MBE) and essay components of the Nevada Bar have little relationship with the assessed lawyering effectiveness of new lawyers, calling into question the usefulness of these tests.

INTRODUCTION

All jurisdictions in the United States require law school graduates to pass a bar examination before being licensed to practice, except Wisconsin.¹⁰ The stated purpose of these exams is to help ensure that only individuals who possess minimum competence to practice law are granted licenses to do so. Bar exams are thus broadly understood as protecting the public from incompetent lawyers. Given this purpose, the primary utility of bar exams should depend on whether exam scores distinguish the competent from the incompetent. Yet, for many people who have taken a bar exam and then practiced law, the relationship between scores and competence appears weak.

Bar exams principally test memorization of legal doctrine, procedural knowledge, analytical reasoning, essay writing, and time management. Although these skills are related to law practice, they are but a small part of what it means to be a competent attorney.

¹⁰ Wisconsin does not require a bar exam for graduates of either law school in the state: University of Wisconsin Law School and Marquette University Law School. Law school graduates from other states wanting to practice law in Wisconsin must pass the Wisconsin Bar Exam, have passed one in a state with which it has an agreement of reciprocity, or be granted admission on proof of practice.

The bar exam also involves, to some degree, assessing test-taking abilities, but these are not requisite skills for being a competent attorney. Although test-taking strategies can be (and often are) taught, some individuals seem to be innately better at taking tests—no matter whether this can be attributed to genetic or environmental factors, or some combination of the two (Plomin, 2023; Plomin et al., 2014; Wadsworth et al., 2002).

Despite their widespread use, the key question of whether bar exams are useful indicators of minimum competence has never been tested, a void this study helps to fill. In the absence of more relevant evidence, proponents of bar exams point out that scores are highly correlated with law school grades and even LSAT scores (Anderson & Muller, 2019; Austin et al., 2017; Kuehn & Moss, 2019). But comparing LSAT scores and law school grades with bar exam scores merely indicates a relationship between test-taking abilities and academic performance, with the latter also largely related to test-taking abilities. Such comparisons are not conclusive of whether scores are measuring lawyering effectiveness. A comparison between performance on the bar exam and performance as an early career lawyer would be most relevant to the issue of effectiveness.

Recognizing the primacy of this comparison, our multi-institutional collaboration was designed to be the first study to assess the relationship between bar exam scores and lawyering effectiveness.¹¹ The multifaceted nature of the

¹¹ This study is a collaboration between the State Bar of Nevada, the Nevada Supreme Court, AccessLex Institute, and the law schools of University of Nevada, Las Vegas and University of California College of the Law, San Francisco (formerly UC Hastings).

collaboration meant that we had access to critical data, including candidates' scores on the Nevada bar exam from 2014 through February 2020, as well as the capacity to analyze the data and interpret the findings. Additionally, we were the first researcher team granted permission by Professors Marjorie Schultz and Sheldon Zedeck to use their innovative instruments for measuring lawyering skills—developed over decades of researching this issue. In the end, we had access to the raw data of bar exam performance to which we could compare validated criteria for lawyering effectiveness.

COMPONENTS OF THE BAR EXAM

The bar exam varies by jurisdiction within the U.S; however, nearly all include the Multistate Bar Examination (MBE), which was first administered in 1972 in just two jurisdictions (Missouri and North Dakota). Developed by the National Council of Bar Examiners (NCBE), the MBE is a six-hour multiple-choice exam, comprising 200 multiple-choice questions on subjects that are typically taught to law students in their first year of law school.

In 2011, the NCBE launched the Uniform Bar Examination (UBE), which consists of the MBE, the Multistate Essay Examination (MEE), and the Multistate Performance Test (MPT). The MEE comprises six, 30-minute essay questions based on content similar to that which is tested on the MBE. The primary distinction between the MBE and the MEE is that the latter requires the examinee to demonstrate a level of ability to communicate effectively in writing.

The MPT comprises two, 90-minute essays that are intended to assess a student’s ability to “practice law.” The objective of these questions is to evaluate an examinee’s ability to accomplish tasks typical of those performed by beginning lawyers. Unlike the MBE and the MEE, the MPT is not intended as a test of substantive knowledge.

The UBE is uniformly administered across the 40 jurisdictions that have adopted it; however, some jurisdictions may also administer additional jurisdiction-specific components.¹²

There are two possible bar exam performance outcomes: candidates either pass or they fail. The jurisdiction’s “cut score” or minimum passing score distinguishes the two outcomes. Candidates who score at or above the cut score pass the exam and those who score below fail the exam. These cut scores are determined as a matter of policy within jurisdictions; they are most typically reflections of values rather than the products of technically rigorous methods. As a result, there is considerable diversity in cut scores nationally, even among UBE jurisdictions (NCBE, 2022).^{13, 14} This means that cut scores are subject to considerable discussion and debate within the legal profession and the legal academy (cf. Winick et al., 2020a, 2020b).

¹² UBE scores are “portable,” meaning that examinees may transfer scores between UBE jurisdictions in seeking bar admission (NCBE, 2022).

¹³ As of January 2023, the lowest minimum passing UBE score is 260 and the highest is 280.

¹⁴ This inconsistency in cut scores across UBE jurisdictions can affect the portability of exam scores. An examinee with a passing score in one jurisdiction may not be able to transfer their score to a jurisdiction with a higher cut score. Alternatively, an examinee that fails the bar exam in one jurisdiction can transfer their score to a jurisdiction with a lower cut score.

Some commentators argue that cut scores are set too high, resulting in minimally competent bar candidates being excluded from the practice (Howarth 2017; cf. Merritt, 2000; Rosin, 2008; Winick et al., 2020a, 2020b). Others argue that cut scores should be maintained at current levels (Anderson and Muller, 2017; cf. Kinsler, 2017; and Patton, 2020). Another thread of commentary regards the variability in cut scores, with some critics asserting that a nationally uniform score is needed, while others see the diversity of cut scores as reflections of jurisdiction-specific needs and contexts (Howarth, 2017; Rosin, 2008). These arguments assume that the bar exam does indeed predict lawyering effectiveness or minimum competence—yet this has never been validated. Prior to discussing what cut score makes the most sense and for whom, it is imperative to understand whether the exam itself predicts whether a lawyer will be effective or ineffective.

The Nevada Bar Exam

Nevada’s approach to the bar exam is distinctive. The State Bar of Nevada has not adopted the UBE but, during the study period, it made use of the MBE and MPT as two components of its bar exam.¹⁵ In place of the MEE, Nevada administers its own jurisdiction-specific seven-question essay examination. Until 2020, these essay scores were scaled to the MBE and then combined to determine

¹⁵ The Nevada Supreme Court has suspended use of the MBE for several consecutive administrations as a result of the COVID-19 pandemic (<https://nvbar.org/july-2022-bar-exam-modified-to-exclude-mbe/>). The Nevada Supreme Court has now ordered that the MBE be used as a knowledge exam and not used to scale the essay questions and the Nevada Performance Test (NPT) essay scores. Beginning in 2020, the Nevada Supreme court replaced the MPT component of the bar exam with the NPT, which are a set of practice based essay prompts created by and scored based on a Nevada-created rubric.

an examinee's total bar score and, ultimately, whether they passed (the scaled score was above the cut score) or failed (the scaled score was below the cut point) the exam.¹⁶ The essay examination includes a question on legal ethics, in addition to those topics also tested by the MBE.

ASSESSING LAWYERING EFFECTIVENESS

The concept of lawyering skills or effectiveness is distinct from minimum competence. Minimum competence is merely a skill *floor* for entry-level lawyers (Kane, 2009). Its possession is expressed as a yes/no binary. Lawyering effectiveness, however, involves levels or degrees. There are gradations to effectiveness, with some lawyers presumably being more effective than others. Only a few studies have attempted to predict lawyer effectiveness, largely due to the difficulty in developing objective criteria by which effectiveness can be measured (Shultz & Zedeck, 2011). Beginning in the 1970s, researchers began cataloging the skills and knowledge necessary for effective lawyering, primarily by surveying practicing lawyers about the skills they valued as important for their work (Benthall-Nietzel, 1974; Schwartz, 1973).

Schwartz (1973) surveyed 1,200 California attorneys, asking them to rate the importance of fifteen skills that he had previously identified through interviews with lawyers. More than half of the responding attorneys viewed

¹⁶ In Nevada, and in virtually every jurisdiction, scaling is used to minimize the risk that examinees are unfairly penalized by variation in exam difficulty or grader leniency. The procedure involves standardizing, separately, the distribution of the MBE and essay scores around its mean with a standard deviation of one. Each examinee's score in standard deviation units is then multiplied by the standard deviation of the scaled MBE scores in the jurisdiction (Case, 2006).

“Analyzing cases,” “Legal research,” “Knowledge of substantive law,” “Investigating facts of client’s case,” and “Counseling clients” as essential skills, while “Memorizing legal concepts” was ranked the least essential. Benthall-Nietzel (1974) found similar results when she administered a survey based on Schwartz’s questions to 959 Kentucky lawyers, with “Knowledge of statutory law subjects,” “Understanding human behavior,” “Organizing facts,” and “Self-confidence” ranking as most important, and “Memorizing legal concepts” ranking as the least important. Notably, in both studies the surveyed lawyers distinguished between *knowledge* of the law and *memorization* of the law (Merritt & Cornett, 2020). Both studies also found some variation in skills rankings depending on the lawyer’s specialization, practice setting, and years of experience (Benthall-Nietzel, 1974; Schwartz, 1973).

Baird et al. (1979) took a different approach in determining which skills lawyers consider important. The researchers sent questionnaires to organizations employing large numbers of lawyers (e.g., law firms, government agencies, Fortune 500 companies) asking them about their performance evaluation practices and requesting that the organizations submit copies of their evaluation forms. Examining these forms, the researchers identified the 20 most commonly rated characteristics. These traits are listed in the third column of Table 1.

In response to public perception that there was a “gap” between the skills taught in law school and those needed for legal practice, the American Bar Association Section of Legal Education and Admissions to the Bar formed a task force to identify those skills (ABA, 1992). The resulting report, commonly called

the “MacCrate Report,” after the task force’s chair, identified ten “fundamental lawyering skills” (see the fourth column in Table 1).

Since the purpose of the bar exam is to test for minimum competence to practice law, state bars and related organizations have also studied what skills and knowledge are necessary for lawyers (Merritt & Cornett, 2020; National Conference of Bar Examiners, 2012; NCBE, 2022; State Bar of California, 2020). The State Bar of California (2020), for example, administered surveys to 16,190 licensed attorneys in the state asking them about the frequency and criticality of the tasks and legal knowledge they encounter in their daily work. They utilized both traditional surveys and experiential sampling method surveys, where emails were sent to participants at random times throughout the day asking them to categorize the work they were doing at that moment. This technique identified six different competencies (see the final column in Table 1).

The most comprehensive effort to catalog effective lawyering skills came from Shultz and Zedeck (2011). Noting that previous research on lawyering skills gave scant attention to determining the best methods for evaluating them, the researchers sought to create an assessment instrument using a behaviorally anchored rating scale (BARS) methodology. Beginning with interviews and focus groups with faculty, students, and alumni of UC Berkeley School of Law, Shultz and Zedeck identified 26 lawyering effectiveness factors (which they grouped into eight thematic categories; see Box 1), along with hundreds of behavioral examples to illustrate how a lawyer would demonstrate various levels of performance within each factor. Shultz and Zedeck then asked 9,555 law school

alumni to review the examples and indicate the extent to which they aligned with each performance level of each skill, finding general agreement between the accuracy of the examples and the levels of lawyering effectiveness they represented. These 26 factors were then validated through a series of surveys of 1,418 law school alumni and ratings of those alumni that were provided by supervisors and peers.

<p>INTELLECTUAL AND COGNITIVE</p> <ul style="list-style-type: none"> ▪ Analysis & Reasoning ▪ Creativity/Innovation ▪ Problem Solving ▪ Practical Judgment 	<p>CLIENT & BUSINESS RELATIONS – ENTREPRENEURSHIP</p> <ul style="list-style-type: none"> ▪ Networking & Business Development ▪ Providing Advice & Counsel & Building Relationships with Clients
<p>RESEARCH & GATHERING</p> <ul style="list-style-type: none"> ▪ Researching the Law ▪ Fact Finding ▪ Questioning & Interviewing 	<p>CHARACTER</p> <ul style="list-style-type: none"> ▪ Passion & Engagement ▪ Diligence ▪ Integrity/Honesty ▪ Stress Management ▪ Community Involvement & Service ▪ Self-Development
<p>COMMUNICATIONS</p> <ul style="list-style-type: none"> ▪ Writing ▪ Speaking ▪ Listening 	<p>WORKING WITH OTHERS</p> <ul style="list-style-type: none"> ▪ Developing Relationships within the Legal Profession ▪ Evaluation, Development & Mentoring
<p>PLANNING & ORGANIZING</p> <ul style="list-style-type: none"> ▪ Strategic Planning ▪ Organizing and Managing One’s Own Work ▪ Organizing and Managing Others 	<p>CONFLICT RESOLUTION</p> <ul style="list-style-type: none"> ▪ Negotiation Skills ▪ Able to see the World through the Eyes of Others

Box 1. Shultz and Zedeck’s (2011) 26 Lawyer Effectiveness Factors by 8 Thematic Categories

Several of these noncognitive skills are also found in major statements of the legal profession’s values and aims, such as the MacCrate Report (American Bar Association, Section of Legal Education and Admissions to the Bar, 1992).

Table 1 below compares the complete list of Shultz and Zedeck’s lawyering skills attributes to those of Baird et al. (1979), the MacCrate Report (1992), and the State Bar of California (2020).

Skill	Shultz & Zedeck (2011)	Baird et al. (1979)	MacCrate Report (1992)	State Bar of CA (2020)
Advising & Counseling/Building Client Relationships	✓	✓	✓	✓
Communication	✓ ^a	✓ ^a	✓	✓
<i>Influencing & Advocating</i>	✓			
<i>Writing</i>	✓	✓		✓
<i>Speaking</i>	✓	✓		
<i>Listening</i>	✓			
Research & Information Gathering	✓ ^a	✓	✓ ^a	✓
<i>Researching the Law</i>	✓		✓	
<i>Fact Finding</i>	✓		✓	✓
<i>Questioning and Interviewing</i>	✓			
Analysis & Reasoning	✓	✓	✓	
Planning & Organizing	✓ ^a	✓	✓	
<i>Organizing One's Own Work</i>	✓			
<i>Organizing and Managing Others</i>	✓	✓		
<i>Strategic Planning</i>	✓			
Creativity/Innovation	✓	✓		
Diligence	✓	✓		
Integrity/Honesty	✓		✓	
Negotiation	✓		✓	
Passion & Engagement	✓	✓		
Practical Judgment	✓	✓		
Problem Solving	✓		✓	
Relationships with Community/ Community Service	✓	✓		
Self/Professional Development	✓	✓		
Developing Others	✓			
Developing Relationships within Legal Profession	✓			
Empathy	✓			
Networking & Business Development	✓			
Stress Management	✓			
Initiative & Independence		✓		

Knowledge of the Law		✓		
Legal Drafting		✓		
Litigation & Alternative Dispute-Resolution Procedures			✓	✓
Quantity of Work		✓		
Relationships within Organization		✓		
Responsibility, Dependability, & Reliability		✓		
Speed, Efficiency, & Timeliness		✓		
Total^b	26	20	10	6

Table 1. Comparison of Schultz and Zedeck Lawyering Skills Attributes

Note. Dark gray shading indicates a skill was identified in all four reports; medium gray, agreement across three reports; light gray, agreement across two reports; no shading, no agreement across reports.

^a Denotes that the report does not explicitly list this skill but does include one of the skills identified within the category. For example, the MacCrate Report does not list Research and Information Gathering, but it does identify two of the skills within the category: Researching the Law and Fact Finding. ^b Total does not include those checkmarks demarked with *a*.

The Relationship Between the Bar Exam and Lawyer Outcomes

Although Shultz and Zedeck (2011) examined the relationship between lawyer effectiveness and law school admission criteria, similar studies have not been performed on the former's relationship with bar examination scores. Instead, much of the research to date has been on the bar exam and lawyer *ineffectiveness*, primarily by examining rates of attorney discipline (e.g., suspension and disbarment; Anderson & Muller, 2019; Kinsler, 2017; Patton, 2020). Kinsler (2017), for example, studied disciplinary data for 7,256 practicing attorneys in Tennessee and found that lawyers who passed the bar exam on their second attempt were twice as likely to be disciplined as those who passed on their first. There are significant limitations to the use of lawyer discipline as proxy for lawyer effectiveness (or ineffectiveness), however. Firstly, rates of lawyer

discipline are exceedingly low. In 2019, approximately one-quarter of one percent of all barred attorneys in the United States were disciplined for unethical conduct—and only one fifth of those who were disciplined were ultimately disbarred (ABA, 2022a). Thus, using lawyer discipline as a measure of effectiveness yields only the blunt conclusion that virtually all lawyers are sufficiently competent (or lucky) enough to not face professional sanctions. But surely there are varying levels of effectiveness even among undisciplined lawyers. This level of granularity is not available in lawyer discipline data. A lawyer is either publicly disciplined or is not.

Secondly, using lawyer discipline requires that equal weight be assigned to each instance of discipline, regardless of its relevance to the type of knowledge or skill tested on bar exams. Consider two hypothetical lawyers who are publicly disciplined for their conduct: Lawyer A is disciplined for failure to keep multiple clients informed about their cases and Lawyer B is disciplined for negligent failure to preserve client property. According to prevailing research frameworks, both lawyers are ineffective (or incompetent) to some degree. The violation committed by Lawyer A, however, may more accurately reflect poor communication skills (a skill not tested on the bar exam) rather than a lack of legal knowledge. Thus, the theoretical link between the exam and the outcome is tenuous. Moreover, focusing on the *types* of discipline imposed on different lawyers adds no useful nuance. If both hypothetical lawyers are suspended from the practice, prevailing research frameworks will once again treat them the same, even though only Lawyer B has been found to lack legal knowledge.

Lastly, rates of lawyer discipline are unevenly distributed across geography, practice area, gender, and race. According to the ABA, “[l]awyer discipline rates vary significantly from state to state” (ABA, 2022a, p. 85). In Alabama and Iowa, 1 out of every 100 lawyers received public discipline in 2019; in the same year, the rate of discipline in Alaska and Rhode Island was 10 times lower (ABA, 2022a). The variability in these rates undermines the reliability of the measure. Additionally, most instances of discipline involve attorneys who are engaged in solo or small private practice (Farkas, 2019; Levin et al, 2013). This suggests that lawyers who serve individual clients (e.g., private practice lawyers) are more likely to have complaints than lawyers who represent corporate or public entities and lawyers who do not have clients at all.

Perhaps due in some part to racial and gender representation among practice areas (National Association for Law Placement, 2022), male attorneys are more likely to be disciplined than females (Hatamyar & Simmons, 2004). Furthermore, the State Bar of California reported that compared to White male attorneys, Black male attorneys in California were more likely to be placed on probation (1.4 percent versus 0.9 percent) or disbarred (1.6 percent versus 1.0 percent), when controlling for the number of complaints (Farkas, 2019).

To better understand the relationship between bar exam scores and lawyer effectiveness, it is necessary to examine new lawyers’ acquisition and development of skills. We utilized assessments of skills possessed by early career lawyers to do so. Due to its comprehensiveness (see Table 1 above) and its widespread impact in the literature, we relied on the Shultz and Zedeck

framework to structure our investigation and were the first research team granted permission by the authors to use those scales and behavioral anchors in a study.

STUDY DESIGN

Data

Our data were obtained from three sources in a three-stage process.

First, we obtained the names and contact information for those lawyers who were admitted to the Nevada Bar between May 2014 and June 2020 and who had not previously been admitted to a state bar in another jurisdiction.¹⁷ This amounted to 1,414 “new lawyers.” For this group of new lawyers, we also received MBE, MPT, Nevada subject essay, and Nevada ethics essay scores from the Nevada State Bar.

Next, in September 2021, we contacted these 1,414 new lawyers via email, soliciting their participation in the study. This involved directly emailing the population of new Nevada lawyers using the survey program Qualtrics[®], asking for their consent to participate and to provide information regarding their demographic background and current employment and specializations. In addition, we asked them to provide contact information for two supervisors, two peers, and one judge who would be qualified to assess their lawyering effectiveness. We also asked them to complete an optional five-item self-assessment of lawyering effectiveness.¹⁸

¹⁷ Those admitted in June 2020 would have taken the bar exam in February 2020, prior to the widespread disruptions caused by the COVID-19 outbreak.

¹⁸ These five items were a subset of the twenty-six items identified by Shultz & Zedeck (2011).

We received usable responses from 524 of the lawyers (hereafter referred to as “participants”)—a response rate of 37 percent.¹⁹ An initial concern was that study participants were systematically different from the overall pool of 1,414 lawyers. However, as shown in the first and second rows of Figure 2, there are only modest differences between participants and nonparticipants in regard to MBE and MPT scores (and their standard deviations) and the proportion that took the bar exam more than once. The average MBE score differs by less than one point, the average essay score differs by about three points, and the proportion of those taking the bar exam more than once differs by less than four percentage points. None of these differences are statistically significant.

Last, in October and November 2021, we contacted the participants’ supervisors, peers, and judges, as provided in their surveys. These external raters were asked to verify their working relationship with the participant and to evaluate the participant’s lawyering effectiveness on the 26 factors published by Shultz and Zedeck (2011).²⁰ Virtually all participants (470 of 524; 89.7 percent) received at least one evaluation from a supervisor, peer, or judge, with an average of 2.18 reviews per participant (see Figure 1).^{21, 22}

¹⁹ To incentivize study participation, the State Bar of Nevada offered participating new lawyers a \$10 gift card and 3 credits toward their annual Continuing Legal Education (“CLE”) requirements.

²⁰ “Working relationship” was defined broadly as time spent time in a professional capacity working with, observing, or otherwise interacting with the study participant in such a way that the evaluator had first-hand experience/knowledge of the participants professional abilities and skills.

²¹ For supervisors, 29 percent received two supervisor reviews, 45 percent received one, and 26 percent received zero; for peers, 22 percent received two, 43 percent received one, and 35 percent received zero; and for judges, 29 percent received one.

²² Some evaluators provided assessments for multiple study participants. This was most notable among judge evaluators.

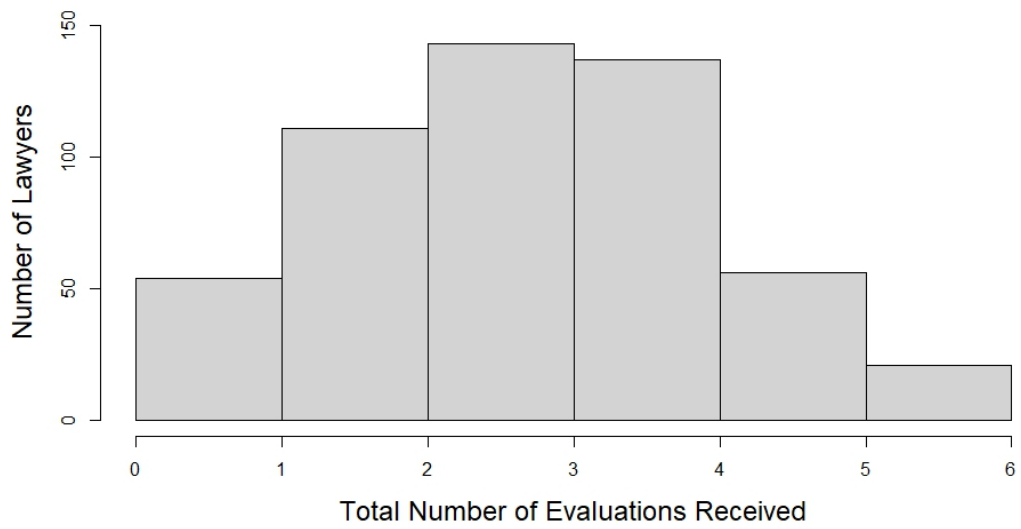


Figure 1. Distribution of Evaluations Received

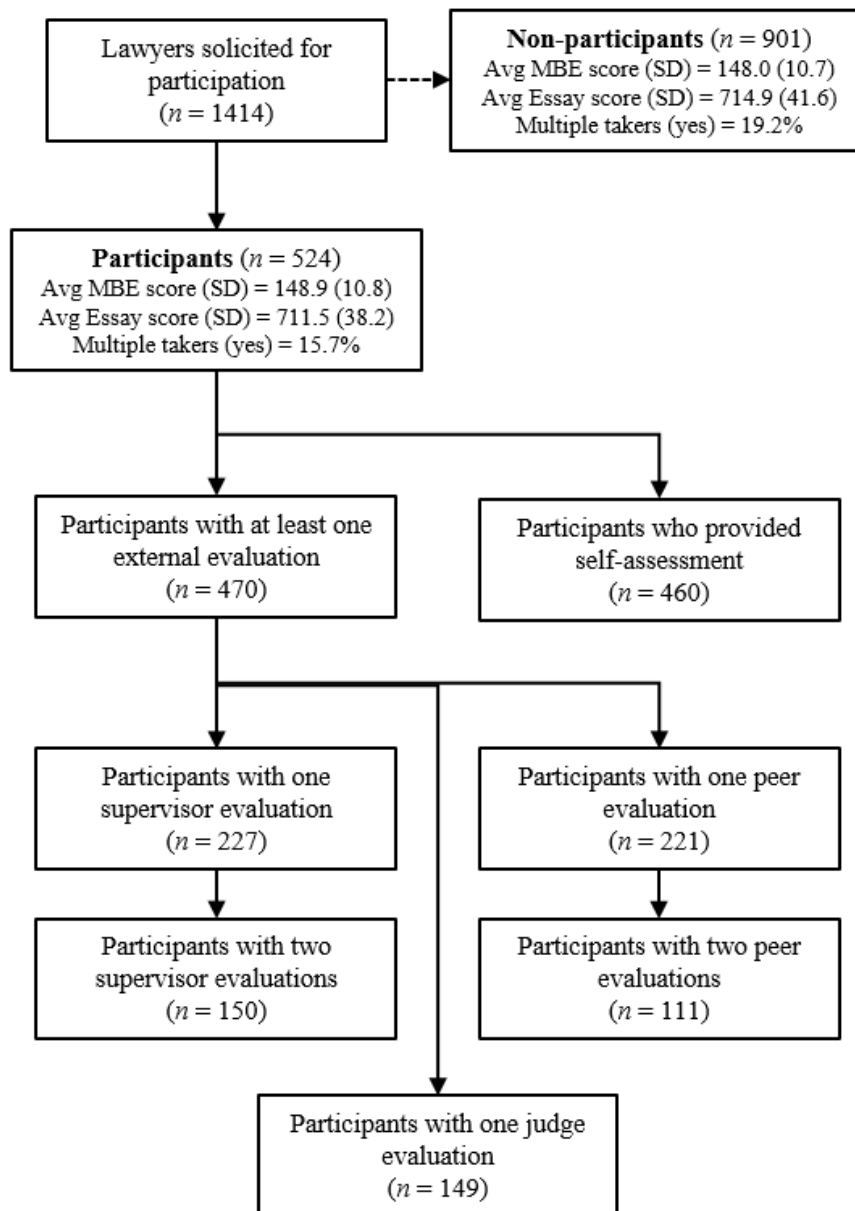


Figure 2. Flow Chart of Study Participation

Measuring Lawyering Effectiveness

Self-assessment

Study participants were asked to self-assess their abilities across five key areas:

- Ability to use analytical skills, logic, and reasoning to approach problems and to formulate conclusions and advice.
- Understanding of legal concepts and utilizing sources and strategies to identify issues and derive solutions.
- Ability to identify relevant facts and issues in a case.
- Ability to generate well-organized methods and work products.
- Ability to write clearly, efficiently, and persuasively.

Each question included a 1–5-point scale, with half-point increments, and an accompanying list of behavioral anchors (examples of behaviors and activity that demonstrate each level along the scale). These behavioral anchors were adapted from those Shultz and Zedeck (2011) created, utilized, and tested.

We elected to limit the number of self-assessment items due to concerns that a long assessment would dampen study participation, especially given the sensitive nature of the information being asked of participants and the duration of the survey. Thus, we selected what we identified as the five (of Shultz and Zedeck’s 26 skills) most broadly relevant and applicable questions across the diverse range of practice areas.

Responses to the self-assessment were not required yet 460 of the 524 participants (88 percent) answered at least one of the self-evaluation items. Of these 460 participants, 444 (85 percent) answered all five questions. And of the 16 participants that did not answer all of the questions, one answered two questions, one answered three questions, and 14 answered four questions. For each

participant, we calculated their average self-assessment rating. As shown in Figure 3, 91 percent of the 460 participants had average self-assessment ratings between 3.0 and 5.0 points. Across the 460 participants, the mean self-assessment rating was 4.20 (SD = 0.43), with a minimum of 2.5 and a maximum of 5.0 (see Table A.2.1).

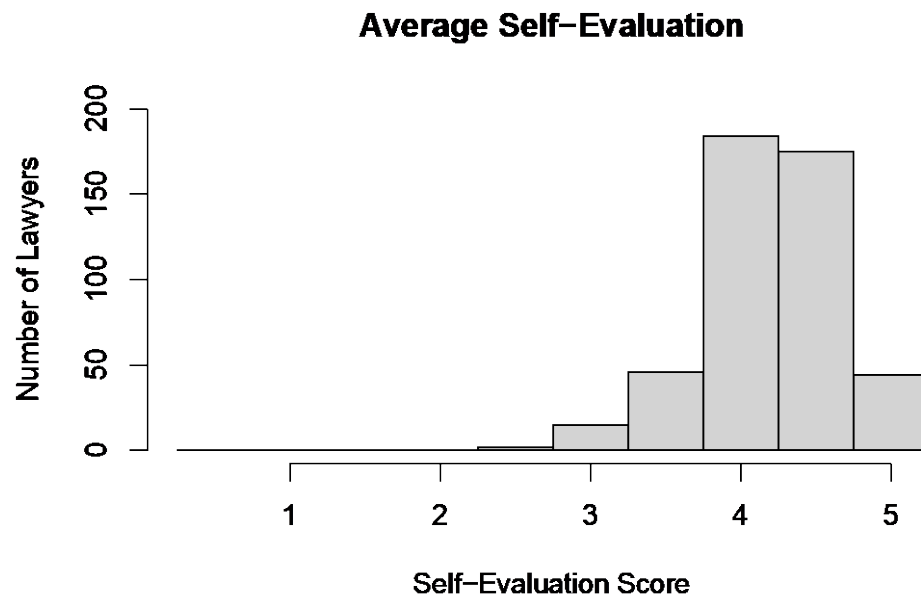


Figure 3. Overall Self-Evaluation of Lawyering Effectiveness Average

Each of the five items has a positive and statistically significant relationship with the others, meaning new lawyers who rated themselves highly on one factor tended to also rate themselves highly on others. Table 3 displays the intercorrelations among the five items on the self-evaluation.

Item	1	2	3	4	5
1 – Intellectual & Cognitive	1.00				
2 – Knowledge of the Law	0.54	1.00			
3 – Research & Info-Gathering	0.44	0.44	1.00		
4 – Planning & Organizing	0.41	0.41	0.42	1.00	
5 – Communications	0.54	0.40	0.36	0.42	1.00

Table 3. Self-Evaluation Intercorrelations

Note. Responses to these items were on a 1–5 scale. Pearson’s r shown. All correlations are calculated from pairwise complete observations and are statistically significant, $p < 0.01$.

Although each of the five items capture different skill constructs, we decided to combine them into one composite variable. This was done to reduce the number of analyses and t -tests performed, which can result in the multiple comparisons problem.²³ We performed principal components factor analysis to test whether this approach was appropriate. As shown in Table A.3.1, the results from this analysis supported our decision: One clear factor was revealed in which each of the five components had a factor loading greater than 0.4 in absolute value (in fact, all had positive loadings).²⁴ This suggested that the five individual items capture our underlying variable of interest—lawyering effectiveness.

²³ The more probability tests are conducted, the greater the chances of Type I error, or the likelihood of finding a statistically significant result that is due to chance.

²⁴ We also examined the five self-evaluation items using exploratory and confirmatory factor analysis to see if one latent variable (or factor) predicted the different constructs. Using exploratory techniques, we found that all five items acceptably loaded onto one factor. We confirmed this was accurate by conducting a confirmatory factor analysis, in which all items were all significantly predicted by one higher order factor. However, when compared to regression models that used: 1. Individual self-evaluation items as outcomes or 2. The average self-evaluation score as the outcome, the factor analysis approach was not a significant improvement in the explanatory power of MBE scores for self-evaluation. Thus, the composite variable of self-evaluation was the most efficient choice for analyses.

External Reviewers

Each supervisor, peer, and judge identified by the study participants was asked to provide an evaluation using a uniform 26-item battery.²⁵ These 26 items corresponded to those skills identified and validated by Shultz and Zedeck, which they categorized by theme (see Box 1 above).²⁶ Each item included a set of behavioral anchors (also developed by Shultz and Zedeck) that describe how a lawyer demonstrates each level of the rating scale (see Box 2).

²⁵ In addition, evaluators were asked to provide demographic and background information, including that related to their working relationship with the participant.

²⁶ Shultz and Zedeck (2011) did not utilize factor analysis in order to define these thematic categories.

Which score best represents $\{e://Field/new_lawyer_first_name\}$'s ability to use analytical skills, logic, and reasoning to approach problems and to formulate conclusions and advice?

5

This attorney takes in a great deal of information and is able to synthesize and refine it so that it is organized and clear.

This attorney surveys and tests out various bodies of law, theories and analogies when deciding how to approach a new or especially difficult problem.

4

When there are multiple issues, this attorney imposes a hierarchical order on them, deciding which are most important, which are potentially significant, and which should be minimized or discarded.

This attorney addresses the legal issues in a case by determining the potential legal positions, the authority for each and the outcome desired.

3

This attorney figures out what answer he/she (or the client) wants and works back from that, looking only for arguments and authority that support that outcome, and trying to fit the facts and law to the desired result.

2

This attorney makes liberal use of his/her organization's brief bank of formula driven, pre-written briefs (or of standard contract forms); cuts and pastes the facts of his/her current case into the form and ignores small differences between issues or facts in his/her case and those in the model.

1

Box 2. Example of Rating Scale and Behavioral Anchors

As shown in Table 4, there is clear variation in score averages across the eight skills themes as well as across the different types of raters.²⁷ Supervisors rated the participants lower than peers and judges across all eight categories.

²⁷ Although study participants were not given the full battery of 26 skills, they were given a subset of 5 of these items, each of which falls under one of the 8 categories. Thus, we are able to compare participants' self-evaluations to the average scores they received from supervisors, peers, and judges for these 5 common items.

Moreover, it is noteworthy that the average self-assessment was 4.20—lower than any of the external reviewers. Thus, it appears reasonable that self-evaluation responses are not overly inflated.

Category	Supervisor(s)	Peer(s)	Judge	Self
Intellectual & Cognitive Factors	4.20	4.39	4.32	4.29
Research & Info Gathering	4.22	4.35	4.29	4.08
Communications	4.14	4.31	4.36	4.29
Planning & Organizing	4.14	4.32	4.36	4.22
Conflict Resolution	4.20	4.37	4.34	— ^a
Client & Business Relationships	4.16	4.39	4.41	4.11
Working with Others	4.33	4.47	4.63	— ^a
Character	4.35	4.47	4.49	— ^a
Overall	4.22	4.38	4.38	4.20

Table 4. Average Evaluations from External Evaluators and Self-Evaluations

^a Participants were not given self-assessment items related to this skill.

As shown in Table 5, each skill category has a positive and statistically significant relationship with the others, meaning a reviewer who rated a participant highly in one area tended to rate them highly on the others as well. Relative to the self-evaluation scores, the intercorrelations for external reviewers are considerably stronger. This trend suggests that each of the skill categories captures the same underlying dimension—lawyering effectiveness.

Thematic Category	1	2	3	4	5	6	7	8
1 - Intellectual & Cognitive	1.00							
2 - Research & Info Gathering	0.86	1.00						
3 - Communications	0.89	0.84	1.00					
4 - Planning & Organizing	0.82	0.77	0.82	1.00				
5 - Conflict Resolution	0.76	0.73	0.78	0.73	1.00			
6 - Client & Business Rel.	0.63	0.60	0.67	0.71	0.62	1.00		
7 - Working with Others	0.59	0.57	0.63	0.66	0.68	0.61	1.00	
8 - Character	0.78	0.72	0.77	0.81	0.70	0.70	0.70	1.00

Table 5. External Evaluations Intercorrelations

Note. Responses to these items were on a 1–5 scale. Pearson’s r shown. All correlations are calculated from pairwise complete observations of averages across all external evaluators and are statistically significant, $p < 0.01$.

As with the self-evaluation scores, we examined whether these items could be combined into fewer latent variables. The differences in average ratings across the eight skill categories by evaluator type suggests that it might be prudent to avoid combining the scores from the supervisors, peers, and judges into one average score. To test this possibility, we assessed the intercorrelations between the self-evaluations and the external evaluations.

Table 6 displays the correlations between the self-evaluation and the overall supervisor, peer, and judge evaluations, both overall and by each thematic category. Each of the evaluation types has only a modest positive and statistically significant relationship with the others. These relatively low intercorrelations among raters could suggest real divergence in perceptions as well as differences in response styles among the three types of raters. With respect to response styles, each of the raters may differ in their interpretation of the wording in individual questions, or the extent to which they strictly apply the behavioral anchors when

choosing a rating. Additionally, supervisors, peers, and judges likely vary in the depth of their experiences with participants across the eight thematic categories. These differences provide additional support to our decision to present separate models and results for each rater type.

We next assessed whether we could combine these variables in order to have one average score for each category of external reviewer. Our results from factor analysis suggested that this approach would best represent the variance in ratings (see Table A.3.2 in the Appendix). We found that all 26 items loaded onto one factor for each type of reviewer. This echoed our earlier finding regarding self-evaluation ratings; taken collectively, all items measure one overall latent factor—lawyering effectiveness. Principal components analysis confirmed this expectation, extracting one single factor, all with highly positive loadings on all subscales.²⁸ We therefore moved forward with regression analyses using average ratings for each type of reviewer as outcome variables.

²⁸ The first extracted factor has an eigenvalue of 15.52, suggesting it captures the variation of roughly fifteen component subscales. The next largest extracted component has an eigenvalue of 1.29.

Overall											
<i>Rater</i>	<i>1</i>				<i>2</i>			<i>3</i>			<i>4</i>
1. Self	1.00										
2. Supervisor	0.19*				1.00						
3. Peer	0.16*				0.19*			1.00			
4. Judge	0.23*				0.22*			0.22*			1.00
Intellectual and Cognitive Factors					Conflict Resolution						
<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>		
1. Self	1.00				1. Self	-					
2. Supervisor	0.10	1.00			2. Supervisor	-	1.00				
3. Peer	0.14*	0.18**	1.00		3. Peer	-	0.18*	1.00			
4. Judge	0.11	0.18	0.21	1.00	4. Judge	-	0.13	0.21*	1.00		
Research and Information Gathering					Client and Business Relations						
<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>		
1. Self	1.00				1. Self	1.00					
2. Supervisor	0.08	1.00			2. Supervisor	0.22***	1.00				
3. Peer	0.14*	0.17*	1.00		3. Peer	0.23	0.12	1.00			
4. Judge	0.10	0.14	0.20*	1.00	4. Judge	0.21	0.29	0.16	1.00		
Communication					Working with Others						
<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>		
1. Self	1.00				1. Self	-					
2. Supervisor	0.19***	1.00			2. Supervisor	-	1.00				
3. Peer	0.18**	0.27***	1.00		3. Peer	-	0.10	1.00			
4. Judge	0.13	0.24*	0.25*	1.00	4. Judge	-	0.19*	0.12	1.00		
Planning and Organizing					Character						
<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>Rater</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>		
1. Self	1.00				1. Self	-					
2. Supervisor	0.09	1.00			2. Supervisor	-	1.00				
3. Peer	0.05	0.26***	1.00		3. Peer	-	0.11	1.00			
4. Judge	0.33***	0.18	0.26*	1.00	4. Judge	-	0.19	0.16	1.00		

Table 6. Inter-rater Relationships Overall and for each Thematic Category

Note. Responses to these items on 1–5 scale, averaged across all included items, Pearson’s *r* shown. Significance level is marked by asterisk(s). If multiple supervisors or peers replied, their scores were averaged. All correlations are calculated on all pairwise complete observations. Conflict Resolution, Working with Others, and Character were not items included in the self-assessment. Significance level is indicated by: **p* < .05, ***p* < .01, ****p* < .001

Research Hypothesis

Given the purpose of the bar exam is to help ensure that only individuals who possess minimum competence to practice law are granted licenses to do so, we hypothesize that if it is indeed serving as prescribed, then those with higher exam scores should be more effective lawyers. If our hypothesis is true, we would expect our results to show some form of positive relationship, where lawyer ratings increase as exam scores increase. However, the presence of a relationship (even a statistically significant one) is not sufficient on its own—there should be some evidence that the relationship has practical significance. If the results fail to meet both criteria, this might suggest that the bar exam, at least in the case of the early career Nevada lawyers participating in this study, is not serving as intended.²⁹

RESULTS

MBE & Lawyering Effectiveness

In this section, we explore the effects of MBE scores on lawyer effectiveness ratings.

In Figure 4, we present descriptive figures plotted using LOESS, or locally estimated scatterplot smoothing. These figures allow us to see the relationships between MBE scores and lawyering effectiveness without any assumptions

²⁹ Since the bar exam is an unvalidated assessment, the most practicable approach for hypothesis testing is:

H₀ = the bar exam does *not* predict lawyering effectiveness

H₁ = the bar exam does predict lawyering effectiveness

regarding linear effects. As shown in the top left of Figure 4, there are modest, positive, and essentially linear relationships between first-attempt MBE scores and the four types of lawyering effectiveness ratings: self, supervisor, peer, judge.³⁰ Overall, these relationships suggest that early career lawyers with higher MBE scores received modestly higher lawyer effectiveness ratings than those with lower scores.

Additionally, there does not appear to be a threshold on the MBE scale above or below which we see marked changes in lawyering effectiveness ratings—the effect remains quite linear throughout the range of MBE scores, except for the slight uptick at the lower end for self- and judge evaluation scores. This is likely due, at least in part, to having only 13 scores below 120.³¹

³⁰ All analyses we present use first-attempt bar scores, which affects only 84 of the total 524 (16 percent) participants who took the bar exam more than once. Using first-attempt bar scores provides us with a wider range of bar exam scores, including those scores that do not constitute passing.

³¹ With any sample of fewer than 20 observations, results may be skewed or spurious due to limited statistical power.

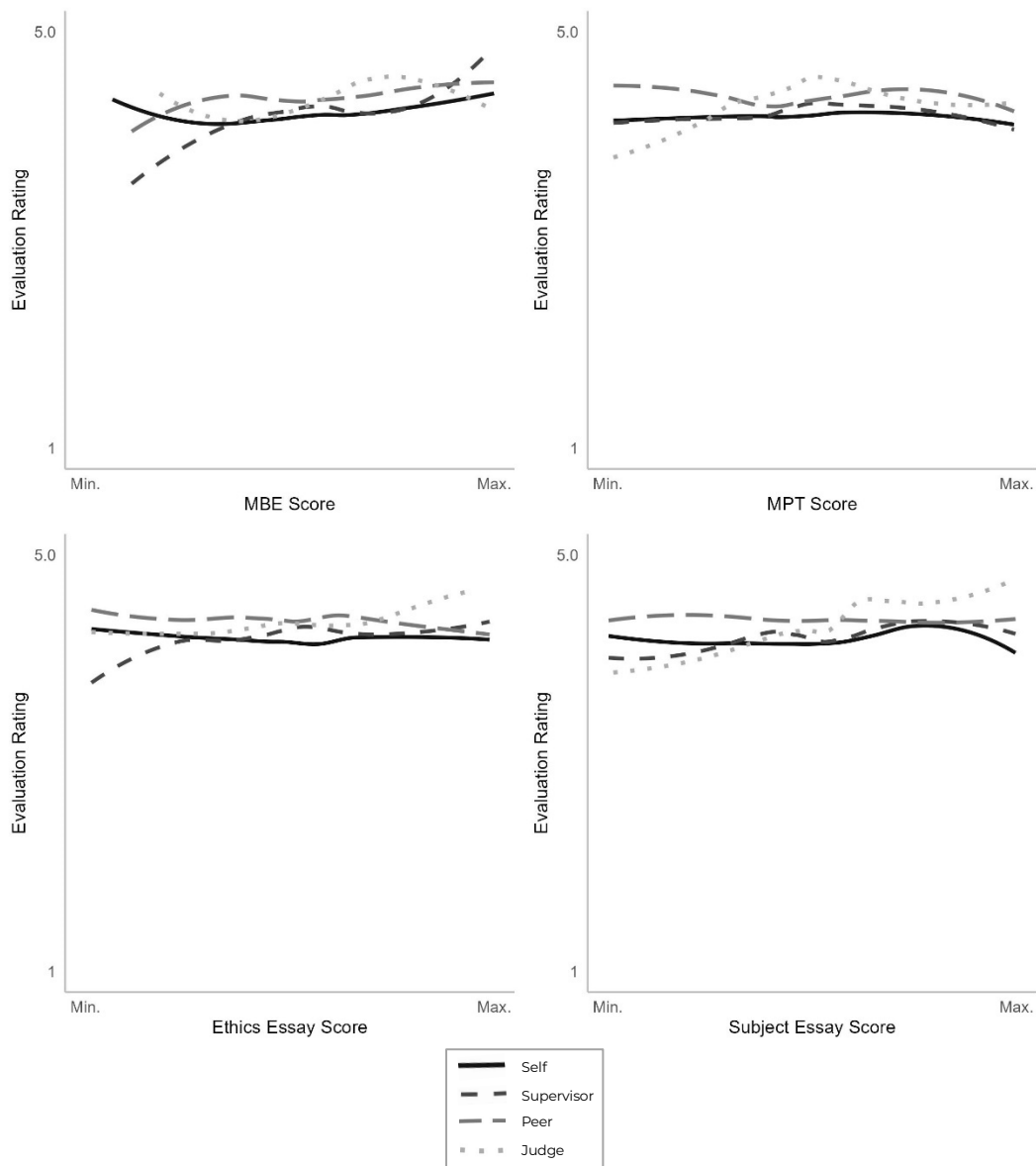


Figure 4. Lawyering Effectiveness Evaluations and First-Attempt MBE, MPT, Ethics, and Subject Essay Scores

Simple bivariate OLS models reveal that MBE scores explain only between 1 and 3 percent of the variation in lawyering effectiveness ratings across

the different types of raters (see table A.3.1).³² In short, almost all of the variation is explained by factors outside of MBE score.

In the ensuing analyses, we present a regression model estimating the extent to which MBE, MPT essay, Nevada subject essay, and Nevada ethics essay scores predict lawyering effectiveness ratings while accounting for possible confounding effects. In this model we control for the following characteristics of the early career lawyers: race, sex, and age; whether the lawyer received their JD from UNLV Boyd School of Law; whether the new lawyer attempted the Nevada Bar Exam more than once; and the amount of elapsed time since bar passage.^{33, 34} Furthermore, recognizing that observations are clustered within bar exam administrations, we utilize cluster-robust standard errors. This allows us to apply a correction to the level of precision (our standard errors) in our estimates. This is particularly useful in this case because the MBE is scaled (like the essay scores) by the scores of those individuals who took the exam in that administration. This means that our model's residuals would not be independent of each other ("autocorrelation").

Across all evaluation types, the results are consistent and clear: for our sample of early career lawyers in this study, MBE scores have a negligible,

³² OLS stands for "ordinary-least-squares" and is the most common type of linear regression.

³³ Nearly half of the study participants (46.6 percent) graduated from UNLV Boyd School of Law. Since UNLV Boyd is the only law school in Nevada, this high proportion of representation is to be expected.

³⁴ It is worth noting that we are careful to include no control variables that would create post-treatment bias in the estimate of the effect of MBE score. If we believe that a new lawyer's MBE score might affect other control variables that are determined after one takes the MBE, we do not include them in the model. One such example is employment type (e.g., solo practice, small/medium/large private firm, business/industry, government, clerkship).

positive relationship with ratings of lawyering effectiveness. Although the results are statistically significant for self-provided and peer evaluations, the effect is not practically significant. We find that going from the lowest MBE score (107) in the sample to the highest (177) is associated with an increase in:

- *Self-evaluation* rating of 0.45 points (on a five-point scale) from an average score of approximately 4.2 to a score of 4.7.
- *Supervisor* rating of 0.24 points (on a five-point scale) from an average score of approximately 4.2 to a score of 4.4.
- *Peer* rating of 0.43 points (on a five-point scale) from an average score of approximately 4.4 to a score of 4.8.
- *Judge* rating of 0.45-points (on a five-point scale) from an average score of approximately 4.4 to a score of 4.9.

Although the absolute value of these changes is substantial, realizing them requires a 70-point increase on the MBE, which approximates five standard deviations.

More practicably, examining these changes in terms of a one standard deviation increase in MBE score (12.34 points), we find that ratings of lawyering effectiveness measured by:

- *Self-evaluations* increase by only 0.07 points (on a five-point scale), from an average score of approximately 4.2 to a score of 4.27 (or roughly 4.3).
- *Supervisor evaluations* increase by only 0.07 points (on a five-point scale), from an average score of approximately 4.2 to a score of 4.27 (or roughly 4.3).

- *Peer evaluations* increase by only 0.08 points (on a five-point scale) from an average score of approximately 4.4 to a score of 4.48 (or roughly 4.5).
- *Judge evaluations* increase by only 0.05 points (on a five-point scale) from an average score of approximately 4.4 to a score of 4.45 (or roughly 4.5).

Table 7 displays the results of our eight Tobit³⁵ regression models: two predicting self-evaluations of lawyering effectiveness, and six predicting supervisor, peer, and judge evaluations of lawyering effectiveness.³⁶

Nevada Essays & Lawyering Effectiveness

In this section, we explore the effect of scores on the three Nevada essay types—MPT, Ethics, and Subject—on effectiveness ratings.^{37,38}

As with our description of MBE scores and represented by the essentially flat horizontal lines in Figure 4, we find no meaningful relationships between essay scores and ratings of lawyering effectiveness. The only possible exception

³⁵ Tobit models are useful when we have left or right censoring for our dependent variable. In this case, because of the 1-5 rating, and its left skew, with most raters choosing a 4 or 5, these models help account for this ceiling effect.

³⁶ While the results reported in Table 7 are generated using Tobit regression, we include in Appendix Table A.5 those results attained from OLS regression to demonstrate that our results are robust to regression methods. We elected to report those results from our Tobit models because these are better able to reduce the risk of bias and correctly estimate standard errors when, as is the case here, there is censoring of data near the top or bottom of the distribution of a variable. Although we include OLS for robustness, given the coarseness of our evaluation scale, it is worth noting that the substantive conclusions we draw from these models is effectively identical across model specification.

³⁷ Since the essay scores are normalized and their scale is arbitrary, scores for each essay type are rescaled to the 0–1 interval.

³⁸ Like before, this allows us to compare the different scales and their estimated effects to each other as they share a common length. Since there are a few extreme outliers in the essay scores, all scores below the 1st percentile or above the 99th percentile (4 scores at each end) were recoded to the 1st or 99th percentile.

is the relationship between subject-based essay scores and judges' ratings of lawyering effectiveness.

Figure 4 displays the relationships between each of the essay types and self-evaluations of lawyering effectiveness. The effects of MPT score on lawyering effectiveness are in the top right, with ethics essay score in the bottom left, and subject essay score in the bottom right. These plots display the relationships between essay scores and self, supervisor, peer, and judge evaluations. For each, we plot the relationship using LOESS methods in order to view the relationships without the assumption that they be linear.

Across each of the LOESS plots in Figure 4, we see remarkably little systematic variation—effectively all lines except for those in the judge evaluations are remarkably flat, suggesting very little variation in lawyering effectiveness explained by the essay score. Interestingly, we see somewhat positive relationships for the judge evaluations, but a non-monotonicity of the relationship above the midpoint of the MPT score (Figure 4, top right).

Simple bivariate OLS models show that MPT, subject-based essay, and ethics essay scores explain between 0.03 and 7.3 percent of the variation in lawyering effectiveness ratings across the different types of raters (see tables A.4.2–A.4.4).

Of course, these descriptive results may be due to chance or confounding relationships with other factors, as discussed above in the MBE results section. For that reason, Table 7 displays the results from two types of Tobit regression

models, one that includes only the bar exam components and one that adds a set of control variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Simple - Self	Full - Self	Simple- Super- visor	Full Super- visor	Simple - Peer	Full - Peer	Simple- Judge	Full - Judge
MBE score (0-1)	0.35 *	0.45 **	0.28	0.24	0.54 **	0.43 *	0.30	0.45
	(0.15)	(0.15)	(0.22)	(0.22)	(0.20)	(0.21)	(0.41)	(0.43)
MPT score (0-1)	-0.06	-0.07	-0.11	-0.04	0.04	0.05	0.14	0.05
	(0.13)	(0.13)	(0.19)	(0.20)	(0.19)	(0.20)	(0.44)	(0.45)
Subject essay score (0-1)	0.13	0.12	0.35	0.11	-0.37	-0.39	0.70	0.72
	(0.15)	(0.17)	(0.23)	(0.26)	(0.23)	(0.25)	(0.51)	(0.55)
Ethics essay score (0-1)	-0.10	-0.14	0.15	0.07	-0.12	-0.11	0.15	0.03
	(0.11)	(0.11)	(0.16)	(0.17)	(0.16)	(0.16)	(0.33)	(0.34)
>1 attempt		0.03		-0.15		-0.07		0.03
		(0.07)		(0.11)		(0.10)		(0.22)
Years since exam		0.01		0.05 **		0.04 *		0.02
		(0.01)		(0.02)		(0.02)		(0.04)
Race - White		-0.09		0.04		0.06		0.16
		(0.09)		(0.13)		(0.12)		(0.33)
Race - Hispanic/Latino		-0.14		0.14		0.11		-0.01
		(0.09)		(0.13)		(0.12)		(0.35)
Race - Black		-0.05		0.04		0.30 *		0.36
		(0.11)		(0.16)		(0.15)		(0.39)
Race - Asian/PI		-0.03		0.07		0.07		0.41
		(0.10)		(0.15)		(0.15)		(0.38)
Race - Other		-0.07		0.07		0.12		0.39
		(0.12)		(0.16)		(0.16)		(0.42)
Female		0.13 **		0.08		0.02		0.17
		(0.05)		(0.07)		(0.07)		(0.14)
Age		-0.00		-0.01		-0.01		-0.01
		(0.00)		(0.01)		(0.01)		(0.01)
UNLV (Boyd) grad		-0.02		0.06		-0.03		-0.05
		(0.04)		(0.07)		(0.06)		(0.13)
<i>N</i>	430	430	338	338	290	290	119	119
<i>Pseudo R</i> ²	0.02	0.06	0.03	0.07	0.03	0.07	0.08	0.12

Table 7. Multivariate Tobit Regression Relationships Between Lawyering Effectiveness and Bar Exam Components.

Note. Standard errors in parentheses. “Simple” models refer to Tobit models run without control variables, while “full” refers to Tobit models run with all relevant control variables.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As shown in Table 7, like the MBE, the essay components are considerably limited in their abilities to predict our four measures of lawyering effectiveness—self-evaluations, supervisor evaluations, peer evaluations, and evaluations by judges. When controlling for the covariates we introduced in the previous section, we see that the three essay scores fail to achieve statistical or practical significance, regardless of who provided the measure of lawyering effectiveness.

Figure 5 shows the comparative effects that MBE, MPT, Nevada subject-based essay, and Nevada ethics essay scores have on predicted ratings of lawyering effectiveness for each evaluator type. As the essentially flat lines indicate, none of these variables has a practically significant influence of ratings of lawyer effectiveness.

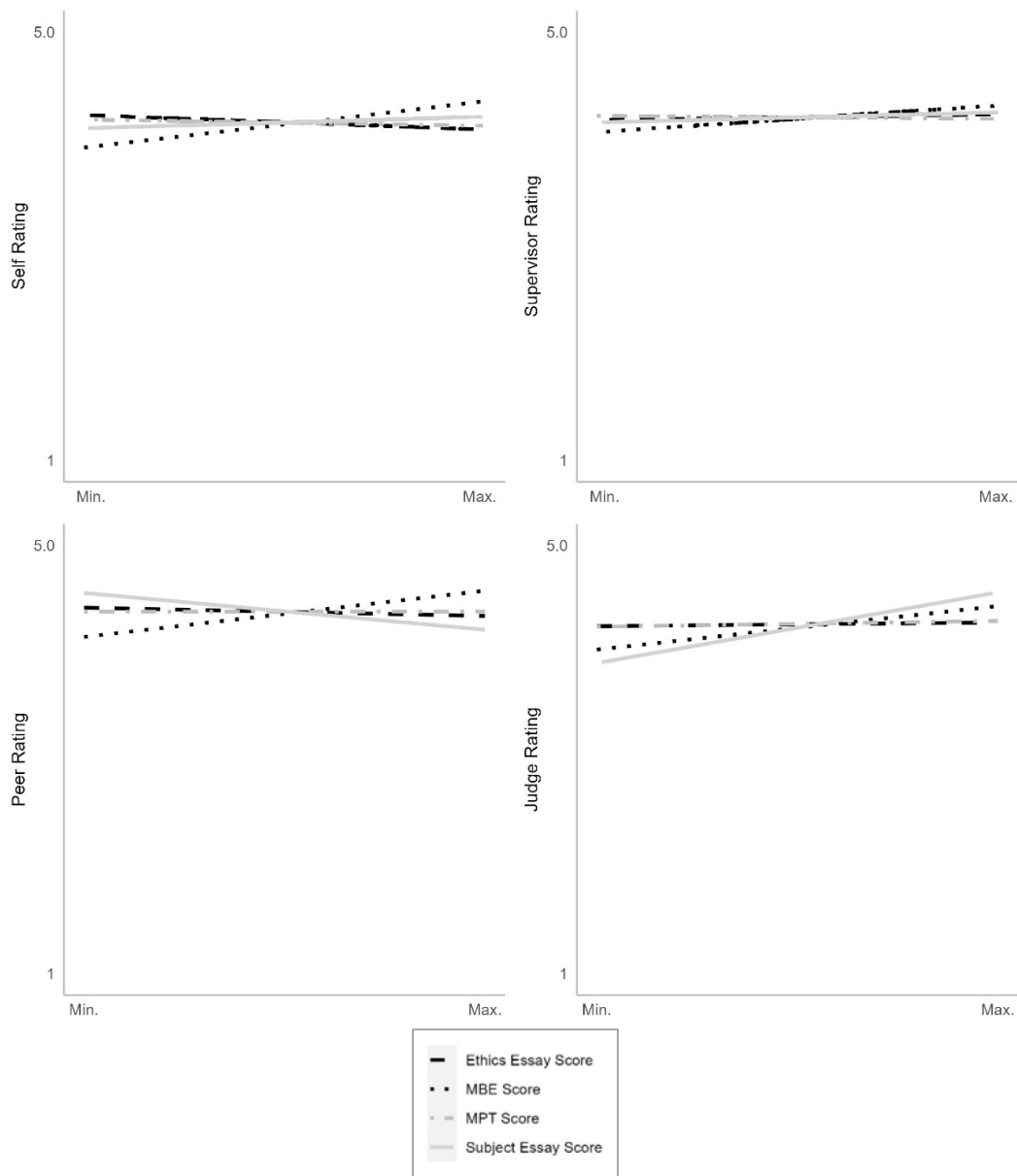


Figure 5. Predicted Evaluation Ratings Based on MBE, MPT, Subject Essay, and Ethics Essay Scores, by Evaluation Type.

DISCUSSION

Data from the ABA National Lawyer Population Survey (2020), which provides the most comprehensive picture of the demographic composition of the legal profession in the U.S., show that, in 2019, only 14.1 percent of practicing attorneys nationwide were people of color. By comparison, people of color

comprised 40 percent of the U.S. population (ABA, 2022a). Clearly, barriers to the legal profession remain.

One such barrier is the bar exam. It is possible that factors such as the Character and Fitness Requirement disproportionately exclude people of color; however, because these requirements vary widely across jurisdictions, these factors are unlikely to explain the widespread disparate exclusion of people of color to the practice of law.³⁹ Furthermore, according to ABA data, in 2021, White law school graduates had first-time pass rates 24 percentage points greater than their Black peers, 13 percentage points greater than their Hispanic peers, and 15 percentage points greater than their Native American peers (ABA, 2022c).

Therefore, given the bar exam’s intent to signal those with the “minimum competence” to practice law and its assumed disparate impact on law school graduates of color, the validity of the exam in serving its stated purpose should be tested. Moreover, since the bar exam has not been rigorously validated, the burden of proof should lay with the bar exam; in order to be a valid gatekeeper, exam scores should have a meaningful and positive relationship with the skills required for effective lawyering—in this case, the 26 lawyering skills developed by Shultz and Zedeck (2011).

As we show above, these assessments of lawyering effectiveness should be treated as presumptively valid. Using principal factor analysis and

³⁹ The Character and Fitness Requirement varies by jurisdiction. It may include a credit check, mental health questions, reference checks, and personal interviews (ABA, 2022b; 2022c).

confirmatory factor analysis testing, we find strong evidence that for each type of rater, taken collectively, the 26 lawyering skills measure the same latent variable—lawyering effectiveness. It appears unlikely that self-assessments are inflated, since average self-assessments are lower than any of those provided by external reviewers. Any differences observed in ratings between supervisors, peers, and judges are likely a reflection of their differing relationships to, and interactions with, the study participants.

Our results suggest that the bar exam does not meaningfully predict the ability of lawyers to perform effectively. We find only modest, positive relationships between first-time MBE scores and all four assessments of lawyering effectiveness (i.e., self, supervisor, peer, and judge).⁴⁰ These effect sizes are not practically significant. Holding constant the number of attempts, years since exam, race, gender, age, and law school attended, an increase of 13 points (one standard deviation) on the MBE is associated with less than a 0.1 increase in ratings of lawyering effectiveness among new lawyers in Nevada. The

⁴⁰ We also investigated whether any of the eight thematic categories of lawyering effectiveness were predicted by MBE, MPT, Nevada subject essay, or Nevada ethics essay scores. MBE score statistically significantly predicted only two of the eight categories: intellectual/cognitive skills and planning/organizing skills. These relationships are negligible but positive. An increase in MBE scores from the minimum to the maximum results in an increase of 0.38 points for intellectual and cognitive ratings and an increase of 0.46 points for planning and organizing ratings. Since the bar exam is explicitly intended to test these constructs, it makes sense that the largest effects are seen here. Note however that while the results are statistically significant, they are not practically significant. In nearly all cases, the effect of the MBE was larger (in absolute terms) than that of the MPT, subject essay, or ethics essay. The exceptions are: MPT negatively predicts research and information gathering and conflict resolution and positively predicts client and business relations skills; subject essay scores negatively predict client and business relations and working with others skills; and ethics question scores positively predict client and business relations.

results are mixed for the MPT, Nevada subject essay, and Nevada ethics essay scores, depending on the rater, the effects may be either positive or negative. But in nearly all cases, the effects do not have practical significance. The only exception is the relationship between subject essay scores and judge's ratings: lawyers that scored higher on the subject essays were rated higher. Although these effect sizes are the largest that we find, their standard errors indicate that there is substantial uncertainty around these estimates.

Notably, the amount of elapsed time since bar passage is a statistically significant control variable, intuitively suggesting that as lawyers gain more experience, their effectiveness increases. It is important to note, however, that this variable is included only to estimate more accurately the magnitude of the relationship between MBE scores and rating of lawyering effectiveness. Any inferences based on time since bar passage should be avoided because, as a control variable, we do not include other potential confounders that might bias the size of its relationship with lawyering effectiveness. This finding might suggest that further research, including whether the adoption of a medical school model with supervised practice, is warranted. Perhaps this experience might ultimately be the best predictor of effective lawyering—evidence suggests that residency is a particularly effective training technique for new medical doctors (Bowen et al., 1999; Klessig et al., 2000; Wright & Schachar, 2020).

Based on our findings, we cannot reject the null hypothesis that the bar exam does *not* validly indicate who is an effective lawyer and who is not—the bar exam, as it was administered in Nevada during the study period, does not

effectively predict who will be an effective lawyer. Although some results achieve statistical significance, this threshold should not be used as the sole determining factor as to whether the bar exam serves its purpose. Several factors are involved when estimating p -values (e.g., sample size, statistical power), which can render statistical significance for even exceptionally small effects. Although p -values are a useful tool, they should never be viewed without context—that is, interpretation of the coefficient and the practical significance of the result.

Sustained interest in defining and assessing minimum competence in legal education means that studies of this kind should be conducted in other jurisdictions. Moreover, further study should focus on identifying other factors that account for the acquisition of skills and noncognitive factors that conduce to lawyering effectiveness. Elapsed time since bar passage is one such possibility, but our findings do not permit us to make any definitive claims. In particular, both the literature and the totality of our findings suggest that noncognitive factors—not easily assessed by a multiple-choice instrument like the MBE, or even constructed-response instruments like the MPT—play a crucial role as determinants of lawyering effectiveness.

The Nevada Supreme Court recently commissioned a group of practitioners and researchers to develop a series of recommendations to restructure the bar exam in Nevada. In brief, the initial recommendations of the commission are to require bar applicants to:

- Following a student's completion of their first year of curriculum, pass a knowledge-based, multiple-choice exam based on the doctrinal coursework taken during this time.⁴¹
- Following graduation, successfully complete a series of practice-based essay questions (the Nevada Performance Test), which are developed by the State Bar and which are scored based on a Nevada-created rubric.
- During their third year or following graduation, successfully complete a period of supervised practice.

Our results support these recommendations in several ways. First, we find that MBE scores are only minimally predictive of lawyering effectiveness. The topics tested on the MBE are typically those covered in law students' first year of studies, which might partly explain its weak relationship with lawyering effectiveness. By testing this acquisition of this knowledge closer to the time at which students learn this material, this component of the bar exam might be a better predictor. In addition, the commission has recommended that those who fail the knowledge test not be allowed to continue their studies until they pass. Such timing would also allow students an earlier chance to gauge their potential for bar admission. This would mean that students could change their career or academic path after only one year. Currently, waiting to take the bar exam until after

⁴¹ This is a significant departure from current practice where no testing by jurisdictions is completed prior to graduation. Students would not be able to progress in their curriculum until passing this exam.

graduation means that those who fail will carry three (or more) years of law school debt without the additional income concomitant with bar licensure.

Second, we find that the MPT is negligibly related to lawyering effectiveness. It is possible that the prompts created by the National Conference of Bar Examiners do not capture the practice of law that is specific to Nevada or that the rubric used to grade the responses are not aligned with best practices in Nevada. The adoption of the Nevada Performance Test may allow for better tailored essay prompts and answer keys that more closely align with practice in Nevada.

Last, as we note above, the control variable *years since passage* might suggest that some form of supervised practice could yield more effective lawyers. If *years since passage* captures the experience of practicing, it is reasonable to assume that with greater practice comes greater effectiveness. It could therefore be a significant boost to lawyering effectiveness to require this type of supervised practice, which is similar to the training of medical doctors.

Additional study when the recommendations have been implemented will be necessary to determine whether the new format better measures and therefore better predicts lawyering effectiveness compared to the administration of the bar exam studied here.

Limitations

As with any empirical study, there are limitations borne out of the data we have available to us for analysis. All analyses of the effectiveness of credentialing or admissions in education suffer from the same significant inferential issue—

censored data. Specifically, while we can observe variation in lawyering effectiveness across the full range of lawyers who eventually passed the bar exam, we cannot observe the potential lawyering effectiveness of those who never successfully became new lawyers.

To some extent, we are able to account for this limitation by using first-time bar exam scores. Of the 524 participants, 84 (16 percent) failed the exam on their first attempt; thus, we are able to examine how the bar exam performs in predicting lawyering effectiveness for those individuals scoring below the cut score. Overall, approximately half of those who failed the bar exam on their first attempt received lawyering effectiveness ratings at or above the average for each evaluation type. Thus, it seems that the bar exam is no more predictive of lawyering effectiveness for those that fail on their first attempt than it is for those that pass.

Nevertheless, it is important to recognize that the participants in our sample did eventually pass the bar and enter the legal profession in Nevada. There are some applicants to the bar who never achieve a passing score. We are unable to take these individuals into consideration in these analyses as we can never know how effective they would have been as lawyers. As such, we do not know the extent to which these individuals differ from those who failed on their first attempt but ultimately passed. These differences could lead to biased estimates, but we cannot know or approximate the direction or the magnitude of the potential bias. It is possible that bar exam scores are more impactful among those that never pass the exam. But it is also possible that they are less so.

Additionally, our ratings of lawyering effectiveness are effectively limited to a range of three points to five points. This data censoring means that there is little room for sizable increases or decreases in ratings as a function of bar exam score. Nonetheless, even when considered relative to the narrower range of rating values, the magnitude of the predictive effect of bar exam scores is small. An argument could be made that the bar exam is responsible for the floor on rating values—that is, those that pass the bar exam are effective lawyers, receiving at least a three on the ratings scales. But, as noted two paragraphs above, this argument is contrary to the data we have for those who failed on their first attempt. Of these individuals, approximately half score at or above the average rating, suggesting that those that fail the bar exam are no less effective.

Finally, we only observe bar performance and lawyering effectiveness in this current study, and do not know how individual metrics that predate bar performance, such as performance in law school, law school admissions metrics, or other characteristics might relate to the outcome of lawyering effectiveness. Future research on the topic should explore what factors *do* predict lawyering effectiveness.

CONCLUSION

Passage of a bar exam is required in nearly all United States jurisdictions in order to practice law. Theoretically, this is because bar results indicate how well a burgeoning attorney will fare in their career, with higher scores indicating a promising future, and those below a particular cut point signaling lack of skills

required for minimum competency. As such, the bar exam should, presumably, significantly predict ratings of lawyering effectiveness. Thus, our analysis builds on Schultz and Zedeck's (2011) earlier work to test whether the relevant bar scores are related to their 26 skills of lawyering effectiveness. We find that while MBE scores are statistically significantly related to some evaluations of lawyering effectiveness, the relationships are small and offer little practical significance. Additionally, the majority of variance remains unexplained, even when including relevant control variables. These findings suggest that the bar exam, as it was administered in Nevada to new lawyers during the study period, may not be the indicator of future career effectiveness for new attorneys that it should be, particularly given equity concerns. More research is needed, but this study finds that while the bar is serving as a significant barrier to the practice of law, there is little indication that it is a robust indicator of what it takes to be a "good" lawyer.

REFERENCES

- American Bar Association (1992). Legal Education and Professional Development – an Educational Continuum, *Report on the Task Force on Law Schools and the Profession: Narrowing the Gap*. [MacCrate Report]
- American Bar Association (2022a). *ABA Profile of the Legal Profession 2022*. <https://www.americanbar.org/content/dam/aba/administrative/news/2022/07/profile-report-2022.pdf>
- American Bar Association (2022b). Mental Health Character & Fitness Questions for Bar Admission. *Commission on Disability Rights Report*.

<https://www.americanbar.org/groups/diversity/disabilityrights/resources/character-and-fitness-mh/>

American Bar Association (2022c). *Legal Education and Admission to the Bar*.

https://www.americanbar.org/content/dam/aba/administrative/legal_education_and_admissions_to_the_bar/statistics/2022/2022-bpq-national-summary-data-race-ethnicity-gender-fin.pdf

Anderson IV, R., & Muller, D. T. (2019). The high cost of lowering the bar. *Geo. J. Legal Ethics*, 32, 307.

Austin, K. A., Christopher, C. M., & Dickerson, D. (2017). Will I Pass the Bar Exam: Predicting Student Success Using LSAT Scores and Law School Performance. *Hofstra Law Review*, 45(3), 753-784

Baird, L. L., Carlson, A. B., Reilly, R. R., & Powell, R. J. (1979). Defining competence in legal practice: The evaluation of lawyers in large firms and organizations. *ETS Research Report Series*, 1979(2), i-69.

Benthall-Nietzel, D. (1974). An empirical investigation of the relationship between lawyering skills and legal education. *Ky. LJ*, 63, 373.

Bowen, J. L., Leff, L. E., Smith, L. G., & Wolfsthal, S. D. (1999). Beyond the mystique of prestige: measuring the quality of residency programs. *American Journal of Medicine*, 106(5), 493-498.

Case, S. (2006). Frequently asked questions about scaling written test scores to the MBE. *The Bar Examiner*, 75(4), 42-44.

<https://thebarexaminer.ncbex.org/wp-content/uploads/PDFs/750406-Testing.pdf>

- Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, research & practice*, 15(3), 241.
- Duncan, G. J. & Murnane, R. J. (Eds.). (2011). *Whither opportunity? Rising inequality, schools, and children's life chances*. Russell Sage Foundation.
- Farkas, G. (2019). *Discrepancies by race and gender in attorney discipline by the State Bar of California: An empirical analysis*. The State Bar of California.
<https://board.calbar.ca.gov/docs/agendaItem/Public/agendaitem1000025090.pdf>
- Foster, S. (2021). Does the Multistate Bar Exam Validly Measure Attorney Competence? *Ohio St. LJ Online*, 82, 31.
- Grenardo, D. A. (2022). The phantom menace to professional identity formation and law success: imposter syndrome. *University of Dayton Law Review*, 47(3), 2.
<https://ecommons.udayton.edu/cgi/viewcontent.cgi?article=1744&context=udlr>
- Hatamyar, P. W., & Simmons, K. M. (2003). Are women more ethical lawyers-an empirical study. *Fla. St. UL Rev.*, 31, 785.
- Howarth, J. W. (2017). The case for uniform cut score. *Journal of the Legal Profession*, 42(1), 69-84.
- Howarth, J. W. (2019). What Law Must Lawyers Know? *Conn. Pub. Int. LJ*, 19, 1.

- Jencks, C., & Phillips, M. (Eds.). (1998). *The Black–White test score gap*. Brookings Institution Press.
- Kane, M. T. (2009). Reflections on bar examining. *The Bar Examiner*, 78(4), 5-20.
- Kinsler, J. S. (2017). Is bar exam failure a harbinger of professional discipline. *John's L. Rev.*, 91, 883.
- Klessig, J. M., Wolfsthal, S. D., Levine, M. A., Stickley, W., Bing-You, R. G., Lansdale, T. F., & Battinelli, D. L. (2000). A pilot survey study to define quality in residency education. *Academic Medicine*, 75(1), 71-73.
- Kuehn, R. R., & Moss, D. R. (2019). A Study of the Relationship Between Law School Coursework and Bar Exam Outcomes. *Journal of Legal Education*, 68(3), 623–649.
- Levin, L. C., Zozula, C., & Siegelman, P. (2013). A study of the relationship between bar admissions data and subsequent lawyer discipline. *Available at SSRN 2258164*.
- Merritt, D. J., & Cornett, L. (2020). Building a Better Bar: The Twelve Building Blocks of Minimum Competence. *AccessLex Institute Research Paper*, (21-02).
- Merritt, D. J., Hargens, L. L., & Reskin, B. F. (2001). Raising the bar: social science critique of recent increases to passing scores on the bar exam. *University of Cincinnati Law Review*, 69(3), 929-968.
- National Association for Law Placement (2022). *Jobs & JDs: Employment and Salaries of New Law Graduates – Class of 2021*.

- National Conference of Bar Examiners (2022). Phase 2 Report of the Testing Task Force. <https://nextgenbarexam.ncbex.org/reports/phase-2-report/>
- National Conference of Bar Examiners (2012). A Study of the Newly Licensed Lawyer: Conducted for the National Conference of Bar Examiners. https://iaals.du.edu/sites/default/files/documents/publications/building_a_better_bar_pre_print.pdf
- Patton, W. W. (2020). Admitting law graduates by bar examination versus by diploma privilege: A comparison of consumer protection. *J. Legal Prof.*, 45, 243.
- Plomin, R. (Accepted/In press). Celebrating a century of research in behavioral genetics. *Behavior Genetics*.
- Plomin, R., Shakeshaft, N. G., McMillan, A., & Trzaskowski, M. (2014). Nature, nurture, and expertise. *Intelligence*, 45, 46–59. <https://doi.org/10.1016/j.intell.2013.06.008>
- Rosin, G. S. (2008). Unpacking the Bar: Of Cut Score and Competence. *Journal of the Legal Profession*, 32, 67-98.
- Schwartz, R. A. (1973). The relative importance of skills used by attorneys. *Golden Gate UL Rev.*, 3, 321.
- Sebert, J. A. (2001). Legal education and professional development—An educational continuum. *MacCrate Report (Japanese edition)*.
- Shultz, M. M., & Zedeck, S. (2011). Predicting lawyer effectiveness: Broadening the basis for law school admission decisions. *Law & Social Inquiry*, 36(3), 620-661.

State Bar of California (2020). *The Practice of Law in California: Findings from the California Attorney Practice Analysis and Implications for the California Bar Exam*.

<https://www.calbar.ca.gov/Portals/0/documents/reports/2020/California-Attorney-Practice-Analysis-Working-Group-Report.pdf>

Wadsworth, S. J., Corley, R. P., Hewitt, J. K., Plomin, R., & DeFries, J. C.

(2002). Parent-offspring resemblance for reading performance at 7, 12 and 16 years of age in the Colorado Adoption Project. *Journal of Child Psychology and Psychiatry*, 43(6), 769-774. <https://doi.org/10.1111/1469-7610.00085>

Wilson, P., Nobles, T., & Scott, J. (2021). Equitable Access to the Legal

Profession Requires Greater Transparency. *AccessLex Institute Research Paper*.

Winick, M., Quintanilla, V. D., Erman, S., Chong-Nakatsuchi, C., & Frisby, M.

(2020a). Examining the California Cut Score: An Empirical Analysis of Minimum Competency, Public Protection, Disparate Impact, and National Standards. *AccessLex Institute Research Paper*.

Winick, M., Quintanilla, V. D., Erman, S., Chong-Nakatsuchi, C., & Frisby, M.

(2020b). A Five-Year Retroactive Analysis of Cut Score Impact: California's Proposed Supervised Provisional License Program. *AccessLex Institute Research Paper*.

Wright Jr., J. R., & Schachar, N. S. (2020). Necessity is the mother of invention:

William Stewart Halsted's addiction and its influence on the development

of residency training in North America. *Canadian Journal of Surgery*, 63(1), E13–E19. <https://doi.org/10.1503/cjs.003319>

APPENDIX

A.1 – Lawyering Effectiveness Scales

Lawyering effectiveness self-assessment scale (Scored 1-5, 5 items total)

1. (Intellectual & Cognitive Factors) Analysis and reasoning: Analytical skills, logic, and reasoning to approach problems and to formulate conclusions and advice
2. (Knowledge of the Law) Specific knowledge: Understanding of legal concepts and utilizing sources and strategies to identify issues and derive solutions
3. (Research & Information Gathering) Fact finding: Ability to identify facts and issues in a case
4. (Planning & Organizing) Organizing and managing one’s own work: Ability to generate well-organized methods and work products
5. (Communications) Writing: Ability to write clearly, efficiently, and persuasively

Supervisors, Peers, & Judge Scale (Scored 1-5, 26 items total for each rater across 8 categories)

Intellectual & Cognitive Factors

1. Analysis and reasoning: Analytical skills, logic, and reasoning to approach problems and to formulate conclusions and advice
2. Creativity and innovation: Thinks “outside the box;” develops innovative approaches and solutions
3. Problem solving: Effectively identifies problems and derives appropriate solutions
4. Practical judgment: Determines effective and realistic approaches to problems

Research & Information Gathering

1. Researching the law: Utilizes appropriate sources and strategies to identify issues and derive solutions
2. Fact finding: Able to identify relevant facts and issues in a case
3. Questioning and interviewing: Obtains needed information from others to pursue issue/case

Communications

1. Influencing and advocating: Ability to persuade others of positions and win support
2. Writing: Writes clearly, efficiently, and persuasively
3. Speaking: Orally communicates issues in an articulate manner consistent with issue and audience being addressed
4. Practical judgment: Accurately perceives what is being said both directly and subtly

Planning & Organizing

1. Strategic planning: plans and strategizes to address present and future issues and goals
2. Organizing and managing one's own work: generates well-organized methods and work products
3. Organizing and managing others: Organizes and manages others' work to accomplish goals

Conflict Resolution

1. Negotiation Skills : resolves disputes to the satisfaction of all concerned
2. Able to see the world through the eyes of others: Understands positions, views, objectives, and goals of others

Client & Business Relationships - Entrepreneurship

1. Providing advice & counsel & building relationships with clients: able to develop relationships with clients that address the clients' needs
2. Networking and business development: develops productive business relationships and helps meet the unit's financial goals

Working with others

1. Developing relationships within the legal profession: establishes quality relationships with others to work toward goals
2. Evaluation, development, and mentoring: manages, trains, and instructs others to realize their full potential

Character

1. Passion and engagement: demonstrates interest in the law for its own merits

2. Diligence: commitment to and responsibility for achieving goals and completing tasks
3. Integrity/honesty: has core values and beliefs; acts with integrity and honesty
4. Stress management: effectively manages pressure or stress
5. Community involvement and service: contributes legal skills to the community
6. Self-development: attends to and initiates self-development

A.2 – Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Self-evaluation score	4.20	0.43	2.50	5.00
Supervisor evaluation score	4.23	0.56	1.71	5.00
Peer evaluation score	4.38	0.49	2.33	5.00
Judge evaluation score	4.37	0.62	1.06	5.00
MBE score	147.50	12.34	107.00	177.00
MPT score	79.61	13.38	51.89	117.60
Subject essay score	626.40	44.77	509.30	736.40
Ethics essay score	77.56	9.49	55.66	100.50

Table A.2.1. Summary Statistics of the Variables of Interest

A.3 – Factor Loadings

Item	Principal Components				
	PC1	PC2	PC3	PC4	PC5
Intellectual & Cognitive	0.49	-0.01	0.41	-0.18	-0.75
Knowledge of the Law	0.45	0.41	0.47	0.50	0.42
Research & Information Gathering	0.42	0.57	-0.50	-0.48	0.11
Planning & Organization	0.43	-0.33	-0.59	0.57	-0.17
Communication	0.44	-0.63	0.14	-0.40	0.47

Eigenvalue	2.74	0.67	0.64	0.54	0.40
Standard Deviation	1.65	0.82	0.80	0.75	0.64
Proportion of Variance	0.55	0.13	0.13	0.11	0.08

Table A.3.1. Principal Components Analysis of Self-Evaluation Items

Principal Components								
Thematic Category	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
<i>Eigenvalues</i>	15.52	1.29	1.18	0.98	0.72	0.64	0.60	0.52
	0.60	0.05	0.05	0.04	0.03	0.02	0.02	0.02
<i>Proportion of variance</i>								
Thematic Category	Item Number	Factor Loading						
<i>Intellectual and Cognitive</i>	<i>Item 1</i>	0.82						
	<i>Item 2</i>	0.83						
	<i>Item 3</i>	0.84						
	<i>Item 4</i>	0.86						
<i>Research and Information Gathering</i>	<i>Item 5</i>	0.83						
	<i>Item 6</i>	0.72						
	<i>Item 7</i>	0.81						
<i>Communications</i>	<i>Item 8</i>	0.78						
	<i>Item 9</i>	0.69						
	<i>Item 10</i>	0.78						
	<i>Item 11</i>	0.80						
<i>Planning and Organizing</i>	<i>Item 12</i>	0.85						
	<i>Item 13</i>	0.84						
	<i>Item 14</i>	0.80						
<i>Conflict Resolution</i>	<i>Item 15</i>	0.75						
	<i>Item 16</i>	0.81						
<i>Client and Business Relations</i>	<i>Item 17</i>	0.73						
	<i>Item 18</i>	0.73						
<i>Working with Others</i>	<i>Item 19</i>	0.71						
	<i>Item 20</i>	0.82						
<i>Character</i>	<i>Item 21</i>	0.74						
	<i>Item 22</i>	0.80						
	<i>Item 23</i>	0.58						
	<i>Item 24</i>	0.74						
	<i>Item 25</i>	0.59						
	<i>Item 26</i>	0.73						

Table A.3.2. Principal Components Analysis for Individual External Evaluation Items

Note: This table represents the principal components analysis for one of the five possible external reviewers (supervisor one) as an example. All eigenvalues and principal component factor loadings are nearly identical for all other external reviewers.

A.4 – Simple Bivariate Regression Outputs

	(1) Self Eval	(2) Sup Eval	(3) Peer Eval	(4) Judge Eval
MBE score (0-1)	0.322** (0.115)	0.604*** (0.167)	0.277 (0.155)	0.580 (0.303)
Intercept	4.011*** (0.0692)	3.870*** (0.103)	4.216*** (0.0934)	4.017*** (0.194)
<i>N</i>	463	364	315	130
<i>R</i> ²	0.0168	0.0346	0.0102	0.0278

Table A.4.1. Bivariate Linear Relationship Between Lawyering Effectiveness and First-Attempt MBE Score

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1) Self Eval	(2) Sup Eval	(3) Peer Eval	(4) Judge Eval
MPT Essay (0-1)	0.0648 (0.102)	0.156 (0.152)	-0.0586 (0.145)	0.668* (0.306)
Intercept	4.166*** (0.0481)	4.158*** (0.0726)	4.401*** (0.0669)	4.060*** (0.145)
<i>N</i>	433	339	292	119
<i>R</i> ²	0.0009	0.0031	0.0006	0.0390

Table A.4.2. Bivariate Linear Relationship Between Lawyering Effectiveness and First-Attempt MPT Essay Score

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1) Self Eval	(2) Sup Eval	(3) Peer Eval	(4) Judge Eval
Ethics Essay (0-1)	-0.0335 (0.0990)	0.254 (0.144)	-0.0792 (0.140)	0.391 (0.287)
Intercept	4.209***	4.102***	4.413***	4.139***

	(0.0532)	(0.0766)	(0.0725)	(0.166)
<i>N</i>	433	339	291	119
<i>R</i> ²	0.0003	0.0091	0.0011	0.0155

Table A.4.3. Bivariate Linear Relationship Between Lawyering Effectiveness and First-Attempt Ethics Essay Score

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1) Self Eval	(2) Sup Eval	(3) Peer Eval	(4) Judge Eval
Subject Essay (0-1)	0.223* (0.105)	0.428** (0.156)	-0.0947 (0.150)	0.926** (0.305)
Intercept	4.077*** (0.0581)	4.001*** (0.0874)	4.423*** (0.0812)	3.842*** (0.177)
<i>N</i>	433	339	291	119
<i>R</i> ²	0.0104	0.0218	0.0014	0.0731

Table A.4.4. Bivariate Linear Relationship Between Lawyering Effectiveness and First-Attempt Subject Essay Score

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.5 – OLS Regression Output Table

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Simple - Self	Full - Self	Simple - Super- visor	Full - Super- visor	Simple - Peer	Full - Peer	Simple - Judge	Full - Judge
MBE score (0-1)	0.34 * (0.14)	0.43 ** (0.07)	0.28 (0.21)	0.25 (0.19)	0.53 ** (0.19)	0.43 * (0.23)	0.28 (0.38)	0.41 (0.71)
MPT Essay score (0-1)	-0.05 (0.12)	-0.06 (0.17)	-0.09 (0.18)	-0.03 (0.19)	0.00 (0.18)	0.00 (0.39)	0.13 (0.40)	0.05 (0.50)
Subject Essay score (0-1)	0.13 (0.15)	0.11 (0.17)	0.32 (0.22)	0.09 (0.18)	-0.33 (0.21)	-0.34 (0.20)	0.65 (0.47)	0.66 (0.85)
Ethics Essay score (0-1)	-0.09 (0.11)	-0.13 (0.10)	0.14 (0.15)	0.05 (0.24)	-0.09 (0.15)	-0.08 (0.20)	0.12 (0.30)	0.04 (0.28)
>1 attempt		0.02 (0.08)		-0.15 (0.10)		-0.07 (0.06)		0.01 (0.26)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Simple - Self	Full - Self	Simple - Super- visor	Full - Super- visor	Simple - Peer	Full - Peer	Simple - Judge	Full - Judge
Years since exam		0.01 (0.01)		0.05 ** (0.01)		0.04 * (0.02)		0.01 (0.03)
Race - White		-0.09 (0.07)		0.04 (0.15)		0.04 (0.15)		0.15 (0.23)
Race - Hispanic/Latino		-0.14 (0.06)		0.12 (0.07)		0.10 (0.13)		-0.03 (0.50)
Race - Black		-0.04 (0.13)		0.02 (0.07)		0.27 (0.08)		0.37 (0.23)
Race - Asian/PI		-0.03 (0.10)		0.09 (0.12)		0.07 (0.12)		0.36 (0.22)
Race - Other		-0.07 (0.07)		0.09 (0.13)		0.12 (0.07)		0.31 (0.23)
Female		0.13 ** (0.04)		0.07 (0.06)		0.01 (0.06)		0.14 (0.13)
Age		-0.00 (0.00)		-0.01 (0.01)		-0.01 (0.01)		-0.01 (0.02)
UNLV (Boyd) grad		-0.02 (0.03)		0.05 (0.10)		-0.03 (0.06)		-0.06 (0.04)
R^2	0.02	0.06	0.03	0.07	0.03	0.07	0.08	0.12
$Adj. R^2$	0.02	0.02	0.02	0.03	0.02	0.02	0.05	-0.00
N	429	424	337	333	289	286	119	118

Table A.5.1. Multivariate OLS Linear Regression Relationships Between Lawyering Effectiveness and Bar Exam Components

NOTE: Standard errors in parentheses. “Simple” models refer to OLS models run without control variables, while “full” refers to OLS models run with all relevant control variables.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$