Technical Report for CSW Computer Simulations LLC

(Computer administrated dental licensing testing developed by Central Regional Dental Testing Service (CRDTS), Southern Regional Testing Service (SRTA), and Western Regional Examination Board (WREB).

March 2007

Written By:
Del Hammond, M.A., Testing Specialist
Chad Buckendahl, Ph.D., Testing Specialist

Introduction

CSW Computer Simulations is a tri-member partnership company that creates and implements computerized periodontal and computerized 3-D prosthodontics dental licensing tests. CSW was officially formed in June of 2004. The CSW home office is located in Phoenix, Arizona.

CSW began administering their tests during the first quarter of 2006 at PearsonVue testing centers nationwide to candidates for dental licensure as part of the SRTA and WREB licensing tests. This report provides documentation describing test development and the post-test analyses from the first year of testing.

Statement of Purpose

The tests were designed to evaluate entry-level candidates' clinical application of knowledge and judgments necessary to provide periodontal and prosthodontic care.

General History

Dentists and dental educators representing CRDTS, SRTA, and WREB (CSW) first met in November, 2001 to develop a joint prosthodontics test. At this meeting, the committee discussed the possible development of a computerized test.

The prosthodontics committee task force decided in February, 2002 to pursue the concept of a computer-based test. In August of 2002 the committee met to evaluate presentations by three testing software developers. A contract for software development was awarded to Zoomorphix, an Australian firm with experience in developing performance test questions that rely on high quality graphics. In August, 2003 the joint CSW Prosthodontics committee met to begin reviewing models and questions that had been previously used independently by the three respective testing agencies.

In November, 2003 a joint CSW periodontal committee comprised of dentists and dental educators representing the three partnering organizations met to begin development of a computerized periodontal test that could be used to support licensure decisions. After the meeting, software and test development continued on both the periodontal and prosthodontic tests with the first tests administered via computer to candidates at PearsonVue testing centers in March, 2006.

Practice Analysis & Test Specifications

An updated practice analysis and test specification was completed by the WREB prosthodontics subcommittee in 1999. The CSW prosthodontics committee comprised of dentists and dental educators from each of the three partnering organizations reviewed the WREB analysis and revised the test specification. The most current test specification for prosthodontics is dated October 27, 2005.

The CSW periodontal committee, that is also comprised of dentists and dental educators from each of the three partnering organizations, analyzed the results from the periodontal practice analysis and developed a test specification in November, 2003. The current test specification for the periodontal test is dated July 10, 2004.

Item Development

The CSW prosthodontics and periodontal committees were composed of subject matter experts that included dentists who are university instructors in dental degree programs and practicing dentists who are dental examiners from the three testing agencies. The items used in the tests were required to be case specific in order to test clinical application skills. General knowledge questions were avoided as they would duplicate information that is gathered in other parts of the licensure testing process such as the national board exam.

The prosthodontics committee developed new models as well as using dental models that had previously been used in testing this discipline in dentistry by the three agencies. CSW's contractor, Zoomorphix, converted the models to three dimensional computer models. The committee reviewed the previous items and item analyses. Items that functioned well (i.e. met CSW's criteria for acceptable item difficulty and discrimination) and were appropriate for the test specification were entered in the item bank. Additional items were developed by the subject matter experts to satisfy the requirements of the test specification. The item pool was intended to be large enough to support multiple forms of the test. Multiple forms are necessary to protect item security and facilitate re-take opportunities for candidates who do not pass on their initial attempt.

The periodontal subject matter experts developed items based upon patient cases that were available at dental schools and private practices. The items for each form of the test utilize a case based on a single patient and cover medical history, oral history, assessment, prognosis, treatment plan, and re-evaluation. As with the prosthodontic test, multiple cases were developed to protect item security and allow for re-take opportunities for candidates who fail on their initial attempt.

The test items for both tests were field tested with junior dental students at various dental schools in each of the partner agencies territories using software

that was similar to that which is used at the testing center. This method of item tryout did not duplicate the testing situation exactly, so the final decision on which items would be scored was postponed until after actual dental licensing candidates had completed the tests. Item functioning was then evaluated and some items were removed from the items to be scored before any decisions were made about candidates' performance. Subsequent revisions to items that were not scored were field tested in unscored portions of the tests during candidate testing to gather information about the quality of the items prior to their operational use.

Standard Setting and Equating

The Ebel method (Livingston & Zieky) of standard setting was employed using the committee subject matter experts as judges. This methodology is a systematic, test-based approach where committee members make independent judgments on each item of the test to provide a recommendation for the passing score appropriate for the test based on its relative difficulty. The passing score recommendations based upon the committee judgments are then sent to the participating agencies and the CSW Management committee for final review and implementation.

Because the average difficulty of the items and the number of scored items for the different forms of each test were not precisely the same, a statistical strategy to insure that each form of the test is equal in terms of the meaning of the passing score was used. Specifically, CSW uses equipercentile equating (Kolen & Brennan, 1995) to provide equal treatment in the scoring of all candidates across forms of each respective examination. This strategy insures fairness to candidates because it means that it does not matter which form a candidate takes, the meaning of the decision is the same.

Results of Post-Test Analysis

CSW contracts Test Specialists to conduct a series of analyses to evaluate the characteristics of items and quality of the tests. The results of these analyses are included as Table 1. Internal consistency analyses are used to evaluate the extent to which the items are inter-correlated. This provides estimates of reliability. These values can range from 0.00 to 1.00 with values above 0.70 generally suggesting acceptable internal consistency reliability estimates. However, one of the conditions for internal consistency estimates such as KR-21 and KR-20 (coefficient alpha) is that there must be a range of ability in the underlying population of examinees in order to estimate reliability. In licensure and certification settings, the range of candidates' abilities tends to be narrower than the range of examinee abilities in other types testing. This lowers the correlation estimates. In these instances, it is more appropriate to use decision consistency estimates rather than internal consistency estimates. Decision consistency estimates are methods that evaluate the extent to which a test user

can be confident in the classification decision that results from an examinee's score relative to the cut score.

CSW Testing Specialists employ two methods for estimating decision consistency, the Hanson and Brennan (1990) method and the Breyer and Lewis (1994) method. Both methodologies rely on information from a single administration of the test and estimate decision consistency using variations of half-tests to estimate the level of confidence in the pass-fail decision on the full-length test. Although these methods rely on slightly different assumptions and calculations, the results are generally comparable. Values on decision consistency estimates can also range from 0.00 to 1.00; however, values of 0.85 or higher are desirable, particularly when there is only one pass-fail decision. In these instances testing agencies want to be very confident in the decisions that are made about a candidate's pass-fail status on a particular test. Readers can see in Table 1 that the decision consistency values generally meet this quality criterion for each form of the periodontics and prosthodontics tests in the CSW test. Note that the values for all forms are acceptable with many being well above .90.

At the item level, a series of classical test theory item analyses were conducted to evaluate whether items should be retained, revised, or removed from the item pool. The two diagnostic characteristics used to review these items are the item difficulty and item discrimination. Item difficulty is the proportion of candidates who respond correctly to the item. Item discrimination is an estimate of how the high scoring candidates performed on an item relative to low scoring candidates. Items that most candidates get wrong and items that most candidates get correct do not provide much information about the examinees. Since most dental candidates are well prepared for their profession and the test items are designed to be at a level that an entry-level dentist is expected to successfully answer, the average difficulty is a higher value (easier) and average discrimination is a lower value (less discriminating) than what is expected in other types of testing. This also lowers variability in scores and consequently lowers reliability estimates.

Summary

Well-qualified individuals have developed the CSW tests in accordance with accepted testing practices. This starts with the statement on purpose and includes: practice analysis, test specification, item development, standard setting, field-testing, administration, and post-test analysis that confirms satisfactory test performance. Test users can be confident that CSW scores provide a basis for valid licensing decisions.

References

Breyer, F. J. & Lewis, C. (1994, July). Pass-fail reliability for tests with cut scores: A simplified method. Princeton, NJ: Educational Testing Service.

Hanson, B. F. & Brennan R. L. (1990). An investigation of classification consistency indexes estimated under alternative strong true score models. Journal of Educational Measurement, 27(4), 345-360.

Livingston, S. & Zieky, M. (1982), *Passing Scores*, Princeton, NJ: Educational Testing Service.

Kolen, M. J. & Brennan, R. L. (1995). *Test Equating: Methods and Practices* (pp.35-47), New York, NY: Springer-Verla

Table 1. Statistics for the Periodontal and Prosthodontic Tests

	Perio 1	Perio 2	Perio 3	Perio 4	Pros 1	Pros 2	Pros 3	Pros 4
Descriptive statistics								
Number of items	35	32	33	30	32	30	29	31
Number of examinees	649	650	632	641	649	650	632	641
Cut score	24	20	22	19	20	20	20	19
Scale mean	29.156	25.846	27.649	25.811	25.037	23.014	23.136	25.109
Scale SD	2.836	3.154	2.858	2.669	3.347	2.983	2.683	2.761
Reliability								
KR-20 (alpha)	0.548	0.589	0.511	0.553	0.569	0.476	0.444	0.451
KR-21	0.406	0.516	0.465	0.511	0.530	0.412	0.363	0.387
Breyer & Lewis (1994): $P_{cc,full}$ (pass-fail classification consistency)	0.960	0.962	0.964	0.986[a]	0.926	0.864	0.890	0.977[a]
Hanson & Brennan (1990): p (classification consistency)	0.948	0.944	0.954	0.977	0.913	0.848	0.853	0.969
Hanson & Brennan (1990): kappa	0.327	0.240	0.224	0.079	0.267	0.301	0.201	0.190

a. average of two consistency indexes computed from two ways of cut score split

Attachment: A

CSW Computer Simulations
Prosthodontics Test Specifications

(Revised Oct 27 05)

2006 Computer Simulation	Complete 26%	Removable 36%	Fixed 26%	Implants 12%
Abutment Selection (anchorage/retention)			1	
Adjustments/Extensions		1		1
Aesthetics- Shape/Form/Size	1	1	1	
Biomaterial	1	1	*(1)	
Framework/Fab/Try-ins/Survey & Design Lab Evaluation/Communication	1	3	1	
Immediate/Overdenture		1		
Impressions			1	
Impression/Altered Cast		1		
Interim (provisional)			1	
Jaw Relations	2	2	*(1)	
Occlusal Vertical Dimension		1		
Occlusion	2	1	1	1
Phonetics	1	1		
Placement	1	1		
Pontic Retainer Design			1	
Preparation		1	4	
Relines/Repairs		1		
Surgical Considerations				1
Tooth Arrangement/Selection	3	2		
Tx Planning/Preprosthetic Considerations Site/Implant Selection	1		2	3
2006/ 50 Questions total	**13**	**18**	**13**	**6**

* Denotes question that could be used as an alternative.

Attachment: B

CSW Computer Simulations
Periodontal Test Specifications
(*revised July 10 2004*)

	Subject	Subcategories
10%	**Medical History and Physical Examination**	Adequate medical history Medications Systemic Influences Habits - smoking, chewing, alcohol use Pregnancy, joint replacement Heart problems, risk factors
4%	**Oral History**	Previous perio treatment Family history Oral lesions Previous endodontic treatment
2%	**Occlusion, Teeth and Restorations**	Bruxism Appliances Overhangs Carries Open contacts Hyper occlusion Alignment Lateral interference
30%	**Assessment**	Inflammation Topography Probing depths and (BOP) Attachment levels (CAL) Recession Furcation Mobility
16%	**Radiograph Evaluation**	Horizontal and vertical bone levels Calculus Furcation involvement Caries, restorations PDL Endodontic problems Anatomic factors
4%	**Diagnosis of Condition**	Distribution - localized/generalized Severity - chronic vs aggressive - All AAP definitions

4%	**Etiology of Condition**	Plaque and calculus Other local factors Systemic factors Iatrogenic
6%	**Prognosis - Response to Treatment**	Realistic Short and long term evaluation Individual and overall
12%	**Treatment Plan**	Logical sequencing Realistic treatment goals Surgical intervention/flap surgery Scaling and root planing Pharmacology modality Implants and periodontal implications Maintenance
12%	**Re-Evaluation of Treatment**	Mechanical and chemical plaque control Appropriate Consultation/Referral Follow up care Endpoint of Treatment