

EXAMINER CALIBRATION TEST RESULTS SUPPORT VALIDITY

THE INCENTIVE FOR ANALYSIS OF CALIBRATION TESTING

An analysis of Western Regional Examination Board (WREB) dental examiners 2008 calibration test scoring was completed in December, 2008. The results validate the quality of the calibration training that was provided for each examiner just prior to every dental examination. Calibration training is specified in the American Association of Dental Examiners (AADE) 2003 *Guidance for Clinical Licensing Examinations in Dentistry*. Adherence to this guidance document, with verification of positive results, then supports the validity of using WREB dental examination scores when granting licenses to dental applicants.

PURPOSES FOR ANALYSIS OF CALIBRATION TESTING

An important purpose of an annual analysis of the calibration tests is to provide information about the calibration test questions. The subcommittees can then determine which questions might be improved. Additionally, the analysis provides information about how well examiners scored overall and how each examiner scored for the year. The information in this report describes the scoring of all examiners in each category of the calibration testing.

ADMINISTRATION AND SCORING OF CALIBRATION TESTS

The tests are administered at a test site to all examiners who will examine at that site. The tests consist of computerized slide presentations that are projected on a screen. All examiners answer the same questions at the same time and the tests are then scored by the floor examiners. The floor examiners do not score candidates during the candidate examinations and consequently are not required to take the calibration tests. If any examiner scores less than 80%, there is immediate retraining and re-testing. The tests include tooth acceptance, restoration and endodontic preparation, restoration finish, endodontic condensation, and case treatment planning.

Except for treatment planning where there are laminated displays for each examiner, all information for the tests is provided in the on-screen presentation. The actual candidate examination scoring is completed by observing the candidates' work. The screen presentation is intended to evaluate the application of scoring criteria, but with the on-screen test, the lack of tactile evaluation and some of the visuals make scoring more difficult than the candidate examination scoring. Consequently, a test question is considered to be answered correctly if the examiner's answer is (on a six point scale) within one point of the answer specified by the subcommittee that developed the test.

RESULTS FROM THE ANALYSIS OF 2008 CALIBRATION TESTING

In acceptance tests, the examiners agreed with the committee decision to accept or reject on 90% of the 2830 accept/reject decisions in operative and on 96% of the 3050 accept/reject decisions in endodontics.

In operative preparation, the examiners ratings were within one point of the committee designated answers in 91% of the 17,100 scores assigned. The examiners were within one point of the committee designated answers in 90.5% of the 15,360 scores assigned in operative finish.

In endodontic treatment, the examiners ratings were within one point of the committee designated answers in 90.4% of the 7,060 scores assigned.

In patient assessment and treatment planning, the examiners ratings were within one point of the committee designated answers in 83.7 of the 3,300 scores assigned.

SUMMARY

The results from the analysis of the calibration test scoring is quite positive with all of the tested sections except one showing examiner agreement with committee designated answers at 90% or greater.

It should be noted that the data analyzed includes all of the testing and includes scores that were not good enough to be considered calibrated. In those cases, the examiners were retrained and re-tested until the test results were satisfactory. The one section where the examiner agreement was less than 90% overall was a new examination section. The criteria and test questions for that section have been revised to improve examiner understanding of both the criteria and the test questions.

These results show that the calibration training has been successful and provides confidence that the examiners understand and can apply the criteria used to score candidate performance. This contributes to the validity of the use of WREB scores by licensing agencies.