



WREB 2015 Technical Report Dental Examination

May 27, 2016



WREB Technical Report

Table of Contents

List of Tables	iii
List of Figures	iii
INTRODUCTION	1
OVERVIEW OF WREB DENTAL EXAMINATION	2
EVIDENCE FOR EXAMINATION VALIDITY	3
Examination Content	4
Construct Definition and Representation	4
Examination Committees	4
Dental Examination Specifications	5
Operative Section Examination Specifications	5
Endodontics Section Examination Specifications	7
Periodontal Section Examination Specifications	9
CTP Section Examination Specifications	9
Examination Administration	13
Examination Timing	13
Accommodations	13
Site Assignments of Examiners	14
Preparation of Candidates Regarding Novel Examination Formats	14
Examination Security	15
Examination Scoring	16
Decision-making Approach	16
Methods of Score Determination	16
Setting of Passing Scores	16
Training and Calibration of Clinical Examiners	18
Score Reporting	19
Penalties, Critical Errors and Unusual Situations	19
Issues Regarding Examination Failure	20
Examination Technical Quality	21
Overview of Methods	21
Tracking and Reporting of Passing Percentages	22
OVERVIEW OF DENTAL EXAMINATION TECHNICAL ANALYSES 2015	24
Criterion Analyses	24
Comparability Analyses	25
Examiner Performance	26
Overall Test Functioning	28
Dental Examination Passing Percentages 2015	29
REFERENCES	31

List of Tables

Table 1	Operative Section Procedure Options	5
Table 2	Operative Section Grading Criteria and Weighting: Preparation	5
Table 3	Operative Section Grading Criteria and Weighting: Finish	5
Table 4	Endodontics Section Required Procedures	7
Table 5	Endodontics Section Grading Criteria and Weighting	8
Table 6	Periodontal Section: Required Procedures	9
Table 7	Comprehensive Treatment Planning (CTP) Section Content Proportions for 2015 Patients Cases	10
Table 8	Five Types of Passing Percentages Tracked by WREB	23
Table 9	Descriptive Statistics of Raw Median Criterion Grades, Dental Examination Sections, 2015.....	24
Table 10	Many-Faceted Rasch Model Criterion Analysis Indicators in Logits, Dental Examination Sections, 2015.....	25
Table 11	Examiner Percentages of Agreement, Harshness, and Lenience, Dental Examination Sections, 2015	26
Table 12	Many-Faceted Rasch Model Examiner Severity Analysis Indicators in Logits, Dental Examination Sections, 2015.....	27
Table 13	Overall Test Summary Statistics for Dental Examination Sections, 2015.....	28
Table 14	Passing Percentages, Dental Examination and Sections 2015.....	29
Table 15	Dental Examination Passing Percentages over Past Seven Years	30

List of Figures

Figure 1	Example of Operative Preparation criteria definitions: Grading criteria definitions for the Preparation stage of the Direct Posterior Class II Composite procedure, 2015	6
Figure 2	Direct Finish criteria definitions: Grading criteria definitions for the Finish stage of the Direct Posterior Class II Composite procedure, 2015	7
Figure 3	Endodontics Section grading criteria definitions, 2015	8
Figure 4	CTP Treatment Plan Global Scoring Criteria, 2015	12
Figure 5	CTP Constructed-Response Scoring Criteria, 2015	12

INTRODUCTION

WREB develops and administers standardized competency assessments to support the licensing of dental professionals by state agencies and dental health care providers. Results from standardized assessments are one source of evidence used by licensing bodies to make decisions about a Candidate's readiness for practice, and must be developed and administered in a valid, reliable, and legally defensible manner. The purpose of this report is to provide test users with descriptive and technical documentation regarding the nature and quality of WREB examinations to support inferences based on examination results. WREB examinations are developed, administered, and scored in accordance with the *Standards for Educational and Psychological Testing* (AERA, APA, NCME; 2014) and *Guidance for Clinical Licensure Examinations in Dentistry* (AADB, 2005). An overview of WREB practices for monitoring and improving assessment quality is provided, as well as psychometric and statistical information that reflects examination quality for the current year. Technical information regarding examination quality is reviewed regularly by WREB's examination development committees, the WREB Dental Examination Review Board, the WREB Dental Hygiene Examination Review Board, WREB dental consultants, and the WREB Board of Directors. Details of additional activities and research studies that support the continued quality and improvement of WREB's examination system are also maintained and available to test users, test takers, and other stakeholders, where applicable.

An overview summarizing the WREB Dental Examination is provided first, followed by four sections describing evidence for examination validity: content, administration, scoring, and technical quality.

- **Examination Content** includes descriptions of the committees that develop, monitor and revise WREB examinations and provides details regarding examination specifications and alignment to analyses of dental practice.
- **Examination Administration** covers policies and practical features of the examination, related to the administration of the examination to candidates.
- **Examination Scoring** addresses standard-setting procedures, technical details of scoring, and issues related to score reporting and failure.
- **Examination Technical Quality** describes psychometric approaches used by WREB to evaluate examination quality.

The report concludes with an overview of Dental Examination technical analyses for 2015. Many technical analyses are conducted routinely and ad hoc but are not summarized in this document. Questions or additional details regarding any aspect of examination policies, procedures, administration or psychometric analyses are available upon request.

OVERVIEW OF WREB DENTAL EXAMINATION 2015

The purpose of standardized assessments that support licensure is to provide a reliable method for identifying practitioners who have met a minimum level of competence in the abilities critical to dental health care practice. Two major assessment approaches are employed to evaluate readiness for practice. One approach involves directly observing the Candidate's performance within an actual or simulated professional encounter. The other approach requires the Candidate to demonstrate professional knowledge, skills, and judgments via responding to a series of tasks or questions. WREB examinations utilize both approaches. A brief overview of the Dental Examination is provided below.

The WREB Dental Examination consists of four sections. Three sections (Operative, Endodontics, and Periodontal) are clinical examination sections that must be attempted initially at the same examination site. One section (Comprehensive Treatment Planning) is computer-based and must be attempted within the same examination season as the initial clinical examination is attempted. Passing the Dental Examination requires Candidates to pass all four sections within twelve months of their initial clinical examination attempt.

Operative Section.

The Operative section is a performance-based clinical examination. The Candidate is required to diagnose correctly and treat two restorative procedures on a Patient. The procedures are chosen from the following four options:

- Direct posterior Class II amalgam restoration (MO, DO or MOD)
- Direct posterior Class II composite restoration (MO, DO or MOD)
- Direct anterior Class III composite restoration (ML, DL, MF, DF)
- Indirect posterior Class II cast gold restoration (up to and including a $\frac{3}{4}$ crown)

Endodontics Section.

The Endodontics section is a performance-based clinical examination. The Candidate is required to perform endodontic treatment on two extracted teeth, mounted in a segmented arch, which is then mounted in an articulated full arch in a mannequin. The teeth must be:

- One anterior tooth
- One multi-canal posterior tooth

Periodontal Section.

The Periodontal section is a performance-based clinical examination. The Candidate is required to diagnose patient need for initial phase periodontal treatment and perform scaling and root planing on at least one quadrant of a Patient's mouth (a minimum of eight surfaces of readily demonstrable subgingival calculus must be present). Eight qualifying surfaces are assessed for errors.

Comprehensive Treatment Planning Section.

The Comprehensive Treatment Planning (CTP) section is a performance-based examination that requires the Candidate to assess three Patient cases of varying complexity, based on case materials presented interactively via computer. The CTP examination requires constructed-responses in the form of a submitted treatment plan, responses to case-related questions and completion of case-related tasks for each of the three cases presented.

EVIDENCE FOR EXAMINATION VALIDITY

Validity is the degree to which inferences and decisions based on test results are supported by evidence that the test is measuring the intended assessment construct and is developed, administered and scored in a manner that ensures reliability and fairness. WREB examinations are intended to measure clinical competence of Candidates seeking licensure in the dental and dental hygiene professions. The results are used by state dental boards and licensing agencies, along with educational requirements, national board test results, and other state requirements to evaluate Candidates and support licensure decisions. To ensure that inferences based on WREB examination results are credible and legally defensible, judgmental and empirical reviews are conducted regularly.

Judgmental review refers to the input, activities, and decisions made by subject matter experts at all levels of examination development and administration. Judgmental review ensures that WREB examinations are measuring dental and dental hygiene clinical competence in accordance with current standards of professional dental practice.

Empirical review refers to the on-going investigation of psychometric, statistical, and qualitative data generated within and by the WREB examination system. Empirical review supports continued quality and improvement and monitors adherence to current standards of educational and psychological testing.

WREB voluntarily undergoes independent external review on a regular basis and at any time upon request by our member states.

A review of WREB examination validity evidence for examination content, administration, scoring, and technical quality follows.

Examination Content

WREB examinations are intended to evaluate dental and dental hygiene clinical skills and abilities, including the ability to make appropriate diagnostic assessments and professional judgments, critical for entry-level practice. WREB has built an infrastructure that supports a broad, active network of subject matter experts. WREB subject matter experts ensure that all test specifications and examination-related content and activities reflect current standards of practice in dental health care. Subject matter experts and WREB staff develop and review test content in accordance with current professional standards and occupational analyses in dentistry and dental hygiene, including the 2005-2006 Survey of Dental Services Rendered (ADA, 2007), the Standards for Clinical Dental Hygiene Practice (ADHA, 2008), the WREB Practice Analysis for General Dentist (WREB, 2007), the WREB Dental Hygiene Practice Analysis Report (WREB, 2009) and the professional standards of practice within member states. A Dental practice analysis was conducted by WREB in 2015, which will support examination development and revision in forthcoming examination seasons.

Construct Definition and Representation

The procedures and tasks assessed within a clinical examination are sampled from the domain of professional practice. Measuring every single practice that entry-level licensees may be expected to perform is not possible. However, very limited assessment requirements can under-represent the domain of interest, leading to limited professional preparation which threatens the validity of inferences made from examination outcomes (Kane, 2006). The requirements of an examination that supports licensure decisions must assess broadly enough from professional practices to ensure adequate representation from the larger domain of all practices. Subject matter experts must review the domain of practices and decide upon a sample of practices for assessment and define criteria for measurement that reflect the judgments and skills expected of a minimally competent entry-level professional. The subject matter experts on WREB examination committees are informed by analyses of professional practices, field-testing, and results of psychometric evaluations to obtain evidence of construct-validity and assess examination quality and dimensionality.

Examination Committees

WREB examination committee responsibilities include on-going evaluation of current professional practices, test specifications, development of examinations and test forms, construction of examination-related informational materials for Candidates, development of Examiner training and calibration materials, monitoring test quality and reviewing examination feedback and suggestions (from Candidates, Patients, and Examiners). All of WREB's examination committees are composed of subject matter experts in dentistry and dental hygiene, representing various WREB member states. At least one member on each committee must be an active educator. The inclusion of an educator is critical because of their familiarity with the population and current dental and dental hygiene curricula. Other committee members must be experienced and licensed practitioners who have served as WREB Examiners (all of whom have served as state board members or designees). Committee membership rotates regularly to ensure regional diversity in representation, while maintaining continuity. Each committee is also supported by professional consultants in examination development and administration and WREB staff, including a professional psychometrician. Significant changes in examination content, administration, or scoring require approval by the Dental Examination Review Board and the WREB Board of Directors, which are comprised of state licensing board representatives from all of WREB's active member states.

Dental Examination Specifications

Examination specifications define the content to be assessed. For clinical examinations, the required clinical procedures are outlined, along with definitions of the specific grading criteria by which Candidate performance is assessed and the relative weighting of each criterion. For computer-based assessment, an outline of major content domains from which the test samples is provided, along with the proportion of assessment items per domain addressed. Current dental terminology (CDT) codes that reflect the range of acceptable procedures attempted are listed for Dental Examination sections.

The Dental Examination has four sections: Operative, Endodontics, Periodontics & Comprehensive Treatment Planning (CTP).

Operative Section Examination Specifications. The Candidate must diagnose correctly and treat two of the four restorative procedure options listed in Table 1. Candidates must submit each Patient, along with a completed medical history, Patient consent form, protective eyewear and radiographs for evaluation of acceptance criteria by Examiners. Acceptance criteria details for the Operative section are provided in the 2015 Dental Exam Candidate Guide (WREB, 2015a).

Table 1. Operative Section Procedure Options

Operative Section Restorative Procedure	CDT Code
Direct posterior Class II amalgam (MO, DO or MOD)	D2150, D2160, D2161
Direct posterior Class II composite restoration (MO, DO or MOD)	D2392, D2393, D2394
Direct anterior Class III composite restoration (ML, DL, MF, DF)	D2331, D2332, D2335
Indirect posterior Class II cast gold restoration (up to and including a $\frac{3}{4}$ crown)	D2780-82, D2520-D2944

Candidate performance on each procedure is graded by three independent and anonymous Examiners and weighted, at preparation and at finish, according to the criteria in Tables 2 and 3.

Tables 2 and 3. Operative Section Grading Criteria and Weighting: Preparation, Finish

Preparation Weighting	
Outline & Extension	46%
Internal Form	39%
Operative Environment	15%

Finish Weighting	
Anatomical Form	36.5%
Margins	36.5%
Finish, Function & Damage	27%

Each grading criterion is defined at five levels of performance for each procedure, with a grade of "3" representing minimal competence. A grade of "5" is defined generally to represent optimal performance, with grades of 4, 3, 2, and 1 corresponding to appropriate, acceptable, inadequate and unacceptable performance, respectively. The detailed definitions, as developed by the examination committee, are critical to guiding Examiner grading. The definitions are used to describe examples of

clinical performance reviewed during Examiner training and calibration, providing performance benchmarks to facilitate Examiner adherence to the criteria and a high degree of Examiner agreement. Figures 1 and 2 provide grading criteria definitions for the Preparation stage (i.e., Outline & Extension, Internal Form, and Operative Environment) of a Direct Posterior Class II Composite restoration and the Finish stage (i.e., Anatomical Form, Margins, and Finish, Function & Damage) for a Direct procedure, from the WREB Dental Exam Candidate Guide (WREB, 2015a) as Operative Section examples. All grading criteria definitions are available in the Candidate Guide.

2015 DIRECT POSTERIOR CLASS II - COMPOSITE PREPARATION
SCORING CRITERIA RATING SCALE

	5	4	3	2	1
OUTLINE & EXTENSION	Outline is generally smooth and flowing, and does not weaken tooth in any manner.	Outline is slightly irregular but does not weaken tooth. Isthmus is slightly wider than required for lesion.	Outline moderately weakens marginal ridge or a cusp. Isthmus is too wide or too narrow for lesion.	Outline severely weakens marginal ridge or a cusp. Outline is mishapen and/or forces improper angle of exit. Unsound demineralized enamel that is tactilely different from the adjacent unaffected enamel is present.	Outline is grossly improper and/or lacks any definite form. Tactilely unsound demineralized enamel penetrates the DEJ. Caries remains in the enamel or is not completely accessed. Unapproved surface prepared.
	Proximal and gingival extensions are visually open and break contact up to 1.0 mm.	Proximal and/or gingival extensions are slightly overextended.	Proximal and/or gingival extensions are moderately overextended.	Proximal and/or gingival extensions are in contact or obviously overextended.	Proximal and/or gingival extensions are grossly overextended.
	Optimal treatment of fissures.	Near optimal treatment of fissures.	Adequate treatment of fissures. Neither the tooth nor restoration is compromised.	Inadequate treatment of fissures will compromise the tooth or restoration.	Lack of treatment of fissures will seriously compromise the tooth and restoration.
	Proximal cavosurface angles are equal to or slightly greater than 90°. The integrity of both tooth and restoration is maintained.	Cavosurface angles are not optimal but do not compromise the integrity of the tooth or restoration. Cavosurface has small areas of minor roughness.	Cavosurface angles possibly compromise the integrity of the tooth or restoration. Cavosurface is moderately rough but will not adversely affect the final restoration.	Improper cavosurface angles or rough cavosurface will cause the final restoration to fail.	Cavosurface angles are grossly improper. Cavosurface has multiple major areas of roughness and/or enamel weakness that will cause the restoration to fail.
INTERNAL FORM	Pulpal floor depth as determined by the lesion or defect does not exceed 2.0 mm from the cavosurface. Enamel may remain on the pulpal floor. Axial wall depth at the gingival floor is 1.0 mm-1.5 mm.	Pulpal floor and/or axial wall is slightly shallow or deep.	Pulpal floor and/or axial wall is moderately shallow or deep.	Pulpal floor and/or axial wall is critically shallow or critically deep. Affected dentin remains.	Walls and/or floors are grossly deep with total lack of concern for the pulp. Caries remains in the dentin or is not completely accessed. Unapproved surface prepared.
	<u>Conventional design:</u> Internal form is smooth and flowing and has no sharp angles that could weaken or cause voids in the final restoration.	<u>Conventional design:</u> Internal form is mostly smooth and flowing, but some minor roughness and/or sharp angles are present.	<u>Conventional design:</u> Internal form is generally smooth and flowing, but some moderate roughness and/or sharp angles are present.	<u>Conventional design:</u> Internal form is rough and unfinished with major areas of roughness or sharp angles that will lead to restoration failure.	<u>Conventional design:</u> Internal form is grossly rough and/or has gross sharp angles that will lead to restoration failure.
	<u>Slot design:</u> Proximal box is present. Proximal line angles are ideal.	<u>Slot design:</u> Proximal box is present. Proximal line angles are slightly more or less rounded than ideal.	<u>Slot design:</u> Proximal box form has moderate variation from ideal.	<u>Slot design:</u> There is excessive rounding of all line angles. Excessive deviation from ideal proximal box form.	<u>Slot design:</u> There is gross lack of internal form.
OPERATIVE ENVIRONMENT	Rubber dam isolation is stable and optimal; the dam is inverted and has no rips, tears, bunching or exposed tissue. The preparation is clean and dry.	Rubber dam isolation is not optimal, but the preparation is clean and dry.	Rubber dam isolation is adequate, but the wrong teeth are isolated. The preparation can be cleaned and dried.	Rubber dam isolation is inadequate. The preparation is difficult to access or visualize due to blood or saliva on the preparation or partial coverage by the dam.	The rubber dam is grossly sloppy and torn, or portions of the preparation are not visible due to blood, saliva, or improper isolation.
	No damage to the adjacent tooth.	Minor damage to the adjacent tooth can be removed by polishing without changing the shape of the contact.	Damage to the adjacent tooth can be removed by polishing, but the shape of the contact will be changed.	Damage to the adjacent tooth will be difficult to polish out and still maintain appropriate proximal contour. The adjacent tooth will likely require restoration.	Damage to the adjacent tooth will definitely require restoration.

Figure 1. Example of Operative Preparation criteria definitions: Grading criteria definitions for the Preparation stage of the Direct Posterior Class II Composite procedure, 2015. Criteria definitions for all operative examination procedures are listed in the WREB 2015 Dental Exam Candidate Guide (WREB, 2015a).

**2015 DIRECT FINISH
SCORING CRITERIA RATING SCALE**

	5	4	3	2	1
ANATOMICAL FORM	Anatomical form is consistent and harmonious with contiguous tooth structure.	Slight variation in normal anatomical form is present.	Moderate variation in normal anatomical form is present. Marginal ridge is improperly shaped.	Anatomical form is improper. Marginal ridge is poorly shaped. Anatomy is too deep or too flat.	There is gross lack of anatomical form.
	Proper proximal contour and shape are restored.	There is slight variation of proximal contour and shape.	There is moderate variation of proximal contour and shape.	Proximal contour is poor. Embasures are severely over or under contoured.	Grossly improper proximal contour or shape.
	Normal proximal contact area and position are restored. Contact is visually closed and resists the passage of lightly waxed floss.	There is slight variation of normal contact area and position. Contact is visually closed and resists the passage of lightly waxed floss.	There is moderate variation of normal contact area and position. Lightly waxed floss will pass through the contact with slight resistance.	Contact is visually open, or floss will not pass through the contact.	Contact is grossly open, or the contact area is bonded to the adjacent tooth.
MARGINS	There are no excesses or deficiencies anywhere along margins.	Slight marginal excesses and/or deficiencies are present.	Moderate marginal excesses and/or deficiencies are present.	A deep open margin is present, or critical excesses or deficiencies are present. A marginal overhang catches floss.	Multiple open margins, or gross excesses or deficiencies, are present. A gross marginal overhang sheds floss.
	FINISH, FUNCTION & DAMAGE	The surface is smooth with no pits, voids or irregularities.	Slight surface irregularities, pitting, or voids are present.	Moderate surface irregularities, pitting, or voids are present.	Critical surface irregularities, pitting, or voids are present.
Occlusion is restored to proper centric with no lateral interferences.				There is severe hyperocclusion in centric or lateral excursions. Occlusal contact marks appear only on the restoration.	Occlusion is grossly inadequate.
There is no damage to hard or soft tissue.		Minor damage to hard or soft tissue is evident.	Moderate damage to hard or soft tissue is evident.	Severe damage to hard or soft tissue is evident.	Gross mutilation of hard or soft tissue is evident.

Figure 2. Direct Finish criteria definitions: Grading criteria definitions for the Finish stage of the Direct Posterior Class II Composite procedure, Direct Posterior Class II Amalgam procedure, and Direct Anterior Class III Composite procedure, 2015. Criteria definitions for all operative examination procedures are listed in the WREB 2015 Dental Exam Candidate Guide (WREB, 2015a).

Endodontics Section Examination Specifications. The Candidate must perform endodontic treatment on two extracted teeth, mounted in a segmented arch, which is then mounted in an articulated full arch in a mannequin. CDT codes that reflect the acceptable procedures attempted are listed in Table 4. Acceptance criteria details for the Endodontics section are provided in the 2015 Dental Exam Candidate Guide (WREB, 2015a).

Table 4. Endodontics Section Required Procedures

Endodontic Procedure	CDT Code
Anterior	D3310
Posterior (Multi-canal)	D3320, D3330

Candidate performance on each procedure is graded by three independent and anonymous Examiners and weighted according to the criteria listed in Table 5.

Table 5. Endodontics Section Grading Criteria and Weighting

Endodontic Weighting	
Access	37.5%
Condensation	62.5%

Similar to Operative Section grading, each Endodontics grading criterion is defined at five levels of performance for each procedure, with a grade of "3" representing minimal competence. Figure 3 provides grading criteria definitions for the Endodontics procedures, as presented in the WREB 2015 Dental Exam Candidate Guide (WREB, 2015a).

2015 ENDODONTIC SCORING CRITERIA RATING SCALE						
		5	4	3	2	1
ACCESS OPENING	Outline	Near ideal shape, size and location. For anteriors esthetics are not affected. If crown is fractured, access is intact or outline and shape can be determined by putting pieces back together.	Some variation in shape, size and/or location. May be slightly over or under extended. For incisors, minor encroachment on incisal edge, but is acceptable for apical instrumentation. If crown is fractured, access is intact or outline and shape can be determined by putting pieces back together.	Shape, size and/or location are functional. May be moderately over or under extended. For anteriors, encroachment on incisal edge is more than necessary for apical instrumentation. If crown is fractured, outline and shape can mostly be determined.	Improper shape, size and/or location (prevents proper instrumentation); or too large (crown is compromised by excessive extension). For anteriors, severe encroachment on the incisal edge inappropriate for apical instrumentation. If crown is fractured, outline and shape can partially be determined.	Grossly improper shape, size or location; crown severely compromised by gross extension. For anteriors, incisal edge is grossly violated, not necessary for apical instrumentation. If crown is fractured, outline and shape cannot be determined.
	Access	No obstructions to canals.	Slight over or under removal of tooth structure. Slight obstruction present.	Moderate over or under removal of tooth structure. Moderate obstruction present.	Excessive over or under removal of tooth structure (prevents proper instrumentation). Filled with gutta percha or other material preventing proper visualization of access.	External crown shape altered. Occlusal surface reduced. Coronal or furcal perforation.
	Chamber Roof/ Pulp Horn	Fully removed.	Not fully removed, a minor tooth ledge.	Not fully removed, moderate tooth ledge.	Not fully removed, excessive tooth ledge. Canal accessed through pulp horn only. Roof remains.	Not entered and canal not accessed.
CONDENSATION	Fill	Gutta-percha fully within root, less than or equal to 1.0 mm from apical foramen. Less than or equal to 1.0 mm of sealer extruded beyond apical foramen.	Gutta-percha fully within root, less than or equal to 1.5 mm from apical foramen. May have more than 1.0 mm but less than or equal to 3.0 mm of sealer extruded beyond apical foramen.	Gutta-percha less than or equal to 2.0 mm from apical foramen, short or long. Sealant extruded more than 3.0 mm beyond the apical foramen.	Gutta-percha less than or equal to 3.0 mm, short or long, from apical foramen.	Gutta-percha more than 3.0 mm short or long from apical foramen or none present; or an unacceptable material used.
	Density	Apical 1/3 dense and without voids. Slight voids in the coronal 2/3 of the fill.	Apical 1/3 dense and without voids. Slight voids in the coronal 2/3 of the fill.	Slight voids in the apical 1/3 or moderate voids in the fill.	Significant void in the fill.	Gross voids in the fill. No evidence of gutta percha condensation or compaction.
	Shape	Smooth and tapered from CEJ to apical foramen.	Smooth and tapered, minor irregularities. Minor under or over instrumentation.	Tapered with moderate irregularities. Moderate under or over instrumentation. Apex transported but less than or equal to 1.0 mm.	Tapered with significant irregularities. Excessive over or under instrumentation. Apex transported greater than 1.0 mm or less than or equal to 3.0 mm, creating an artificial canal.	Root perforation due to stripping. Apex transported greater than 3.0 mm creating an artificial canal.

A separated file in the canal will be scored based on established WREB criteria.
A root fracture can score no higher than a 3 for condensation.

Figure 3. Endodontics Section grading criteria definitions, 2015.

Periodontal Section Examination Specifications. The Candidate must diagnose Patient need for initial phase periodontal treatment and perform scaling and root planing on at least one quadrant of a Patient's mouth. Candidates must submit each Patient, along with a completed medical history, Patient consent form, protective eyewear and radiographs for evaluation of acceptance criteria by Examiners. Patient criteria for acceptance includes sufficient numbers of teeth, a minimum of eight surfaces of readily demonstrable subgingival calculus, and a minimum of two 5mm pocket depths, recorded by Candidates at submission and independently determined by grading Examiners. If additional teeth beyond one quadrant are submitted to meet the Patient acceptance criteria, all teeth in the second quadrant must be treated also. CDT codes that reflect the range of acceptable procedures attempted are listed in Table 6. Acceptance criteria details for the Periodontal section are provided in the 2015 Dental Exam Candidate Guide (WREB, 2015a).

Table 6. Periodontal Section: Required Procedures

Periodontal Treatment	CDT Code
Scaling and Root Planing (minimum eight qualifying surfaces)	D4341, D4342

Eight qualifying surfaces are assessed for errors. Points are deducted for calculus remaining, validated by at least two out of three independent and anonymous Examiners, proportional to surfaces treated, (i.e., 12.5% of points are deducted for each of eight treated surfaces validated by two or more Examiners to have calculus remaining).

Comprehensive Treatment Planning (CTP) Section Examination Specifications. The Candidate is required to assess three Patient cases of varying complexity, based on Patient case materials (i.e., Patient information, medical history, radiographic images, intraoral and extraoral photographs, dental and periodontal charts, clinical findings) presented interactively via computer.

Candidates assess Patient information, plan treatment and respond to related constructed-response questions for three Patient cases presented via computer. The CTP examination is designed to integrate the various disciplines of dentistry as conducted in practice. Appropriate treatment plans and question responses require the identification and planning of dental services that sample from a broad array of diagnostic, preventive, restorative, endodontic, periodontal and prosthodontic procedures, including procedures appropriate for the pediatric Patient, where applicable. The following list indicates the areas of dentistry that are assessed on the examination:

- **Restorative Treatment**
 - Single Units/Operative
 - Multiple Units
 - Fixed Prosthodontics
 - Interim Restorations
 - Removable Prosthodontics
 - Partial Dentures
 - Complete Dentures
 - Implant-Supported Restorations

- **Periodontal Treatment**
 - Phase I (Non- Surgical) Therapy
 - Re-evaluation
 - Surgery/referral
 - Maintenance
- **Endodontic Treatment**
- **Surgery**
 - Exodontia
 - Pre-prosthetic
 - Periodontal
 - Implant Placement
- **Prescription Writing**
 - Pharmacy
 - Dental Laboratory
- **Follow-up/Prognosis/Maintenance**
- **Diagnosis, Etiology and Treatment Planning** is integrated throughout the exam and overlaps the test specifications listed above. Also included are principles of pediatric dentistry, orthodontics, pharmacology, and specialist referrals where appropriate.

Table 7 provides the proportion of procedures, by category, that could be expected to be addressed in optimal treatment plans and constructed responses, distributed across all Patient cases for the current season. Examples of CDT codes that correspond to procedures addressed are also provided. The broad content category “Diagnosis, Etiology and Treatment Planning” is not listed in the content proportions in Table 7, but is integrated throughout all treatment plans and is assessed in addition to other content areas on almost 40% of all case-related questions.

Table 7. Comprehensive Treatment Planning (CTP) Section Content Proportions for 2015 Patient Cases, with Example CDT Codes (where applicable).

Procedure Category	Example CDT Codes	Proportion of Content Addressed
Restorative	D2140, D2150, D2160, D2161, D2330-D2332, D2335, D2391-D2394, D2510-D2664, D2710-D2799, D5110-D5140, D5211-5281, D5410-D5422, D5820-21, D6010-D6199, D6205-D6634, D6710-92, D6794	23%
Endodontics	D2430, D3220, D3310, D3320, D3330, D3333, D3410-D3426	25%
Periodontics	D4249, D4341, D4342, D4355, D4910	17%
Prescription Writing	-	13%
Other (Surgery, etc.)	D1110, D1120, D1203, D1204, D1351, D7111, D7140, D7210-D7240, D7320, D7880, D8010-90, D8210-20, D8660, D9940, D9972-73	21%

The Candidate is required to submit an appropriate treatment plan and construct responses to case-related questions and tasks for each of the three cases.

Each treatment plan must address the following requirements:

- Address appropriately the Patient's chief complaint or concern.
- Include appropriate treatment modifications if there are medical conditions that may affect the delivery of dental care. If medications are required, the plan must include drug, dose, and directions for use.
- Recommend additional diagnostic tests or specialist referrals as part of the treatment plan, if indicated. If referring to a specialist, a diagnosis and proposed treatment must be indicated.
- Contain a comprehensive appropriately-sequenced list of procedures that address the Patient's dental needs.
- Be succinct, organized, and readily interpretable.

Constructed responses to questions should address the question or task within the context of the Patient case and be clear, succinct, and easily understood by the Examiners; question responses do not need to be in full sentences. Example responses are included in the CTP Exam Candidate Guide.

All completed treatment plans and constructed responses are graded by three independent and anonymous Examiners according to scoring criteria published in the CTP Candidate Guide. Treatment plan global scoring criteria reflect the elements essential to the quality of any dental treatment plan and the constructed-response scoring criteria reflect treatment elements expected to address appropriately each case-related question or task. Each scoring criterion is defined with up to five levels of performance for each graded element, with a grade of "3" representing minimal competence. Figures 4 and 5 provide criterion definitions for the treatment plan global scoring criteria and constructed response scoring criteria, respectively, as presented in the WREB CTP Exam Candidate Guide (WREB, 2015b).

CTP GLOBAL SCORING CRITERIA					
	5	4	3	2	1
Treatment Modification	All needed modifications are appropriately noted (medication, referral, etc.). Specific medication and dosage is required.	The modifications are not optimal, but appropriate.	Not all modifications are noted, but the patient's health is not compromised. Modifications are noted at the minimally acceptable level.	Not all modifications are noted, and or the modifications are incorrect or incomplete. The patient's health may be compromised.	Modifications are ignored or incorrect and the patient's health is compromised.
Is the Treatment plan Inclusive?	All items that must be addressed are addressed. Chief complaint correctly addressed.	Most items that need to be addressed are addressed. Those that may be missing have little or no impact on the well-being of the patient. Chief complaint correctly addressed.	Missing items (one or more) do not pose a short term threat. Missing items (one or more) might affect patient well-being if the next regular periodic exam is missed. Chief complaint not fully addressed.	Important items (one or more) are missing. If not corrected, patient well-being is at risk in the near term. Chief complaint wrongly addressed.	Critical items (one or more) are missing. Patient well-being is currently at risk. Chief complaint not addressed.
Does the Treatment plan exhibit overtreatment?	Only those item(s) that must be addressed are addressed.	There are some (one or more) items that do not need to be addressed, but do not pose a risk to the patient.	There are some (one or more) items for which justification is questionable, but that pose little risk to the patient.	There are some (one or more) items that are not justified; if performed, they would result in limited physical damage to the patient.	There are multiple items that are not justified; if performed they would damage the patient.
Is the Treatment sequence appropriate?	The sequence is optimal.	The sequence is not optimal but will accomplish treatment goals.	The sequence is not correct, but can be corrected as treatment progresses. Rationale for the proposed sequence is unclear.	The sequence has definite flaws that are likely to result in backtracking and additional treatment.	The sequence has serious flaws and will not accomplish treatment goals.
Is the Treatment plan concise, well organized and easily interpreted?	The treatment plan is concise, well organized, and easily interpreted.		The treatment plan, as presented, may be confusing, but can be interpreted.		The interpretation of the plan cannot be determined.

Figure 4. CTP Treatment Plan Global Scoring Criteria, 2015.

CTP CONSTRUCTED-RESPONSE SCORING CRITERIA				
5	4	3	2	1
Optimal	Appropriate	Acceptable	Inadequate	Unacceptable
All essential elements are addressed and/or no errors are present. Response reflects optimal planning, rationales, and/or procedures.	Most essential elements are addressed and/or minor errors are present. Response is not optimal, but reflects an appropriate level of planning, rationales, and/or procedures.	Response is missing more than one essential element or contains moderate errors, but reflects planning, rationales, and/or procedures that are not likely to put the patient at risk.	Response is missing several essential elements or contains significant errors. Response reflects planning, rationales, and/or procedures that are likely to put the patient at risk.	Response is absent, missing most or all essential elements, or is mostly in error. Response reflects planning, rationales, and/or procedures that will harm the patient.

Figure 5. CTP Constructed-Response Grading Criteria, 2015.

Examination Administration

Standardization of examination administration and testing conditions ensures that all Candidates have an equivalent opportunity for success. WREB adheres to, and reviews regularly, examination administration policies and procedures that guarantee consistency and fairness of the examination experience for all Candidates. Examples of administration issues essential for standardization are reviewed briefly here, and include examination timing, accommodations, site assignments of Examiners, preparation of Candidates regarding novel examination formats, and examination security. Additional details of examination administration are available in the WREB 2015 Dental Exam Candidate Guide (WREB, 2015a), the WREB 2015 CTP Exam Candidate Guide (WREB, 2015b), on the WREB website (<http://www.wreb.org>), and in WREB Dental Policies and Procedures (WREB, 2015c).

Examination Timing

WREB examinations are administered within standardized time frames that provide adequate time for Candidates to complete the task and/or assessment. Speed of response is not an aspect of the assessment domains, so time limits are reasonable and set in accordance with Standard 4.14 of the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014). Details of time frames and limits are provided within each examination Candidate guide (WREB, 2015a & 2015b). The amount of time allowed for each examination is the same for all Candidates, unless an accommodation for additional time (applicable to computer-based tests) is granted (*Standards for Educational and Psychological Testing*, AERA, APA, & NCME, 2014; Americans With Disabilities Act, 1990).

The dental examination consists of one computer-based examination and three clinical examinations. The computer-based Comprehensive Treatment Planning (CTP) examination is administered by Pearson VUE at testing centers around the country. Dental Examination Candidates are allowed up to three hours to complete the CTP section, and provided with an additional fifteen minutes to complete a tutorial prior to beginning the examination. At the clinical examination, Candidates must complete the three clinical sections within 2½ days (8:00am to 4:30pm on days one and two and, if necessary, 8:00 to 11:00am on day three). Candidates are pre-assigned to a 4½ hour block of time during day one or two, for completing the Endodontics section. They must organize their time to complete two Operative section procedures and the Periodontal section, around their assigned Endodontics section block. While most Candidates do not need the additional morning, the third day is provided to allow for flexibility with scheduling Patients and to accommodate unexpected situations.

Accommodations

WREB makes every reasonable effort to offer examinations in a manner which ensures the comparability of scores for all Candidates, as per the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) and the Americans with Disabilities Act (1990). If an examination accommodation is requested and supported by documentation from an appropriate professional, WREB attempts to make the necessary provisions for the accommodation unless providing such would fundamentally alter the measurement of skills and knowledge the examination is intended to test or would provide an unfair advantage to the Candidate.

Site Assignments of Examiners

In addition to ensuring that grading Examiners are trained and calibrated to WREB grading criteria prior to every examination, the composition of the examining team for each clinic-based examination is planned with attention to several factors. Restriction or limits on participation by Examiners that belong to certain categories are followed, to prevent conflict of interest or to enhance grading quality. For example, an Examiner who is a dental educator may not examine in the state where he or she teaches and Examiners with connections to the examination site's host school may not participate in that examination. Examiners from member states are also prioritized in Examiner assignments; WREB requires member states to be involved in all aspects of examination administration, development, and review.

Examiner teams are also planned to ensure a very high level of calibration to WREB grading criteria. For example, Examiner teams may contain only one new Examiner, to allow maximum oversight and guidance of the new Examiner by the Examiner team captain. Site assignments are also planned to guarantee that all teams are interconnected to a degree that allows stable estimation of Examiner severity within statistical analyses of Examiner performance across the entire Examination season and across the entire Examiner pool.

Experienced Examiners are chosen for leadership roles, such as Team Captain, Floor Examiner, Endodontics Floor Examiner and Chief Floor Examiner. The Chief Floor Examiner ensures that the examination proceeds in accordance with established WREB policies and oversees the Examiner Orientation and Calibration Session. Grading Examiners never have contact with Candidates to guarantee anonymity in scoring. The only Examiners who have contact with the Candidates are Floor Examiners, which includes the Chief Floor Examiner, and do not function in a grading capacity. Floor Examiners must have experience as an Examiner, as they assist Candidates on the clinic floor and act as liaison between the Candidates and Grading Examiners. The Endodontics Floor Examiner oversees the Endodontics laboratory and performs the role of Floor Examiner for the Endodontic examination section. Team Captains are Grading Examiners who are also responsible for overseeing WREB procedures within the grading area, answering Grading Examiner questions and acting as primary contact with the Floor Examiners.

Preparation of Candidates Regarding Novel Examination Formats

Lack of familiarity with an examination format can be a source of construct-irrelevant variance, placing Candidates who are not familiar with the format at an unfair disadvantage. While most Candidates have had previous experience with computer-based assessment and interactive computer-based environments, WREB provides clear descriptions of computer-based examination features and multiple opportunities for Candidate review and practice, where applicable. For example, WREB's Comprehensive Treatment Planning (CTP) computer-based examination has features such as the ability to navigate among different screens, click on some screens to access additional information via "pop-ups" or enlarge images, and access additional resources available on different tabs. Prior to the examination, Candidates are provided with opportunities to participate in on-line tutorials to become familiar with the test's format and interface navigation options and receive examination descriptions, examples and example screen images in the CTP Candidate Guide. At the time of administration, Candidates also participate in an interactive instructional tutorial immediately prior to the examination.

Examination Security

WREB engages in practices and procedures which ensure the security of examination materials and the integrity of the examination process. A primary concern for computer-based tests is unauthorized exposure of assessment material, including details regarding simulated Patient cases and other assessment stimuli. WREB continually develops new testing materials to support multiple test forms. In addition, all Examiners, staff, and observers at examinations, as well as subject matter experts who participate on examination development committees, must sign a non-disclosure agreement regarding all secure examination material and information.

A primary concern for clinic-based examinations is Candidate identification. Candidates must confirm that all school credentials, personal identification documents, and photographs submitted in support of the examination application are authentic and unaltered, as well as agree to not disclose test questions or other examination-related materials.

WREB reviews security practices regularly from several perspectives: administrative, technological, legal, and psychometric. Potential threats to examination security are identified and prevention and response strategies are discussed (e.g., increasing educational efforts regarding appropriate test preparation practices to Candidates and educators).

Examination Scoring

WREB ensures that all examinations are scored accurately, fairly, and in accordance with the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014). Practices relevant to examination scoring include the decision-making approach; methods of score determination; setting passing scores; training and calibration of clinical Examiners; score reporting; penalties, critical errors, and unusual situations; and issues regarding examination failure.

Decision-making Approach

Information from multiple assessments, such as WREB's four Dental Examination sections, may be combined using one of two basic approaches, conjunctive or compensatory. WREB employs a conjunctive approach with regard to separate sections of the same examination. A conjunctive approach requires that performance on each element must meet or exceed a standard set for that element. In contrast, a compensatory approach combines scores for one final overall score; higher performance on one element may "compensate" for lower performance on another. In 2009, WREB moved from a partially conjunctive approach (i.e., performances on separate sections were compensatory, but only above lower bound limits set within each section) to a full conjunctive model. Candidates must meet the passing score for each examination section, set by examination committees within the conjunctive framework, to pass the examination.

Methods of Score Determination

The pass or fail decision regarding Candidate performance on examination sections is based on the final score, which is derived from a raw score. Raw scores for most WREB clinical and performance-based assessments are calculated by summing and averaging the median of ratings, or "grades," assigned by the grading Examiners on each scoring criterion. The raw score for the Dental Periodontal section is based on the percentage of Examiner-validated error-free tooth surfaces.

Where applicable, raw scores are scaled and/or equated to facilitate interpretability and to ensure comparability of scores on different test forms and across years. For example, the raw passing score on a difficult form of a test may be lower than the raw passing score on a less challenging form of the test. Scaling and equating procedures allow for unambiguous interpretation of comparable performance on each form, where a scale score of say, "75," represents passing on each form. Scaling is simply a linear or proportional conversion to another, more interpretable, numeric score scale. Linear equating or Rasch model equating is conducted to address variations in the difficulty level of multiple test forms or Patient cases. Pass or fail decisions based on final scores, after applicable weighting, equating, and scaling, reflect accurately the passing standards set by examination committees and ensure that Candidates of comparable proficiency will be equally likely to pass the examination, regardless of test form or date of administration.

Setting of Passing Scores

The process of setting the passing standard must be credible, legally defensible, and well-informed, in order to protect the public as well as the rights of Candidates. The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) state that passing standards should be high, in order to protect the public and the profession by excluding unqualified individuals, but not so high as to "unduly restrain the right of qualified individuals to offer their services to the public" (p.175).

Standard 11.16 in the current *Standards for Testing* states that the "level of performance required for passing a credentialing test should depend on the knowledge and skills necessary for credential-worthy performance in the occupation or profession and should not be adjusted to regulate the number or proportion of persons passing the test" (p. 182; AERA, APA, & NCME, 2014). The passing standards set by WREB examination committees are set in accordance with the *Standards for Testing* and are absolute, or criterion-referenced. An absolute, or criterion-referenced, standard is set to reflect a standard of knowledge and practice, meaning that, theoretically, all Candidates could pass or all could fail when compared to an absolute standard. In practice, pass rates of 100% and 0% are unlikely when a credible and defensible passing standard has been set. For many credentialing examinations, the vast majority of Candidates are very well-prepared, so relatively high pass rates are not unusual.

Passing scores on WREB examinations are set, and reviewed regularly, by WREB examination committees. WREB's examination committees determine passing scores based on professional standards of content and practice, even when arbitrary cut scores have been legislated, such as "75%." A passing score should reflect minimal competence, not an arbitrary percentage. Setting a passing score at 75% without evidence to support that the level of performance corresponds clearly to minimal competence is not a credible, defensible standard for a credentialing test; 75% of a difficult test is not comparable to 75% of a less challenging test. Some states have acknowledged that setting a percentage for passing is not appropriate. For example, California has stated that "Boards, programs, bureaus, and divisions that have laws or regulations requiring a fixed passing percent score should seek to change the law or regulation to require a criterion-referenced passing score that is based on the minimal competence criteria" (California Department of Consumer Affairs, 2000, p. 6). Until all states reject arbitrary fixed passing percentages, WREB continues to re-scale some examination passing scores to be interpreted as "75"; however, the scores reflect the defensible passing standard set by each professional examination committee. For performance-based tests, the examination committees define levels of performance with respect to critical aspects of clinical practice. The level of performance that reflects minimal competency (e.g., an average grade of "3.00" out of 5) is the passing score.

The standard-setting process for selected-response examinations, e.g., WREB's Dental Hygiene Local Anesthesia written examination, involves committee judgments of each item on the exam, according to Ebel's method (Ebel, 1972; Zieky, Perie, and Livingston, 2008). Each committee member must assign each test item to a category that reflects degree of professional relevance (e.g., essential) and degree of difficulty (i.e., the estimated probability of correct response by a minimally competent Candidate or empirical values of proportion correct if available). Estimated probability values are weighted by relevance and applied to the test form to set a raw passing standard. Raw scores may be further scaled to equate among test forms of differing difficulty with 75 as the scaled passing score for each form.

Standards set for performance-based examinations are based on definitions of professional behavior and performance, agreed upon and written by the examination committees. The committee defines minimally competent performance, and where applicable, defines additional levels of possible performance that exceed or fall below minimal competence. Definitions are developed to be as unambiguous as possible to facilitate a high degree of Examiner agreement. Committees determine whether a critical scoring criterion requires a dichotomous judgment (e.g., determining the presence

or absence of calculus remaining for the Periodontal section examinations), or a judgment aligned with multiple levels of performance quality (e.g., rating scales of 5 points for most Dental clinical sections). For example, on the Dental Operative section, each grading criterion is defined at five levels of performance for each procedure, with a grade of "3" representing minimal competence. A grade of "5" is defined generally to represent optimal performance, with grades of 4, 3, 2, and 1 corresponding to appropriate, acceptable, inadequate, and unacceptable performance, respectively. An example of the detailed definitions for Operative grading criteria are displayed on pages 6-7, in Figures 1 and 2. All grading criteria definitions for the Dental examination are available in WREB Candidate Guides (WREB, 2015a & 2015b).

Training and Calibration of Clinical Examiners

Clinical examination scores are dependent upon the judgments of grading Examiners. A high degree of Examiner agreement is critical to assessing Candidate ability in a reliable and fair manner. Ratings by a lenient Examiner for one Candidate cannot be compared meaningfully to ratings by a harsh Examiner for a second Candidate. Most examination judgments in WREB examinations are made by three independent Examiners. The median of the three grades assigned contributes to the Candidate's score. The median is more robust than the mean to extreme grades assigned. Situations where two Examiners may be involved in a decision that impacts the Candidate's score include evaluation of Patients for acceptance and clinical materials, and detection of conditions or behaviors that may result in a penalty; in these situations, at least two Examiners must validate on the same rationale for rejection or penalization, respectively.

Having multiple Examiners helps to moderate the effects of varying levels of Examiner severity; however, it is essential that all Examiners are trained and calibrated to an acceptable level of agreement with respect to the scoring criteria for the examinations in which they participate. Examiners are required to complete a series of tutorials and self-assessments prior to each examination. For each examination, Examiners spend approximately eight to ten hours of preparation time at home with WREB secure online training materials. Examiners must also attend orientation and calibration sessions that take place before every examination. New Examiners are also required to participate in an additional, earlier session to discuss their preparation with the Team Captain. During calibration, Examiners take assessments in which they grade examples of clinical performance according to the grading criteria. Their judgments are compared to scores that have been previously selected by the examination committees as representative of the defined levels in the criteria. The Examiner team completes calibration tests until they have all reached an acceptable level of agreement. All calibration tests are reviewed regularly for content and psychometric quality by WREB examination committees.

Most Examiners are members or designees of their state boards. Approximately ten percent of Examiners are dental educators; the proportion of educators is limited to prevent conflict of interest. All Examiners must be actively licensed and in good standing, with no license restrictions, submitting proof of license renewal annually. Most Examiners participate directly in grading, while some highly experienced Examiners participate in leadership roles, such as Chief Floor Examiner. Examiners receive regular feedback on their performance. Examiners with low percentages of agreement, high percentages of harshness or lenience, or erratic grading patterns are remediated and monitored to ensure increased understanding of criteria definitions. Continued lack of agreement may result in dismissal from the examination pool.

Score Reporting

WREB ensures that examination results are available to Candidates as soon as possible. Candidates are notified via electronic mail when they are able to access their official results at their secure WREB login online. Dental Candidates generally have access to their results within days after completing the examination. For computer-based examination sections, timing may be longer in the earliest part of the examination season, until a sufficient quantity of data has been collected to confirm the adequacy of equating.

WREB results focus on the Pass/Fail decision, e.g., Dental examination results show "Pass" or "Fail" for each of the four Dental examination sections. Pass/Fail decisions need to distinguish between Candidates who are minimally competent to practice the profession and those who are not. From a legal perspective, higher scores on a licensure examination do not reflect enhanced qualifications when the passing standard is developed to assess minimum, entry-level competence, consistent with statutory public protection obligations (Atkinson, 2012). The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) do not dictate the level of detail that a test user in licensing and certification applications is obligated to provide, other than whether the decision is passing or failing. While no obligation exists to report total scores or category subscore details, WREB recognizes that there is often a desire by Candidates for performance details beyond passing or failing. WREB Candidates who have been unsuccessful receive additional details regarding their performance, but they are encouraged to consider all content categories and criteria in their preparation for re-take, as performance within each category is likely to vary more than overall section score across subsequent performances. Detailed score reports are available to successful Candidates upon request.

Penalties, Critical Errors, and Unusual Situations

Some errors, as defined in the Candidate Guides (WREB, 2015a & 2015b) may result in point deductions on each of the examination sections, e.g., late penalties or if a modification request submitted during the Operative section is validated as not appropriate. Penalty values are set to discourage inappropriate behaviors or to reflect aspects of inadequate performance, not to diminish the intention of the pass/fail outcome that results from the grading of examination criteria. The impact of penalties is reviewed regularly to ensure that penalties rarely make the difference between passing and failing outcomes. The evaluation of proposed changes to penalty values includes the estimation of the impact that the proposed change will have on Candidate pass/fail outcomes.

Each examination section can result in loss of points or failure if a Candidate commits a critical error that is validated by Examiners. For example, a validated finding of caries remaining results in failure of the Operative section, a validated finding of major tissue trauma results in failure of the Periodontal section and a validated finding that a response reflects life-threatening harm, such as planning to administer a lethal dosage of local anesthetic to a pediatric Patient, results in failure of the CTP section.

Rarely, a Candidate may be dismissed from an examination because of an unusual situation. If a Candidate engages in improper performance relative to procedural skills or clinical judgment or exhibits unethical conduct he or she may be dismissed from the examination resulting in examination

failure and must obtain permission from the WREB Board of Directors to become eligible for re-examination.

Penalty and critical error details, definitions, possible point deductions, and examples of improper performance and unethical conduct can be found in the WREB Candidate Guides (WREB, 2015a & 2015b).

Issues Regarding Examination Failure

All test scores are subject to random error. Many sources of testing error, or construct-irrelevant variance, can be identified, addressed and minimized via best practices in psychometric analysis, regular review by subject matter experts and standardization of administration procedures. Construct-irrelevant variance may also stem from a Candidate's lack of information about examination logistics. To ensure Candidate knowledge of examination logistics WREB encourages Candidates to participate in multiple opportunities provided to review examination logistics through detailed Candidate guides, website resources and tutorials, pre-Candidate orientations, and Candidate orientations at each examination. WREB staff members also respond to Candidate questions via telephone and email communications. Other sources of construct-irrelevant variance include Candidate physical illness or anxiety, which can reduce the potential of the examination score to estimate accurately his or her actual level of ability or skill. Allowing an unsuccessful Candidate to attempt the examination again is reasonable and appropriate. WREB currently adheres to all testing standards relevant to informing Candidates about their results, as well as their rights and responsibilities with respect to examination failure and the opportunity to appeal an examination result and/or retake the examination.

A Candidate may appeal a failing examination result on a WREB examination. All procedures for filing an appeal, including criteria for consideration and related policies, are available on the WREB website (<http://www.wreb.org>). WREB maintains an Appeals Committee that is comprised of Examiners from WREB's Board of Directors appointed by the President. Members of the Appeals Committee must be current WREB Examiners. The committee provides anonymous, impartial, and timely examination appeal consideration to any Candidate who requests its services.

Candidates may retake failed examinations and examination sections; details regarding eligibility for re-examination and applicable remediation requirements are provided in the Candidate Guides for the Dental examination (WREB, 2015a & 2015b). If remediation is required before the Candidate may attempt the examination again, WREB notifies the Candidate of the required hours of remediation. Individual states may have additional requirements regarding remediation. Remediation must be completed at an accredited dental school in the United States or Canada and must include practical experience.

Examination Technical Quality

Empirical review of WREB examination quality is conducted throughout all stages of development, field-testing, revision, and operational administration. Results are reviewed with subject-matter experts from WREB examination committees and reported to WREB examination review boards. An overview of methods and quality indicators follow.

Overview of Methods

Analyses of graded elements and overall test functioning are conducted routinely on examination data. Methods are based on classical test theory and Rasch/item response theory (IRT) methods. Classical item analysis statistics reviewed include proportion per rating scale point; rating-measure correlations, c.f., point-biserial; and conventional descriptive statistics on graded elements (i.e., mean, standard deviation, etc.). Classical indicators of overall selected-response test performance reviewed include overall means, standard deviations, medians, standard errors of measurement, internal consistency reliability estimates, visual inspection of score distributions, as well as conditional standard errors of measurement at raw score passing cuts.

The Rasch model (Rasch, 1960/1980), c.f., one-parameter logistic IRT model, is the model chosen for the majority of psychometric applications in reviewing WREB examination data. The Rasch model is well-suited for monitoring and improving assessments because requirements of the basic model include data properties consistent with optimal test design (e.g., unidimensionality). Indicators of item and test performance under the Rasch model reflect the degree of departure from outcomes that would be expected given optimal item and test functioning. The basic Rasch model for dichotomous responses can be expressed as follows,

$$\log(P_{ni} / P_{ni} - 1) = B_n - D_i, \quad (1)$$

where P_{ni} is equal to the probability of correct response by a person n on a given item i , which is a function of the difference between the person's ability, B_n , and the item's difficulty, D_i . Rasch model analysis item statistics reviewed include parameter estimates of item difficulty, infit and outfit mean-square fit statistics, discrimination estimates and other statistics, where applicable (e.g., displacement values, when anchoring for pre-equating). For most analyses, means of all parameter estimates, except Candidate ability, are constrained at zero, to allow estimation of Candidate ability relative to item difficulty. Parameter estimates are reported in log-odds units, or logits, which can range from negative ∞ to positive ∞ , but usually do not exceed $|5.0|$. Lower, negative parameter estimates correspond to lower Candidate ability and lower levels of item difficulty. Higher, positive parameter estimates correspond to higher Candidate ability and higher levels of item difficulty. Fit statistics should generally fall between 0.5 and 1.5 logits, with a range of 0.8 to 1.2 logits considered reasonable for high-stakes selected-response tests (Wright and Linacre, 1994). Mean-square statistics that exceed 2.0 may reflect distortion in the measurement system and prompt close review. Discrimination values within the range of 0.5 to 1.5 provide reasonable fit to the Rasch model. The person separation reliability value is also noted, as it is similar to Cronbach's alpha internal consistency reliability estimate coefficient, except that it is calculated without the inclusion of perfect or zero scores. Rasch model indicators of overall selected-response test performance include model statistics, mean parameter estimates of Candidate difficulty, and review of item and Candidate score distributions via construct maps, also called Wright maps (Wilson, 2005).

Percentages of Examiner agreement, harshness, and lenience, are examined, by criterion or subset of criteria, where applicable. The many-faceted Rasch model (Linacre, Engelhard, Tatum, and Myford, 1994), an extension of Rasch ordered-category and partial credit models (Andrich, 1978; Masters, 1982; Rasch, 1960/1980), is applied to rating scale data to assess the effect of Examiners, as well as other potential sources of construct-irrelevant variance. The analysis applies a many-faceted conjoint measurement model which can estimate simultaneously Candidate ability and task difficulty while accounting for the degree of Examiner severity and other facets, where applicable. The many-faceted Rasch model is applied to all Examiner-graded assessments. For example, one of the models applied to the analysis of the Dental CTP section data is a four-facet model (i.e., Candidate, Examiner, CTP Patient Case and CTP Grading Criterion) that can be expressed as follows:

$$\log(P_{mnijk} / P_{mnijk-1}) = C_m - E_n - P_i - G_j - T_{kj}, \quad (2)$$

where P_{mnijk} is equal to the probability of Candidate m being rated k on Grading Criterion j within Patient Case i , by Examiner n . $P_{mnijk-1}$ is equal to the probability of Candidate m being rated $k-1$ on Grading Criterion j within Patient Case i by Examiner n . C_m is the ability of Candidate m , E_n is the severity of Examiner n , P_i is the difficulty of Patient Case i , G_j is the difficulty of Grading Criterion j , and T_{kj} is the difficulty of rating threshold k , relative to rating threshold $k-1$, for Grading Criterion j . The inclusion of the threshold parameter reflects a partial credit model, where estimates of rating category thresholds may vary within each item, and allows inspection of category functioning within each Grading Criterion.

Model statistics, including mean-square fit statistics (infit and outfit) and person separation reliability indices where applicable, are examined for Candidate, Examiner, scoring criterion, and other applicable facets. Parameter estimates, as with other Rasch analyses, are reported in logits, with lower estimates corresponding to lower Candidate ability, Examiner lenience, and lower levels of criterion difficulty. Higher, positive parameter estimates correspond to higher Candidate ability, Examiner harshness, and higher levels of criterion difficulty. As with the analyses of selected-response tests, fit statistics should generally fall between 0.5 and 1.5 logits. Wright and Linacre (1994) have suggested a range of 0.5 to 1.7 as reasonable for clinical observations and 0.4 to 1.2 logits as reasonable for tests that involve judgments. Category response thresholds are also examined in accordance with guidelines for optimizing rating scale effectiveness outlined by Linacre (2002).

Tracking and Reporting of Passing Percentages

Tracking the proportion of successful Candidates, e.g., over time, across examination sections, or among different test forms, is another component of technical review. Unexpected changes in trends over time or among Candidate subpopulations can reveal dramatic curricular shifts, threats to examination security or other phenomena that may warrant immediate investigation or pose a threat to examination validity. Reporting passing percentages provides a context for stakeholders (e.g., Candidates, state licensing Boards, educational institutions) with respect to the impact of examination outcomes.

Passing percentages can be computed and reported in different ways and for different purposes. Five types of passing percentages tracked at WREB are described below in Table 8.

Table 8. Five Types of Passing Percentages Tracked by WREB.

Type of Passing Percentage
All Examination Attempts
First Attempts
Retakes
By Individual Candidates at End of Season
Over Time (multiple years)

All Examination Attempts. The percentage of successful examination attempts out of all attempts, including all retakes, for a particular examination or section provides context for organizational planning and examination scheduling.

First Attempts. The percentage of successful first-time attempts provides Candidates, state licensing boards and educators with a context for the initial preparedness of the Candidate population.

Retakes. The percentage of successful retakes can provide comparison to first-attempt performance, which, particularly over time, should show that the likelihood of success decreases with subsequent attempts. All pass/fail tests, theoretically, misclassify some examinees (i.e., false negatives and false positives), particularly for observed scores that are close to the passing score. Providing appropriate retake opportunities allows a Candidate who was misclassified hypothetically in their examination outcome but may be truly minimally competent an opportunity to demonstrate minimal competence upon retake. However, the probability that a competent Candidate would be theoretically misclassified (i.e., false negative) upon third or higher retake becomes very low and decreases with the number of retakes (Clauser & Case, 2006).

By Individuals at End of Season. The individual passing percentage counts each individual Candidate’s final outcome for the examination season only, regardless of whether the Candidate passed upon first attempt or after two or more attempts. The individual passing percentage provides context for state licensing boards and the public regarding how many Candidates have met the clinical examination requirements for licensure within a given year.

Over Time (multiple years). Tracking passing percentages over time involves counting each individual Candidate’s final outcome at the end of a specified multi-year period. WREB longitudinal passing percentages are conducted every year for the past seven or more years. Failing percentages over time provide context for how many individual Candidates, even after multiple attempts and multiple remediation efforts, remain unsuccessful or never returned to participate in the retake process.

OVERVIEW OF DENTAL EXAMINATION TECHNICAL ANALYSES 2015

Basic Dental Examination analyses of grading criteria, comparability, Examiner performance and overall test functioning are summarized in this section. Passing percentages for 2015 and combined for the past seven years follow. Many other technical analyses are conducted routinely and ad hoc in addition to the analyses summarized here. Questions or additional details regarding any aspect of psychometric and statistical analyses are available upon request.

Criterion Analyses. Table 9 provides basic descriptive statistics for the raw means of all medians computed from the three sets of Examiner grades for each criterion (e.g., Operative procedures are graded on six criteria per procedure; three criteria for Preparation and three criteria for Finish).

Table 9. Descriptive Statistics of Raw Median Criterion Grades, Dental Examination Sections, 2015

Indicator	Operative Grading Scale: 1 to 5	Endodontics Grading Scale: 1 to 5	Periodontal Grading Scale: 0 or 1	CTP Grading Scale: 1 to 5
Raw Score Means (Range)	3.57 – 4.29	3.65 – 3.88	0.95 – 0.99	2.38 – 4.90
Raw Score Standard Deviations (Range)	0.61 – 0.73	0.82 – 0.83	0.11 – 0.22	0.49 – 1.70
Minimum; Maximum	1; 5	1; 5	0; 1	1; 5
Number of Criteria Graded	Six (by Two Procedures)	Two (by Two Procedures)	One (by Eight Surfaces per Attempt)	Nine to Thirteen (by Three Cases per Form)
Number of Graded Attempts	2,295	2,337	2,227	2,419
Total Number of Graded Procedures/Cases	4,590	4,674	2,227	7,257

Table 10, provides summary results from many-faceted Rasch model analyses for graded criteria in logit, i.e., log-odds, values. Mean-square fit statistics and discrimination parameter estimates are within suggested ranges. Criteria with multi-point rating scales are assessed for category functioning, as well, in accordance with Linacre’s (2002) rating scale guidelines (additional details are available upon request).

Table 10. Many-Faceted Rasch Model Criterion Analysis Indicators in Logits, Dental Examination Sections, 2015

Indicator	Operative (N = 2295)	Endodontics (N = 2337)	Periodontal (N = 2227)	CTP (N = 2419)
Criterion Measure Logit (Range)	-0.73 – 0.37	-0.32 – 0.38	-0.45 – 0.51	-1.53 – 0.72
Standard Error (Range)	0.01 – 0.01	0.02 – 0.02	0.01 – 0.01	0.01 – 0.02
Criterion Measure Logit Mean ^a	0.0	0.0	0.0	0.0
Criterion Measure Logit Standard Deviation	0.43	0.29	0.33	0.42
Many-Facet Point-Biserial Correlation ^b (Range)	0.22 – 0.28	0.36 – 0.37	0.13 – 0.23	0.10 – 0.21
2pl Discrimination Estimate ^c (Range)	0.94 – 1.06	0.97 – 1.03	0.97 – 1.03	0.83 – 1.19
Infit Mean-Square (Range)	0.94 – 1.05	0.98 – 1.03	0.98 – 1.03	0.93 – 1.09
Outfit Mean-Square (Range)	0.95 – 1.07	0.97 – 1.03	0.90 – 1.18	0.89 – 1.11

^a Mean constrained at 0 for criterion parameter estimation

^b Correlation between observations and corresponding average observations, excluding current observation

^c Estimate of discrimination parameter, as in two-parameter logistic IRT model; Rasch model fit requires values close to 1.00 (i.e., 0.5 to 1.5 logits)

Comparability Analyses. For the Operative and CTP sections, comparability of procedure combinations and test forms are evaluated. The Operative section allows the Candidate to choose different combinations of procedures, which has the potential to be a source of construct-irrelevant variance. In 2015, three combinations comprised all graded Operative submissions: a) Amalgam Direct Class II Posterior and Composite Direct Class II Posterior, b) Composite Direct Class II Posterior and Composite Direct Class III Anterior, and c) Amalgam Direct Class II Posterior and Composite Direct Class III Anterior. No Cast Gold Indirect Class II Posterior procedures were submitted in 2015. No significant difference ($\alpha = 0.05$ for all analyses reported) in Pass/Fail outcome was found among completed two-procedure graded submission combinations ($\chi^2 (2, N=2279) = 3.51, p = 0.17$).

The CTP section had eight different test forms comprised of six different Patient cases of varying complexity. Forms were assembled to ensure each form had one complex case, one moderate case,

and one pediatric case. Cases were designed to be highly comparable with respect to level of challenge to ensure that all Candidates engaged in comparable assessment experiences. Patient cases and forms were also equated to ensure that Candidates of comparable ability would obtain comparable results. No significant difference in Pass/Fail outcome was found among the eight CTP 2015 test forms ($\chi^2 (7, N=2419) = 8.00, p = 0.33$).

Examiner Performance

Examiner Agreement. One approach used to assess Examiner performance is to calculate the percentage of assigned grades in exact or adjacent agreement with the other two Examiners per graded element. Examiners may assign several hundred or more individual grades within an examination season. Each grade is compared to the mean of the other two grades assigned and if the difference exceeds 1.00, that grade is considered either Harsh or Lenient depending on the direction of the difference. Examiners are expected to be in exact or adjacent agreement in over 80% of assigned grades. Average percentages of Examiner agreement, harshness and lenience and ranges across individual Examiners are provided in Table 11. The few Examiners with percentages of agreement below 80% are less experienced Examiners and consistent with trends across years. Examiners with lower percentages of agreement and/or high percentages of harshness or lenience are remediated and monitored to ensure increased understanding of criteria definitions. Lenience tends to be very low for the Periodontal section, due to the high number of perfect and high scores. The many-faceted Rasch analysis provides additional insight into Examiner performance for the Periodontal examination.

Table 11. Examiner Percentages of Agreement, Harshness, and Lenience, Dental Examination Sections, 2015

Indicator	Operative	Endodontics	Periodontal	CTP
Agreement Percentage* Weighted Average	89.6%	90.3%	94.8%	83.5%
Agreement Percentage (Range)	80.4 – 94.8%	78.9 – 100.0%	81.5 – 100.0%	72.6 – 88.8%
Harshness Percentage Weighted Average	5.3%	4.8%	4.4%	8.6%
Harshness Percentage (Range)	0.7 – 15.7%	0.0 – 16.3%	0.0 – 17.9%	1.6 – 25.0%
Lenience Percentage Weighted Average	5.1%	4.8%	0.8%	7.9%
Lenience Percentage (Range)	0.4 – 15.2%	0.0 – 21.1%	0.0 – 6.5%	2.3 – 21.9%

*Agreement is exact and adjacent agreement for multi-rating sections: Operative, Endodontics and CTP; agreement is exact for the Periodontal section

Examiner Severity Estimation. The other approach used to assess Examiner performance is the estimation of Examiner severity within the many-faceted Rasch model, with high negative logits reflecting more lenience and high positive logits reflecting more harshness. Table 12 provides summaries of results in logit units. Most Examiners fall within one logit unit of the mean; Examiners at the extremes of each examination section range are reviewed for possible remediation and monitoring. Examiner severity estimates are highly correlated with Examiner agreement; however, the Rasch analysis allows Examiner performance to be compared across all Examiners across all examination sites which can temper the effects of specific groupings of three Examiners where one highly calibrated Examiner could be assessed as harsh, when compared to two Examiners that may be somewhat lenient. Most Examiners fall within recommended ranges with respect to infit and outfit mean-square fit statistics. While most high values of mean-square fit statistics are also associated with harshness or lenience, occasionally a high value can reveal erratic or inconsistent grading, which may be overlooked when reviewing conventional Examiner agreement statistics. Examiner teams are also compared within the Rasch framework as well as comparing weighted averages of agreement to assess comparability of examination sites. Details of exam site comparability analyses are available upon request.

Table 12. Many-Faceted Rasch Model Examiner Severity Analysis Indicators in Logits, Dental Examination Sections, 2015

Indicator	Operative ($N_E = 110$)	Endodontics ($N_E = 110$)	Periodontal ($N_E = 110$)	CTP ($N_E = 110$)
Severity Measure Logit (Range)	-0.98 – 1.12	-1.60 – 0.92	-2.24 – 1.59	-0.74 – 0.51
Standard Error (Range)	0.03 – 0.16	0.06 – 0.17	0.11 – 0.49	0.01 – 0.05
Severity Measure Logit Mean ^a	0.0	0.0	0.0	0.0
Severity Measure Logit Standard Deviation	0.41	0.43	1.09	0.21
Infit Mean-Square (Range)	0.58 – 1.87	0.47 – 2.01	0.86 – 1.24	0.67 – 1.63
Outfit Mean-Square (Range)	0.59 – 1.87	0.50 – 1.88	0.31 – 1.82	0.68 – 1.75

^a Mean constrained at 0 for criterion parameter estimation

^b Correlation between observations and corresponding average observations, excluding current observation

^c Estimate of discrimination parameter, as in two-parameter logistic IRT model; Rasch model fit requires values close to 1.00 (i.e., 0.5 to 1.5 logits)

Overall Test Functioning. Table 13 provides summary statistics for Dental examination section test functioning. The Operative, Endodontics and CTP sections have small Conditional Standard Errors of Measurement (CSEM) and moderately high Rasch person separation reliability estimates, which are similar to Cronbach’s alpha coefficient of internal consistency reliability, but exclude zero and perfect scores. The Periodontal section has a relatively high CSEM and a very low reliability estimate, which is due to the highly skewed distribution of scores. Most Periodontal Candidates perform very well or obtain perfect scores; however, a small percentage of Candidates fail the Periodontal section, even upon multiple retake. Reliability is often attenuated in criterion referenced credentialing assessment because of the high level of Candidate preparedness. Trends in passing percentages over time become critical for characterizing the quality of the Periodontal section and providing evidence of test validity.

Table 13. Overall Test Summary Statistics for Dental Examination Sections, 2015

Indicator	Operative (Max Score 5)	Endodontics (Max Score 5)	Periodontal (Max Score 100)	CTP (Max Score 5)
<i>N</i> Attempts	2267	2333	2227	2354
Score Mean	3.67	3.79	97.2	3.63
Score Standard Deviation	0.39	0.57	7.41	0.32
Minimum; Maximum	0.58; 4.77	0.65; 5.00	25.00; 100.00	2.21; 4.49
Conditional <i>SEM</i> at Passing Score	0.018	0.013	6.69	0.019
Indicators below are reported in logits:				
Candidate Ability Estimate Logit Mean	1.27	1.11	3.95	0.52
Candidate Ability Estimate Logit <i>SD</i>	0.68	0.96	1.13	0.29
Logit Minimum; Maximum	-2.85; 3.53	-3.78; 5.79	-0.15; 6.34	-0.72; 1.45
Person Separation Reliability Estimate*	0.87	0.81	0.10	0.87

*Comparable to alpha coefficient internal consistency reliability estimate (Cronbach, 1951) with zero and perfect scores excluded

Dental Examination Passing Percentages 2015

Dental Examination passing percentages are shown for:

- All attempts – includes all examination attempts including all retakes.
- First attempts – counts only initial examination attempts
- Retakes – counts only re-examination attempts (i.e., second or higher attempts). For Overall Dental, retakes can include between one and all four sections; most retakes involve one- or two-section re-examination attempts.
- Individual Candidates at End of Season – counts each Candidate’s final result at the end of the examination season, i.e., each Candidate is counted only once, even if they engaged in one or more retakes
- Individual Candidates at End of 2009 to 2015 – counts each Candidate’s final result at the end of the seven-year period from 2009 to 2015, i.e., each Candidate is counted only once, even if they engaged in multiple retakes across years

The first four types of passing percentages are provided in Table 14. Note that the Overall Dental passing percentages are lower than the passing percentages for each section because all sections must be passed to pass the overall Dental Examination. Also note that Candidate numbers for the CTP section are about 5% higher than other sections due to incomplete examinations (e.g., Candidates that took the CTP section early in the season but chose to not take the clinical section because of acceptance into post-graduate programs).

Table 14. Passing Percentages, Dental Examination and Sections, 2015

Section	All Attempts (Includes Retakes)		First-time Attempts		Retakes		Individual Candidates (End of season result)	
	% Passing	<i>N</i>	% Passing	<i>N</i>	% Passing	<i>N</i>	% Passing	<i>N</i>
Operative	94.3%	2,295	94.9%	2,175	82.5%	120	98.1%	2,204
Endodontics	92.2%	2,337	92.7%	2,175	94.8%	162	98.0%	2,198
Periodontal	98.4%	2,192	98.5%	2,172	96.4%	55	99.8%	2,189
CTP	94.3%	2,419	98.8%	2,309	87.3%	110	98.7%	2,309
Overall Dental	83.4%	2,562	83.5%	2,176	82.9%	386	96.4%	2,217

Passing percentages, broken down by first, second, and higher attempts, for the seven-year period from 2009 to 2015 are provided in Table 15. Over the past seven years, almost 15,000 individual Candidates engaged in almost 18,000 Dental examination attempts. The proportion of individual Candidates who remain unsuccessful over time continues to fall between 2 and 3% upon each seven-year period update.

Table 15. Dental Examination Passing Percentages over Past Seven Years, 2009 – 2015

Attempt Category	Total <i>N</i>	Passing <i>N</i>	Failing <i>N</i>	Passing Percentage
First Attempts	14,766	12,150	2,616	82.3%
Second Attempts	2,573	2,087	486	81.1%
Third Attempts	459	316	143	68.8%
Fourth Attempts	69	37	32	53.6%
Fifth or higher Attempts	45	18	27	40.0%
All Dental Examination Attempts	17,912	14,608	3,304	81.6%
Individual Candidate Results After Seven Years	14,915	14,598	317	97.9%

REFERENCES

- American Association of Dental Boards (2005). *Guidance for Clinical Licensure Examinations in Dentistry*. Chicago, IL: AADB.
- American Dental Association. (2007). *The 2005-06 Survey of Dental Services Rendered*. Chicago, IL: ADA.
- American Dental Hygienists' Association. (2008). *The Standards for Clinical Dental Hygiene Practice*. Chicago, IL: ADHA.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Americans With Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 328 (1991).
- Andrich, D. (1978). A rating formulation for ordered categories. *Psychometrika*, 43, 357- 374.
- Atkinson (2012) Legal issues and considerations for standard setting in professional licensure and certification examinations. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Innovations* (pp. 502-534). New York: Routledge.
- California Department of Consumer Affairs (2000). *Examination Validation Policy*. Sacramento, CA: CDCA.
- Clauser, B. E., & Case, S. M. (2006). Testing for Licensing and Certification in the Professions. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 701-731). Westport, CT: American Council on Education, Praeger.
- Ebel, R. L. (1972). *Essentials of Educational Measurement* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education, Praeger.
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.
- Linacre J. M., Engelhard, G., Jr., Tatum, D. S., & Myford, C. M. (1994). Measurement with judges: Many-faceted conjoint measurement. *International Journal of Educational Research*, 21(6), 569-577.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.

- Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research, 1960. Expanded edition, Chicago: The University of Chicago Press, 1980.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Western Regional Examining Board. (2007). *WREB Practice Analysis for General Dentist*. Phoenix, AZ: WREB.
- Western Regional Examining Board. (2009). *WREB Dental Hygiene Practice Analysis Report*. Phoenix, AZ: WREB.
- Western Regional Examining Board. (2015a). *2015 Dental Exam Candidate Guide*. Phoenix, AZ: WREB.
- Western Regional Examining Board. (2015b). *2015 CTP Exam Candidate Guide*. Phoenix, AZ: WREB.
- Western Regional Examining Board. (2015c). *Dental Policies and Procedures*. Phoenix, AZ: WREB.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8, 370. Retrieved November 29, 2009 from <http://www.rasch.org/rmt/rmt83b.htm>.
- Zieky, M.J., Perie, M., & Livingston, S. L. (2008). *Cutscores: A Manual for Setting Standards of Performance on Educational and Occupational Tests*. Princeton, NJ: Educational Testing Service.

Sharon E. Osborn Popp, Ph.D.
WREB Testing Specialist/ Psychometrician

Copyright 2016
WREB - A National Dental and Dental Hygiene Testing Agency
23460 N. 19th Avenue, Suite 210 - Phoenix, Arizona 85027

phone: 602-944-3315 fax: 602-371-8131

www.WREB.org
generalinfo@wreb.org • dentalinfo@wreb.org • hygieneinfo@wreb.org