

## Examination Review for 2009-10 Testing Year

### Board of Certification (BOC) Certification Examination for Athletic Trainers

Stephen B. Johnson, Ph.D.

April 20, 2010

**Prepared for:**

Board of Certification, Inc.  
1415 Harney St., Ste 200  
Omaha, Nebraska 68102

## 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, which had then existed for nearly 20 years for entry-level athletic trainers and recertification standards for certified athletic trainers (AT). 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 services. The examinations are designed to test an individual's knowledge across the practice of athletic training based on a defined test blueprint.

In order to attain certification, an individual must complete an entry-level athletic training education program accredited by the Commission on Accreditation of Athletic Training Education (CAATE) and pass the BOC certification exam. In order to qualify as a candidate for the BOC certification exam, an individual must meet the following requirements:

- Endorsement of the exam application by the recognized Program Director (PD) of the CAATE accredited education program.
- Proof of current certification in emergency cardiac care (ECC) (*Note: ECC certification must be current at the time of initial application and any subsequent exam retake registration.*)

The BOC testing year runs 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 January/February. 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.

### *Field-Testing and Scoring*

Part 1 of each test form contains 30 field-test items. In March/April and May/June 2009, these items included traditional multiple-choice items placed throughout the 125 scored items. Beginning in August 2009, part 1 included 15 traditional multiple-choice items placed throughout the 125 scored items, as well as 15 field-test items placed at the end of each set of five-option multiple-choice items (item locations 141 to 155). Each form contained two field-test FT of five items each (item locations 141 to

150), followed by five field-test alternative item types not associated with a testlet (item locations 151 to 155). All alternative items and focused testlets were unscored.

#### *Number of Test Forms*

Two core multiple-choice sections of 125-scored items (form family A and B) were assigned six different experimental sets for the year, creating 12 different test forms. The two sets of 125-scored items had 65 items (52%) in common.

In March/April 2009, two of the forms (3618 and 3619) contained five additional experimental items, administered to a small group of candidates. These forms are reported separately.

Test forms in form family A were administered in March/April and November 2009. Form family B test forms were administered in May/June and July/August 2009 as well as January/February 2010.

#### *Equating Test Forms*

Upon introduction of a new test form, the performance of candidates on the new form is equated to performance of candidates on a prior test form. The BOC equating follows the protocols for *common items non-equivalent groups design* using the Levine True Score Method Applied to Observed Scores using internal anchors (Kolen & Brennan, 2004). This design compares the performance of one group of test takers on one examination form to another group of test takers on an earlier examination form with a known cut-score. Ultimately, all equating is compared to the performance standard established for the base form (342) used for the current role delineation/practice analysis.

#### *Use of Scale Scores*

Since examination versions are possibly of different difficulty, providing raw scores can be misleading. As a result, many programs, including the ACT® and SAT® examinations, use scale scores. Scale scores are particularly useful at providing the basis for long-term, meaningful comparisons of results across different administrations of an examination.

Scale scores are used because, over the life of every testing program, there are situations when changes in test length occur, for example, when a decision is made to assess more or fewer areas; when the numbers of items that are scored versus experimental (field-test) changes; or different examination versions of different difficulty are being compared.

For scale scores, the passing standard (number correct) on any examination version is always reported as the same scale score.

The equated scores for the BOC examination are converted via linear transformation so that the cut-score for all test forms are reported to candidates as 500 on a scale of 200 to 800.

## ANALYSIS OF THE EXAMINATION

This section is broken into a number of parts. The first section provides a review of candidate performance for three different cohorts of candidates. The last section details form-level statistics for the examination program.

### *Candidate Performance*

Statistics reported refer to the performance of ‘analyzed’ candidates. Statistical reports are generated for a particular time (e.g., an examination window). Some candidates are excluded from the pool of analyzed data, specifically those candidates who completed less than 25% of their examinations. It is likely that such candidates experienced problems such as being late to the site or other issues, and therefore, their data is problematic. 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 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 via paper-and-pencil or those candidates with incomplete data. As a result, the number of candidates analyzed for this report may not match the number of candidates who took the BOC athletic trainer examination. Data from previous years may only include two of the three cohort groups.

Data prior to the introduction of the two-part examination (April 2007) are excluded from the remainder of this report, except where noted, because the program used to assess candidates is not equivalent to the revised BOC testing program.

There were 6,171 analyzable administrations of the BOC examination in 2009-10. This total was similar to the number of administrations for the previous test year, 2008-09 (6,135).

Of the 6,171 results analyzed, 2,854 (46%) were administered to first-time candidates, consistent with 2008-09.

### *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.

In 2005-06 and 2006-07, the BOC examination consisted of three components (simulation, multiple-choice, and practical), of which candidates were required to pass all three elements. From 2007-08 onward, candidates were required to pass one test administration as documented above (part 1 and part 2). Table 1 provides annual pass rates for BOC administrations from 2005-06, but only reports the pass rates for 2005-06 and 2006-07 that are associated with the multiple-choice element.

**Table 1:** Number of Candidates in Three Cohorts and Pass Rate for BOC Examinations, 2005-06 to 2009-10 (2005-06 and 2006-07 are for the Multiple-Choice Element Only).

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%
2009-10	2,852	1,235	43.3%	3,319	1,120	33.7%	6,171	2,355	38.2%

For 2005-06 and 2006-07, the three-component examination resulted in a pass rate for first-item test takers of 26.2% in 2005-06 and 31.5% in 2006-07. This was substantially lower than the pass rates for the combined examination protocol used since 2007-08.

A test of proportions indicated that the pass rate for all examinations administered in 2009-10 is significantly higher than the percentage that passed the 2007-08 examination ( $z=2.16$ ,  $p < .05$ ), but significantly lower than 2008-09 ( $z=6.08$ ,  $p < .05$ ). The pass rate for retake candidates is significantly higher for 2009-10 than the previous year ( $z=2.63$ ,  $p < .05$ ). Table 2 details the pass rates for each form by testing window.

**Table 2:** Passing Rates for Each Test Form for All Candidates for BOC Examinations, 2009-10.

Form	Frequency			Percent	
	Fail	Pass	Total	Fail	Pass
3618	59	22	81	72.8%	27.2%
3619	42	28	70	60.0%	40.0%
3620	570	297	867	65.7%	34.3%
3621	569	258	827	68.8%	31.2%
<i>March/April</i>	<i>1,240</i>	<i>605</i>	<i>1,845</i>	<i>67.2%</i>	<i>32.8%</i>
3622	425	417	842	50.5%	49.5%
3623	410	419	829	49.5%	50.5%
<i>May/June</i>	<i>835</i>	<i>836</i>	<i>1,671</i>	<i>50.0%</i>	<i>50.0%</i>
3624	330	152	482	68.5%	31.5%
3625	338	178	516	65.5%	34.5%
<i>July/August</i>	<i>668</i>	<i>330</i>	<i>998</i>	<i>66.9%</i>	<i>33.1%</i>
3626 <sup>i</sup>	364	130	494	73.7%	26.3%
3627	342	161	503	68.0%	32.0%
<i>November</i>	<i>706</i>	<i>291</i>	<i>997</i>	<i>70.8%</i>	<i>29.2%</i>
3626 <sup>ii</sup>	2	1	3	66.7%	33.3%
3628	173	155	328	52.7%	47.3%
3629	192	137	329	58.4%	41.6%
<i>January/February</i>	<i>367</i>	<i>293</i>	<i>660</i>	<i>55.6%</i>	<i>44.4%</i>
<b>ALL</b>	<b>3,816</b>	<b>2,355</b>	<b>6,171</b>	<b>61.8%</b>	<b>38.2%</b>

### *Distribution of Candidate Scores*

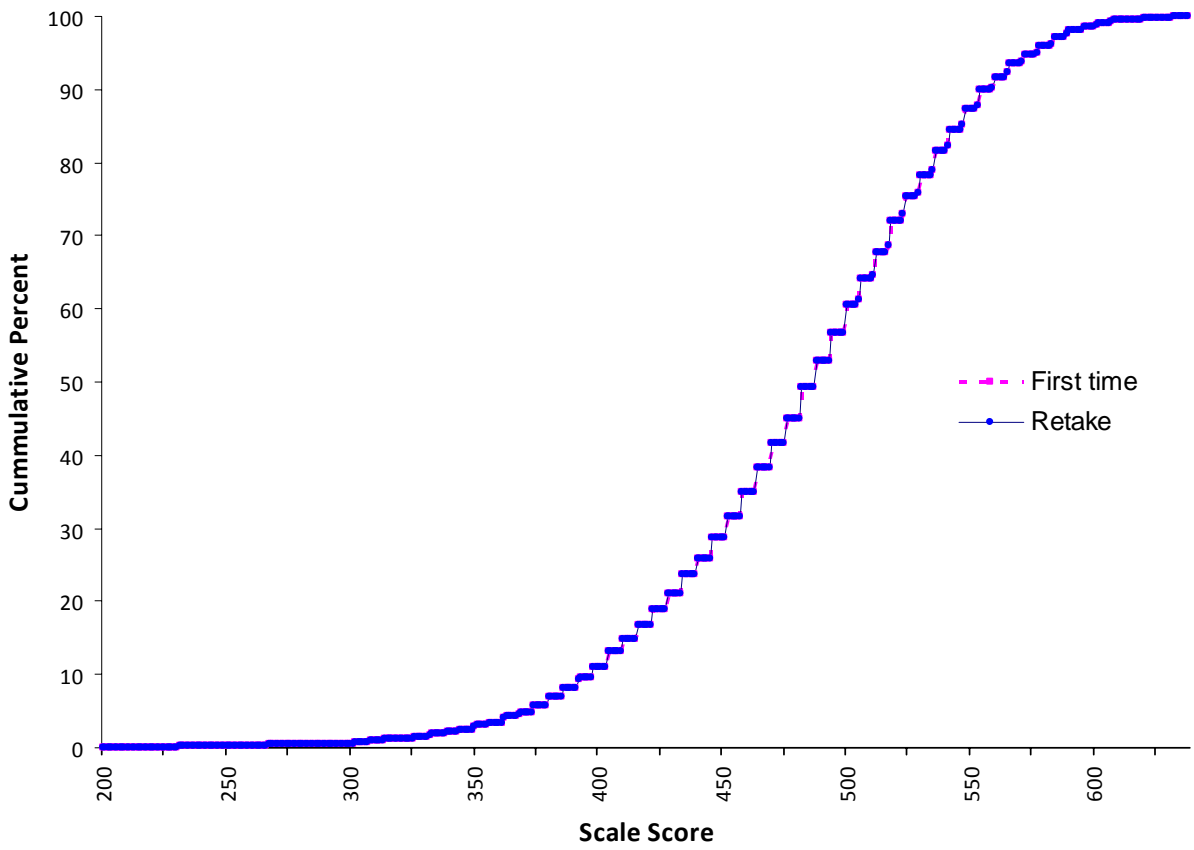
Table 3 details the overall scale score performance for the program for 2009-10 with a comparison the performance of 2008-09 candidates.

**Table 3:** 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, 2009-10.

Cohort	N	Avg.	Median	Std. Dev.	Min	Max
First-time	2,852	481	488	63	200	638
Retake	3,319	471	476	53	230	620
All 2009-10	6,171	476	482	58	200	638
All 2008-09	6,135	473	476	79	200	686

Similar to 2007-08 and 2008-09, a Univariate General Linear Model (GLM) test determined that there was a statistically significant, but very small, difference in the scaled scores of retake and first-time candidates [ $F(1, 6169) = 48.25, p < .001, \eta = .01$ ]. First-time candidates scored on average 10 scale points higher than retake candidates (481 compared to 471).

For 2009-10, the score distributions of first-time and retake candidates were almost identical despite this small statistical difference. This was confirmed by an examination of the distribution of scaled scores for first-time and retake candidates for whom no difference in the score distributions can be noted (Figure 1).



**Figure 1:** Cumulative Percentage of First-time New Graduates and Retake Candidates by Scaled Score, BOC 2009-10.

*Test Form Summary Statistics*

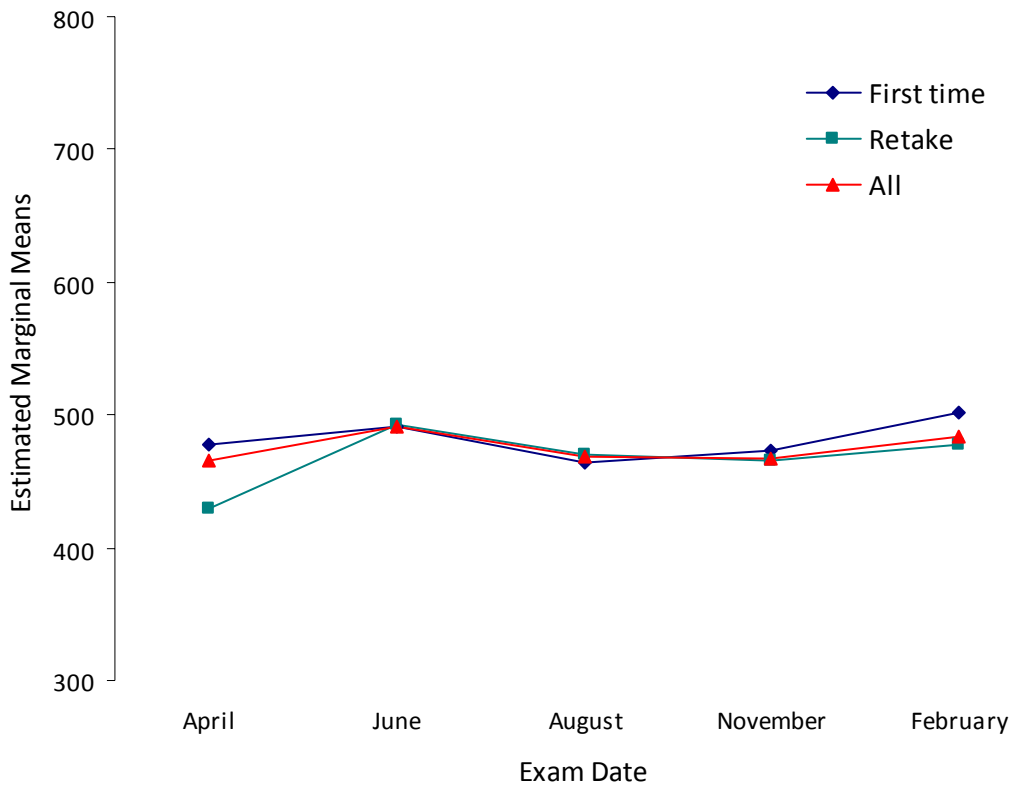
Table 4 provides form descriptive statistics for each test administration period.

**Table 4:** Summary Test Form Statistics in Scaled Scores for All Candidates for BOC Examinations, 2009-10.

Exam	N	Mean	Median	Std. Dev.	Min	Max
3618	81	466	458	61	350	626
3619	70	481	488	57	308	590
3620	867	465	470	63	200	620
3621	827	466	470	60	230	614
<i>March/April</i>	<i>1845</i>	<i>466</i>	<i>470</i>	<i>61</i>	<i>200</i>	<i>626</i>
3622	842	493	494	53	302	638
3623	829	491	500	60	248	638
<i>May/June</i>	<i>1671</i>	<i>492</i>	<i>500</i>	<i>56</i>	<i>248</i>	<i>638</i>
3624	482	466	470	56	231	613
3625	516	471	476	57	249	619
<i>July/August</i>	<i>998</i>	<i>469</i>	<i>470</i>	<i>57</i>	<i>231</i>	<i>619</i>
3626	494	465	470	53	248	631
3627	503	471	476	54	225	607
<i>November</i>	<i>997</i>	<i>467</i>	<i>476</i>	<i>54</i>	<i>225</i>	<i>631</i>
3626	3	462	482	56	398	505
3628	328	485	494	52	315	600
3629	329	482	488	54	321	631
<i>January/February</i>	<i>660</i>	<i>483</i>	<i>488</i>	<i>53</i>	<i>315</i>	<i>631</i>
<b>ALL</b>	<b>6171</b>	<b>476</b>	<b>482</b>	<b>58</b>	<b>200</b>	<b>638</b>

As shown in Table 4, there appears to be some difference in the scaled scores for each month and for each form. A statistical test (a Univariate General Linear Model) was undertaken to examine whether there was any statistical difference in the scaled scores for candidates based on the month they tested, whether they were retake or first-time candidates, and the form they received. There was a significant, though small, interaction between the candidate's month and retake status [ $F(4,6161) = 101.43$ ,  $p < .05$ ,  $\eta = .06$ ]. The major difference was in the month the candidate took the examination [ $F(4,6161) = 49.05$ ,  $p < .05$ ,  $\eta = .03$ ]. As with previous years, first-time candidates who took the examination earlier in 2009 or 2010 outperformed other candidates. (See Figure 2 for comparison of scale score means.)





**Figure 2:** Scale Score Means for First-time, Retake, and All Candidates for Each of Five Testing Windows, BOC 2009-10.

*Difficulty and Discrimination*

Average difficulty and discrimination was computed for all test forms. Table 5 contains the average, minimum, and maximum values for difficulty and discrimination.

**Table 5:** Average Difficulty and Discrimination Statistics for Each Test Form for All Candidates for BOC Examinations, 2009-10.

Test Form	Statistic	Number of Candidates	Average	Min	Max
3618	Difficulty	81	0.70	0.06	1.00
	Discrimination	81	0.21	-0.21	0.54
3619	Difficulty	70	0.72	0.07	1.00
	Discrimination	70	0.18	-0.30	0.51
3620	Difficulty	867	0.70	0.06	0.99
	Discrimination	867	0.22	-0.06	0.50
3621	Difficulty	827	0.70	0.05	0.99
	Discrimination	827	0.20	-0.18	0.41
3622	Difficulty	842	0.70	0.23	0.99
	Discrimination	842	0.17	-0.21	0.42
3623	Difficulty	829	0.70	0.24	1.00
	Discrimination	829	0.21	-0.13	0.46
3624	Difficulty	482	0.66	0.06	0.99
	Discrimination	482	0.18	-0.17	0.39
3625	Difficulty	516	0.67	0.05	1.00
	Discrimination	516	0.19	-0.08	0.46
3626	Difficulty	494	0.66	0.04	0.99
	Discrimination	494	0.18	-0.15	0.43
3627	Difficulty	503	0.67	0.04	1.00
	Discrimination	503	0.19	-0.03	0.42
3628	Difficulty	328	0.69	0.22	0.99
	Discrimination	328	0.17	-0.04	0.46
3629	Difficulty	329	0.69	0.21	1.00
	Discrimination	329	0.19	-0.07	0.53
<b>ALL</b>	<b>Difficulty</b>	<b>6,169</b>	<b>0.69</b>	<b>0.04</b>	<b>1.00</b>
	<b>Discrimination</b>	<b>6,169</b>	<b>0.20</b>	<b>-0.17</b>	<b>0.54</b>

Analysis was conducted on the difficulty and discrimination statistics obtained for the first large-scale administration of each of the two sets of 125-scored items (forms 3620 and 3622). A Multivariate GLM using Wilks' criterion test showed no significant difference in the discrimination between the two sets of 125-scored items [ $F(1, 248) = 0.10, p = 0.75, \eta = 0.00$ ]. Discrimination was statistically different [ $F(1, 248) = 6.06, p = 0.02, \eta = 0.69$ ], with the 125 scored items on form 3620 being slightly more discriminating than the scored items on form 3622.

Overall throughout the year, discrimination statistics for the items were within an acceptable range of 0.1 to 0.3. Average difficulty for the forms was slightly high for a five-option multiple-choice examination (ideally it should be around 0.6).

### *Domain Performance*

Test validity is a concept that refers to how well a test measures what it is designed to measure. Test forms for the BOC examinations were constructed according to test specifications that were based on the results of a role delineation study completed in 2004. This study was undertaken to define the job-related activities, knowledge, and skills required of entry-level athletic trainers. To ensure that test items account for the content areas presented in the test specifications, each item has been classified by content experts according to its application to the practice domains and tasks of the role delineation study.

Each multiple-choice test item has been linked to a specific content area of the test specifications, and items meet minimum standards of criticality to entry-level work as an athletics trainer. Thus, the procedures used to construct the test support the inference that the test has been built to achieve its stated purpose. Consistent with the objectives of the BOC examination program, the test is designed to separate candidates into two distinct groups: candidates whose knowledge and skill levels are deemed acceptable for entry-level certification as a practitioner and candidates whose level of knowledge falls below the minimum requirements for certification. The BOC examinations are not intended as predictors of future success within the profession.

There are six performance domains in the content framework for the BOC examination, consistent with the role delineation study upon which the examination is based (2004). The domains are *Prevention; Clinical Evaluation and Diagnosis; Immediate Care; Treatment, Rehabilitation, and Reconditioning; Organization and Administration; and Professional Responsibility*. Table 6 reports descriptive statistics at the domain level for the multiple-choice/part 1 examinations using raw score.

**Table 6:** Multiple-Choice/Part 1 Portion Domain Level Statistics for Each Test Form for All Candidates for BOC Examinations, 2009-10 (Based on Raw Scores).

Form	Statistic	Prevention	Clinical Evaluation and Diagnosis	Immediate Care	Treatment Rehabilitation and Reconditioning	Organization and Administration	Professional Responsibility
3618	N	81					
	Minimum	7	15	11	11	6	1
	Maximum	18	29	22	28	14	10
	Mean	11.8	21.7	16.7	18.8	10.7	7.4
	Std. Deviation	2.3	3.5	2.5	3.7	1.6	1.6
3619	N	70					
	Minimum	8	9	9	11	8	3
	Maximum	17	28	21	27	14	10
	Mean	12.6	22.9	16.5	19.3	11.2	7.4
	Std. Deviation	2.1	3.5	2.6	3.5	1.4	1.7
3620	N	867					
	Minimum	3	8	7	5	3	2
	Maximum	19	29	22	27	14	11
	Mean	12.1	21.8	16.2	18.5	10.7	7.5
	Std. Deviation	2.4	3.6	2.6	3.5	1.8	1.6
3621	N	827					
	Minimum	4	8	8	7	3	2
	Maximum	18	29	22	28	14	11
	Mean	12.2	21.9	16.2	18.5	10.7	7.6
	Std. Deviation	2.3	3.4	2.6	3.4	1.8	1.6
3622	N	842					
	Minimum	5	12	7	10	3	2
	Maximum	19	29	22	27	14	10
	Mean	13.2	22.5	16.2	18.5	10.5	6.9
	Std. Deviation	2.2	3.3	2.6	3.2	1.7	1.4
3623	N	829					
	Minimum	5	8	5	7	3	2
	Maximum	19	30	22	28	14	10
	Mean	13.1	22.4	16.0	18.4	10.5	7.0
	Std. Deviation	2.3	3.4	2.8	3.6	1.8	1.4
3624	N	482					
	Minimum	5	8	6	4	3	3
	Maximum	18	29	21	26	14	11
	Mean	11.6	20.9	15.4	17.3	10.4	7.3
	Std. Deviation	2.2	3.5	2.6	3.3	1.8	1.7

Form	Statistic	Prevention	Clinical Evaluation and Diagnosis	Immediate Care	Treatment Rehabilitation and Reconditioning	Organization and Administration	Professional Responsibility
3625	N	516					
	Minimum	4	7	8	7	3	2
	Maximum	18	29	22	26	14	11
	Mean	11.8	21.1	15.5	17.4	10.5	7.3
	Std. Deviation	2.3	3.5	2.7	3.3	1.8	1.6
3626	N	494					
	Minimum	4	12	5	6	3	3
	Maximum	19	28	21	25	14	11
	Mean	11.5	20.8	15.5	17.2	10.2	7.3
	Std. Deviation	2.2	3.3	2.5	3.2	1.9	1.5
3627	N	503					
	Minimum	4	5	7	7	4	2
	Maximum	18	29	22	27	14	11
	Mean	11.8	21.1	15.7	17.1	10.4	7.5
	Std. Deviation	2.2	3.4	2.6	3.3	1.8	1.6
3628	N	328					
	Minimum	5	10	8	11	4	3
	Maximum	18	29	21	27	14	10
	Mean	13.1	22.2	15.7	18.1	10.5	7.0
	Std. Deviation	2.0	3.5	2.6	3.1	1.7	1.3
3629	N	329					
	Minimum	8	9	8	8	2	3
	Maximum	19	30	21	25	14	10
	Mean	13.0	22.2	15.6	17.6	10.4	7.1
	Std. Deviation	2.1	3.4	2.5	3.3	1.8	1.5
<b>All</b>	<b>N</b>	<b>6171</b>	<b>6171</b>	<b>6171</b>	<b>6171</b>	<b>6171</b>	<b>6171</b>
	<b>Minimum</b>	<b>3</b>	<b>5</b>	<b>5</b>	<b>4</b>	<b>2</b>	<b>1</b>
	<b>Maximum</b>	<b>19</b>	<b>30</b>	<b>22</b>	<b>28</b>	<b>14</b>	<b>11</b>
	<b>Mean</b>	<b>12.35</b>	<b>21.78</b>	<b>15.89</b>	<b>18.03</b>	<b>10.55</b>	<b>7</b>
	<b>Std. Deviation</b>	<b>2.36</b>	<b>3.48</b>	<b>2.65</b>	<b>3.41</b>	<b>1.81</b>	<b>1.56</b>

Correlations in candidate performance between the six domains ranged from 0.18 to 0.50, indicating that the domains were assessing somewhat different constructs. These correlations are consistent with the results obtained for 2008-09.

*Test Form Internal Reliabilities*

Reliability is assessed using the Brennan-Kane statistic (Brennan & Kane, 1977), a measure typically used for estimating decision consistency for criterion referenced tests, and the Standard Error of Measurement (presented in Scale Score units). The Brennan-Kane reliability estimate accounts for the more constrained dispersion of candidate scores and the use of a passing standard and is consistent with reporting standards for accreditation purposes (Table 7).

**Table 7:** Internal Reliability Estimates for Multiple-Choice Section for Each Test Form for All Candidates for BOC Examinations, 2009-10.

Form	N	Std. Error	Brennan-Kane Estimate
3618	81	8.19	0.98
3619	69	8.05	0.98
3620	867	9.41	0.98
3621	827	8.92	0.98
3622	841	9.43	0.97
3623	828	10.32	0.97
3624	482	12.70	0.95
3625	516	12.79	0.95
3626	497	11.43	0.95
3627	504	12.60	0.95
3628	328	13.98	0.93
<b>ALL</b>	<b>6169</b>	<b>10.97</b>	<b>0.96</b>

Data presented in Table 7 show that the multiple-choice portion for each testing window meets general guidelines for a Brennan Kane statistics of greater than 0.70 and is consistent with previous years. Standard Errors of Measurement also are consistent with previous years.

*Summary test form data*

Data presented in the following table summarizes the performance of the test forms used for the BOC examination and is consistent with reporting requirements for NCCA/ICE Accreditation (Table 8). The data is also presented for each form family (A and B) that represents the common set of 125 scored items.

**Table 8:** Summary Statistics for the 2009-10 Administrations of BOC Athletic Trainer Test Forms.

Form #	Total # of Candidates Tested	% of Candidates Passing Each Form	Passing Point	Average Score	Standard Deviation	Standard Error of Measurement	Reliability Estimate	Total # of Items on Form
<b>Form Family A</b>								
3618	81	27.2%	500	466	61.2	8.19	0.98	160
3619	69	39.1%	500	481	56.9	8.05	0.98	160
3620	867	34.3%	500	465	63.3	9.41	0.98	155
3621	827	31.2%	500	466	59.8	8.92	0.98	155
3626	497	26.4%	500	465	52.8	12.70	0.95	155
3627	504	32.1%	500	471	54.4	12.79	0.95	155
Total A	2845	31.5%		467	58.8	10.22	0.97	
<b>Form Family B</b>								
3622	841	49.5%	500	493	53.2	11.43	0.97	155
3623	828	50.6%	500	491	59.6	12.60	0.97	155
3624	482	31.5%	500	466	56.1	9.43	0.95	155
3625	516	34.5%	500	471	57.2	10.32	0.95	155
3628	328	47.3%	500	485	51.6	13.98	0.93	155
3629	329	41.6%	500	482	53.8	13.08	0.94	155
Total B	3324	43.8%		483	55.8	11.57	0.96	
<b>TOTAL</b>	<b>6169</b>	<b>38.2%</b>	<b>500</b>	<b>476</b>	<b>57.1</b>	<b>11.20</b>	<b>0.96</b>	

Data presented Table 8 is in scale score units for passing point, average score, standard deviation, and standard error of measurement.

## Conclusion

Statistics concerning the quality of the BOC examination as a measurement device indicate that 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 very strong. Likewise, candidate performance on all parts of the examination is consistent with the public protection mission of the BOC.

<sup>i</sup> One candidate was administered the form in paper-and-pencil format and was excluded from this analysis.

<sup>ii</sup> This form was administered to three candidates in paper-and-pencil format.

## 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.
- Brennan, R. L., & Kane, M. T. (1977). An index of dependability for mastery tests. *Journal of Educational Measurement, 14*, 277–289.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.
- 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.
- 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.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*, 151–160.