

PRIMER

TOWARD MEANINGFUL REFORMS IN THE BAR EXAMINATIONS*

*Vicente V. Mendoza***

1. *Why is there a need to institute reforms in the bar examinations?*

The bar examination system was established in the Philippines more than a century ago. Since then the system has undergone only minor changes, mostly procedural, rather than substantive or methodological, in nature. Over the years, there have been growing concerns regarding its reliability as a gateway to the legal profession. These concerns include the following:

- The reliability of the bar examinations as a test of professional competence when examinees are tested solely on the basis of essay-type questions.
- The lack of consistency in grading procedures.
- The question of subjectivity or bias arising from the formulation of test questions and the reading and marking of test answers by a single individual (the bar examiner) in each test area.
- Significant year-to-year variations in the level of difficulty of the examinations which, in effect, “penalize” or “reward” examinees, as the case may be.
- The time it takes to administer the bar examinations and release its results.
- The need for computerization or automation to facilitate testing, grading, and reporting of test results.

* Justice Mendoza's extended article on this topic will be published in the next issue of the Philippine Law Journal.

** Associate Justice, Supreme Court of the Philippines and Chair, Committee on the 2002 Bar Examinations. University of the Philippines College of Law, LL.B., 1957; Yale Law School, LL.M., 1971.

- The insufficient determination of an examinee's character and fitness for admission to the Bar.
- The need for a permanent bar examining authority with a tenured membership.

2. *What types of reforms are being proposed to address these issues?*

The proposed reforms can be grouped into three broad categories:

A. Structural and policy reforms, particularly

- the appointment of a tenured Board of Bar Examiners in lieu of *ad hoc* Committees on Bar Examinations appointed every year;
- the creation of readership panels for each subject area to address the issue of bias or subjectivity and facilitate the construction of test questions and the correction of examination booklets;
- the creation of an advisory committee to assist the Board and the Supreme Court and to address related issues in legal education to better prepare law school graduates in taking the bar examinations; and
- the provision for character and fitness investigation as a prerequisite for taking the bar examinations.

B. Changes in the design and construction of test questions, particularly

- the introduction of objective multiple-choice questions in the bar examinations;
- the formulation of essay test questions and "model" essays by more than one bar examiner; and
- the introduction of performance testing by way of revising and improving the essay examination on Legal Ethics and Practical Exercises.

C. Methodological reforms, particularly

- the adoption of the calibration method to correct variations in the level of test difficulty and grader leniency;
- consideration of alternative grading methodologies, such as scaling, to promote test equity and further standardize levels of test difficulty and grade interpretation; and
- further computerization or automation of the bar examinations to facilitate application, testing, and reporting procedures.

3. *How would these reforms speed up the administration of the bar examinations?*

The proposed reforms would facilitate each phase of the examination process (from the filing of test applications to the reporting of test results) through the following:

- The provision for character and fitness investigation, which would disqualify from taking the bar examinations those who fail it and delay the examination of applicants who are under investigation. This would reduce the number of bar examinees in any given year.
- The introduction of a multiple-choice section which can be corrected in a faster and automated manner.
- The appointment of readers to assist the examiner in reading and grading the essay test booklets for each subject.
- The adoption of the calibration method which would facilitate the correction of essay test booklets based on a uniform and defined set of grading standards and criteria for each test subject.
- The computerization and automation of application processing, test construction, scoring, and reporting phases of the bar examinations.

4. *How would these reforms make the bar examinations a more reliable gauge of legal competency?*

- By giving equal emphasis to merit and fitness. The new provision for a thorough character and fitness screening process *before* any examinee is allowed to take the bar examinations underscores the fact that admission to the bar cannot be satisfied simply by passing the written examinations, but that the determination of good moral character and mental and emotional stability of the applicant to practice law is equally important. In other countries, such as the United States, the written examinations and the character and fitness requirements constitute the twin pillars of the bar admission process.
- By introducing objective multiple-choice questions. This section of the examinations would require an examinee to choose the best possible answer out of four possibly correct answers to each question. This would, in effect, simulate the competing options and strategies that a lawyer frequently has to choose from in taking a position as an advocate or in analyzing the legal relationships arising from a factual situation. Indeed, the multiple-choice questions can be an excellent exercise for developing precision on the part of the examinees.
- By introducing a performance test. Performance testing would assess the examinee's ability for case planning, problem solving, factual investigation, and other skills and methods that are critical to the competent practice of law but are not necessarily measured by the traditional essay and multiple-choice questions. This will not only improve the overall validity of the bar examinations by testing for a broader range of lawyering skills, but it will also emphasize skills training as a necessary part of the education of every lawyer.
- By calibrating the essay test answers of the bar examinees so that grading procedures and practices will be consistent and equitable.
- By proposing alternative scoring methods such as scaling. A scaled score, in contrast to a raw score or weighted average, takes into consideration the varying levels of test difficulty from one administration of the bar examination to another so that a particular scaled score is indicative of approximately the same level of proficiency as the identical scaled score on any bar examination. This would allow for a correct interpretation of the grades obtained by examinees in any bar examination, which is not reflected under the present system.

By establishing a Bar Admissions Advisory Committee to advise and work with the Board of Bar Examiners in ensuring that the questions asked in the bar examinations reliably measure the basic standards and requirements of legal competency and in designing other pertinent reforms in the bar admissions process.

5. *How different would the proposed Board of Bar Examiners be from the present Committee on Bar Examinations?*

Quite different. The Board of Bar Examiners will consist of bar examiners and readers who will be appointed for a fixed term of office to ensure continuity of policy and consistency of examination practices, in other words, to professionalize the administration of the bar examinations. The Board of Bar Examiners will be vested with additional powers and responsibilities necessary to implement the proposed reforms in the bar examinations in collaboration with the Bar Admissions Advisory Committee. Decision-making within the Board would also be collegial.

The identity of the bar examiners need no longer be confidential. Secrecy serves no useful purpose and only breeds unwarranted speculation, especially on the part of the examinees. It may be mentioned in this connection that the members of examining boards for various professions under the Professional Regulation Commission are publicly known, and no complaint has been heard that this makes them susceptible to influence. The selection of examiners who are known for their probity is the best guarantee of the integrity of the bar examinations.

The Board will remain as an agency of the Supreme Court and continue to be chaired by a member of the Court.

6. *How would the structure of the redesigned bar examinations look?*

The written bar examinations will continue to be administered annually and the anonymity of the examinees will be preserved.

There will be two main sections in the bar examinations: a multiple choice section and an essay section. There will be over a hundred multiple choice questions covering various subjects, including those which may not be covered or represented

by the essay subjects. The essay portion shall consist of the current eight law subjects. The Board of Bar Examiners shall determine whether a particular subject will have one or more essay test questions.

7. *What would be the weight assigned to the multiple-choice and essay sections?*

The Board of Bar Examiners will determine the weight or percentage to be assigned to each section of the examination as well as the weight that each essay question shall carry. Typically, the multiple-choice section is assigned less weight (e.g., 40 percent) than the essay section, which includes the performance test (e.g., 60 percent).

8. *How would the multiple-choice section be corrected?*

For the multiple choice section, examinees will be asked to choose the best answer from four stated alternatives. Credit will be given only if the examinee has selected the best answer. Examinees shall be asked to mark only one answer for each question, and multiple answers will not be counted. A correct answer to a multiple-choice section will yield a single point for the examinee.

The test booklets for the multiple-choice questions will be corrected separately from the essay section. Although this section can be corrected manually by using an answer key developed and approved by the Board of Bar Examiners and the Bar Admissions Advisory Committee, test correction and scoring will be greatly facilitated by using computer key pads.

9. *Would not multiple-choice questions encourage memorization, rather than problem-solving, on the part of the bar examinees?*

No. There need be no fear that multiple-choice questions will easily lend themselves to memorization, or even guesswork. To the contrary, this type of questions can be a combination of rigor and precision as long as the examiners are properly trained and they take corresponding care and precision in the construction of the test questions. A given multiple-choice question may involve a choice of the applicable statute, theory of liability, or comparable principle of law. Questions can be so designed as to require the examinee to analyze the legal relationships arising from a factual situation or to take a position as an advocate. Some questions could also be designed to call for suggestions about interpreting, drafting, or counseling that might lead to more effective structuring of a transaction. In short, multiple-

choice questions can be designed to measure an examinee's problem-solving skills rather than his/her ability to simply recall pertinent legal information.

10. *How would the performance test be constructed and administered?*

The performance test will be in the form of an essay test question prepared by the bar examiner assigned to the Legal Ethics and Practical Exercises panel. It is proposed that this redesigned test subject be given greater weight or percentage of an examinee's final grade.

The performance test question will consist of performing an assigned task using given resource materials. The task may include the preparation of a brief in support of a motion, a memorandum evaluating grounds for objecting to the probate of a will, a simple complaint, or some other legal document. The assignment may raise an ethical issue as well. The difference between the performance test and the essay test questions in the seven other subjects is that it will be an open book examination. For this purpose, examinees will be provided with the "record" of a hypothetical case containing the facts and a "library" containing cases, statutes, or regulations.

11. *What changes would be instituted for the seven other essay subjects?*

The test question(s) for each essay subject will be developed and reviewed by more than one individual (the bar examiner). The Bar Admissions Advisory Committee will provide a critique of the proposed test question.

The "model" essay answer will be "refined" further after the sampled examinee essay test responses are calibrated (see # 12). After the examinations have been given, the Committee will critique the "model" essay answer.

The essay questions should be designed to test the examinee's ability to analyze a given set of facts, to identify the issues involved and the applicable principles of law, and to reason to a sound conclusion. The examinee should be asked to state his conclusion at the beginning of his answer, followed by his analysis and reasoning. The essay question may cover two or more subject areas, whether or not such areas are covered in any of the traditional eight test subjects in the bar examinations.

12. *How would the calibration of the essay questions be performed?*

Calibrating essay test answers is widely practiced by bar examining authorities in the United States where it originated. Similar to the issue it seeks to address in our country's bar admissions process, its adoption there was in response to the acknowledged lack of mathematical precision in grading the relatively subjective essay section.

Calibration refers to the process of estimating the level of difficulty of an essay test question (including the performance test question) by initially determining the representative or "benchmark" examinee test answers before all the essay test booklets are read and marked. Point values are also determined and distributed by the readers based on the "calibrated" essay answers.

13. *How would calibration help make the bar examinations a more equitable measure of legal competency?*

First, calibration procedures place equal emphasis on what the graders and experts in the field determine to be the "model answers," and what can be reasonably expected of an examinee based on a representative sampling or unbiased selection of examinee test booklets. Second, the assignment of several readers to each test subject to perform the calibration exercise eliminates subjectivity and "tailor-fitting" (*i.e.*, the tendency of examinees to speculate on the identity of the bar examiners and to try to fit their answers to test questions to the style and inclination of the perceived examiners). Third, the calibration method has a built-in "appeals" or "resolution" process for the benefit of examinees whose final grades fall within the "twilight zone," *i.e.*, whose grades are slightly below the predetermined "cut-off" or passing standard.

The expected outcome of the calibration process is that the passing standards in any bar examination will be reasonable and representative of demonstrated performance without sacrificing the acceptability of these standards to the bar examining authority.

14. *How would point values be distributed for a "calibrated" model essay?*

The Board of Bar Examiners, or its readership panels, should choose between two essay grading methods: the analytic and the holistic. Either method is fine as long as the essay question is properly constructed.

(1) In the analytic method, the calibrated model essay is broken down into its key elements or factors. The total score for an essay is distributed among these elements or factors. The examinee receives the corresponding point value if his/her essay contains any of these elements.

(2) In the holistic method, a scale or continuum of points is initially constructed. The scale indicates increasing levels of an examinee essay's overall quality relative to the calibrated model essay. Hence, a examinee could receive only a point if his essay is very poor compared to the calibrated model essay, while another examinee could get the highest grade of 5 points for a comparatively excellent essay.

15. *What if the sampled examinee test booklets for a given essay(s) do not yield any "majority" trend for calibration purposes?*

That is unlikely to happen. If it does, the model essay answer developed by the readership panel(s) at the polling stage should serve as the standard for purposes of correcting the examinee essays and for distributing point values.

16. *Would the calibration exercise result in the lowering of standards for correcting or scoring the essay test questions?*

Not at all. Calibration is not simply a matter of determining a model essay answer and assigning point values to its key components on the basis of what the representative or sampled essay test booklets yield. Rather, it requires that the sampled essay test responses be initially "matched" against the grading parameters or guidelines previously established by the readership panels during a so-called polling stage. The readership panels will accept into the calibrated model essay only those aspects or items in the sampled essays that meet (or exceed) these predetermined guidelines. Thus, there could even be instances where the calibrated model essay is superior to the initial guidelines set forth by the readership panels during the polling phase.

17. *Would there be a passing grade required for each essay test question or for each section of the bar examinations?*

No. There is no minimum grade or passing grade for each essay question

or for either the multiple-choice or essay sections of the bar examinations. The examinee's final grade is based on the combination of scores from the multiple-choice and essay (which includes the performance test) portions of the bar examinations.

18. *How is the passing standard in the bar examinations determined?*

Either arbitrarily by setting a predetermined "cut-off" score to meet or based on a statistical process of determining the distribution or spread of the examinee scores. The latter approach, which is preferred on account of its objectivity, is usually performed with the help of a consulting statistician or psychometrician (a statistician who specializes in educational testing and measurement).

Once the passing or "cut off" mark is established, all examinees whose scores are equal to or greater than the "cut off" mark will be deemed to have passed the bar examinations. The Board of Bar Examiners, either arbitrarily or with the help of a statistician/psychometrician, shall determine a range of scores below the "cut-off" mark which shall constitute a "Conditional" category (or "twilight zone"). The essay test booklets of examinees in this category shall be set aside for regrading. Finally, those examinees whose scores still fall below the "Conditional" category will be deemed to have failed the bar examinations. The results of the bar examinations will not be released until after the internal resolution or "appeals" process for examinees classified under the "Conditional" category or "twilight zone" is completed.

19. *How does the built-in appeals process work?*

Given the absence of precision or total objectivity in marking the essay portion of the examinations, the purpose of having an internal appeals or resolution process is to give examinees whose scores fall only slightly below the passing mark another chance of passing the bar examinations.

All the essay test booklets of examinees falling within the "Conditional" category will automatically be reread and regraded by another panel reader who is unaware of the grade given by the first grader. The first and second grades are then averaged to produce the final grade for a particular essay and then recomputed with the other essay grades to produce the new score for the essay section.

If the essay grade given by the second reader differs from the first reader's

by more than 1 or 2 points (to be determined by the Board of Bar Examiners), the essay will be forwarded to the appropriate bar examiner, as panel chair, for a third and final reading. The "resolution grade" given by the bar examiner will be the final grade for the essay, and it will be recomputed with the other essays. The bar examiner cannot give a grade which is either above the higher of the two grades or below the lower of the two grades.

20. *After the essay test booklets of the "Conditional" examinees have been reread and regraded, how is the passing standard determined?*

The recomputed final grades of "Conditional" examinees will be held against a revised Pass/Fail standard, which would only apply to them. The new Pass/Fail categories are based on the average of the lower and upper ranges of the former "Conditional" category.

21. *What type of score will be used and reported for the bar examinations?*

The Board of Bar Examiners will determine what type of score to report. There are different types of scores that may be used for reporting purposes. These include raw scores, weighted scores and averages (as at present reported), scaled scores, percentile ranks, or a combination thereof.

A raw score is generated from the number of questions that were correctly answered. A weighted score, similar to the weighted averaging system currently used in the Philippine bar examinations, involves the multiplication of raw scores by an assigned factor that is indicative of the value or weight given to a particular portion of the examinations. A percentile rank indicates how an examinee performed in relation to all others taking the test over a given administration of the bar examinations. Finally, a scaled score is produced by statistically adjusting examinee scores against a uniform standard or scale to account for variations in the difficulty of different administrations of the examination, so that any particular scaled score will represent the same level of performance from one administration of the bar examinations to another.

While slight numerical differences in subtotals create obvious differences in examinee raw and weighted scores, a bar examinee need not answer correctly every question in order to attain the highest or "perfect" scaled score or the highest percentile score. Scaling and percentile rankings tend to ignore or disregard minimal differences in raw scores in the process of converting them into scaled scores or percentile ranks.

22. *How is scaling done and why is it a more equitable scoring procedure?*

Scaling seeks to address the inequity arising from variations (oftentimes sharp fluctuations) in the level of test difficulty from one administration of the bar examinations to another. For this reason, tests which scale the scores of examinees are referred to as standardized tests.

Scaling involves the statistical process of adjusting the raw scores (frequently in the multiple-choice section) of bar examinees by comparing their performance as a group with prior examinee pools. A scaled score is this standardized score which has been adjusted to account for the differences in difficulty of the questions appearing in different administrations of the bar examinations. Total essay scores can then be scaled against the multiple-choice section scale. Scaling is thus designed to ensure that test standards remain as constant as possible from one examination to the next, so that an examinee is neither "rewarded" nor "penalized" for having taken one examination rather than another. This is because the questions asked and the grades given out in one examination may prove to be, on the average, tougher than those in other administrations of the bar examinations.

23. *What aspects of the bar administration will require computerization and automation?*

Technological enhancements form an integral part of the bar examination reform proposals. Computerization is essential in order to reduce paperwork associated with test applications and character investigations and public hearings, to reduce manual errors, and to speed up test correction (e.g., automated answer keys for the multiple choice section) and reporting of test results.

Appropriate statistical software programs are available, particularly to support the various aspects of the calibration method such as sample selection, score analysis and classification, score conversion, and scaling.

The Board of Bar Examiners has the additional responsibility of taking the necessary steps to initiate the computerization and automation of the bar examinations along these lines.

24. *What would it take to implement these reforms?*

Initially, the following:

- Approval by the Supreme Court of the reforms deemed feasible for immediate implementation after the interested parties shall have been consulted and briefed.
- A set of rules, approved by the Supreme Court, to cover these proposals. These rules may be either embodied in the Rules of Court or may constitute new legislation governing admissions to the bar.
- Selection of the initial membership of the Board of Bar Examiners, as outlined in the concept paper.
- Training of bar examiners and readers.
- Establishment of procedures and practices for inquiring into an examinee's character background, including the adoption of a standard character and fitness questionnaire and provision for public hearings in aid of investigations made pursuant thereto.