STATE OF CALIFORNIA
DEPARTMENT OF CONSUMER AFFAIRS

STATE AND CONSUMER SERVICES AGENCY • GOVERNOR EDMUND G. BROWN JR.
ACUPUNCTURE BOARD
1747 N. Market Blvd, Suite 180, Sacramento, CA 95834
P (916) 515-5200 F (916) 928-2204 www.acupuncture.ca.gov

August 2012 California Acupuncture Licensing Examination (CALE) Investigation Findings Report

Conducted by

Terri Thorfinnson, J.D., Executive Officer
Acupuncture Board
February 19, 2013

Executive Summary

At the November 15, 2012 meeting of the California Acupuncture Board (hereafter referred to as Board) there was discussion about the August 2012 California Acupuncture Licensing Examination (CALE). Concerns were expressed by attendees regarding the quality of the August 2012 CALE and the methodology that was used to determine the cut score/passing score that resulted in a low pass rate for candidates. The Board ordered the Executive Officer to conduct an investigation into these concerns.

The investigation into the August 2012 CALE began with several key questions:

1) Was the exam validated and what does that mean?
2) Did the exam follow the exam plan set by the occupational analysis?
3) Is the item bank adequate and how was it impacted by the "compromised" questions?
4) Was there anything different about the exam development compared with past CALEs?
5) Was the exam reliable in predicting and ensuring minimum acceptable competence?
6) Was the cut score set correctly or should it be changed and why?

Definitions

In reviewing the concerns, it is evident that there is confusion about the examination terminology and concepts that must be clarified. Therefore, the following are definitions of examination terminology and key concepts. This information is derived in large part from the Licensure Examination Validation Policy (OPES 12-01) from the Office of Professional Examination Services (OPES) at the Department of Consumer Affairs (DCA). These definitions are based on or quoted from published national psychometric standards:

Content Domain is the "set of behaviors, knowledge, skills, abilities, attitudes or other characteristics to be measured by a test, represented in a detailed specification, and often organized into categories by which items are classified."[i]

For the CALE, those content domains are:

1. Patient Assessment,
2. Developing a Diagnostic Impression,
3. Providing Acupuncture Treatment,
4. Prescribing Herbal Medicinals, and
5. Regulations for Public Health and Safety.

This information is available on the Board's Web site and contained in the CALE August 2012 preparation guide provided to candidates.[ii]

Occupational analysis is the method for identifying the tasks performed in a profession and the knowledge, skills, and abilities required to perform those tasks. For occupational licensing, the term occupational analysis is preferred over job analysis or practice analysis because the scope of the analysis is across a profession, not an individual job.

Reliability is the "degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable, and repeatable for an individual test taker, the degree to which scores are free of errors of measurement for a given group."[iii]

A **reliability coefficient** is a "unit-free indicator that reflects the degree to which scores are free of measurement error. The indicator resembles (or is) a product-moment correlation coefficient. In classical test theory, the term represents the ratio of true score variance to observed score variance for a particular examinee population. The conditions under which the coefficient is estimated may entail variations in test forms, measurement occasions, raters, scorers, or clinicians, and may involve multiple examinee products or performances. These and other variations in conditions give rise to qualifying adjectives, such as alternate-form reliability, internal consistency reliability, test-retest reliability, etc." [iv]

Validation is the "the process by which evidence of validity is gathered, analyzed, and summarized."[v]

Validity is the "degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses of a test."[vi] Validity is not a property inherent in a test. It is the degree to which the decisions in all phases of exam development are accurate. For licensure examinations, validity is interpreted as correctly differentiating between persons who are qualified to safely practice a profession from those who are not.

Content-related evidence of validity is the evidence that shows the extent to which the content domains of a test are based upon tasks performed in practice and the knowledge, skills and abilities required to perform those tasks. This information is acquired through the occupational analysis.

Examination outline (examination plan) is the detailed description for an examination that specifies the number or proportion of items required to assess each contact domain.

Minimum acceptable competence is the level of knowledge, skill, and ability required of licensees that, when performed at this level, would not cause harm to the public health, safety, and welfare. Minimum acceptable competence is an absolute standard and is determined by a group of licensees serving as Subject Matter Experts (SMEs). These SMEs review the occupational analysis task and knowledge statements and determine the level of performance on them that is required for minimum acceptable competence in the profession. This description serves as the criterion upon which the passing score for a licensure examination is based.

Criterion-referenced passing score is the score on a licensure examination that establishes minimum acceptable competence. Since the difficulty level of the test questions on an examination vary from test form to test form, the passing score that identifies minimum acceptable competence varies according to the difficulty of the specific examination. The passing score is not dependent upon the performance of the candidates who sit for the examination. Arbitrary fixed passing scores are not considered legally defensible[vii].

Examination development specialists are individuals who are trained, experienced, and skilled in licensure-related occupational analysis; licensure-related examination planning, development, validation, administration, scoring, and analysis; and the professional and technical standards, laws, and regulations related to these tasks. OPES employs the examination development specialists with whom the Board contracts through an Intra-Agency Contract agreement to develop its licensure exams.

Passing score (cut score): The primary professional standards followed by OPES, *Standards for Educational and Psychological Testing,* defines a "cut score" (or passing score) as a "specific point on a scale, such that scores at or above that point are interpreted or acted upon differently from scores below that point."[viii] For licensure examination purposes, the criterion-referenced passing score is the point that determines minimum acceptable competency. For the August 2012 CALE, the passing score was set at 133 out of a possible 175 points (i.e., 76%).

The **pass rate** is the proportion of candidates who achieve the cut score or higher on the exam. The pass rate for the August 2012 CALE was 38.5%.

Subject Matter Experts (SMEs) are practitioners currently possessing an active license in good standing, who are active in their practice, and are representative of the diversity of the professional population in terms of years licensed, practice specialty, ethnicity, gender, and geographic area of practice.

Legal Authority

Business and Professions Code section 4938 (c) mandates that the acupuncture licensure exam be developed by OPES.

Business and Professions Code section 139 (a), (b), (c), (d) sets out the exam requirements and standards with which OPES must comply including occupational analyses, psychometric evaluation and exam validation.

Findings

A thorough review of the facts and concerns was conducted and the following findings were concluded:

1. **The August 2012 CALE was validated by OPES.**
2. **The August 2012 CALE is a reliable measure of minimum acceptable competence.**
3. **The passing score (cut score) was accurately set for the August 2012 CALE.**

Finding #1: The August 2012 CALE was validated by OPES.

The Acupuncture Board is required by law to contract with OPES to develop the California Acupuncture Licensing Examinations (CALEs.)[ix]

OPES' mission is to protect the interests of consumers by supporting the Department of Consumer Affairs and its regulatory entities in their commitment to establish and maintain licensure examination programs that are fair, valid, and legally defensible.

OPES' guiding principles are to:

- Develop and implement quality licensure examination programs;
- Promote the sound, ethical, and fair use of licensure examinations;
- Base all licensure examinations on current and valid occupational analyses;
- Verify that any national licensure examination used in California is job-related and valid for State licensees; and
- Promote innovative technological applications to improve examination-related services.

OPES provides professional examination services to the boards, bureaus, and committees within DCA on a fee-for-service basis through Intra-Agency Contract agreements, or IACs. An IAC is developed prospectively by mutual agreement between OPES and the board, bureau, or committee. The IAC defines the activities, roles, and responsibilities of each party to the agreement, and a summary outline of the processes and benchmarks.

Currently, OPES is performing examination-related work for 43 projects, including, but not limited to, the CA Acupuncture Board, CA Architects Board, Board of Behavioral

Sciences, Cemetery and Funeral Bureau, Court Reporters Board, Dental Board of CA, Dental Hygiene Committee of CA, Board for Engineers, Land Surveyors, and Geologists, Board of Guide Dogs for the Blind, Landscape Architects Technical Committee, Optometry Board, Pharmacy Board, Physician Assistant Committee, Board of Psychology, Bureau for Security and Investigative Services, Speech-Language Pathology, Audiology, and Hearing Aid Board, CA Veterinary Medical Board, and Board of Vocational Nursing and Psychiatric Technicians.

The purpose of a licensure examination is to protect consumers by verifying that new licensees possess the minimum acceptable knowledge and experience necessary to perform tasks on the job safely and competently. Examination questions are designed to test the application of knowledge and tasks related to the practice of acupuncture rather than simply the knowledge of acupuncture. This is an important distinction in understanding the nature and structure of the CALE.

The OPES examination validation policy states that "All aspects of the test development and test use, including occupational analysis, examination development, and validation, should adhere to accepted technical and professional standards to ensure that all items on the examination are psychometrically sound, job-related, and legally defensible."[x]

OPES adheres to the above *Standards* and *Principles* in developing, analyzing and validating the CALE. In addition, OPES follows the Equal Employment Opportunity Commission (EEOC) Uniform Guidelines on Employee Selection Procedures (1978); Business and Professions Code section 139 (a), (b), (c), and (d); Business and Professions Code section 101.6; Government Code section 12944 (a) of the Fair Employment and Housing Act; and the Civil Rights Act of 1964, as amended to ensure that the examination is legally defensible.

There are two distinct phases to the CALE examination development process. In the first phase, an occupational analysis and content outline is required to be developed. OPES follows the national exam industry standard and recommends that an occupational analysis be conducted every five years to be considered current (unless there is a significant change in the profession that warrants a more frequent occupational analysis be done). The occupational analysis for acupuncture was conducted in 2008 and a summary is posted on the Board's Web site.[xi] It is anticipated that the next acupuncture occupational analysis will be conducted in 2013.

The minimum requirements for psychometrically sound occupational analyses are as follows:

- Adhere to the *Standards for Educational and Psychological Testing* and the *Principles for Validation and Use of Personnel Selection Procedures* or other psychometrically sound examination method as reference in a recognized professional source.
- Gather data from a sample of current licensees in the State of California that represents the geographic, professional, and other relevant categories of the profession.

- Develop an examination outline from the occupational analysis. [xii]

The last stage of the first phase of exam development is the creation of an examination outline or plan. The occupational analysis provides the guide for required knowledge that must be understood and the tasks that must be performed by licensed Acupuncturists in order to demonstrate minimum acceptable competency. The exam plan is developed based on the results of the data gathered from the sample of licensees in the occupational analysis. The exam outline or plan is the content guide for developing exam questions.

The most recent report documenting the content-related evidence of the validity of the CALE is the "Validation Report: Acupuncturist" published in 2008. This report details 152 separate Acupuncturist job task statements organized into five major content areas and 260 knowledge statements that are required for performance of the tasks. The linking of these knowledges to the job tasks and job task content areas, along with the weightings of these components in the licensure examination are documented in the "Examination Outline" published in this report. A summary version of this report is available on the Acupuncture Board website at http://www.acupuncture.ca.gov/students/oareport2008.pdf.

The exam plan determines the content of each exam. Each of the 175 scored and 25 pretest (unscored) test questions on the August 2012 CALE is linked, by multiple groups of Licensed California Acupuncturists serving as SMEs in OPES-facilitated workshops, to one or more of the job task statements and its related knowledge statement(s). The number of test questions measuring each job task/knowledge combination is specified in the examination outline. The OPES validation report provides strong evidence for the content-related validation of the CALE.

The second phase in exam development is for SMEs to develop exam questions. This process has several stages that include writing questions for and selecting questions from the test item bank, reviewing existing questions to determine if they need to be modified or used in the exam; and constructing the exam with new, modified or existing questions.

The participation of SMEs is essential to the development of licensure exams, and ensures that the exams accurately assess whether candidates possess the minimally acceptable knowledge, skills, and abilities necessary to perform tasks on the job safely and competently.

The selection of SMEs by boards, bureaus, and committees of DCA critically affects the quality and defensibility of their licensure examinations, and is based on the following minimum criteria:

- Reflect the profession in specialty, practice setting, geographic location, ethnicity, and gender.
- Represent the current pool of practitioners.
- Possess current skills and a valid license in good standing.

- Articulate specialized technical knowledge related to a profession.

In addition, at least half of the six to ten SMEs in each workshop should be licensed about five to seven years to ensure an entry-level perspective is represented.

The minimum requirements for psychometrically sound examination development and validation are as follows:

- Adhere to the *Standards for Educational and Psychological Testing* and the *Principles for Validation and Use of Personnel Selection Procedures.*
- Document the process following recommendations in the *Standards* and *Principles.*
- Conduct with a trained examination development specialist in consultation with SMEs.
- Use an examination outline and psychometrically sound item-writing guidelines.
- Follow established security procedures.

OPES completes a detailed analysis of the entire test and each question as part of its examination development and validation process. Adhering to testing industry standards and principles ensures the credibility of the CALE as a licensing exam that evaluates minimum acceptable competency in acupuncture and protects the public health and safety. A validated exam also produces statistical data to demonstrate that the quality of the CALE itself is valid, reliable, and legally defensible.

As part of the validation process, the number of questions in an examination should be sufficient to ensure content coverage and provide reliable measurement, including the results from an occupational analysis, item analysis, and examination analysis. The standard for having a sufficient number of test questions in the item bank is: 1) At least one new form of the examination could be generated if a security breach occurred; and 2) Questions are not exposed too frequently to repeating examinees.[xiii]

Concern was expressed about the number of questions in the CALE item bank and whether it contains a sufficient number of questions. There is a misperception that a significant number of “compromised” questions were removed from the item bank. In fact, these questions were not removed, so there was no decrease in the number of questions in the item bank.

The misconception about item bank maintenance related to exam questions that were deemed to be “compromised” by appearing in an unauthorized study guide sold on the street needs further clarification. After a close review of the study guide, OPES determined there were specific questions identified as compromised, so they were flagged for revision and not to be used without revision in future exams, including the August 2012 CALE. These compromised questions were not technically removed from the item bank. They are being revised as needed for future use. The nature of exam development includes creating both new questions and revised questions that evaluate the correct application of knowledge and tasks and modifying existing questions. Thus,

the existing CALE item bank was not diminished by the flagging and modifying of the compromised questions.

According to OPES, they conduct workshops year round to create new questions to ensure each content category has a wide variety of questions available for future exams. The item bank has sufficient questions in each content category to allow for new or modified questions to be used for each exam. The CALE item bank has a sufficient number of questions regardless of the compromised questions. Thus, the compromised questions did not have any impact in the development of the August 2012 exam. However, modifying the phrasing of the compromised questions could have impacted the examinees' performance as reflected in the pass rate, particularly for those individuals who chose to memorize test questions rather than learn the knowledge required to perform the required job tasks.

Conclusion: The August 2012 CALE was developed according to nationally recognized testing industry standards. The exam plan was used in developing and validating the exam. The same identical and well-documented process was adhered to as has been done with past CALEs—there was no deviation in the exam development process from the process used in past CALEs.

OPES is in compliance with the testing industry *Standards* and *Principles* and has been fully transparent with publishing its standards, exam policies, and occupational analysis. The August CALE is valid, accurate, and legally defensible. The data show that the August 2012 CALE performed extremely well.

Finding #2: The August 2012 CALE is a reliable measure of minimum acceptable competence.

OPES performs detailed psychometric analysis of each test question checking for whether the test accurately measures that those who understand the concept the question is supposed to test in fact answer the question correctly. OPES analyzes each question for whether the answers vary by language reflecting some advantage or disadvantage in the question wording in each language. OPES also analyzes the questions to ensure that the answer to one question does not provide a clue to the answer to another question in the exam. In analyzing the August 2012 CALE, OPES found the questions that were scored accurately measured the application of Acupuncture knowledge. They also found that language was not a factor in whether someone answered questions correctly or incorrectly, thus each language version of the exam was deemed equivalent to each other in its ability to test for required knowledge.

The national standards followed by OPES states the following: "The inferences made from the resulting scores on a licensing examination are validated on a continuous basis. Gathering evidence in support of an examination and the resulting scores is an on-going process. Each examination is created from an examination outline that is based upon the results of a current occupational analysis that identifies the job related

critical tasks, and related knowledge, skills, and abilities necessary for safe and competent practice."[xiv]

OPES validates each exam through detailed question-by-question analysis: Does the exam adhere to the exam plan? Do questions overlap or provide clues to answers to other questions in the exam? Is there any variation in the how exam takers perform on each question? Is there any variation in how exam takers perform on each question by language?

The Council on Licensure, Enforcement, and Regulation states that *". . .* reliability is an index of the stability of test scores. Reliability indices range between 0 and 1.00, with higher numbers being associated with a greater level of score stability. Reliability indices above 0.90 are considered very acceptable for most purposes, while indices less than 0.70 usually indicate an unacceptable level of score stability."[xv]

Cronbach's alpha reliability coefficient is a measure of the internal consistency reliability of an examination. The Cronbach's Alpha reliability coefficient for the 175 scored test questions on the August 2012 CALE is 0.914. This value is above the standard of 0.90 noted above.

The standard error of measurement statistic is an estimate of the degree of accuracy of any particular score on a test. This estimate is calculated using an estimate of the test's overall reliability value. The smaller the value of the standard error of measurement, the more accurate is any particular score on the exam. The standard error of measurement for the August 2012 CALE is 5.556 raw score points. This is a typically small standard error of measurement for OPES examinations.

The point biserial correlation coefficient (Rpb) is a mathematically simplified calculation of the Pearson Product Moment correlation coefficient between the proportion of candidates who get an individual test question correct and their respective total scores on that test.[xvi] The correlation, which can range from -1.00 through zero to +1.00, indicates how closely the performance on an individual test question is related to overall performance on the test. The importance of the Rpb for the item analysis of an examination goes beyond reaching the level of statistical significance, especially for examinations with larger numbers of applicants. When there are 400 candidates, the critical value for statistical significance is approximately 0.10. At or beyond this value the Rpb can be considered statistically significant.

However, for examination item analysis, any Rpb of approximately zero indicates that the specific test question under investigation is not contributing to an accurate identification of minimum acceptable competence. In addition, any negative Rpb indicates that the test question under investigation is "working backward" to the degree indicated by the negative value. For example, an Rpb of -0.01 indicates that the test question is practically of no value in determining whether candidates are at the level of minimum acceptable competence. Furthermore, any Rpb value of -0.20, if it occurred, would indicate that candidates who get the test question *incorrect* actually tend to get

higher scores on the examination as a whole. Such a test question, if it had been present, would need to be corrected or dropped from the examination.

For the August 2012 CALE, there were no negative Rpb values. All scored items had Rpb values in the desired range for statistical significance and correlation analysis.

Conclusion: The psychometric analysis performed by OPES on the August 2012 CALE determined that the CALE was reliable in its predictability for evaluating minimum acceptable competency. This reliability, in turn, contributed to the exam's validity and credibility as an exam based on sound testing industry standard psychometric analysis and evaluation. The August 2012 CALE was developed in the same manner as previous examinations, using the same processes.

Finding #3: The passing score (cut score) was accurately set for the August 2012 CALE.

439 candidates took the August 2012 CALE. There were 175 possible points for the exam. The cut score was 133. 169 (38.5%) candidates achieved a passing score of 133 out of 175 points. As a comparison, the cut score for the February 2012 CALE was 128 and the pass rate was 68%.

The concern that led to this investigation focused on the cut score and the low pass rates. There has been significant confusion about the cut scores and pass rate and both terms have been incorrectly used interchangeably. By definition, the cut score is determined by extensive psychometric analysis of individual exam questions in workshops facilitated by an OPES' Examination Development Specialist with SMEs recommended by the Board. The pass rate is simply the percentage of candidates that achieved a passing score.

OPES employs a criterion-referenced passing score methodology called the "modified Angoff technique" for determining licensure examination passing scores. The criterion applied is minimum acceptable competence to practice the profession. A criterion-referenced passing score maximizes the likelihood that candidates who pass the licensure examination have sufficient knowledge and experience to practice safely and competently.

Criterion-referenced standard setting begins with the establishment of a minimally acceptable level of competence for safe practice that candidates must possess in order to pass the examination. The group of licensed Acupuncturists serving as SMEs developed common definitions of different levels of candidate performance by identifying critical work behaviors that contrast the highly competent, the minimally competent, and the incompetent candidate.

Because licensing examinations are known to vary in difficulty from one examination form to another, a fixed passing score or percentage such as 70% does not represent

the minimally acceptable competence for all administrations of an examination. Therefore, arbitrary passing scores are not considered accurate or legally defensible.[xvii]

By applying a criterion-referenced methodology, a passing score is lowered for an examination containing a large number of difficult items (questions) and raised for an examination containing a small number of difficult items. Candidates who take a more difficult test would be placed at a disadvantage unless a criterion-referenced passing score is established. Thus, the passing score provides safeguards to both the candidate and the consumer affected by the particular profession.

This criterion-referenced passing score development methodology is independent of the performance of other candidates who take the examination at the same time. The passing score is not based on performance with respect to the group. Rather, the passing score is based upon minimum acceptable competence as it relates to the difficulty of the particular set of items within the examination form.

The passing score standards for licensure examinations must: 1) Follow a process that adheres to accepted technical and professional standards; 2) Adheres to a criterion-referenced passing score methodology that uses minimum competence at the entry level to the profession.[xviii] OPES adheres to these test industry standards in setting passing scores for the CALE.

Eight California-licensed acupuncturists served as SMEs in the passing score workshop for the August 2012 CALE conducted in the OPES offices on August 15 and 16, 2012 under the direction of a test development specialist. The process included a number of newly licensed practitioners to ensure participation from entry-level licensees. Each of the SMEs had participated in previous CALE exam development workshops at OPES. These SMEs had an average of 9.5 years of licensed experience and ranged from a minimum of 6.3 to a maximum of 12.3 years of experience.

The SMEs were trained and calibrated in minimum acceptable competence for acupuncturists and trained in the modified Angoff technique for setting passing scores. Following the training, the SMEs independently estimated the passing score for each of the 200 (scored and unscored) questions on the examination. When passing rate estimates provided by individuals SMEs had differences greater than 20%, raters discussed the differences and resolved them if possible. None of the scoreable items had differences in estimated passing scores greater than 20%. This is one of the many layers of analysis and checks and balances OPES employs to control individual bias and ensure accuracy.

The workshop facilitator also monitored actual test question difficulty of each test question based on the item analysis for the August 2012 CALE. Over all SME raters and all scoreable test questions, the difference between the average actual item difficulty and the average estimated passing score was 4.65 points. This is a very small, statistically and practically insignificant difference.

An important assessment of the quality of the decisions made by the licensed Acupuncturists serving as SMEs in the passing score workshop is the inter-rater reliability or inter-rater consistency. One measure of this consistency is the **Intraclass Correlation Coefficient**. This correlation coefficient describes the degree of consistency among a group of independent raters. Values range from -1.00 through zero to +1.00. Positive values indicate a consistent relationship among the raters, with higher positive values indicating a stronger relationship.

The Intraclass correlation coefficient among the eight SMEs serving in the passing score workshop was 0.838. This value is highly statistically significant ($p<.000$). More importantly, this value indicates that the set of SME raters who determined the passing score for the August 2012 CALE were highly consistent with one another.

These data show that the SMEs serving as passing score workshop participants were consistent with one another and closely paralleled the actual average item difficulty of the scored questions on the CALE.

One other measure, the classical test theory-based **Conditional Standard Error Measurement** (CSEM) indicates a high degree of precision (accuracy) of the passing score used for the August 2012 CALE. The CSEM is parallel in meaning to the Standard Error of Measurement discussed above, except it more accurately measures the degree of consistency of decisions made at the cut score for the examination. The conditional standard error of measurement for the August 2012 CALE is 5.407 raw score points and is more precise than the value of the standard error of measurement across the full range of scores on the CALE (SEM = 5.556).

The above excerpts of the various psychometric and statistical analysis employed by OPES to ensure the exam tests minimum acceptable competency and that its results are reliable have been included in this report to demonstrate the scientific, evidence-based analysis that is conducted for each and every Acupuncture Licensing Exam developed by OPES. The August CALE and other exams developed by OPES are developed with national industry testing standards. The extensive analyses indicate all of the CALE are in fact accurate and can be backed up with statistical data demonstrating its accuracy and reliability.

There is a misperception that the passing score should be a fixed score; but, in fact, the passing score is set based solely on whether the cut score reflects minimally acceptable competence to practice acupuncture, not a fixed score. The goal of the exam is to test minimum acceptable competency to protect the public health and safety of consumers. An arbitrary fixed passing score or percentage, such as 70 percent, does not represent minimally acceptable competence.[xix] Arbitrary passing scores are not legally defensible.[xx]

The advantage of using criterion-referenced methodology is that the passing score is independent of the performance of other candidates who take the examination at the same time. The passing score is not based on performance with respect to the

candidates. Rather, the passing score is based upon the difficulty of the items within the examination.

Claims that the August 2012 CALE examination results are problematic due to the fluctuating passing scores from exam to exam or the fact that the passing score is not fixed are inaccurate and do not represent the industry testing scoring standards where the Angoff method is utilized. The CALE must adhere to testing standards and principles and not be changed to accommodate requests for score alteration. Additionally, if the August 2012 CALE passing score were changed, such change must be based on sound testing standards and principles. Since OPES adhered to the required testing industry standards and principles, any change in the cut score without evidence would render the examination invalid, inaccurate, and legally indefensible.

Another incorrect conclusion made through public comment was that a low pass rate was a reflection of the poor quality of the exam itself. The validity and reliability of the exam in evaluating minimum competency is based on detailed exam evaluation standards. Whether an exam is valid includes an evaluation of whether it adheres to the exam plan that is guided by the occupational analysis. Validation also includes psychometric analysis of each question and whether the answers to each question predicts reliably that those who answer a question correctly actually understand and can apply the content knowledge the question is testing. Conclusions regarding the quality of an exam are based on this validation process and standards, not on the pass rates of candidates.

The pass rate is a function of the candidates taking the exam and not the exam itself. Since the exam is a measure of how many candidates possess minimum acceptable competence, the pass rate is a measure of how many who took the exam possesses minimum acceptable competence. The exam is developed from the occupational analysis of what clinical knowledge and tasks candidates must know in order to practice with minimum acceptable competence in acupuncture, not simply academic knowledge of acupuncture. A low pass rate reflects the percentage of candidates who do not have minimum acceptable competence. It is important to understand that a licensure exam is developed to test for minimum acceptable competence based on what acupuncture practitioners need to know to practice so they do not harm the public's health and safety. This clinical knowledge of the practice is determined by the occupational analysis.

There were also concerns expressed about candidates not being provided enough guidance for exam preparation. As a convenience, and to assist candidates prepare for the CALE, the Board provides candidates with an examination preparation guide in English, Chinese and Korean. All August 2012 CALE candidates were mailed this preparation guide.

This comprehensive, Board-authorized exam preparation guide contains information about the examination, including the purpose, development, and passing score; examination and site security; examination administration details, including scheduling,

directions to the site(s), ADA accommodations, and testing in Chinese or Korean; and the check-in and testing process. Appendices include a detailed CALE content outline, sample questions by content area, and lists of acupuncture pulses, points, herbs and herbal formulas.

A review of the preparation guide provides all candidates with an awareness of what content areas they need to know to successfully pass the CALE. The licensure exam is not the same as a school exam that tests academic knowledge related to recognition and recall. A licensure exam tests on how to apply the knowledge[xxi]. The candidates who failed the August 2012 CALE did not demonstrate a high enough proficiency in understanding and applying minimum acceptable competence in the required content areas.

Conclusion: There are no anomalies in the exam scoring process for the August 2012 CALE. The exam plan was utilized in scoring and validating the exam. The same identical and well-documented process was adhered to, as has been done with past CALEs—there was no deviation in the exam scoring process from the process used in past CALEs.

OPES is in compliance with the testing industry *Standards* and *Principles* and has been fully transparent with publishing its standards, exam policies, and occupational analysis on its internal Web site. The August CALE cut score is accurate and legally defensible. Extensive analysis indicates that the quality of the August 2012 CALE is excellent, as evidenced by the data cited within this report.

---

[i] American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, *Standards for educational and psychological testing,* Washington, D.C.,1999, p. 174.

[ii] See http://www.acupuncture.ca.gov/students/exam_content.shtml

[iii] American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, *Standards for educational and psychological testing,* Washington, D.C.,1999, p.180.

[iv] Ibid. p.181.

[v] Society for Industrial Organizational Psychology, Inc., *Principles for the Validation and Use of Personnel Selection Procedures*, Bowling Green, OH, 2003, p. 72.

[vi] American Educational Research Association, p.184.

[vii] The AERA, APA, NCME, *Standard for Educational and Psychological Testing* (1999) in discussing fixed passing scores on licensure examinations state the following (on page 157): "Legislative bodies sometimes attempt to legislate a cut score, such as 70%. Arbitrary numerical specifications of cut scores are unhelpful for two reasons. First, without detailed information about the test, job requirements, and their relationship, sound standard setting is impossible. Second, without detailed information about the format of the test and the difficulty of items, such numerical specifications have little meaning."

[viii] American Educational Research Association, p. 175.

[ix] California Code of Regulations, Business and Professions Code section 4938 (c).

[x] OPES Licensure Examination Validation Policy, 2012, p.4.

[xi] See http://www.acupuncture.ca.gov/students/exam_content.shtml

[xii] OPES Licensure Examination Validation Policy, 2012, p.5

[xiii] Ibid., p.5

[xiv] Ibid. p. 4.

[xv] See also Nunnally and Bernstein's Psychometric Theory (3rd Edition, 1994), p. 265 for additional support for these recommended values.

[xvi] To avoid spuriously increasing the correlation, the analyzed item's score is not added to the calculation of the total score for that particular point biserial analysis. Hence, that item's correlation with the total score is not inappropriately influenced by adding that items values to the total score.

[xvii] The AERA, APA, NCME, *Standard for Educational and Psychological Testing* (1999) in discussing fixed passing scores on licensure examinations state the following (on page 157): "Legislative bodies sometimes attempt to legislate a cut score, such as 70%. Arbitrary numerical specifications of cut scores are unhelpful for two reasons. First, without detailed information about the test, job requirements, and their relationship, sound standard setting is impossible. Second, without detailed information about the format of the test and the difficulty of items, such numerical specifications have little meaning."

[xviii] OPES Licensure Examination Validation Policy, 2012, p. 6.

[xix] Ibid., p. 6

[xx] Ibid., p. 6