

Annual Summary for 2008 Testing Year

Board of Certification (BOC) Certification Examination for Athletic Trainers

CASTLE Worldwide, Inc.

May, 2009

Introduction

The Board of Certification (BOC) is a nonprofit credentialing agency that provides certification for the athletic training profession. The BOC was incorporated in 1989 to govern the certification program that had existed then for nearly 20 years for entry-level athletic trainers and recertification standards for Certified Athletic Trainers. The entry-level certification program is designed to establish a common benchmark for entry into the athletic training profession. The BOC serves the public interest by developing, administering, and continually reviewing a certification process that reflects current standards of practice in athletic training.

In order to develop a credible and valid examination, the BOC contracts with CASTLE Worldwide, Inc. (CASTLE) a certification and licensure design, development, and administration service company, to develop and review the form and item statistics for the currently administered BOC examinations. CASTLE follows and recommends widely accepted standards and regulations (e.g., *Standards for Educational and Psychological Testing*, American Educational Research Association, 1999; *Uniform Guidelines on Employee Selection Procedures*, EEOC, 1978; *Standards for the Accreditation of Certification Programs*, National Commission for Certifying Agencies, 2005) for the development and analysis of the BOC examination.

The aim of BOC's certification is to establish that individuals have the skills and knowledge necessary to create and implement safe and effective athletic training programs. The examinations are designed to test an individual's knowledge across the practice of athletic training based on a defined test blueprint.

The BOC testing year occurs from March 1 to February 28/29 of the following year. The BOC offers candidates five two-week testing windows during the testing year: March/April, May/June, July/August, November, and February, during the testing year. During each testing

window two forms of the examination are delivered. These two forms consist of items in common with an anchor form. Candidates who fail are not restricted in their retakes during the testing year.

In May/June 2007 the BOC introduced a combined test that consisted of two components, a multiple-choice portion with 125 scored items and a hybrid portion with two scored problems. The hybrid problems consisted of 12 to 17 items per problem. Examinations are conducted in one session and candidates are allotted a period of four hours. Candidates complete the 125 multiple-choice questions then transition to the hybrid portion.

For 2007-08, no multiple-choice experimental items were placed on the test forms, and two experimental hybrid problems were placed on each test form. The questions are not grouped by subject area or domain in the examinations.

Due to differences in length, the hybrid problems assigned to a test form are scaled so that each problem is equivalent in its impact on candidate total scores. This score is then equated to the multiple-choice anchor form for the current test blueprint, using common-item non-equivalent group equating procedures (Kolen & Brennan, 2004). The equated scores for the examinations are converted via linear transformation so that the raw score passing standard is reported to candidates as 500 on a scale of 200 to 800.

Exam Development

There are a total of 15 hybrid problems both active and in development, in the hybrid item bank. Currently BOC has 1,708 5-option multiple-choice items in the item bank. Of this group, 877 items are active or retired and 831 items are either in development or experimental.

A total of 55 stand-alone questions are in development using the alternative item types (e.g., multi-select/n-wise, drag and drop, n-wise, hot spot) currently used in the hybrid portion of the examination. These items are designed to assess the potential to expand the item types available for the traditional stand-alone multiple-choice items.

In November 2008, the BOC item writing team began development of focused testlets for experimental purposes. Each focused testlet consists of a scenario followed by five questions utilizing the alternative item types (e.g., multiple-choice, drag and drop, n-wise, hot spot) used in the hybrid portion. The testlets focus on asking questions about critical skills/knowledge required

by the entry-level athletic trainer associated with the scenario. Currently there are seven testlets in development.

Candidate Performance

Statistics reported refer to the performance of ‘analyzed’ candidates. Statistical reports are generated for a particular time period (e.g., an examination window), but on occasion, candidates who arrive at the test location do not take the examination. For these and other reasons, some candidates are excluded from the pool of analyzed data. As a consequence, the number of ‘analyzed’ candidates may not match the number of reported candidates. As of 2007, the three cohorts of candidates reported for the BOC examinations are:

1. First-time candidates – candidates reported as first-time test takers and/or recent college graduates from athletic training education programs accredited by the Commission on Accreditation of Athletic Training Education programs (CAATE).
2. Retakes – candidates who re-sat the examination one or more times.
3. All – candidates who tested.

Candidates Excluded from this Report

The report does not include, except where noted, those candidates who were administered the examination as paper-and-pencil. The April 2007 candidate performance data are excluded from the remainder of this report, except where noted, as the program used to assess these candidates is not equivalent to the revised BOC testing program.

There were a total of 6,135 administrations of the examination in 2008-09, a 31% increase from the 4,691 administrations in 2007-08. The number of first-time candidates almost doubled from 2007-08 and is again in line with previous years.

Pass Rates

In 2005-06 new test specifications and the associated passing standard were introduced. All later forms of the examination are equated back to this standard. Pass rates for the current examination can be compared to the pass rates for the multiple-choice examination used for

2005-06 to 2006-07. Table 1 provides the annual pass rates for BOC multiple-choice test administrations from 2005-06 for the three cohort groups described above.

Year	First-time	Pass	% Pass	Retake	Pass	% Pass	All	Pass	% Pass
2005-06	2,074	968	46.7%	3,017	660	21.9%	5,091	1,628	32.0%
2006-07	2,322	1,125	48.4%	3,549	1,076	30.3%	5,871	2,201	37.5%
2007-08	1,495	584	39.1%	3,196	1,073	33.6%	4,691	1,657	35.3%
2008-09	2,762	1,423	51.5%	3,373	1,035	30.7%	6,135	2,458	40.1%

Table 1: Number of Candidates in Three Cohorts and Pass Rate for BOC Examinations, 2005-06 to 2007-08 (2005-06 and 2007-08, for the Multiple-Choice Examination Only).

The pass rates in Table 1 for the first two years are for candidates who passed the multiple-choice (MC) component. From 2007-08 the pass rate is for candidates who passed the combined examination (multiple-choice and hybrid).

As can be seen from Table 1, the pass rates for the first-time candidates on the combined examination increased substantially for 2008-09 when compared to the prior multiple-choice component only. However, the pass rate for all candidates is in line with fluctuations in the pass rates for the prior year's multiple-choice examination only due to a drop in the retake pass rate.

Distribution of Candidate Scores

Candidate performance for each of the three cohorts is detailed below in the tables and figures. Overall, there was an increase in the average scale scores for all three cohorts when compared to 2007-08. Table 2 details the overall scale score performance for the program for 2008-09.

Cohort	N	Avg.	Median	Mode	Std. Dev.	Min	Max
First-time	2,762	490	500	500	86	200	686
Retake	3,373	489	464	500	69	200	680
All	6,135	473	476	500	79	200	686

Table 2: Number of Candidates in Three Cohorts, Minimum, Maximum and Average Scaled Score, Median and Mode Scaled Score, and Standard Deviation (Scaled Score) for BOC Examinations, 2008-09.

Similar to 2007-08, a Univariate General Linear Model (GLM) test determined that there was a statistically significant, but small, difference in the scaled scores of retake and first-time candidates [$F(1, 6133) = 255.47, p < .001, \eta = .04$]. First-time candidates were more widely distributed along the score scale than retake candidates. This difference was confirmed by an examination of the distribution of scaled scores for first-time and retake candidates (Figure 1).

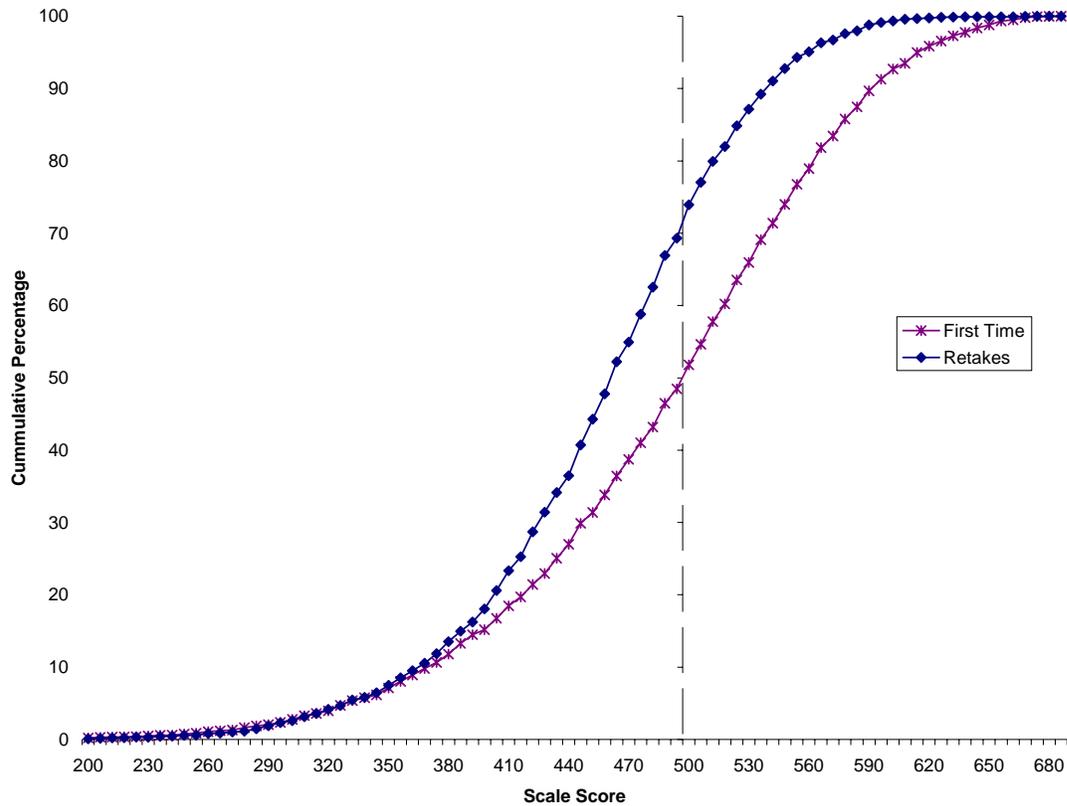


Figure 1: Cumulative Percentage of First-time New Graduates and Retake Candidates by Scaled Score, BOC 2008-09.

Test Form Summary Statistics

For 2008-09 there were four forms of the BOC examination. Two core multiple-choice components (3614 and 3615) were assigned different hybrid problems for each test window, creating the four different test forms. Table 3 provides form descriptive statistics for each test administration period (see Appendix A for information on the statistics reported).

Exam	N	Mean	Median	Mode	Std. Dev.	Min	Max
3614	953	494	500	506	77.8	224	686
3615	962	495	500	518	78.4	206	674
<i>March</i>	<i>1915</i>	<i>495</i>	<i>500</i>	<i>506</i>	<i>78.0</i>	<i>206</i>	<i>686</i>
3616	783	463	470	470	80.7	200	674
3617	762	465	464	452	81.2	200	680
<i>May</i>	<i>1545</i>	<i>464</i>	<i>470</i>	<i>512</i>	<i>80.9</i>	<i>200</i>	<i>680</i>
3614	536	461	461	506	74.4	230	680
3615	542	457	464	470	69.3	230	668
<i>July</i>	<i>1078</i>	<i>459</i>	<i>464</i>	<i>446</i>	<i>71.9</i>	<i>230</i>	<i>680</i>
3616	429	461	458	458	76.7	218	668
3617	490	455	464	476	77.6	200	632
<i>November</i>	<i>919</i>	<i>458</i>	<i>464</i>	<i>476</i>	<i>77.2</i>	<i>200</i>	<i>668</i>
3614	307	477	482	500	74.4	236	656
3615	370	474	476	464	76.8	206	674
3617	1	554	554	554	NA	554	554
<i>February</i>	<i>678</i>	<i>475</i>	<i>482</i>	<i>500</i>	<i>75.6</i>	<i>206</i>	<i>674</i>
TOTAL	6,135	473	476	500	78.9	200	686

Table 3: Summary Test Form Statistics in Scaled Scores for All Candidates for BOC Examinations, 2008-09.

There was a significant difference based on candidate's month and retake status. First-time candidates who took the examination earlier in 2008 outperformed other candidates [F (3,6114) = 102.67, p = .0021, $\eta = .90$].

Test Form Internal Reliabilities

As noted earlier, the number of individual items on a hybrid problem varies. If the items included on the hybrid problem were added to each of the form's scored multiple-choice items, the total number of elements that form an analysis for reliability would differ slightly from form to form. As a consequence, reliability estimates were made separately for each component of the examination, and a study was undertaken to assess the relationship between the two components. For the hybrid problems, as the scores are not dichotomous data (0 and 1), Cronbach's alpha was used to determine reliability (Cronbach, 1951). For review purposes, Cronbach's alpha was also computed for the multiple-choice elements.

Table 4 details the K-R20, Standard Error of Measurement and Cronbach's alpha for the multiple-choice portion (MC), the hybrid portion, and the combined test.

N	MC K-R20	Std. Error	MC alpha	Hybrid alpha	Combined alpha
6135	0.87	4.68	0.89	0.31	0.45

Table 4: Internal Reliability Estimates for Multiple-Choice for All Candidates for BOC Examinations, 2008-09

Conclusion

Statistics concerning the quality of the BOC examination as a measurement device indicate that the multiple-choice portion of the examination complies with psychometric requirements that pertain to certification and licensure tests. Notably, estimates of reliability and equivalence across forms for the various parts of the examination are quite strong. Likewise, candidate performance on all parts of the examination is consistent with the public protection mission of the BOC.

REFERENCES

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: AERA.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.
- Kolen, M.J., & Brennan, R.L. (2004) *Test Equating, Scaling and Linking: Methods and Practices Statistics for Social Science and Behavioral Sciences* (2 ed.). Springer-Verlag New York Inc..
- Equal Employment Opportunity Commission (EEOC), U.S. Civil Service Commission, U.S. Department of Labor, and U.S. Department of Justice. (1978). Uniform Guidelines on Employee Selection Procedures. *Federal Register*, *43* (166), 38290-38315.