

## Guidelines for a Licensing System Based on Supervised Practice

Logan Cornett,<sup>1</sup> Danette McKinley,<sup>2</sup> and Deborah Jones Merritt<sup>3</sup>

Licensing systems protect the public by certifying only candidates who are minimally competent to practice the profession. Many systems use written exams, along with other requirements, to measure that competence. Measurement experts, however, have long recognized that assessments of workplace performance can also form the core of a licensing system. “The time-honored way to find out whether a person can perform a task,” one group of highly regarded experts wrote, “is to have the person try to perform the task.”<sup>4</sup>

These experts come from the field of psychometrics, an academic field that focuses on techniques for measuring knowledge, skills, attitudes, and other facets of human cognition. Assessment experts set three primary criteria for use in the selection of licensing instruments: reliability, validity, and fairness.<sup>5</sup>

- **Reliability** means that an assessment produces consistent results. A reliable bathroom scale registers the same weight if you step on the scale twice within a minute.
- **Validity** means that evidence establishes a link between the assessment outcomes and the purpose for which those outcomes are used. It is valid to use bathroom scale readings to determine the body weight of able-bodied people, but not to determine what they ate for breakfast. Psychometricians stress that validity is not a property of the assessment itself, but of the interpretations made based on the scores.
- **Fairness** means that the assessment does not discriminate, explicitly or implicitly, based on characteristics that are irrelevant to the quality being measured. Fairness also requires assessors to treat candidates with respect. A bathroom scale is not a fair measure of body weight for a person wearing a heavy leg brace who cannot stand without that brace.

---

<sup>1</sup> Director of Research, IAALS (Institute for the Advancement of the American Legal System).

<sup>2</sup> Director of Diversity, Fairness, and Inclusion Research, National Conference of Bar Examiners (NCBE).

<sup>3</sup> Distinguished University Professor and John Deaver Drinko/Baker & Hostetler Chair in Law Emerita, The Ohio State University.

<sup>4</sup> Michael Kane, Terence Crooks & Allan Cohen, *Validating Measures of Performance*, EDUC. MEASUREMENT: ISSUES & PRACTICE, Summer 1999, at 5. Dr. Kane holds the Samuel J. Messick Chair in Test Validity at the Educational Testing Service. He previously served as Director of Research at NCBE and held faculty positions at several universities.

<sup>5</sup> AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION, NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING (2014).

Scholars have identified two other criteria that may affect the choice of licensing instruments: alignment and feasibility.<sup>6</sup>

- Alignment means that education and licensing align to produce new professionals with the knowledge, skills, and judgment needed for entry-level practice. For example, if negotiation is an important skill in law practice, then including negotiation tasks in a licensing assessment should motivate students to work on their negotiation skills.
- Feasibility means that the instrument used should be affordable and efficient; the test should not be too costly for either administrators or test takers. Feasibility, however, should not impede change: new approaches often seem more costly than established ones because the costs of the latter have become normalized.

A substantial body of psychometric literature shows that assessments of workplace performance can offer a reliable and valid means for determining competence. Fairness, alignment, and feasibility have been less extensively studied in the context of performance assessments, but an emerging literature suggests that these criteria can also be satisfied in that context. This handout offers some guidelines from the educational literature for designing a licensing system for lawyers based on post-graduate supervised practice. Similar principles would guide design of a system based on an experiential education path, but a future handout will address that pathway.

1. Begin with an Evidence-Based Definition of Competence. A sound assessment system rests on an evidence-based definition of the quality being measured. For a licensure system, that quality is minimum competence to practice the profession. In law, this aspect of the validity argument is based on identifying the tasks newly licensed lawyers perform, as well as the knowledge, skills, and judgment they need to perform those tasks. A national study by IAALS (Institute for the Advancement of the American Legal System), *Building a Better Bar: The Twelve Building Blocks of Minimum Competence*, identifies the competencies that newly licensed lawyers and their employers associate with entry-level practice.<sup>7</sup> A recent practice analysis by the National Conference of Bar Examiners (NCBE), the *Testing Task Force Phase 2 Report*, offers another helpful overview of the competencies and tasks newly licensed lawyers perform.<sup>8</sup> Some states have conducted similar analyses within their own jurisdiction. For example, the State Bar of California did so with a practice analysis published in 2020.<sup>9</sup>

---

<sup>6</sup> John J. Norcini & Danette W. McKinley, *Assessment Methods in Medical Education*, 23 TEACHING & TEACHER EDUC. 239 (2007).

<sup>7</sup> The report is available at <https://iaals.du.edu/publications/building-better-bar>.

<sup>8</sup> This report is available at <https://nextgenbarexam.ncbex.org/reports/phase-2-report/>.

<sup>9</sup> California's report is available at <https://www.calbar.ca.gov/Portals/0/documents/reports/2020/California-Attorney-Practice-Analysis-Working-Group-Report.pdf>.

Identifying the knowledge, skills, and judgment that new professionals use to perform common tasks is essential to define *what* needs to be measured. The definition of competence also provides guidance for *how* the requisite knowledge and skills should be measured.

2. Gather Rich Data for Each Candidate. The Uniform Bar Exam (UBE) achieves high reliability in part because it gathers many data points: answers to 175 scored multiple-choice questions, six essay responses, and two documents created for performance tests. These varied data points also support interpretations made about scores. A licensing system based on supervised practice can also achieve sufficient reliability, while supporting broader claims of competence, if it collects many data sources to support interpretations of the candidate's performance. There are several instruments that have been used in the health professions, teacher certification, and human resources field to evaluate workplace performance. This research suggests that an assessment system based on supervised practice should:

- Collect information from multiple sources (supervisors, peers, staff, the candidate, and potentially clients);
- Measure a range of relevant knowledge and skills;
- Use different formats; and
- Rely upon multiple observations.

Materials like these can be assembled into a portfolio. The richness of the data addresses reliability and allows collection of sufficient evidence of validity. Candidates can be assessed on many competencies and tasks, providing evidence to support interpretations made about competence. Multiple assessments from a wide number of raters can be monitored for potential bias, promoting fairness. Because the measures sample the full range of knowledge and skills, the multiple data sources contribute to the reliability of workplace-based assessment, supporting the decisions made about performance. With a large number of data points, errors associated with a single assessment are reduced. The portfolio as a whole is likely to paint a reliable portrait of the candidate's competence.

3. Give Candidates Feedback and Opportunities to Improve. The supervision period should provide opportunities for candidates to grow and learn from their mistakes. Supervised practice is an important part of the education process, and it is essential to provide guidance towards improvement. Initial months should focus on formative assessments designed to provide feedback rather than summative assessments used for decision-making. That focus benefits both the candidate and the employer, because the candidate will learn to provide better service.

Implementation of this guideline is particularly important in promoting fairness and reliability. Providing similar opportunities for recent graduates to hone their competence, especially for

skills and knowledge that are not taught in law school, addresses aspects of fairness in education and employment. Research, meanwhile, suggests that supervisors are more comfortable providing negative feedback if they know that candidates will have an opportunity to improve. This safety net enables assessors to give negative feedback when warranted—and to fail candidates who have not responded adequately to that feedback. Training and feedback during supervised practice also aligns with cognitive science on the development of expertise;<sup>10</sup> candidates are provided with opportunities to progress towards goals through practice and the provision of feedback.

4. Take Certification Out of the Workplace. Supervisors, peers, and other members of a candidate's workplace offer essential information about the candidate's competence. Indeed, an employer's willingness to allow a candidate to handle client matters offers strong evidence of the candidate's competence. It is best, however, for an independent decision maker to make the final decision about a candidate's eligibility for a license. This reinforces the acceptance of supervised practice as a learning opportunity for candidates, and supports the focus on formative feedback and the ongoing development of knowledge and skills.

This structure also lends credibility to the assessment system. Equally important, it allows candidates to change supervisors during the assessment period, which helps protect them from abusive or unethical supervisors. In these ways, the structure promotes fairness, reliability, and validity.

There are several models for independent certification. In the simplest model, a trained examiner reviews the candidate's portfolio, using established rubrics and standards, and determines whether the portfolio establishes minimum competence. If the examiner has doubts—or rules against the candidate—then a panel of examiners reviews the portfolio and reaches a consensus decision on the candidate's competence.

5. Use Credible Supervisors and Examiners. The reliability, fairness, and validity of a licensing system depend greatly on the credibility of the professionals administering the system. In a system that relies upon supervised practice, the supervisors should be licensed lawyers within the state who have demonstrated an interest in training and mentoring new lawyers. Their disciplinary records should be clean or demonstrate clear rehabilitation. A supervisor should have at least three years of experience practicing law, with at least two of those years in the state. Greater seniority does not necessarily spell better supervisory competence. Lawyers in government agencies, nonprofit organizations, and small law firms assume primary

---

<sup>10</sup> K. Anders Ericsson & Kyle W. Harwell, *Deliberate Practice and Proposed Limits on the Effects of Practice on the Acquisition of Expert Performance: Why the Original Definition Matters and Recommendations for Future Research*, FRONTIERS IN PSYCH. 10:2396 (2019).

responsibility for client matters very early in their careers. These junior lawyers may also be more in touch with new practice methods than their senior colleagues.

Examiners, similarly, should have experience practicing law in the state. It may be appropriate to require more years of experience (five or seven years) for an examiner than for a supervisor. It is best for examiners to devote only part of their time to portfolio examination, while they maintain positions practicing law or teaching in clinics. The latter work keeps them rooted in the profession, enhancing their credibility.

Diversity, equity, and inclusion are key components of credibility for both supervisors and examiners. The state's high court should assemble a group of examiners who are demographically diverse, represent different parts of the state, work in different practice areas, and come from a variety of organization types. Supervisors should be similarly diverse and should work in organizations that have demonstrated a commitment to diversity, equity, and inclusion.

6. Provide Training and Support for Supervisors and Examiners. Good supervisors and examiners will bring their own expertise to the licensing system. It is not necessary to erase differences of opinion among these experts or train them to reach agreement on every aspect of a candidate's portfolio. One of the strengths of a licensing system based on supervised practice is that it recognizes nuances and differences in approaches: These are the hallmarks of a profession.

Supervisors and examiners, however, will benefit from regular training on matters such as the purpose of the assessments used; the scoring of those assessments; methods to avoid implicit bias; approaches to providing constructive feedback; and ways to inform candidates that their work is not competent. It is also useful for examiners to meet periodically and review sample portfolios together; this can help individual examiners expand their perspectives. Supervisors may also appreciate a support network that allows them to share tips on providing constructive feedback and discuss other aspects of their role.

7. Make the System User Friendly. Training new lawyers, providing feedback, and rating performance take time, but a user-friendly system will greatly reduce that time commitment. System designers should solicit input from supervisors and new lawyers to understand the tools that will help users provide efficient, effective feedback. Some workplaces may have existing practices that will inform the new system. Best practices in performance review could help in implementing the system.

Supervisors and candidates often find online systems convenient. These systems also allow for ready sharing of materials with examiners. Some fields have even developed apps for providing feedback or assessments through smart phones. Designers, however, may prefer to use a paper

system during the early phase of implementation. This allows users to identify flaws and suggest improvements before putting the system online.

8. Design High-Quality Feedback and Assessment Tools. A good licensing system will use a variety of tools for feedback and assessment—just as the bar exam uses varied question formats. The literature offers these tips for designing fair and reliable tools that contribute to valid uses:

- Tie feedback and assessment to specific components of competence.
- Operationalize those components with prompts that remind supervisors of actions signifying levels of developing competence.
- Avoid numerical ratings on the forms: supervisors find them difficult to assign and err towards high ratings. Candidates also receive little guidance from numbers. Scores, however, may be built into the system for review by examiners.
- Encourage narrative feedback and ensure that the feedback is shared with the candidate.

The final pages of this handout show two sample feedback forms following this guidance. The forms are based on competencies identified by the *Building a Better Bar* report. The sample forms focus on a client interview; other forms would focus on different tasks and highlight different competencies.

The sample forms offer just two options for providing feedback and assessing competence. A candidate's final portfolio should include multiple feedback forms of different types, reflections from the candidate, and samples of the candidate's work product (with client-identifying information deleted).

9. Be Transparent. Transparency is essential to fairness. Candidates should know who will evaluate them, how they will be evaluated, and the criteria for successful performance. Transparency also enhances reliability and valid uses of assessment because the system is open to criticism and improvement.

10. Evaluate and Adjust. Assessment systems evolve as they benefit from experience and encounter new challenges. A good system will provide for regular evaluation and feedback from stakeholders. For a lawyer licensing system, those stakeholders include practicing lawyers and judges, legal educators, law students, recent graduates, clients, and the public at large. Regular assessment will assure that the licensing system retains credibility and protects the public.

## Bibliography

David A. Cook et al., *When Assessment Data Are Words: Validity Evidence for Qualitative Educational Assessments*, 91 ACAD. MED. 1359 (2016).

David A. Cook et al., *A Contemporary Approach to Validity Arguments: A Practical Guide to Kane's Framework*, 49 MED. EDUC. 560 (2015).

Anthony A. Donato et al., *Validity and Feasibility of the Minicard Direct Observation Tool in 1 Training Program*, J. GRAD. MED. EDUC., June 2015, at 225.

Marjan Govaerts & Cees P.M. van der Vleuten, *Validity in Work-Based Assessment: Expanding Our Horizons*, 47 MED. EDUC. 1164 (2013).

Michael T. Kane, *Validating the Interpretations and Uses of Test Scores*, 50 J. EDUC. MEASUREMENT 1 (2013).

Michael Kane, Terence Crooks & Allan Cohen, *Validating Measures of Performance*, EDUC. MEASUREMENT: ISSUES & PRACTICE, Summer 1999, at 5.

Nyoli Valentine et al., *Fairness in Human Judgment in Assessment: A Hermeneutic Literature Review and Conceptual Framework*, 26 ADVANCES IN HEALTH SCIENCES EDUC. 713 (2021).

Nyoli Valentine et al., *Making It Fair: Learners' and Assessors' Perspectives of the Attributes of Fair Judgment*, 55 MEDICAL EDUC. 1056 (2021).

John J. Norcini & Danette W. McKinley, *Assessment Methods in Medical Education*, 23 TEACHING & TEACHER EDUC. 239 (2007).

Cees P.M. van der Vleuten & Lambert W.T. Schuwirth, *Assessing Professional Competence: From Methods to Programmes*, 39 MEDICAL EDUC. 309 (2005).

Cees P.M. van der Vleuten, *Revisiting "Assessing Professional Competence: From Methods to Programmes,"* 50 MEDICAL EDUC. 885 (2016).

John Q. Young & Matthew McClure, *Fast, Easy, and Good: Assessing Entrustable Professional Activities in Psychiatry Residents with a Mobile App*, 95 ACAD. MED. 1546 (2020).

**Feedback: Initial Client Interview**

**Date**\_\_\_\_/\_\_\_\_/\_\_\_\_

**Candidate:**\_\_\_\_\_

**Observer:**\_\_\_\_\_

Circle points that the candidate does competently. Place an X over those needing work. Write "NA" for any that do not apply.

PROFESSIONALISM

Explained Unlicensed Status

Explained A/C Privilege

Explained any Fees

Treated Client with Respect

Identified Client Needs, Interests, and Concerns

Provided Contact Information

Promised Follow Up within Specified Time

**Comments:**

-----

CLIENT INTERACTION & COMMUNICATION

Established Rapport

Let Client Tell Story

Avoided Jargon

Explained Legal Concepts

Invited Questions

Answered Questions

Described Next Steps

**Comments:**

-----

KNOWLEDGE & ISSUE IDENTIFICATION

Identified Issues

Stated Law Correctly

Acknowledged Need to  
Check/Research New Points

**Comments:**



**Client Interview**

**Date:**

**Candidate Name:**

**Supervisor Name:**

<b>Competency</b>	<b>Requires Corrective Response</b>	<b>Developing</b>	<b>Competent</b>
<b>Professionalism</b>	Does not disclose intern status	Discloses status but does not invite questions	Discloses status and invites questions/concerns from client
	Does not mention attorney client privilege	Mentions attorney client privilege but does not discuss nuances	Explains attorney client privilege fully, including when disclosure is required
	Does not explain fees (if applicable)	Mentions fees but does not give sufficient detail or invite questions	Describes fees in detail and invites questions
	Does not treat client with respect	Generally treats client with respect	Treats client with respect throughout interview
	Seems uninterested in client needs, interests, and concerns	Shows interest in client but does not sufficiently follow up on needs, interests, and concerns	Fully engages with client needs, interests, and concerns
	Provides no contact information	Provides information, but does not invite contact	Provides information and invites contact
	Does not mention follow up	Mentions follow up but does not give specifics	Promises specific follow up within specified time
<b>Client Interaction and Communication</b>	Questions client without first attempting to establish rapport	Establishes some rapport with client but the relationship is strained	Establishes sufficient rapport that client evidences trust in lawyer
	Uses highly structured questions rather than allowing client to tell story	Allows client to tell story but interrupts often	Lets client tell full story with only essential interruptions
	Repeatedly uses jargon without explanation	Occasionally uses jargon but clarifies in response to client questions	Uses jargon only when essential and explains without client prompting
	Uses legal concepts without explaining them	Explains some but not all legal concepts	Explains all legal concepts
	Does not invite any questions from client	Invites questions but seems unreceptive	Repeatedly invites questions and shows interest in those questions
	Ignores questions posed by client	Answers some, but not all, questions posed by client	Answers all questions posed by client
	Does not describe next steps	Mentions next steps, but does not explain them	Describes next steps in terms client readily understands
<b>Knowledge and Issue Identification</b>	Does not identify legal issues raised by the interview	Raises some, but not all, legal issues raised by the interview	Properly identifies all legal issues raised by the interview
	States several points of law incorrectly	States one point of law incorrectly	States all points of law correctly
	Claims full knowledge when research is necessary	Skirts over need for research rather than acknowledging need directly	Acknowledges need to check or research unfamiliar points of law

**Comments:**