WREB
2016 Technical Report
Dental Hygiene Examinations

July 12, 2017

WREB Technical Report

Table of Contents

List of Tables iii
List of Figures iii
INTRODUCTION 1
OVERVIEW OF WREB DENTAL HYGIENE EXAMINATIONS 2016 2
EVIDENCE FOR EXAMINATION VALIDITY 3
Examination Content 4
Construct Definition and Representation 4
Examination Committees 4
Dental Hygiene Examinations Specifications 5
Dental Hygiene Examination Specifications 5
Local Anesthesia Examination Specifications 6
Restorative Examination Specifications 9
Examination Administration 11
Examination Timing 11
Accommodations 11
Patient Safety and Comfort 12
Infection Control 12
Site Assignments of Examiners 13
Examination Security 13
Examination Scoring 15
Decision-making Approach 15
Methods of Score Determination 15
Setting of Passing Scores 15
Training and Calibration of Clinical Examiners 17
Score Reporting 18
Penalties and Unusual Situations 18
Issues Regarding Examination Failure 19
Examination Technical Quality 20
Overview of Methods 20
Tracking and Reporting of Passing Percentages 21
OVERVIEW OF DENTAL HYGIENE EXAM TECHNICAL ANALYSES 2016 23
Dental Hygiene Examination 23
Local Anesthesia Examination 25
Restorative Examination 29
Examiner Performance 31
Dental Hygiene Examinations Passing Percentages 2016 34
REFERENCES 36

List of Tables

Table 1 Treatment Evaluation for Dental Hygiene Examination 5
Table 2 Local Anesthesia Written Section: Domains Assessed 6
Table 3 Eight Aspects of Injection Evaluated on the Local Anesthesia Clinical Section 7
Table 4 Restorative Section Grading Criteria Weighting 9
Table 5 Five Types of Passing Percentages Tracked by WREB 22
Table 6 Dental Hygiene Examination 2016 Graded Elements: Descriptive Statistics..... 23
Table 7 Dental Hygiene Examination 2016 Graded Elements: Many-Faceted Rasch Model Analysis Indicators in Logits 24
Table 8 Dental Hygiene Examination 2016: Penalties Applied 24
Table 9 Overall Test Summary Statistics for Dental Hygiene Examination, 2016 25
Table 10 Local Anesthesia Written 2016 Forms: Mean Number Correct and Mean Number Correct and Standard Deviation by Content Area 26
Table 11 Local Anesthesia Written 2016 Forms: Indicators of Overall Test Functioning by Form 26
Table 12 Local Anesthesia Clinical 2016: Validated Critical Errors per Attempt ……..... 27
Table 13 Local Anesthesia Clinical 2016: Number of Validated Errors per Injection Aspect 27
Table 14 Local Anesthesia Clinical Examination, 2016 Critical Errors: Many-Faceted Rasch Model Analysis Indicators in Logits 28
Table 15 Restorative Examination 2016 Graded Elements: Descriptive Statistics 29
Table 16 Restorative Examination 2016 Graded Elements: Many-Faceted Rasch Model Analysis Indicators in Logits 30
Table 17 Overall Test Summary Statistics for Restorative Examination 2016 31
Table 18 Examiner Percentages of Agreement, Harshness, and Lenience: Dental Hygiene, Local Anesthesia Clinical and Restorative Examinations, 2016 32
Table 19 Many-Faceted Rasch Model Examiner Severity Analysis Indicators in Logits: Dental Hygiene Examinations, 2016 33
Table 20 Passing Percentages, Dental Hygiene Examinations and Sections, 2016 34
Table 21 Dental Hygiene Examinations Passing Percentages Over Past Seven Years, 2010 – 2016 35

List of Figures

Figure 1 2016 Restorative Examination Grading criteria 10

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 2016. 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 and dental hygiene 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 Hygiene 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 Hygiene Examination technical analyses for 2016. Many additional 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 HYGIENE EXAMINATIONS 2016

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.

WREB administers three Dental Hygiene Examinations: Dental Hygiene, Local Anesthesia and Restorative. Candidates may take one or more of the three examinations, depending on the requirements for licensure in each state. A brief overview of each Dental Hygiene Examination is provided below. Additional details are provided under Dental Hygiene Examinations Specifications on pp. 5 – 10 of this document and in the WREB Candidate Guides available at www.wreb.org.

Dental Hygiene Examination
The WREB Dental Hygiene examination is a performance-based clinical examination in which the Candidate is required to perform calculus removal and a series of periodontal assessments on a patient. The Candidate is evaluated on the following:

- Patient Selection
- Extraoral and intraoral examination
- Diagnostic quality of radiographs
- Calculus detection and removal
- Tissue management
- Acciracy of periodontal pocket measurement and recording
- Accuracy of gingival recession assessment and recording

Local Anesthesia Examination

Written Section. The Candidate is required to respond to a series of discipline-based and case-based selected-response questions.

Clinical Section. The Candidate is required to demonstrate injection technique by administering two block injections on a patient. The required injections are:

- Inferior Alveolar (IA) Nerve Block
- Posterior Superior Alveolar (PSA) Nerve Block

Restorative Examination
The Restorative Examination is a performance-based clinical examination requiring the Candidate to place, carve, and finish two restorative procedures on dentoform teeth. The procedures required are:

- Class II amalgam restoration
- Class II composite restoration

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 current Dental Hygiene practice analysis is in development.

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 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). 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 Candidate population and with 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 are reviewed by the Dental Hygiene Examination Review Board and the WREB Board of Directors, which are comprised of state licensing board representatives from WREB's active member states.

Dental Hygiene Examinations Specifications

Dental Hygiene Examination

The purpose of the WREB Dental Hygiene examination is to evaluate the ability of a Candidate to utilize professional judgment and clinical competency in providing oral health care to a Patient. The Candidate is required to perform calculus removal and complete a series of periodontal assessments on one quadrant of a qualifying patient's mouth (a minimum of twelve surfaces of readily demonstrable subgingival calculus must be present). The Candidate is evaluated on the following:

- Submission of a Patient that reflects accurate and effective interpretation of published qualification criteria
- Extraoral and Intraoral Examination on the Patient
- Performance on the entire assigned quadrant (including additional teeth where applicable) with evaluation of the treatment listed in Table 1

Table 1. Treatment Evaluation for Dental Hygiene Examination.

Domain	Proportion
Calculus detection and removal; twelve qualifying surfaces are assessed for remaining calculus	75%
Tissue management	
Periodontal pocket measurement and recording; eighteen qualifying surfaces are assessed for probing depths	25%
Gingival recession assessment and recording; six qualifying surfaces are assessed for recession	

Candidates submit a Patient and designate one quadrant (with one to four additional teeth optional) for review. Examiners evaluate the Patient submitted and select twelve surfaces of qualifying, subgingival calculus. If the submission does not have twelve qualifying surfaces, the submission is not accepted. Up to three submissions are allowed. Subsequent submissions may be the same quadrant with additional teeth, a different quadrant with or without additional teeth or a different patient.

Three independent, anonymous and calibrated Examiners evaluate the Patient after treatment. Points are deducted from the total possible points (i.e., 100) if an error is validated by two or more Examiners. A score of 75 or higher is required to pass. Validated errors or penalties that result in point deductions include:

- Non-qualifying submission penalties
- Calculus errors (i.e., subgingival and/or supragingival, detectable with explorer or air, burnished, spicule)
- Tissue trauma (i.e., any iatrogenic damage to extraoral or intraoral tissues, tissue tags, lacerations, ultrasonic burns, amputated papilla)
- Radiographic penalties (i.e., non-diagnostic)
- Probing depth errors (i.e., within a 1 mm margin of error)
- Gingival recession errors (i.e., within a 1 mm margin of error)
- Late penalties

Local Anesthesia Examination

The purpose of the WREB Local Anesthesia examination is to evaluate the ability of a Candidate to utilize professional judgment and knowledge to safely and competently administer a posterior superior nerve block injection and an inferior alveolar nerve block injection to a Patient. The WREB Local Anesthesia Examination consists of two sections: a computer-based selected-response examination (Local Anesthesia Written) and a performance-based clinical portion (Local Anesthesia Clinical). Candidates must pass the Local Anesthesia Written section prior to attempting the Local Anesthesia Clinical section. Successful completion of the Local Anesthesia Examination requires passing scores on both the Written section and the Clinical section within twelve months. Candidates who do not successfully pass both sections within that time must retake both the Written and Clinical sections.

Local Anesthesia Written Section. The Candidate is required to respond to a series of discipline-based and case-based selected-response questions that assess knowledge, application, and problem-solving regarding local anesthesia. The Local Anesthesia Candidate Guide provides recent academic references that Candidates may consult to review relevant content in local anesthesia and medical emergencies. Content domains and the proportions of the test dedicated to each domain are provided in Table 2.

Table 2. Local Anesthesia Written Section: Domains Assessed.

Domain	Proportion
Medical History: interpretation, prevention, recognition, management of possible complications, life support	30%
Pharmacology: anesthetic agents and vasoconstrictors, clinical actions and maximum recommended doses	30%
Delivery: methods of delivery of local anesthesia, armamentarium, injection type selection, administration technique	20%
Anatomy and Physiology: anatomical factors (head and neck), recognition and management of systemic complications	20%

The Local Anesthesia written section is computer-based and administered in a standardized manner through Prometric testing centers throughout the country. Candidates choose the time and location of the test administration.

Multiple forms of the Local Anesthesia written section test are administered to maintain examination security. Between 5% and 10% of questions on every test form are field-test questions that are not scored to allow empirical review of each question's performance prior to acceptance into the item bank. Field-test questions are embedded throughout each test form. Test forms are equated to ensure that Candidates of comparable proficiency will be equally likely to pass the examination. Local Anesthesia written section test scores are reported on a scale where 75 points is the passing score.

Local Anesthesia Clinical Section. The Candidate is required to demonstrate clinical competency by administering two nerve block injections on a patient who meets the acceptance criteria published in the Candidate guide regarding age, dentition and current health status. Both injections must be performed to examination specifications to pass.

The required injections are:

- Inferior Alveolar (IA) Nerve Block (The lingual and long buccal injections are not included in the IA section of the examination)
- Posterior Superior Alveolar (PSA) Nerve Block

The Candidate is not required to describe technique as each injection is performed; however, the Candidate is required to stop and inform the Examiners at four critical times during each injection until being instructed to proceed: 1) initial penetration, 2) optimal angle and depth, 3) aspiration and whether it is negative or positive, and 4) deposition. If a positive aspiration occurs, the Candidate must use professional judgment to determine how to handle a positive aspiration appropriately.

Two independent, calibrated Examiners observe the Candidate's technique and evaluate each injection based on the eight aspects listed in Table 3. After the completion of both injections, the Examiners independently record their grades. If the Examiners validate on one or more critical errors, or on three less-critical errors, the result is examination failure.

Table 3. Eight Aspects of Injection Evaluated on the Local Anesthesia Clinical Section.

Aspects	Criticality

(1) Proper Utilization of Medical History, Anesthetic and Syringe Selection
• Medical History
  o No contraindication(s) to local anesthetic
  o No health history contraindications
• Anesthetic Appropriate
  o No long-acting anesthetics and high concentration vasoconstrictor
• Syringe Type is Correct
	Critical

(2) Syringe Preparation and Handling
• Armamentarium
  o Errors: Appropriate protective eyewear not worn by Clinician or Patient; No hemostat or locking forceps present on tray; Expired local anesthetic
• Syringe Properly Prepared
  o Errors: Harpoon is not securely engaged; Bubbles are not expelled from the cartridge prior to injection; Expelled solution is more than the width of a stopper
• Syringe Handling
  o Error: Syringe in Patient's direct line of vision
	Less Critical

(3) Penetration Site • Needle Contamination o The needle is contaminated if it touches any surface, facial anatomy or intra oral object (gauze, glove) prior to needle penetration. Proceeding will result in failure of the injection. • Three Penetrations Allowed o Three (3) penetrations are allowed to reach the optimal angle and depth. If the third penetration attempt results in a first positive aspiration, a fourth penetration is permitted. • Optimal IA Penetration Site o The penetration site is at the area bordered medially by the pterygomandibular raphe, laterally by the internal oblique ridge, and at the height of the coronoid notch. o Errors: Too superior; Too inferior; Too medial; Too lateral • Optimal PSA Penetration Site o The penetration site is at the height of the vestibule in the mucobuccal fold posterior to the zygomatic process of the maxilla (visually approximates distal facial root of 2nd molar). o Errors: Too anterior; Too posterior; Not in mucobuccal fold	Critical
(4) Optimal Angle and Depth • IA o At the point when optimal depth is achieved, the barrel of the syringe is positioned over the premolars on the contralateral side and the needle is parallel to the occlusal plane of the mandibular teeth. Depth of insertion is 20-25 mm (approx. 2/3 the length of a long needle or 4/5 the length of a short needle). o Errors: Barrel too distal; Barrel too mesial; Angle too high; Angle too low; Too shallow; Too deep • PSA o The depth of insertion is approximately 16 mm (about 1/2 the length of a long needle or 3/4 the length of a short needle). o Errors: Needle not at 45° angle toward midline; Needle not at 45° angle to occlusal plane; Too shallow; Too deep	Critical
(5) Aspiration • Large window visible o Prior to aspirating, the large window must be toward the operator • Aspiration observed • Proper Handling of Positive Aspiration o Any sign of a positive aspiration must be observed and handled appropriately	Critical
(6) Amount and Rate • Deposition of Anesthetic Prior to Aspiration o Amount of local anesthetic deposited is less than $1/4^{th}$ (one-fourth) of the cartridge. • Rate of Administration is Acceptable o Acceptable rate approximately fifteen seconds to administer two stopper widths	Critical
(7) Tissue Management • The needle is visibly bent upon removal from tissue. • The degree of bowing would likely result in excessive submucosal soft tissue injury. • Visible laceration of tissue upon approach or withdrawal of needle. • The safety and well being of the Patient is compromised. See pp. 5-10 in Policy Guide (2016).	Critical
(8) Handling of Sharps • Proper Recapping Technique o A single-handed method is required when recapping the needle. Once the needle is protected within the cap, the needle must be secured. Needles and cartridges must be disposed of properly o Errors: Two handed recapping; Hand anywhere on safety shield during recapping; Holding needle cap during recapping • Proper Disposal of Sharps o Errors: Improper disposal of sharps; After the completion of both injections, sharps and cartridges are not disposed of in the appropriate container(s) and according to school policy. Improper handling of Sharps results in failure of both injections. Refer to the WREB website for Exam Site Information with site specific information regarding disposal of sharps and cartridges.	Critical

Restorative Examination

The purpose of the Restorative Examination is to evaluate a Candidate's ability to utilize professional judgment and competency in providing restorative procedures as allowed by state statute. Candidates for the Restorative Examination include qualifying dental hygiene graduates, school-certified senior dental hygiene students, and expanded-function dental assistants. The Restorative Examination is a performance-based clinical examination that requires a Candidate to place, carve, and finish two restorative procedures on dentoform teeth. The procedures required are:

- Direct Posterior Class II Amalgam Restoration
- Direct Posterior Class II Composite Restoration

Candidates receive two simulated teeth prepared for restoration. The two teeth will include one maxillary and one mandibular, and one distal-occlusal and one mesio-occlusal preparation. Several combinations of teeth are possible. The teeth assigned to each examination group are announced at the onsite examination orientation and received as each group begins their treatment. Restorative material is randomly assigned via computer; one is required to be restored with amalgam and the other with composite.

The Candidates must follow universal precautions and work with the typodont positioned on the rod post to mimic a natural treatment position of a patient. Candidate performance on each procedure is graded by three independent, anonymous and calibrated Examiners and weighted according to the criteria in Table 4.

Table 4. Restorative Section Grading Criteria Weighting.

Grading Criterion	Weight
Occlusal	30%
Margins	35%
Proximal	35%

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. Figure 1 provides grading criteria definitions, as published in the 2016 Restorative Examination Candidate Guide.

RESTORATIVE GRADING CRITERIA

	5	4	3	2	1	
OCCLUSAL (30%)	Replicates proper anatomy (restoring harmonious form). Ridges and marginal ridge present and properly formed. Smooth surface, no pits or voids. Functional occlusion.	Slight variation of harmonious form. Major grooves formed and positioned correctly. Supplemental grooves may or may not be present. Fossae present. Ridges and marginal ridge present, but slight variation. Slight surface irregularities (pitting or voids). Functional occlusion.	Moderate variation of harmonious form. Anatomy adequate. Marginal ridge height has functional contour. Anatomy or marginal ridge can be corrected with minimal polishing and finishing. Moderate surface irregularities (pitting or voids). Functional occlusion.	Anatomy inadequately carved and/or over/ under carved. Marginal ridge contour/height improper. Pits and fossae improperly placed (would alter occlusion or cause food impaction). Any ridges, oblique and/or transverse, improperly placed. Critical surface irregularities or defects (pitting or voids). Placement of glazing/unfilled resin over finished restoration. Hyperocclusion; contact marks appear only on restoration.	Incorrect anatomy (would significantly alter function). Incorrect marginal ridge contour/height >1 mm (would significantly impact function). Pits and fossae incorrectly placed. Any ridges, oblique and/or transverse not present. Improper manipulation or trituration of material. Severe surface irregularities or defects (pitting, voids, and/or fractures). Uncured resin. Hyperocclusion; contact marks appear only on restoration.	OCCLUSAL (30%)
MARGINS (35%)	Minimal variation of cavosurface margin (+). Slight scarring of tooth structure (one area).	Slight variation of cavosurface margin (+ or -). Slight scarring of tooth structure (multiple areas).	Moderate variation of cavosurface margin + or - (can be corrected with minimal polishing and finishing). Moderate scarring of tooth structure. Alters tooth structure without compromising normal tooth function or restoration.	Critical variation of cavosurface margin, open >.5 mm; integrity compromised (not correctable). 1 mm excess - would compromise periodontal health. Critical damage of tooth structure affecting normal function or restoration. Placement of glazing/unfilled resin over finished restoration.	Severe variation of cavosurface margin open >1 mm (not correctable). Greater than 1 mm excess - would compromise periodontal health. Severe damage of tooth structure affecting normal function or restoration (multiple areas).	MARGINS (35%)
PROXIMAL (35%)	Replicates proper anatomy (restoring harmonious form). Optimal contact - will allow waxed floss to pass through contact with proper resistance. Smooth surface, no pits or voids.	Slight variation of proximal contour, shape, and/or position of contact area. Nearly optimal contact - will allow waxed floss to pass through contact with near proper resistance. Slight surface irregularities (pitting or voids).	Moderate variation of proximal contour, shape, and/or position of contact area. Gingival-occlusal embrasures not defined, but functional. Barely adequate contact, will allow waxed floss to pass through contact with slight resistance or moderately tight resistance (may shred waxed floss). Moderate surface irregularities (pitting or voids).	Critical variation of proximal contour, shape, and/or position of contact area. Improper contact - tight (may break waxed floss). Improper contact (+/slightly open with light resistance). Critical surface irregularities (pitting or voids). Placement of glazing/unfilled resin over finished restoration.	Severe variation of proximal contour, shape, and/or position of contact area. Incorrect contact-cannot get floss through contact. Open contact. (+/slightly open with no resistance). Improper manipulation or trituration of material. Severe surface irregularities or defects (pitting, voids, and/or fractures). Uncured resin.	PROXIMAL (35%)

Figure 1. 2016 Restorative Examination Grading Criteria

Examiners also evaluate the hard and soft tissue surrounding the preparations and adjacent teeth for damage. If Examiners validate on the detection of damage a penalty is incurred. The following types of tissue damage, if validated, will result in point deductions:

- Damage to soft tissue is trauma in excess of 3 mm
- Damage to hard tissue is trauma in excess of 1 mm on the assigned preparations or adjacent teeth

Point deductions due to late penalties may also apply if a Candidate continues to work after the announcement to stop has been made.

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 and safety are reviewed briefly here, and include examination timing, accommodations, patient safety and comfort, infection control, site assignments of Examiners, and examination security. Additional details of examination administration are available in the WREB 2016 Dental Hygiene Examination Candidate Guide (WREB, 2016a), the WREB 2016 Local Anesthesia Examination Candidate Guide (WREB, 2016b), the WREB 2016 Restorative Eamination Candidate Guide (WREB, 2016c), on the WREB website (http://www.wreb.org), and in the WREB 2016 Policy Guide (WREB, 2016d).

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; 2016a, 2016b & 2016c). 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).

Dental Hygiene. The Dental Hygiene examination is administered in pre-assigned morning or afternoon groups. Candidates are randomly assigned to an examination day and group upon registration, but may request a specific day or time, e.g., to accommodate Patient scheduling. Once a Candidate's Patient submission has been accepted, they have up to two hours to complete treatment.

Local Anesthesia. The computer-based Local Anesthesia Written examination is administered by Pearson VUE at testing centers around the country. Candidates are allowed one hour to complete the examination. No strict time-limit is enforced on the Local Anesthesia Clinical examination. Candidates are scheduled at times that provide approximately 20 minutes to complete the required injections, but the time to complete the examination may be shorter or longer, depending on the features of each injection, e.g., if a positive aspiration leads to cartridge replacement or a Candidate proceeds with a different needle following contamination.

Restorative. The Restorative examination is adminstered in pre-assigned morning or afternoon groups. Candidates are randomly assigned to an examination day and group upon registration, but may request a specific day or time. Once each group enters the clinic, Candidates have one and one-half hours to complete the two procedures.

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.

Patient Safety and Comfort

Guidelines and requirements regarding Patient safety and comfort are addressed throughout WREB Candidate Guides and reinforced throughout each examination or examination section that involves patient treatment. For each Patient, a Candidate must complete and submit a Patient Consent form, a Patient Medical History form, and a Follow-up Care Agreement, all signed by the Patient. The WREB Candidate Guide lists medical conditions and other factors to consider when selecting a Patient to participate in the examination and describes expectations for Candidates regarding Patient care and comfort during the examination, such as nourishment, breaks, and administration of appropriate local anesthesia as needed. In the clinic, every Patient is greeted and assessed by a Chief Examiner (Dental Hygiene, Local Anesthesia), who reviews the Patient's Medical History, radiographs and current vital signs.

WREB Candidate Guides, Examiner training materials, and staff training emphasize Patient safety, including the review of infection control guidelines and current recommendations of the American Heart Association. Members of each Examiner team also meet with a representative(s) of the exam site's host school prior to the examination to review medical emergency protocols. WREB includes an automated external defibrillator (AED) with every set of materials sent to an examination that will be available in addition to host-site equipment. The Candidate Guide also describes situations where the health of the Patient may require additional treatment or follow-up care. Examples, procedures, and details are provided for Postoperative Care, Instructions to Candidate, and the Referral Needed form.

All Examiners and WREB staff are trained to treat Patients with care and respect. Instruction includes the review of WREB's zero-tolerance policy regarding actions or conduct that could be viewed as sexual harassment and sexual misconduct (Title VII of the Civil Rights Act of 1964). Conduct of this nature will result in immediate dismissal from an exam or removal for cause. The Candidate Guide also describes examples of improper performance that may result in a Candidate's dismissal from an exam, including disregard for Patient welfare and/or comfort and failure to recognize or respond to conditions which may jeopardize Patient health.

All Patients receive a Patient Information and Questionnaire handout which provides information regarding the purpose of the examination and a description of what the Patient can expect while at the examination. The handout includes a brief survey where Patients can evaluate how they were treated by the Candidate, by Examiners, and WREB staff, as well as provide any comments or suggestions. Patient responses are reviewed after every examination.

Infection Control

WREB Candidate Guides, Examiner training materials, and staff training emphasize adherence to published clinical treatment guidelines and standards for infection control procedures. Procedures regarding proper infection control protocol, compliance with OSHA guidelines for proper clinic attire, protection from contaminated instruments and proper disposal of biohazardous and pharmaceutical materials and sharps, are addressed. Candidates, Examiners, WREB staff and observers are required to adhere to examination site host-school policies and procedures as well as Centers for Disease Control

and Prevention (CDC) guidelines. For examinations that involve treatment within a simulated treatment setting (i.e., no patient), universal precautions and infection control procedures still apply. Failure to maintain acceptable standards of infection control and mercury hygiene may result in examination failure or dismissal.

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 or dental hygiene educator may not examine at the school in which 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 and Chief Examiner. The Chief Examiner ensures that the examination proceeds in accordance with established WREB policies and oversees the Examiner Orientation and Calibration Session. Grading Examiners for the Dental Hygiene and Restorative examinations never have contact with Candidates to guarantee anonymity in scoring. The only Examiners who have contact with the Candidates at Dental Hygiene and Restorative examinations are Chief Examiners, who do not function in a grading capacity. Chief 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. 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 Chief Examiner. The clinical section of the Local Anesthesia examination is the only WREB clinic-based examination with limited anonymity; two independent Examiners observe directly the Candidate's injection technique as performed on a patient. If necessary, Examiners can intervene immediately and stop any procedure that could pose a health or safety risk to the Patient. Procedures are followed to ensure as much anonymity as possible including the assignment of Candidate identifiers that do not reflect Candidate name, school or region, assigning different identifiers for Candidates that may participate in more than one examination at that site and ensuring that Examiners do not participate at host-school sites where they have a history of affiliation.

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 items. WREB continually develops and field-tests new testing items 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 and unusual situations; and issues regarding examination failure.

Decision-making Approach

Information from multiple assessments, such as WREB's Local Anesthesia Written and Clinical examinations, 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 section scores for one final overall score; higher performance on one element may "compensate" for lower performance on another. Candidates must meet the passing score for each examination section, set by examination committees within the conjunctive framework, to pass the examination. The Local Anesthesia examination has the additional requirement of success on the Local Anesthesia Written examination prior to challenging the Clinical examination.

Methods of Score Determination

The pass or fail decision regarding Candidate performance on most WREB examinations is based on the final score. Final scores for the Dental Hygiene examination are calculated by applying point-deductions from the total points possible for any Examiner-validated errors or penalties. Final scores for the Local Anesthesia Written examination are calculated by re-scaling the sum of correct responses to a percentage-like scale of 0 to 100. Final scores for the Restorative examination are calculated by summing the weighted median ratings or “grades” assigned by the grading Examiners on each scoring criterion and then averaging the scores of the two preparations treated. The Local Anesthesia Clinical examination does not generate a final score; each injection is graded as passing or failing. Both injections must be passing to pass the Local Anesthesia Clinical examination.

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. 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, 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 WREB examinations that assess mulitple levels of performance per grading criterion, the examination committees define each level of performance with respect to critical aspects of clinical practice. The level of performance that reflects minimal competency on the Restorative examination (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 Dental Hygiene examination or judging passing or failing of an injection on the Local Anesthesia Clinical section), or a judgment aligned with multiple levels of performance quality (e.g., rating scales of 5 points for the Restorative examination). For example, on the Restorative examination, 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. Grading criteria definitions for the Restorative examination are available in the WREB Candidate Guide (WREB, 2016c).

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. Examination judgments for the Local Anesthesia Clinical examination are made by two independent Examiners; both must validate on the observation of the same critical error to have an impact on an injection pass/fail outcome.

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. All 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 licensing boards. Approximately fifteen percent of Examiners are 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. Dental Hygiene and Local Anesthesia Clinical Candidates receive their provisional results onsite, after completing the examination. All Candidates are notified via electronic mail when they are able to access their official results at their secure WREB login online. Restorative and Local Anesthesia Written Candidates receive their results within about one week of the examination.

WREB results focus on the Pass/Fail decision and are intended 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 and Unusual Situations

Some errors, as defined in the Candidate Guides (WREB; 2016a, 2016b & 2016c) may result in point deductions on each of the examination sections. Many penalties are set to reflect aspects of performance that are directly related the content being assessed and have been performed inadequately or reflect unsafe or harmful behavior, e.g., tissue trauma penalties. Other penalties, such as late penalties, are set to discourage inappropriate behaviors, and 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 certain 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.

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 details, definitions, possible point deductions, and examples of improper performance and unethical conduct can be found in the WREB Candidate Guides (WREB; 2016a, 2016b & 2016c).

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 retake the examination and/or appeal an examination result.

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; 2016a, 2016b & 2016c). 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 hygiene 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,

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

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 Hygiene Restorative examination data is a four-facet model (i.e., Candidate, Examiner, Restorative Material and Restorative Grading Criterion) that can be expressed as follows:

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

where P_{mnijk} is equal to the probability of Candidate m being rated k on Grading Criterion j within Restorative Material i, by Examiner n. $P_{mnijk} - 1$ is equal to the probability of Candidate m being rated $k - 1$ on Grading Criterion j within Restorative Material i by Examiner n. C_m is the ability of Candidate m, E_n is the severity of Examiner n, R_i is the difficulty of Restorative Material 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 5.

Table 5. 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 HYGIENE EXAMINATION TECHNICAL ANALYSES 2016

Analyses of graded elements, penalties, grading criteria, comparability of forms and overall test functioning are summarized in the first part of this section for the Dental Hygiene, Local Anesthesia and Restorative Examinations. Analyses of Examiner performance for the three examinations follows. Finally, passing percentages for the three 2016 examinations and combined for the past seven years are presented. 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.

Dental Hygiene Examination

Graded Elements and Penalties – Dental Hygiene. Table 6 provides basic descriptive statistics for graded elements of the Dental Hygiene examination, based on the sum of raw means of medians computed from the three sets of Examiner no-error grades per tooth surface evaluated. Out of 2,035 examination attempts in 2016, 2,007 attempts were graded on calculus removal, probing depths, and gingival recession recording. Twenty-eight attempts were unsuccessful due to multiple unacceptable Patient submissions and were not graded.

Table 6. Dental Hygiene Examination 2016 Graded Elements: Descriptive Statistics, $N = 2{,}007$.

Statistic	**Calculus Removal** 12 Surfaces Possible Range: 0 to 12	**Probing Depths** 18 Surfaces Possible Range: 0 to 18	**Gingival Recession** 6 Surfaces Possible Range: 0 to 6
Mean	10.92	17.79	5.67
Standard Deviation	1.59	0.69	0.99
Minimum; Maximum	1; 12	8; 18	0; 6

Table 7 provides summary results from a many-faceted Rasch model analysis for graded elements in logit, i.e., log-odds, values. The ranges of logits show distinct differences in degree of challenge among the skills assessed. Calculus removal is consistently more challenging for Candidates across all graded surfaces (i.e., higher positive logit parameter estimates), Probing Depths is consistently less challenging (i.e., negative logit values) and the recording of Gingival Recession measurements falls between the other graded elements. Little or no variation in standard error values exists across surfaces within each and is not unexpected due to surfaces not being identified and broken out by tooth number or surface location within this analysis. Additional analyses that assess the differential level of challenge by all teeth (i.e., 1 through 32) and surface location (i.e., distal, mesial, etc.) are conducted and reviewed with the Dental Hygiene examination committee. Point-biserial values are low, particularly for Probing Depths, given the limited degree of variation within Candidate performance. Almost 30% of all examination attempts receive perfect scores, with 87% receiving no validated errors on the eighteen Probing Depths measured. All mean-square fit statistics and discrimination parameter estimates are within suggested ranges.

Table 7. Dental Hygiene Examination 2016 Graded Elements: Many-Faceted Rasch Model Analysis Indicators in Logits, $N = 2{,}007$.

Indicator	**Calculus Removal** 12 Surfaces	**Probing Depths** 18 Surfaces	**Gingival Recession** 6 Surfaces
Logit (Range)	0.83 – 1.21	-1.73 – -0.21	0.30 – 0.47
Logit Mean[a]	0.97	-0.77	0.37
Logit Standard Deviation	0.12	0.44	0.07
Standard Error Mean	0.04	0.09	0.05
Standard Error (Range)	0.04 – 0.04	0.07 – 0.13	0.05 – 0.05
Many-Facet Point-Biserial Correlation[b] (Range)	0.17 – 0.22	0.06 – 0.11	0.18 – 0.20
2pl Discrimination Estimate[c] (Range)	0.99 – 1.01	0.97 – 1.00	1.01 – 1.03
Infit Mean-Square (Range)	0.98 – 1.01	1.00 – 1.03	0.97 – 0.99
Outfit Mean-Square (Range)	0.90 – 1.04	0.90 – 1.38	0.80 – 1.01

[a] Mean of all three categories of graded elements 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)

Table 8 lists the number of penalties applied and the number attempts receiving no penalties for Dental Hygiene examination attempts in 2016. Most Candidate examination attempts (i.e., 88% to 97%) do not incur penalties. Most penalties are applied once with multiple penalties being less frequent as the number of penalties increases.

Table 8. Dental Hygiene Examination 2016: Penalties Applied, Total attempts $N = 2{,}035$.

Penalty	**One Penalty**	**Multiple Penalties**	**Count (and Percentage) of Attempts with No Penalties**
Patient Submission Rejection	168	82	1,785 (87.7%)
Radiographic Penalties	92	81	1,862 (91.5%)
Late Penalties	196 (1 to 5 minutes)	32 (More than 5 minutes)	1,807 (88.8%)
Tissue Trauma	44	20	1,971 (96.9%)

Overall Test Functioning – Dental Hygiene. Table 9 provides summary statistics, the standard error of measurement (SEM), conditional standard error of measurement at the passing cut (CSEM) and the Rasch person-separation reliability estimate, for the 2016 Dental Hygiene examination. Person-

separation reliability is comparable to Cronbach's alpha coefficient of internal consistency reliability, but excludes zero and perfect scores. The reliability estimate is an underestimate, due to a negatively skewed distribution of scores. Reliability is often attenuated in criterion referenced credentialing assessment because of the high level of Candidate preparedness. Most Dental Hygiene Candidates perform very well or obtain perfect scores, although a small percentage continues to be unsuccessful upon retake. Trends in passing percentages over time and the degree of Examiner agreement are other sources of validity evidence characterizing the reliability and quality of the Dental Hygiene examination. Details of passing percentages are provided on pp. 34-35 and details regarding Examiner performance are provided on pp. 31-33.

Table 9. Overall Test Summary Statistics for Dental Hygiene Examination, 2016.

Indicator	**Dental Hygiene** (Maximum Possible Score 100)
N Attempts	2,035
Final Score Mean	91.00
Final Score Standard Deviation	11.06
Minimum; Maximum	13.75; 100.00
Standard Error of Measurement (*SEM*)	6.54
Conditional *SEM* at Passing Score	3.57
Candidate Ability Estimate Logit Mean	3.42 logits
Candidate Ability Estimate Logit *SD*	1.11 logits
Logit Minimum; Maximum	0.21; 5.44[a] logits
Person Separation Reliability Estimate[b]	0.65

[a]Maximum logit value for non-perfect scores
[b]Comparable to alpha coefficient internal consistency reliability estimate (Cronbach, 1951) with zero and perfect scores excluded

Local Anesthesia Examination

Local Anesthesia Examination - Written Section

Content Areas by Form - Local Anesthesia Written Section. Local Anesthesia Written examination forms are developed to be equivalent in content, level of challenge and length of time needed to complete the test. Table 10 provides the mean number correct and standard deviation, by test form, for the four content areas assessed on the Local Anesthesia Written examination. Final scores are based on all items; however, Candidates who are not successful receive details broken out by content area, with a caution to consider all content areas in their preparation for retake since a moderately high degree of relationship exists among the four content areas (i.e., intercorrelations from 0.25 to 0.41).

Table 10. Local Anesthesia Written 2016 Forms: Mean Number Correct and Standard Deviation by Content Area.

	Mean Number Correct (*SD*)			
	Medical History (15 Items)	**Pharmacology** (15 Items)	**Delivery** (10 Items)	**Anatomy and Physiology** (10 Items)
Form A	12.81 (*1.60*)	12.61 (*1.70*)	8.29 (*1.25*)	8.52 (*1.20*)
Form B	13.04 (*1.54*)	12.62 (*1.79*)	8.19 (*1.35*)	8.50 (*1.32*)
Form C	13.06 (*1.54*)	12.61 (*1.76*)	8.20 (*1.45*)	8.36 (*1.35*)

Overall Test Functioning – Local Anesthesia Written Section. Summary statistics, SEMs, CSEMs, indicators of reliability and passing percentages by test form are presented in Table 11. CSEM values are reported for the passing cut-score for each form. Estimated values of Cronbach's alpha coefficient of internal-consistency reliability (Cronbach, 1951) are shown for each form. Alpha reliability estimates depend upon sample variability and may be attenuated due to the high level of Candidate preparedness in criterion-referenced credentialing assessment. Many Candidates perform very well on several test items. While eliminating these items can increase the alpha estimate, they are included because subject matter experts have determined that the information assessed is essential to minimal competence. Similarly, adding additional items, especially more challenging items, can increase the estimate of alpha, but are not included since the purpose of the examination is to assess minimum competence, not to discriminate among Candidates with very high levels of knowledge and ability. Other indicators, such as Peng-Subkoviak P_0 estimates of classification consistency (Peng & Subkoviak, 1980) and the Brennan-Kane $\Phi(\lambda)$ index of dependability (Brennan & Kane, 1977), provide insight into the reliability of pass-fail outcomes. Estimates of alpha are moderately high, with a range of 0.64 to 0.67. Dependability index values, which take item variance into account, are relatively high, with a range of 0.84 to 0.85, while classification consistency values are even higher, with a range of 0.88 to 0.89, since mean scores are far above the passing cut-score, making misclassification less likely. Passing percentages by form range from 84.8% to 88.2%. A chi-square analysis was conducted to assess pass/fail outcome by form. No significant difference in pass/fail outcome was found among forms (χ^2 (2, N=1,317) = 2.26, p = 0.32).

Table 11. Local Anesthesia Written 2016 Forms: Indicators of Overall Test Functioning by Form.

2016 Test Form	***N***	**Scale Score Mean (*SD*)**	**CSEM**	**α Reliability Estimate**	**Φ(λ) Index of Dependability**	***P₀* Classification Consistency**	**Passing Percentage** (All Attempts)
Form A	434	84.5 (*8.1*)	2.92	0.64	0.84	0.89	87.1%
Form B	441	84.7 (*8.4*)	2.90	0.67	0.85	0.89	88.2%
Form C	442	84.5 (*8.4*)	2.92	0.67	0.84	0.88	84.8%

Local Anesthesia Examination - Clinical Section

Injection Type: Local Anesthesia Clinical Section. Candidates are slightly more likely to be successful on the IA injection (87.3%) than the PSA injection (82.3%). Just over three-quarters (75.8%) of full examination first attempts are successful on both injections, with 17.9% failing one injection upon first attempt and 6.3% failing both. Candidates may retake the failed injection(s) onsite.

Validated Errors: Local Anesthesia Clinical Section. A validated critical error results in failure of the Local Anesthesia Clinical examination. Most unsuccessful examination attempts incur only one validated error. A very small number of unsuccessful attempts (i.e., nine) incurred four or more critical errors, out of 1,403 total attempts in 2016. The number of validated critical errors per attempt in 2016 are provided in Table 12. A very small number of successful attempts incurred one or more validated *less*-critical errors with six incurring one and six incurring two validated less-critical errors, out of 1,403 total attempts in 2016.

Table 12. Local Anesthesia Clinical 2016: Validated Critical Errors per Attempt.

Number of Validated Errors Incurred	Number (Percentage) of Attempts
0	1,063 (75.8%)
1	235 (16.7%)
2	74 (5.3%)
3	22 (1.6%)
4 or more	9 (0.6%)

Table 13 provides the number and percentage of validated critical and less-critical errors within each of the eight injection aspects described in the Candidate Guide and on pp. 7 – 8 of this document. The aspects of injection with the most validated critical errors are Penetration Site and Optimal Angle and Depth. The errors validated most frequently within Penetration Site include Needle Contamination and Too Lateral and Too Superior, associated with optimal IA injection penetration. The errors validated most frequently within Optimal Angle and Depth include Too Shallow (associated with both IA and PSA) and Needle not at 45° angle toward midline or occlusal plane (associated with PSA).

Table 13. Local Anesthesia Clinical 2016: Number of Validated Errors per Injection Aspect.

Aspects of Injection	Number (Percentage) of Validated Errors
(1) Proper Utilization of Medical History, Anesthetic and Syringe Selection	11 (0.4%)
(2) Syringe Preparation and Handling*	27 (0.9%)
(3) Penetration Site	188 (9.2%)
(4) Optimal Angle and Depth	140 (8.0%)
(5) Aspiration	89 (3.0%)
(6) Amount and Rate	12 (0.4%)
(7) Excessive Trauma	13 (0.4%)
(8) Handling of Sharps	45 (1.7%)

*Less-critical errors: three must be validated to result in failure.

Table 14 provides summary results from a many-faceted Rasch model analysis for critical and less-critical errors in logit values. Possible critical errors from the first injection aspect category, Proper Utilization of Medical History, Anesthetic and Syringe Selection, are not included since the Candidate does not proceed with the injection for any validated error in this category (note that validation on certain medical conditions, e.g., high blood pressure, does not result in failure, as per the Candidate Guide, pp.6 – 7). The logit means show distinct differences in degree of challenge among injection aspects, with Penetration Site and Optimal Angle and Depth having the highest positive logit parameter estimates, indicating higher degree of challenge. Standard error values are relatively consistent across possible errors and injection aspects, with the exception of a few rarely validated errors (e.g., there was only one validated instance of Barrel Too Distal within Optimal Angle and Depth). Point-biserial values are very low, given the very limited degree of variation within Candidate performance. Almost 77% of all examination attempts were successful, with no validated critical errors. Mean-square fit statistics and discrimination parameter estimates are within suggested ranges, with the exception of underutilized critical errors, which yielded slightly higher outfit mean square values, but no values exceeded 1.50 logits.

Table 14. Local Anesthesia Clinical, 2016 Critical Errors: Many-Faceted Rasch Model Analysis Indicators in Logits, $N = 1{,}397$.

Indicator	**(2) Syringe Prep…** 3 Possible Errors	**(3) Penetration Site** 9 Possible Errors	**(4) Optimal Angle…** 8 Possible Errors	**(5) Aspiration** 3 Possible Errors	**(6) Amount and Rate** 2 Possible Errors	**(7) Excess. Trauma** 1 Possible Error	**(8) Handling of Sharps** 2 Possible Errors
Logit (Range)	-0.66 – -0.12	-0.86 – 1.34	-1.88 – 2.07	-0.80 – 0.53	-1.12 – -0.95	–	-1.21 – 0.19
Logit Mean[a]	-0.47	0.34	0.31	-0.11	-1.04	-0.39	-0.51
Logit Standard Deviation	0.30	0.78	1.43	0.67	0.12	–	0.99
Standard Error Mean	0.16	0.16	0.19	0.14	0.21	0.15	0.27
Standard Error (Range)	0.14 – 0.18	0.08 – 0.30	0.07 – 0.50	0.10 – 0.19	0.20 – 0.22	–	0.18 – 0.36
MF Pt-Bis. Corr[b] (Range)	0.05 – 0.06	0.00 – 0.06	0.00 – 0.28	0.05 – 0.07	0.07 – 0.07	0.07	0.00 – 0.00
2pl Discrimination Estimate[c] (Range)	0.97 – 1.02	0.97 – 1.04	0.98 – 1.04	0.99 – 1.02	1.01 – 1.02	1.01	0.99 – 1.04
Infit Mean-Square (Range)	0.99 – 1.01	0.95 – 1.03	0.97 – 1.02	0.99 – 1.01	0.99 – 1.00	0.99	0.97 – 1.01
Outfit Mean-Square (Range)	0.77 – 1.50	0.81 – 1.44	0.75 – 1.31	0.71 – 0.96	0.63 – 0.80	0.85	0.36 – 1.19

[a] Mean of all categories of graded elements constrained at 0 for criterion parameter estimation

[b] Many-Facet Point-biserial 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 – Local Anesthesia Clinical Section. Conventional estimates of error such as the SEM and CSEM are not applicable to the pass or fail judgments of the Local Anesthesia Clinical examination, as there is no summated final score. Converting findings of no-error into "points" yields a mean final percentage of 99.3% for 2016, since each injection is evaluated for errors on many critical features (26 errors are possible on the IA injection and 23 errors are possible on the PSA) and so many Candidates are well-prepared to demonstrate competence that few errors are committed. Most Local Anesthesia examination attempts are passing (76.6% in 2016), which are "perfect" scores, with very few exceptions for validated less-critical errors, resulting in 75.8% perfect scores, i.e., no errors assessed. Trends in passing percentages over time can demonstrate the effectiveness of the Local Anesthesia Clinical examination to identify Candidates that do not perform at a minimally competent level of practice. Unsuccessful Candidates that have failed both injections are significantly less likely to pass upon retake and a very small percentage of Candidates continue to be unsuccessful upon multiple retakes and remediation. Details regarding passing percentages are provided on pp. 34 – 35. The degree of Examiner agreement is another source of validity evidence that characterizes the quality of the Local Anesthesia Clinical examination. Examiner agreement is assessed using the many-faceted Rasch model analysis, which treats the total sum of no errors across the injections as a kind of score, but also accounts for the judging behavior of the Examiners, yielding results that can discriminate among Examiner performance. Details regarding Examiner performance are provided on pp. 31 – 33.

Restorative Examination

Graded Elements and Penalties – Restorative Examination. Table 15 provides means and standard deviations for each Restorative examination grading criterion and overall, based on the raw means of medians computed from the three sets of Examiner grades by criterion, arch and material. The Occlusal criterion is more challenging than Margins and Proximal, the Maxillary arch is more challenging than Mandibular and Composite is more challenging than Amalgam. Differences are consistent across criteria, arches and material.

Table 15. Restorative Examination 2016 Graded Elements: Descriptive Statistics, $N = 466$ (932 Treated Preparations) by Criterion, Arch and Material.

Criterion	Overall Mean (*SD*)	Arch		Material	
		Maxillary Mean (*SD*)	**Mandibular** Mean (*SD*)	**Amalgam** Mean (*SD*)	**Composite** Mean (*SD*)
Occlusal (30%) Grading Scale: 1 to 5	3.02 (*0.54*)	2.95 (*0.68*)	3.09 (*0.65*)	3.06 (*0.70*)	2.98 (*0.64*)
Margins (35%) Grading Scale: 1 to 5	3.33 (*0.51*)	3.28 (*0.67*)	3.39 (*0.64*)	3.48 (*0.62*)	3.19 (*0.66*)
Proximal (35%) Grading Scale: 1 to 5	3.36 (*0.47*)	3.28 (*0.61*)	3.43 (*0.58*)	3.40 (*0.61*)	3.32 (*0.58*)
Overall	3.25 (*0.41*)	3.18 (*0.51*)	3.31 (*0.48*)	3.32 (*0.50*)	3.17 (*0.49*)

Table 16 provides summary results from many-faceted Rasch model analyses for graded criteria in logit 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 16. Restorative Examination 2016 Graded Elements: Many-Faceted Rasch Model Analysis Indicators in Logits, $N = 466$.

Indicator	Occlusal		Margins		Proximal	
	Maxillary	Mandibular	Maxillary	Mandibular	Maxillary	Mandibular
Logit[a]	0.20	0.14	-0.05	0.10	-0.07	-0.32
Standard Error	0.04	0.04	0.04	0.04	0.05	0.05
Many-Facet Point-Biserial Correlation[b]	0.26	0.26	0.24	0.26	0.24	0.23
2pl Discrimination Estimate[c]	1.04	0.98	1.02	1.00	1.01	0.96
Infit Mean-Square	0.96	1.02	0.99	1.01	0.99	1.05
Outfit Mean-Square	0.95	1.01	0.99	1.02	0.98	1.05

[a] Mean of all graded elements 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)

Eleven of 466 examination attempts (2.4%) had validated tissue damage penalties applied in 2016. No wrong material or late penalties were applied.

Overall Test Functioning – Restorative. Table 17 provides summary statistics, the standard error of measurement (SEM), conditional standard error of measurement at the passing cut (CSEM) and the Rasch person-separation reliability estimate, for the 2016 Restorative examination. The person-separation reliability estimate of 0.87 is relatively high for a performance-based assessment and is equivalent to Cronbach's alpha coefficient internal reliability consistency estimate, since there were no perfect or zero scores on the Restorative examination in 2016.

Table 17. Overall Test Summary Statistics for Restorative Examination 2016.

Indicator	**Restorative** (Score Range 1 - 5)
N Attempts	466
Final Score Mean	3.25
Final Score Standard Deviation	0.41
Minimum; Maximum	1.68; 4.50
Standard Error of Measurement (*SEM*)	0.148
Conditional *SEM* at Passing Score	0.075
Candidate Ability Estimate Logit Mean	0.51 logits
Candidate Ability Estimate Logit *SD*	1.09 logits
Logit Minimum; Maximum	-3.27; 3.32[a] logits
Person Separation Reliability Estimate[b]	0.87

[a]Maximum logit value for non-perfect scores
[b]Comparable to alpha coefficient internal consistency reliability estimate (Cronbach, 1951) with zero and perfect scores excluded

Examiner Performance

Examiner Agreement. Evaluating Examiner performance is critical to assessing examination quality for performance-based assessments, since outcomes are based on Examiner judgments. Examiner performance is also important in the collection of validity evidence for criterion-referenced tests in which most Candidates are well-prepared. On examinations where many Candidates perform at the highest possible level, such as the Dental Hygiene and Local Anesthesia Clinical examinations, other indicators may under-estimate assessment quality, given the limited degree of variation in Candidate performance.

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, for all three clinic-based Dental Hygiene examinations, are provided in Table 18. Averages are weighted by the number of grades assigned by each examiner, as the number of examinations in which a grading Examiner participates may vary. Examiners for all Dental Hygiene examinations had percentages of agreement well above 80% in 2016, with most over 90%. Note that the percentage of non-validated grades assigned is reported for the Local Anesthesia Clinical examination, where two Examiners evaluate Candidate performance. Examiner Harshness or Lenience cannot be determined for an individual non-

validated grading instance with only two Examiners; however, the many-faceted Rasch analysis, reported in the next section, provides additional insight into Examiner performance for the Local Anesthesia examination.

Table 18. Examiner Percentages of Agreement, Harshness, and Lenience: Dental Hygiene, Local Anesthesia Clinical and Restorative Examinations, 2016.

Indicator	**Dental Hygiene** (N_E = 74)	**Local Anesthesia Clinical** (N_E = 28)	**Restorative** (N_E = 19)
Agreement Percentage[a] Weighted Average	96.1%	99.0%	92.6%
Agreement Percentage (Range)	91.1 – 98.0%	97.9 – 99.6%	88.0 – 97.0%
Harshness Percentage Weighted Average	2.9%	1.0%[b]	3.6%
Harshness Percentage (Range)	0.8 – 6.4%	0.4 – 2.1% [b]	1.0 – 10.2%
Lenience Percentage Weighted Average	1.0%	1.0%[b]	3.8%
Lenience Percentage (Range)	0.2 – 4.4%	0.4 – 2.1% [b]	0.3 – 8.3%

[a]Agreement is exact for the Dental Hygiene and Local Anesthesia Clinical examinations; agreement is exact and adjacent agreement for Restorative, which employs multiple-level ratings
[b]Percentage non-validated is reported for Local Anesthesia Clinical examination

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 19 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, especially if they demonstrate extreme performance in conventional agreement statistics, as well. 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, e.g., a set 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 19. Many-Faceted Rasch Model Examiner Severity Analysis Indicators in Logits: Dental Hygiene Examinations, 2016.

Indicator	**Dental Hygiene** (N_E = 74)	**Local Anesthesia Clinical** (N_E = 28)	**Restorative** (N_E = 19)
Severity Measure Logit (Range)	-0.90 – 0.79	-0.75 – 0.71	-1.05 – 0.96
Standard Error (Range)	0.05 – 0.29	0.09 – 0.36	0.05 – 0.19
Severity Measure Logit Mean[a]	0.0	0.0	0.0
Severity Measure Logit Standard Deviation	0.40	0.39	0.46
Infit Mean-Square (Range)	0.90 – 1.11	0.92 – 1.10	0.74 – 1.27
Outfit Mean-Square (Range)	0.53 – 1.95	0.66 – 1.92	0.74 – 1.27

[a] Mean constrained at 0 for rater severity parameter estimation

Dental Hygiene Examinations Passing Percentages 2016

Five types of passing percentages from the 2016 Dental Hygiene, Local Anesthesia and Restorative Examinations are provided in this section. The five types are listed below and described in additional detail on pp. 21 – 22 of this document.

- 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 2010 to 2016 – counts each Candidate's final result at the end of the seven-year period from 2010 to 2016, 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 20. Note that the Overall Local Anesthsia passing percentages show only all attempts and end of season results by individual Candidates; first attempts and retakes are shown for the two Local Anesthesia examination sections, Written and Clinical.

Table 20. Passing Percentages, Dental Hygiene Examinations and Sections, 2016.

Examination	All Attempts (Includes Retakes)		First-time Attempts		Retakes		Individual Candidates (End of season result)	
	% Passing	*N*	% Passing	*N*	% Passing	*N*	% Passing	*N*
Dental Hygiene	89.9%	2,035	91.3%	1,832	77.3%	203	98.4%	1,859
Local Anesthesia Written	86.7%	1,317	88.6%	1,141	74.4%	176	97.9%	1,167
Local Anesthesia Clinical	76.6%	1,403	77.8%	1,109	72.1%	294	96.4%	1,115
Local Anesthesia Overall	74.4%	1,445	-	-	-	-	94.3%	1,140
Restorative	80.0%	466	79.9%	399	80.6%	67	91.4%	408

Passing percentages for the seven-year period from 2010 to 2016 are provided in Table 21. Passing percentages for all attempts include all initial attempts and retakes. The passing percentage for individuals counts each Candidate only once, regardless of whether the Candidate challenged the examination only once or engaged in repeated retakes. Candidates that have been unsuccessful multiple times must submit documentation of remediation to retake the examination. For the Dental Hygiene and Local Anesthesia examinations, the proportion of individual Candidates who remain unsuccessful over time continues to fall between 2 and 3% upon each seven-year period update, which is consistent with findings for the WREB Dental examination. The Restorative examination is an elective examination for many Candidates, with less unsuccessful Candidates returning for retakes. In states where successful completion of the Restorative examination is required for Dental Hygiene practice, the passing percentages are higher, with 72.9% of all attempts passing and 96.7% of individuals succeeding within the seven-year time frame.

Table 21. Dental Hygiene Examinations Passing Percentages Over Past Seven Years, 2010 – 2016.

	All Attempts (Includes Retakes)		**Individual Candidates (End of Seven-yearResult)**	
Examination	**% Passing**	***N***	**% Passing**	***N***
Dental Hygiene	89.3%	13,554	98.2%	12,326
Local Anesthesia Overall	78.0%	9,770	97.6%	7,802
Restorative	66.5%*	3,708	85.6%*	2,877

*72.9% for All Attempts and 96.7% for Individuals, over seven years, where required for Dental Hygiene practice.

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.

Brennan, R. L., & Kane, M. T. (1977). An index of dependability for mastery tests. *Journal of Educational Measurement, 14,* 277-289.

California Department of Consumer Affairs (2000). *Examination Validation Policy*. Sacramento, CA: CDCA.

Civil Rights Act of 1964, Pub.L. 88-352, 78 Stat. 241 (1964).

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.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,16*, 297-334.

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.

Peng, C-Y. J., & Subkoviak, M. J. (1980). A note on Huynh's normal approximation procedure for estimating criterion-referenced reliability. *Journal of Educational Measurement, 17,* 359-368.

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. (2016a). *2016 Dental Hygiene Examination Candidate Guide*. Phoenix, AZ: WREB.

Western Regional Examining Board. (2016b). *2016 Local Anesthesia Examination Candidate Guide*. Phoenix, AZ: WREB.

Western Regional Examining Board. (2016c). *2016 Restorative Examination Candidate Guide*. Phoenix, AZ: WREB.

Western Regional Examining Board. (2016d). *2016* WREB *Policy Guide*. 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

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
dentalinfo@wreb.org • hygieneinfo@wreb.org