JCNDE JOINT COMMISSION ON NATIONAL DENTAL EXAMINATIONS

Dental Licensure Objective Structured Clinical Examination (DLOSCE)

Technical Report

 | 800.232.1694 | ADA.org/DLOSCE

Executive Summary
Technical Report: Dental Licensure Objective Structured Clinical Examination

The Technical Report for the Dental Licensure Objective Structured Clinical Examination (DLOSCE) is the main source of validity evidence available to state licensing boards and other users of DLOSCE results. Validity is the most important consideration for any examination program. For the DLOSCE, validity refers to the degree to which logic and evidence support the use and interpretation of examination results for making pass/fail decisions affecting candidates for licensure to practice dentistry. The technical report contains both direct evidence and references to other documents and sources of information that contribute to this body of validity evidence. The background and historical information in this report allow users to understand the development of this program.

The content of the Technical Report is presented to address professional standards regarding the validity of credentialing examinations (American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME), 2014). Some of the principal information presented in the Technical Report is summarized below.

Purpose: The purpose of the DLOSCE program is to measure whether a candidate possesses the clinical judgment and skills required for the safe, independent practice of entry-level general dentistry.

Content: Content specifications for the DLOSCE are based on subject matter expert judgment, and validity studies involving practice analyses. Test constructors are responsible for recommending minor modifications during the interim period between practice analyses. The Joint Commission on National Dental Examinations (JCNDE) approves all changes to the content specifications.

Item and Examination Development: Test construction teams are responsible for the development of items and forms/editions of the DLOSCE using JCNDE guidelines for writing high-quality items.

Standard Setting and Scoring: The DLOSCE standard is criterion-referenced (not norm-referenced). This means examination results are determined by specific criteria and not by the process sometimes known as “grading on a curve.” A panel of expert educators and practitioners recommend the minimum passing score, which was ultimately established by the JCNDE. The DLOSCE standard is maintained across examination forms through the use of equating procedures designed to control for small differences in the difficulty of items from one examination form to another. The equating process places examination results on a common metric regardless of which particular examination form was administered.

Administration: The JCNDE maintains a high level of security on all examination materials. Strict precautions in place at the Joint Commission’s offices and testing centers help ensure test content remains secure. The Joint Commission offers the DLOSCE via computer at Prometric professional level testing centers throughout the United States, and its territories. Once eligible, candidates can schedule an examination for any business day, conditional on testing center availability.

In addition to the items above, this report provides information on the history of the examination program, reliability of results, and examination security, among other matters. A copy of this Technical Report is available for download on the JCNDE website, ada.org/JCNDE.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC: Author.

Table of Contents

1. Introduction

Purpose of the DLOSCE Technical Report

High-stakes examination programs must be concerned with validity. Validity refers to the degree to which logic and evidence support the use and interpretation of examination results in accordance with the purpose of the examination (AERA, APA, & NCME, 2014). The Joint Commission has an obligation to inform dental boards and communities of interest concerning its efforts to provide the highest quality examination programs possible. Established professional standards provide useful guidance to improve the quality of examinations. Testing programs must adhere to these standards and provide evidence their policies and procedures conform to them to help ensure confidence in the examination program.

The *Standards for Educational and Psychological Testing*, most recently published by AERA, APA, and NCME in 2014, provide professional standards for testing organizations. Chapter 7 of the *Standards* describes the importance of documented validity evidence in technical reports so examination users can evaluate the validity of examination results they interpret and use. The overarching standard for Chapter 7 is as follows:

Standard 7.0
Information relating to tests should be clearly documented so that those who use tests can make informed decisions regarding which test to use for a specific purpose, how to administer the chosen test, and how to interpret test scores (AERA, APA, NCME, 2014, p. 125).

This technical report provides a comprehensive summary of DLOSCE validation efforts, as well as background information which allows the reader to understand the program's development to its present state. The Joint Commission endeavors to provide the highest quality examination programs possible.

The Joint Commission on National Dental Examinations

The Joint Commission on National Dental Examinations is the agency that oversees DLOSCE examination design, administration, scoring, and reporting. The ADA's Department of Testing Services (DTS) provides operational and technical support for the corresponding outlined activities. The mission of the Joint Commission is as follows:

> *Protecting public health through valid, reliable, and fair assessments of knowledge, skills, and abilities to inform decisions that help ensure safe and effective patient care by qualified oral healthcare team members.*

The Rules of the JCNDE provide descriptions of Joint Commission membership and their core responsibilities.

2. DLOSCE Overview

The first and most fundamental step in the development of any examination program is to establish a purpose as described in Standard 1.1.

Standard 1.1
The test developer should set forth clearly how test scores are intended to be interpreted and consequently used. The population(s) for which a test is intended should be delimited clearly, and the construct or constructs that the test is intended to assess should be described clearly (AERA, APA, NCME, 2014, p. 23).

The purpose of the DLOSCE program is to measure whether a candidate possesses the clinical judgment and skills required for the safe, independent practice of entry-level general dentistry. The intended examinee population for the DLOSCE consists of candidates who are seeking a license to practice general dentistry in any state, district or other jurisdiction of the United States. The intended interpretation of DLOSCE results concerns the candidate's ability to apply clinical judgment and skills in the provision of patient care. A passing score on the DLOSCE indicates that a candidate is able to apply the aforementioned judgment and skills at the level required for the safe, independent practice of entry-level general dentistry. DLOSCE results are used by dental boards in determining qualifications of dentists who seek licensure to practice in any state, district or other jurisdiction of the United States, which recognizes the DLOSCE.

3. Historical Perspective

In 2016, the ADA's Council on Dental Education and Licensure (CDEL) requested that the ADA's Department of Testing Services develop a business plan for development and implementation of a Dental Licensure OSCE. CDEL reviewed the plan at their December 2016 meeting and recommended the ADA's Board of Trustees provide funding to develop the DLOSCE. In January 2017 a National Licensure Task Force, jointly sponsored by the ADA and ADEA, unanimously endorsed the development of the DLOSCE. During its February 2017 meeting, the ADA Board of Trustees discussed the DLOSCE business plan written by DTS. At that time, the Board of Trustees authorized the formation of a DLOSCE Steering Committee charged with developing and validating the DLOSCE. Dr. Gary L. Roberts, ADA President, appointed a set of highly qualified individuals to the Steering Committee based on criteria established by the Board of Trustees. The DLOSCE Steering Committee held its inaugural meeting in July 2017 at the ADA headquarters in Chicago.

Throughout the development and validation process—and particularly during its first meetings—the Committee devoted considerable time and energy to discussions concerning the establishment of the DLOSCE content domain, and in what form and by which methods DLOSCE content should be presented to candidates. Subsequent to thorough review, in March 2018, the DLOSCE Steering Committee determined that the DLOSCE should be a virtual (i.e., computer-based) examination that would not directly measure psychomotor skills. The Committee also approved preliminary content areas and test specifications for the DLOSCE. Detailed information concerning the factors considered are provided in Section 6, "*Content Basis for the Examination,*" and in a published article entitled "*The Dental Licensure OSCE: A Modern Licensure Examination for Dentistry*" (Ziebert and Waldschmidt, 2020). At this time, the Committee also authorized the formation of a DLOSCE Working Committee composed of dental

subject matter experts, to recommend structures for DLOSCE test construction teams, and to guide the development of DLOSCE content during test construction meetings.

The first DLOSCE test construction meeting took place in November 2018, and a large number of additional test construction meetings were held in the six month period that followed. In 2019, the DLOSCE Steering Committee approved modifications to the initial test specifications, and determined that the examination would contain questions involving lifelike, three-dimensional (3D) models that could be interacted with and manipulated (magnified, moved, and rotated) by the candidate. Development of the 3D models began shortly thereafter.

In January 2020, the ADA Board of Trustees approved the JCNDE as the governing body for the DLOSCE Program – an action consistent with the wishes of both the DLOSCE Steering Committee and the JCNDE, as expressed through formal communications between the two groups beginning in 2018 and continuing through 2019. In February 2020, the JCNDE voted to accept governance responsibilities pertaining to the DLOSCE, and the DLOSCE Steering Committee became an *ad hoc* Committee of the JCNDE at that time. In February 2020, the DLOSCE Steering Committee and several DTS staff travelled to a Prometric testing center to review an initial completed version of the DLOSCE. Subsequent to their review, the Steering Committee members expressed overwhelmingly positive feedback regarding the quality of the examination. In April 2020, the JCNDE announced that the DLOSCE would be made available for use by dental boards in the United States, beginning in June 2020. Shortly thereafter, the JCNDE published a summary of validity evidence supporting the intended use and interpretation of DLOSCE results, and conducted a series of webinars for dental board members, dental educators, and dental students. Dental boards from a number of US states subsequently indicated they would accept DLOSCE results as either fully or partially fulfilling their clinical examination requirement. The DLOSCE was updated to incorporate minor changes recommended by the Committee in February 2020, and then administered for the first time from June 15 through July 17, 2020. Results from the first administration were released to candidates, dental boards, and dental schools in August 2020.

4. Professional Test Standards

Large testing organizations responsible for developing, administering, and scoring examinations need criteria, or standards upon which to judge their effectiveness. Three professional organizations – AERA, APA, and NCME – joined forces and resources to create the latest version of the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 2014). These standards provide useful information to guide testing organizations in the validation of their test score interpretations and uses. Throughout this technical report, validity evidence is identified and connected to testing standards. Many sections of this technical report correspond to chapters in the *Standards* (AERA, APA, NCME, 2014). Where applicable, the standards are referenced in this document.

5. Overview of Validity

Validity is defined in the *Standards* as “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (AERA, APA, & NCME, 2014, p. 11). Validation involves the investigative process of creating a validity argument and collecting evidence relevant to this argument, the examination purpose, and the intended interpretation of results (Kane, 2016. When acquired validity evidence reveals weaknesses or deficiencies, the

testing organization is expected to take steps to address the deficiencies to strengthen the validity of the test.

The intended interpretation of DLOSCE results concerns the candidate's ability to apply clinical judgment and skills in the provision of patient care. A passing score on the DLOSCE indicates that a candidate is able to do so at the level required for the safe, independent practice of entry-level general dentistry. DLOSCE results are used by dental boards in determining qualifications of dentists who seek licensure to practice in any state, district or other jurisdiction of the United States, which recognizes the DLOSCE. This technical report presents validity evidence and additional references that support the intended interpretation and use of DLOSCE results, as suggested by Standard 1.0.

Standard 1.0
Clear articulation of each intended test score interpretation for a specified use should be set forth, and appropriate validity evidence in support of each intended interpretation should be provided (AERA, APA, & NCME, 2014, p. 23).

This report is organized to address major categories of validity evidence. Each section contains narrative and validity documentation. In some instances, data are provided, as appropriate. The report addresses the following important categories of validity evidence, presented with corresponding section numbers:

6. Content Basis for the Examination
7. Test Design and Development
8. Scoring and Equating Methods
9. Standard Setting
10. Reliability
11. Test Administration
12. Results Reporting
13. Convergent Validity Evidence
14. Test Security
15. Rights and Responsibilities of Test Takers
16. Candidate Performance

The information provided in this technical report covers the entire span of DLOSCE development through December 2023.

Legal Issues

All examination programs where results are used for high-stakes decisions run the risk of legal challenge based on validity. As a result, examination programs must be designed to withstand legal challenges.

This technical report represents an effective way to present the examination validity argument and corresponding validity evidence. This document organizes, describes, and presents a large amount of validity evidence. In so doing, boards can have confidence that the Joint Commission has acted responsibly in its duty to develop and administer an examination program capable of fulfilling its intended purpose.

6. Content Basis for the Examination

Content-oriented validity evidence is a critical source of validity evidence supporting the interpretation and use of DLOSCE results. The *Standards* indicate that developers of licensure examinations should provide a thorough description of the examination's content domain, along with evidence that the domain reflects the requirements of the profession for which candidates are seeking licensure, as discussed in Standards 11.2 and 11.3 (AERA, APA, NCME, 2014, p. 178-179).

Standard 11.2
Evidence of validity based on test content requires a thorough and explicit definition of the content domain of interest (AERA, APA, NCME, 2014, p. 178).

Standard 11.3
When test content is a primary source of validity evidence in support of the interpretation for the use of a test for employment decisions or credentialing, a close link between test content and the job or professional/occupational requirements should be demonstrated. (AERA, APA, NCME, 2014, p. 178-179).

This chapter details the DLOSCE content domain and describes the theoretical rationale and empirical evidence that support it. In short, the content domain for the DLOSCE consists of the clinical tasks that a dentist performs while providing direct, chair-side treatment to patients. The content domain is formalized in the DLOSCE test specifications, which were established using the methods and procedures described below.

Establishing the DLOSCE Test Specifications

In 2018, the DLOSCE Steering Committee convened a review panel of subject matter experts to recommend test specifications for the DLOSCE. The recommended test specifications would describe the topic areas the DLOSCE should cover and the percentage of test items that should be allocated to each topic area. The review panel consisted of 11 dental subject matter experts, including general dentists, and specialists with expertise in the following areas: Prosthodontics, periodontics, oral radiology, oral diagnosis, oral surgery, endodontics, behavioral science, orthodontics, pharmacology, and dental anesthesia. The panel included three members of the DLOSCE Steering Committee as well as an individual who had served on the Committee for an Integrated Examination (CIE) (the committee that developed and validated the Joint Commission's Integrated National Board Dental Examination (INBDE)). The panel met for 1½ days at the ADA building in Chicago.

As a first step, the panelists studied the results of a dental practice analysis survey conducted by the JCNDE in 2016. The purpose of the practice analysis survey was to gather feedback from a nationally representative sample of practicing dentists, concerning the importance of various tasks that general dentists perform. The first section of the survey gathered information about the dentist and their practice environment. The second section consisted of a list of 56 clinical content areas (see Appendix A). In this section, responding dentists were asked to rate each clinical content area with respect to its importance to patient care, and its frequency of use in patient care. The levels of the rating scale were defined as follows:

Importance to Patient care:

5. Extremely important
4. Very important
3. Important
2. Somewhat important
1. Not important

Frequency of Use in Patient Care:

6. More than 5 times per day
5. 3-5 times per day
4. 1-2 times per day
3. 1-4 times per week
2. Less than once per week
1. Never

The JCNDE distributed the practice analysis survey to a total of 34,441 dentists. Among those, 2,542 (7.4%) provided valid responses. The mean importance rating and mean frequency rating were calculated for each clinical content area. The mean importance ratings across clinical content areas ranged from 3.22 to 4.82. The mean frequency ratings ranged from 1.7 to 5.92. The multiplicative model (Kane, Kingsbury, Colton, & Estes, 1989) was used to provide an overall index of importance for each clinical content area.

The members of the DLOSCE test specifications review panel studied the 56 clinical content areas from the practice analysis survey, along with the mean importance rating and mean frequency rating for each area. The panel members then engaged in a group discussion through which they 1) established a preliminary list of topic areas that the DLOSCE should cover (e.g., endodontics, periodontics), and 2) made a preliminary determination regarding the percentage of test items that should be allocated to each topic area. Once the panel had established the preliminary percentages as a group, each panelist separately reviewed the percentages, and suggested changes as needed. The recommended changes were then summarized across the panelists, and presented to the group for consideration.

A key step in establishing the recommended DLOSCE test specifications involved articulating areas of commonality and important differences between the DLOSCE and the JCNDE's Integrated National Board Dental Examination (INBDE). As part of this discussion, the review panel established a preliminary scope and boundaries for skills it felt the DLOSCE should assess (i.e., the DLOSCE skill domain). A statement concerning this scope is provided below.

> *The DLOSCE covers the clinical tasks that a dentist performs while providing direct, chair-side treatment to patients in a clinical environment. This includes addressing issues that arise during the performance of a dental procedure.*

To further clarify the DLOSCE content domain, the panelists scrutinized the 56 clinical content areas included on the Joint Commission's practice analysis survey, which were broken down into three component sections: 1) Diagnosis and Treatment Planning, 2) Oral Health Management, and 3) Practice and Profession. Panel members then selected the clinical content areas they believed fit within the description of the DLOSCE skill domain they established in the previous step. Each panelist did this separately. Then, the panel worked as a group in an attempt to reach a consensus for each clinical content area.

Next, the panelists completed an exercise that required them to link the preliminary topic areas to the clinical content areas from the JCNDE practice analysis. As part of this exercise, panelists were asked to identify the topic areas that were related to each of the 56 clinical content areas. Results of the linking exercise demonstrated a strong relationship between the preliminary topic areas and the tasks that entry level general dentists perform, as indicated by the 56 clinical content areas from the practice analysis. This provided support for the appropriateness and comprehensiveness of the DLOSCE topic areas identified by the panel.

Just prior to the close of the meeting, panelists were given an opportunity to recommend changes to the percentages allocated to the established DLOSCE topic areas. Five of the eleven panelists recommended no changes to the percentages. No panelist recommended a change larger than two percent for any topic area. After discussing their individual recommendations as a group, the panel decided to retain the original percentages. The resulting topic areas and corresponding percentages represented the review panel's recommendation concerning preliminary test specifications for the DLOSCE. The DLOSCE Steering Committee reviewed and approved the review panel's recommendation in March 2018. This established the preliminary test specification for the DLOSCE.

In 2019 the DLOSCE Steering Committee revisited the preliminary DLOSCE test specifications and made modifications based on feedback from the DLOSCE Working Committee, DLOSCE test constructors and DTS staff. In 2021, the JCNDE reviewed and approved the test specifications, which appear in Appendix B. Each DLOSCE form is built to meet the specifications, ensuring that candidates who attempt the DLOSCE encounter an examination that is comprehensive and parallel in its coverage of the content domain. The JCNDE conducts comprehensive practice analyses on a periodic basis, and will continue to use practice analysis results, in combination with subject matter expert judgments, to ensure that the DLOSCE test specifications reflect clinical dental practice. In the time period between practice analyses, DLOSCE test constructors evaluate the specifications and – accompanied by appropriate justification – recommend minor changes as needed, for consideration by the JCNDE.

DLOSCE Content and the Question of Psychomotor Skill Evaluation

The preceding discussion focuses largely on the procedures used to determine the DLOSCE test specifications. An important question present throughout DLOSCE development involves whether to include or not include a psychomotor skill evaluation component within DLOSCE administrations. The DLOSCE Steering Committee specifically and carefully considered this important question. In so doing, the following factors were thoroughly discussed:

- research evidence as it pertains to current clinical licensure examinations, which include both patient-based and manikin components
- research from the National Dental Examining Board (NDEB) of Canada, which had for decades utilized an Objective Structured Clinical Examination (OSCE) instead of a patient-based clinical examination
- the fidelity of existing manikin and dental simulation technology, as it relates to the day-to-day experience of practicing dentists
- ethical considerations pertaining to the use of patients in clinical examinations for licensure purposes
- the pre-eminent role of clinical judgment as it relates to the application of psychomotor skills

- the standards for dental education as promulgated by the Commission on Dental Accreditation (CODA), and the corresponding dental subject matter expert site visitors who scrutinize the quality of educational training provided at US dental schools
- the educational training provided by dental schools in accordance with CODA standards
- the evaluative tools and methods used by dental schools to understand whether a given student has demonstrated the necessary level of clinical judgment and skills
- the focal reasons for board disciplinary actions
- the applied experience of dental educators
- the need for comprehensive assessment of a candidate's clinical knowledge, skills, and abilities, at the time of licensure
- the validity and reliability of available and proposed solutions, from a rigorous psychometric perspective and in accordance with professional standards and guidelines

The DLOSCE Steering Committee fully acknowledged the critical importance of psychomotor skills in dental practice. Dentists rely heavily on psychomotor skills in treating their patients. Having noted this, the Committee was dismayed to see the dearth of research evidence supporting the validity of current clinical licensure examinations, whose focus primarily rested upon the measurement of these psychomotor skills (see, for example, Chambers, 2011; Formicola et al., 1998; Gadbury-Amyot et al., 2014; Hangorsky, 1981; Ranney et al., 2004). The Committee noted in particular a published editorial appearing in the Journal of Dental Education, offered by Dr. Steven Friedrichsen, Dean and Professor of the College of Dental Medicine of the Western University of Health Sciences. Consistent with the Committee's findings, Dr. Friedrichsen (2016) indicated the following:

> "There is no peer-reviewed scientific evidence that correlates [clinical licensure examination] outcomes with other validated assessments of clinical competence … the process yields no verifiable value in its ultimate objective of providing for the protection of the public." (p. 640)

The Committee was acutely aware of the essentiality of validity, particularly in high-stakes licensure testing in health care, where the public health is at risk. The following opening statements from the first chapter of the *Standards* were germane:

> "Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing tests and evaluating tests.
>
> …
>
> Evidence of the validity of a given interpretation of test scores for a specified use is a necessary condition for the justifiable use of the test." (p. 11).

In considering these matters, the Committee's review yielded the following core findings:

- peer-reviewed research evidence fails to provide adequate support for the use of patient-based and manikin-based clinical licensure examinations (Chambers, 2011; Formicola et al., 1998; Gadbury-Amyot et al., 2014; Hangorsky, 1981; Ranney et al., 2004); these examinations unfortunately do not appear to protect the public.
- peer-reviewed research evidence has supported use of the NDEB Canada OSCE (see Gerrow et al., 2003; Gerrow et al., 2006); additionally, Canada has relied on their OSCE and written examination for decades, without apparent issue.

- existing dental simulation technology was interesting but did not yet possess the level of fidelity necessary to warrant application in licensure testing; this technology should continue to be monitored and considered in the future.
- current manikins also lacked reasonable fidelity from the Committee's perspective. Manikin utilization was regarded as perhaps useful in the early stages of dental education, but represented a step backward when used for licensure purposes. As one member of the DLOSCE Steering Committee noted "drilling on plastic teeth just shows that an individual can drill on plastic teeth."
- in the past, there were varying points of view regarding the perceived rigor of the CODA accreditation process, and questions were present concerning the scope and rigor of school-based assessment procedures. However—thanks to over 25 years of hard work and the adoption and evolution of competency-based education in accredited dental schools, as well as the identification of new, effective pathways for dental clinical assessment—the situation has changed and strong accreditation standards are now in place and uniformly enforced throughout the US (American Dental Association, American Dental Education Association, American Student Dental Association, 2018).
- consistent with CODA standards, dental students are currently evaluated on their psychomotor skills and performance on hundreds of occasions during their enrollment in dental programs accredited by CODA (American Dental Association, American Dental Education Association, American Student Dental Association, 2018). This evaluation can take place using a variety of proven methods of skill evaluation and assessment, including patient-based, manikin-based, simulations, and OSCEs.
- current and former dental board members serving on the DLOSCE Steering Committee indicated that board disciplinary actions appear to predominantly focus upon issues involving mistakes arising from poor clinical judgment, substance abuse, and ethical failures, as opposed to deficiencies in psychomotor skills
- numerous dental educators have indicated to Committee members and staff that current clinical examinations appeared to be failing candidates arbitrarily:
    - the strongest students sometimes failed clinical licensure examinations while less skilled students passed without issue
    - virtually all students who failed a clinical licensure examination passed on their next attempt, in many cases without any remediation
- with respect to ethical issues, the Committee noted that the American Dental Association, American Dental Education Association, and American Student Dental Association had all adopted policies seeking to end the use of patients in dental clinical licensure examinations. These three associations in turn formed a joint Task Force on Assessment of Readiness for Practice, and issued a report (Sept 2018) indicating the following:

> "… the Task Force opposes single, encounter, procedure-based examinations on patients, which virtually all states currently use to fulfill the clinical examination requirement. This approach has been demonstrated to be subject to random error; does not have strong validity evidence; is not reflective of the broad set of skills and knowledge expected of a dentist; and poses ethical challenges for test-takers, dental schools, and the dental profession … this single focus is typically in lieu of the patient's comprehensive and most severe or urgent needs, resulting in a standard of care that may well be below today's acceptable level … the Task Force calls upon state dental boards to eliminate the single encounter,

procedure-based patient exams, replacing these with clinical assessments that have stronger validity and reliability evidence." (p2)

This task force set the stage for the Coalition for Modernizing Dental Licensure, which has moved forward to help achieve the desired changes.

- current clinical examinations are not adequately comprehensive, focusing only on a narrow set of procedures conducted on an extremely small number of patients (e.g., often just two or three) with an extremely limited sample of performance obtained.

In light of these findings, the Steering Committee determined that the new examination should be computer-based and not directly measure psychomotor skills (i.e., due to the unfortunate deficiencies associated with current methods of psychomotor skill evaluation in dentistry). Given the positive research findings associated with OSCEs, the Committee felt that a hybrid or "virtual OSCE" should be pursued, with lifelike 3D models to emulate the experience in the dental clinic. Utilization of 3D models in lieu of live patients could provide further benefits through increased standardization of the testing experience, improving the reliability and validity of the examination with respect to its intended purpose. Extended multiple choice questions to accompany these 3D models could reduce the impact of guessing and provide candidates with a simulated clinical situation possessing greater fidelity and requiring sound clinical judgment, to truly understand whether a candidate "thinks like a doctor." In short, the Committee determined that the public would be far better served by a comprehensive examination focused upon clinical judgment. The quality of clinical judgments made by practicing dentists have a causal effect on patient outcomes, and psychomotor behaviors themselves. The Steering Committee concluded that utilization of manikins and patient-based demonstrations of performance should be unnecessary given the questionable evidence that is present even after decades of use of these examinations by boards.

7. Test Design and Development

Having established the content basis for the DLOSCE, the next considerations involved test design and development. The overall design of an examination is a crucial step in test development. Standard 4.0 describes the importance of documenting the test design process.

Standard 4.0
Tests and testing programs should be designed and developed in a way that supports the validity of interpretations of the test scores for their intended uses. Test developers and publishers should document steps taken during the design and development process to provide evidence of fairness, reliability, and validity for intended uses for individuals in the intended examinee population.

The DLOSCE is designed with the full participation of content expert teams and supervised by staff specialists working in the Department of Testing Services test development area. This process ensures that the expertise of highly qualified, licensed dentists is brought to bear during the examination design process. Joint Commission staff in the Department of Testing Services provide technical support and guidance to help ensure the desired technical qualities are achieved during the examination design phase.

Examination Format

The DLOSCE is a comprehensive examination consisting of 150 items. This includes 148 multiple choice items, and two prescription tasks. Pre-examination materials (e.g., the DLOSCE Candidate Guide) provide candidates with information concerning the format and scoring rules for each item type.

Multiple choice items. Multiple choice items appearing on the DLOSCE represent clinical problems that the candidate must solve. Each multiple-choice item consists of a stem, which poses a clinical problem, followed by a list of possible answers. The stem of an item is usually either a question or an incomplete statement. The two types of multiple choice items that appear on the DLOSCE are described below.

Single correct answer (i.e. single select). These multiple choice items consist of a stem, which poses a clinical problem, followed by a list of possible options. A candidate can only select one option, and only one of the possible options is correct. If the candidate selects the correct option, they earn full credit for the question; otherwise they earn no credit.

One or more correct answers (i.e., multi-select). These multiple choice items consist of a stem, which poses a clinical problem, followed by a list of possible options. One or more of the possible options is correct. To earn full credit, a candidate must select all of the correct options and avoid selecting any of the incorrect options. A candidate who selects an incorrect option automatically earns no credit for the item. A candidate can earn partial credit if they select one or more of the correct options and avoid selecting any of the incorrect options. When multiple correct options are present within an item, DLOSCE test constructors assign a point value to each correct option. Some options may also be designated as unscored. Candidates neither gain credit nor lose credit for selecting an unscored option. An option can be designated as unscored, for example, if it cannot be judged definitively based on the information presented in the item, or if subject matter experts disagree on whether or not it is correct. DLOSCE test constructors determine which options, if any, will be unscored.

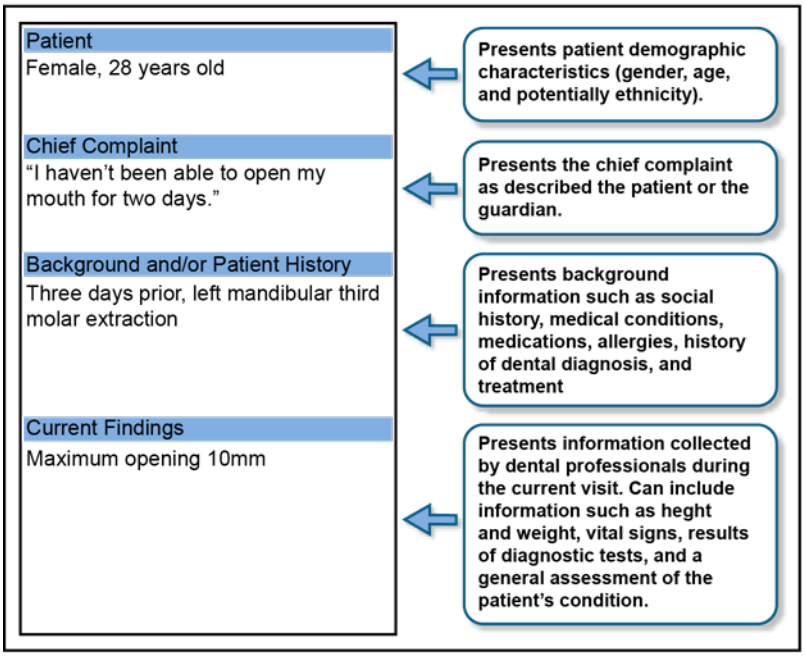
Prescription Tasks. As noted previously, the DLOSCE contains two prescription tasks. These tasks require a candidate to review a Patient Box and determine an appropriate prescription for the patient described therein. For each prescription task, a candidate must

1. review a Patient Box, which provides information about the patient for whom the prescription will be written;
2. select appropriate medication(s) from a list;
3. specify the strength of the tablet/capsule (e.g., 500 mg);
4. specify the total number of tablets/capsules that should be dispensed;
5. identify the number of tablets/capsules that should be taken per administration;
6. specify whether or not the patient should take a loading dose; and
7. identify the frequency of administration (e.g., once a day until finished, twice a day as needed)

Prescription task responses are evaluated against a scoring key established by subject matter experts. Based on the combination of responses selected by the candidate, it is possible for the candidate to receive no credit, partial credit, or full credit for each prescription task.

Three-dimensional models. The DLOSCE contains items involving three-dimensional (3D) models that can be interacted with and manipulated (magnified, moved, and rotated). The JCNDE has made an online tutorial available, so that candidates can practice interacting with a sample 3D model before they attempt the examination. A tutorial provided at the beginning of the examination instructs examinees on how to manipulate the model. Items involving 3D models include a help feature that displays similar instructions for the candidate to reference during the examination.

Patient Box. Many DLOSCE items include a Patient Box. The Patient Box presents information available to the dentist at the time of the visit. The elements of the Patient Box are described below.



There are a number of benefits associated with using the Patient Box format to present patient information. Specifically, the Patient Box:

- permits the candidate to focus on the concept tested, as opposed to question wording (thereby reducing construct-irrelevant variance),
- simplifies the item writing process for test constructors, allowing them to focus on concepts for evaluation, and
- presents concepts to be tested within the context of an actual patient, thereby increasing the correspondence between test content and the actual experiences of practicing entry-level dentists.

In short, the Patient Box is intended to maximize construct-relevant variance and minimize construct-irrelevant variance. Candidates are instructed to always consider the Patient Box in their responses, and a tutorial provided at the beginning of the examination instructs examinees on how to appropriately interpret information provided in the Patient Box. Similarly, pre-examination materials (e.g., the DLOSCE Candidate Guide) also includes information concerning the Patient Box.

DLOSCE Test Constructors

The Joint Commission relies on subject matter expert test constructors to develop and review DLOSCE items and examination forms. The role of test constructors is fundamental to the examination's validity argument. Test constructors are responsible for developing a clear, precise, and comprehensive set of items for each examination form; in accordance with established test specifications and utilizing rigorous procedures. Together these efforts providing content-related validity evidence in support of test usage. The *Standards* indicate that examination developers should describe the qualifications and characteristics of test constructors, and provide information about the training and materials test constructors receive, as described in Standard 4.8 (AERA, APA, NCME, 2014, p. 88).

Standard 4.8
The test review process should include empirical analyses and/or the use of expert judges to review items and scoring criteria. When expert judges are used, their qualifications, relevant experiences, and demographic characteristics should be documented, along with the instructions and training in the item review process that the judges receive. (AERA, APA, NCME, 2014, p. 88).

The section below presents this information, as it pertains to the DLOSCE.

Test constructors meet in teams to engage in test development activities. Test constructors use their subject-matter expertise—including their experience and understanding of dental practice and familiarity with the curriculum in accredited dental schools—to create, review, and finalize examination content. The following is a list of the responsibilities of every test constructor.

- Submit test development materials (e.g., images, items), in compliance with JCNDE guidelines, within the designated time frame. The number of materials that test constructors are expected to submit varies according to the needs of the examination program.
- Attend each test construction meeting for the duration of the session.
- Construct examination forms according to JCNDE guidelines, test specifications, and content outlines, within the designated time frame.
- Construct additional items for JCNDE item banks as necessary.
- Assign ownership of all examination materials to the ADA and JCNDE, by agreeing to the terms of the copyright assignment.
- Inform the Joint Commission of changes in dental practice, dental procedures, and dental education curricula, suggesting modifications to the test specifications as appropriate.
- Consider special issues and make recommendations at the request of the JCNDE.

- Safeguard the security and confidentiality of the examination by declining offers to assist with review courses and examination preparation materials while serving as a test constructor, and for at least one year following the final term of their appointment.
- Comply with the ADA's policy on professional conduct. This policy includes prohibitions against sexual harassment and other forms of unlawful conduct.

The DLOSCE test specifications provide core information to new test constructors. New test constructors receive an orientation which provides information about the DLOSCE program, and the item development and review process.

Test Construction Teams. DLOSCE test constructors work in teams, referred to as Test Construction Teams (TCTs), to develop DLOSCE items. In 2018, the DLOSCE Steering Committee authorized the formation of a DLOSCE Working Committee, composed of dental subject matter experts, to recommend structures for DLOSCE TCTs, and to guide the development of DLOSCE content during test construction meetings. The Working Committee proposed structures for DLOSCE TCTs, based on the DLOSCE test specifications and the needs of the examination program, and the structures were accepted by the DLOSCE Steering Committee and implemented shortly thereafter. Most DLOSCE TCTs meet multiple times per year, with most meetings approximately 2½ days in duration. TCT meetings are typically facilitated by one or more members of the DLOSCE Working Committee, in collaboration with DTS staff facilitators. The main categories of DLOSCE TCTs are described in detail below. Additional teams may also be created on an *ad hoc* basis to meet the targeted needs of the examination program.

Item Writing and Review Teams. Item Writing and Review teams typically consist of three to ten test constructors. Depending upon item development needs, multiple teams may be formed. Each team is responsible for developing items and reviewing newly written items to ensure content accuracy, currency, and validity, as well as adherence to the test specifications and item guidelines outlined by the Joint Commission. Item Writing and Review teams are typically organized according to the major areas of the DLOSCE test specifications (e.g., Oral Surgery, Periodontics). In order to serve on an Item Writing and Review team, a test constructor must be a dentist, educator and/or practitioner, and in the case of special areas of dentistry, a graduate of an accredited advanced education program in the specialty area for which they develop items.

Clinical Relevance Review Teams. Each Clinical Relevance Review team consists of at least 8 General Dentists. This team confirms the appropriateness of examination items in terms of their relevance to day-to-day clinical practice. The teams are also responsible for the final categorization of items, relative to the DLOSCE test specifications and in support of the general needs of the DLOSCE program. In order to serve on a Clinical Relevance Review Team, a test constructor must be a graduate of a dental program accredited by either the Commission on Dental Accreditation (CODA) or the Commission on Dental Accreditation of Canada (CDAC), a full-time or part-time practitioner or clinician/scientist with at least five years of experience in the United States or Canada, and currently licensed as a dentist in the United States or Canada.

Form Review Teams. Form Review teams provide a final review of DLOSCE items and images identified for placement on examination forms, with respect to clinical relevance and the activities of a practicing general dentist. They ensure that the content being tested is

comprehensive, meets the test specifications, and that there is no unintended overlap among the items included on each form.

Test Constructor Selection. On an annual basis the Joint Commission advertises and promotes its need for test constructors. A letter explaining the online application materials and selection criteria is emailed to dental schools, dental boards, constituent dental societies, and other institutions and individuals each year. All applications are processed by staff and presented to the Joint Commission's Test Constructor Recommendation Committee, which is responsible for recommending individuals to serve in the DLOSCE Test Constructor Pool.

On an annual basis the Joint Commission approves and reapproves test constructors to serve in the DLOSCE Test Constructor Pool. An individual who has completed five years of service in the pool may reapply and be considered for re-approval as dictated by the needs of the examination program. DTS staff place approved test constructors onto specific TCTs based on the expertise of the individual, and the needs of the TCT and examination program. A team is formed for each specific meeting, and disbands at the end of that meeting. These teams are flexible and may or may not consist of the same test constructors each year. Teams may be rearranged as needed in the event that a given volunteer is not able to attend. If a volunteer is invited but is unable to attend, an alternate volunteer may be identified and invited. Additionally, if a volunteer is invited to attend a meeting and does not respond in a timely manner, an alternative volunteer may be identified and invited to attend the meeting. This process helps ensure teams have a sufficient number of volunteers with the required expertise, so that meeting goals can be accomplished efficiently and effectively.

Item Validation

Standard 4.7 indicates that examination developers should document the process used to develop, review, and evaluate items (AERA, APA, NCME, 2014, p. 87-88). This section of the Technical Report describes the item validation process that the JCNDE has implemented for the DLOSCE.

Standard 4.7
The procedures used to develop, review, and try out items and to select items from the item pool should be documented. (AERA, APA, NCME, 2014, p. 87-88).

Content accuracy review. During content accuracy review, test constructors review items for accuracy and currency. In some cases, this review is conducted by the members of the original Item Writing and Review team who developed the item. In other cases, the review is conducted by test constructors who are external to the original Item Writing and Review team.

Item classification. Item classification review is performed to specify the areas of content expertise identified for the item. This review is similar to how a librarian classifies material into subject areas using a defined taxonomy. The classification review includes the review or specification of all metadata for the item.

Editorial review. During editorial review, items are reviewed for grammar, style, formatting, and alignment with DTS item writing guidelines. Similarly, item stimulus materials are reviewed to ensure the information is of diagnostic quality and in accordance with modern dental practice.

Legal/intellectual property (IP) review. Joint Commission staff seek counsel from the ADA Division of Legal Affairs concerning the articulation of guiding principles that might inform procedures and help avoid legal issues involving examination content. This includes, for example, issues arising around privacy and the use of intellectual property. Individuals who submit images and materials to the Joint Commission are responsible for verifying intellectual property rights.

Clinical relevance review. The DLOSCE is designed for licensure purposes, to help state boards understand whether a candidate possesses the necessary clinical judgment to enter the profession and safely practice dentistry. The general dentist is thus of focal importance to the DLOSCE Program. During the clinical relevance review, corresponding review teams scrutinize items to help confirm item content is clinically relevant and applicable to the work of practicing dentists. This review helps reduce the likelihood of an examination form containing trivial and/or esoteric content.

Fairness and sensitivity review. DLOSCE items are reviewed based on fairness and cultural sensitivity considerations, in alignment with the item writing guidelines. A fairness and sensitivity review takes place as part of the original item development process. A supplemental review may also take place to improve items from this perspective as well. In 2022, the JCNDE approved an updated fairness and sensitivity review process and directed staff to pilot the process in 2023. Through the updated review process, Fairness and Sensitivity Reviewers evaluate examination content and presentation through the lens of diversity, equity and inclusion, to help ensure that test questions are respectful of the diversity of perspectives present. In pursuing this charge, Fairness and Sensitivity Reviewers facilitate the accurate and unbiased measurement of candidate knowledge skills, and abilities. Fairness and Sensitivity Reviewers are asked to focus specifically on diversity, equity and inclusion considerations during their review—as well as potentially sensitive subject matter—and to avoid focusing on other factors that would distract from and dilute the fairness review.

Item performance review. Items that survive the reviews described above are eligible to be placed on examination forms. Once an item has been placed on a form, and once the form has been administered to a sufficient number of candidates, DTS calculates the following statistics for each item:

1) The mean score for the item; and
2) The item-total correlation, defined as the Pearson correlation between performance on the item and performance on the examination.

The mean score for the item is an indicator of an item's difficulty, and the item-total correlation is an indicator of an item's ability to discriminate among candidates of different ability levels. Items that fail to discriminate among candidates of different ability levels are scrutinized by staff and then routed to appropriate test constructors for review. Items that do not perform appropriately are removed from scoring. Subsequent to administration, test constructors review examination content and relevant psychometric information and determine whether items can and should be revised. The revision process could, for example, involve rewording the stem or changing the distractors. If an item is revised, it is returned to the item bank where it becomes a candidate for placement on a future examination form. If test constructors determine that an item cannot be improved through revision, the item is designated as unusable.

8. Scoring and Equating Methods

Scoring Approach

There are two common approaches to scoring licensure examinations. Under the first approach, the pass/fail decision is based on a single score that is determined from the candidate's performance on the entire examination. Under the second approach, an examination is divided into separately scored sections, and the candidate must pass each section in order to pass the examination. The latter approach is often used when the topic areas on an examination are substantially distinct from one another and candidate competence on each topic area must be verified separately. When examination topic areas are highly correlated, on the other hand, the former approach is often used, because a single score based on all the test items will be more reliable than the scores determined for the individual topic areas.

Analysis of data from DLOSCE administrations strongly suggests that a candidate's DLOSCE result (i.e., pass or fail) should be based on a single score derived from the candidate's performance on the entire examination. Exploratory factor analysis of candidate scores on the nine DLOSCE topic areas suggested they were indicators of a common underlying ability that can be well represented with a single score; the ratio of the first to second eigenvalue from the factor analysis was 9.8, and a parallel analysis (Horn, 1965), scree test (Cattell, 1966), and the application of Kaiser's (1960) criterion all converged on the finding that a single underlying factor was present. Based on this, the JCNDE adopted a scoring approach for the DLOSCE whereby a candidate's result is determined based on their overall performance on the examination. The section below describes how the overall score and corresponding pass/fail result is determined for each DLOSCE candidate.

Scoring Methods

DLOSCE results are determined through a multi-step process. In the first step, non-performing items are identified and removed so that they do not count toward candidate scores. Non-performing DLOSCE items are identified based on two statistics: the mean score for the item, which is an indicator of an item's difficulty, and the item-total correlation, which is an indicator of the item's ability to discriminate among candidates of different ability levels.

In the second step, a *raw score* is determined for each candidate. The raw score represents the total number of points the candidate earned on the examination, after removing the non-performing items. Each raw score is also expressed as a *percent correct* score, which is calculated as the raw score divided by the total number of points possible. As mentioned previously, the DLOSCE contains three types of items: 1) single-select items 2) multi-select items, and 3) prescription tasks. Candidates can earn a maximum of 1 point for each single-select and multi-select item, and a maximum of 4 points for each prescription task. Partial credit is possible for multi-select items and prescription tasks. When multiple forms of the DLOSCE are administered, raw scores are adjusted through a psychometric process known as equating, to statistically adjust for any differences in the difficulty of the examination forms (for details, see the section titled Equating Methods). The equating process helps ensure that all DLOSCE candidates are held to the same performance standard, regardless of which examination form they attempt.

In the third step, each candidate's equated raw score is converted to a scale score. DLOSCE scale scores can range from 49 to 99 and are expressed as whole numbers (e.g., 49, 50, 51). A scale score of 75 represents the minimum level of clinical judgment and skills required for the safe, independent practice of entry-level general dentistry, as determined through standard setting activities (see Chapter 9). A candidate must earn a scale score of 75 or higher to pass the DLOSCE. Candidates who receive a scale score of 75 or higher receive a status of "Pass," while candidates who receive a scale score below 75 receive a status of "Fail."

Equating Methods. Multiple forms of the DLOSCE are available for administration. The JCNDE takes care to ensure that all DLOSCE forms meet the DLOSCE test specifications and are as parallel as possible. However, because the forms contain different items, small form-to-form differences in difficulty are typically present. The JCNDE uses a process called *equating* to statistically adjust for these differences. The equating process helps ensure that all DLOSCE candidates are held to the same performance standard, regardless of which examination form they attempt. Standard 5.12 and 5.13 indicate that test developers should provide evidence supporting the claim that results from different forms of an examination may be used interchangeably (AERA, APA, NCME, 2014, p. 105). The discussion provided herein is intended to help provide that supporting evidence.

Standard 5.12
A clear rationale and supporting evidence should be provided for any claim that scale scores earned on alternate forms of a test may be used interchangeably (AERA, APA, NCME, 2014, p. 105).

Standard 5.13
When claims of form-to-form score equivalence are based on equating procedures, detailed technical information should be provided on the method by which equating functions were established and on the accuracy of the equating functions.
(AERA, APA, NCME, 2014, p. 105).

The JCNDE uses a common-item nonequivalent groups design to equate DLOSCE scores. In the common-item nonequivalent groups design, there are two samples of candidates, each of which is administered a different form of the examination. There are also some items that are common to both examination forms. The common items comprise an *anchor test* which ultimately forms the basis for the score adjustments. Due to their importance, the anchor test items are carefully chosen based on the guidelines described in Kolen and Brennan (1995). According to these guidelines, the anchor test should meet the test specifications proportionally, and have a sufficient number of items (e.g., 20 percent of the length of a full examination form, or at least 30 items in the case of long examinations).

The *Tucker linear method* (Angoff, 1971) is used to place scores from different DLOSCE forms on the same measurement scale. The Tucker linear equating method is intended for use when a common-item nonequivalent groups equating design is employed. Under the Tucker method, scores from Form X are placed on the Form Y scale using the following equation:

$$Y' = \frac{\sigma(Y_T)}{\sigma(X_T)}[X - \mu(X_T)] + \mu(Y_T), \qquad 8.1$$

where Y' is the Form X score expressed on the Form Y scale, and $\mu(X_T)$, $\mu(Y_T)$, $\sigma(X_T)$, and $\sigma(Y_T)$ are the means and standard deviations of the scores on Forms X and Y for the combined population of candidates (i.e., the candidates who take Form X combined with the candidates who take Form Y). Because the candidates who take Form X do not have scores on Form Y, and vice versa, the means and standard deviations in Equation 8.1 are estimated using information about candidate performance on the anchor test items (for the algebraic formulas used to estimate the parameters in Equation 8.1, see Angoff, 1971 or Kolen, 1985).

Quality Assurance. The *Standards* indicate that those responsible for scoring examinations should establish and document quality assurance measures (AERA, APA, NCME, 2014, p. 118).

Standard 6.9
Those responsible for test scoring should establish and document quality control processes and criteria. Adequate training should be provided. The quality of scoring should be monitored and documented. Any systematic source of scoring errors should be documented and corrected (AERA, APA, NCME, 2014, p. 118).

Accordingly, the JCNDE has established strict quality control measures to facilitate accurate scoring of the DLOSCE. At the close of each DLOSCE administration window, a roster of candidates scheduled to complete the DLOSCE is compared with the candidates appearing in result files, to ensure no result files are missing. Examinations are independently scored by two separate DTS analysts, and the resulting scores are compared against one another to ensure they are identical before results are released to candidates. DTS staff maintain documentation related to the examination scoring process, and corresponding quality assurance procedures.

9. Standard Setting

A critical step in the development of any licensure examination involves the establishment of the cut score that separates passing and failing candidates (AERA, APA, NCME, 2014, p. 100-101). The *Standards* indicate that subject matter experts involved in setting cut scores should be qualified, and that the process for setting the cut score should be well described and documented (AERA, APA, NCME, 2014, p. 107-108). The information provided below is presented in fulfilment of this requirement.

Standard 5.21
When proposed score interpretations involve one or more cut scores, the rationale and procedures used for establishing cut scores should be documented clearly (AERA, APA, NCME, 2014, p. 107-108).

Standard Setting Procedures

In August 2020, the JCNDE convened a standard setting panel to recommend a performance standard (i.e., cut score) for the DLOSCE. The panel identified its recommended cut score using a modified version of the Bookmark standard setting method (Lewis, Mitzel, Mercado, & Schulz, 2012). The modified method was inspired by Angoff's (1971) "Yes/No" method of cut score establishment (see Impara & Plake, 2006), and had been successfully implemented previously

as reported by Buckendahl et al. (2006). The standard setting activities involved the following steps:

1. A standard setting panel was convened. The panel was composed of seven members who were diverse with respect to practice experience, gender, areas of specialized knowledge, and geographic region.

2. The panel members received a thorough overview of the purpose and content of the DLOSCE. This included a description of the test specifications, test construction methods, and scoring methods. As a reference, panel members were also provided with information concerning recent failure rates for several existing dental licensure examinations.

3. Prior to the meeting, panel members completed an abbreviated version of the DLOSCE that was approximately representative of a full version of the DLOSCE with respect to content, difficulty level, timing, and item formats. During the meeting, panel members self-scored their abbreviated examinations and subsequently discussed the items as a group.

4. The panel members engaged in a complete and thorough discussion of the characteristics and skills of the "just qualified" (i.e., minimally competent) candidate, focusing on candidate skills in the specific topic areas covered on the DLOSCE.

5. Following the discussion phase, panel members were trained in the Bookmark standard setting method and given an opportunity to practice the method using provided practice materials.

6. Next, panel members reviewed a large set of examination items that had been placed into an Ordered Item Booklet (OIB) assembled as follows:
    - Each page of the OIB contained one item.
    - The items included in the OIB spanned a representative range of difficulty levels.
    - Items within the OIB were presented in ascending order of difficulty such that the item on the first page was the least difficult and the item on the last page was the most difficult.
    - Single-select items appeared once within the OIB. The "success criteria" for these items involved the candidate answering the question correctly, thereby earning the candidate full credit for the item.
    - Each partial credit item represented in the OIB appeared twice within the booklet:
        1) When an item involving partial credit appeared in the OIB for the first time, the difficulty value for that item was based on the proportion of candidates who earned at least partial credit for the item.
        2) When an item involving partial credit appeared for the second time, the difficulty value for that item was based on the proportion who earned full credit for the item.
    - Given the preceding—and for purposes of the OIB—partial credit items therefore involved two separate "success criteria" levels (i.e., partial credit and full credit).

7. After reviewing the OIB, each panel member was asked to independently "bookmark" the page number in the OIB of the last item for which a minimally competent candidate

would have at least a 50 percent probability of meeting the aforementioned OIB item success criteria. The cut score associated with the bookmarked OIB page was then defined as the score earned by a hypothetical candidate who succeeded on all of the items up to and including the marked page, and failed on all of the items thereafter. The median OIB page placement (across panelists) and corresponding cut score was used to represent the group's recommendation.

8. After making their judgments, panel members engaged in group discussion regarding their bookmark placements and the rationales for their judgments. During this phase panel members were provided with information about the bookmark placements of the other panel members, and the anticipated impact of using the cut score associated with the median bookmark placement (i.e., the percent of candidates who would fail under that cut score).

9. Steps 7 and 8 as described above were repeated three times. After each replication, panel members were provided an opportunity to ask questions, request clarification, express any concerns, and engage in group discussion. Subsequently, each panel member was asked to provide a final recommended OIB page placement. The final recommended cut score for the examination was based on the median of the panelists' page placements.

10. At the conclusion of the activities, panel members were asked to complete a questionnaire regarding their impressions of the process. Most panel members strongly agreed with the following statement: "Overall, I support the final group-recommended cut score as fairly representing the appropriate performance standard for the DLOSCE." On a five-point rating scale, ranging from 1=Strongly Disagree to 5=Strongly Agree, the mean rating for this question was 4.14.

The recommended performance standard resulting from the 2020 DLOSCE standard setting activities was accepted and implemented by the Joint Commission in August 2020. As a matter of practice, the JCNDE revisits performance standards periodically, conducting new standard setting activities as needed (e.g., if substantial changes are made to the DLOSCE test specifications, and as dental practice evolves in substantive ways).

10. Reliability

Score reliability is an important indicator of examination quality. Test developers strive to ensure test scores provide a stable and precise measurement of a candidate's knowledge, skills, and abilities. Despite efforts to eliminate possible sources of measurement error, random factors can affect candidate performance and subsequent examination results. Reliability indices assess the degree to which random error affects scores. When scores on an examination demonstrate low reliability, they are strongly influenced by random sources of measurement error. Conversely, when scores on an examination demonstrate high reliability, they are less subject to random sources of error. The *Standards* highlight the importance of reporting the reliability of test-based decisions for high stakes licensing examinations (AERA, APA, NCME, 2014, p. 46-47) in Chapter 2. The overarching standard for this chapter is provided below:

Standard 2.0
Appropriate evidence of reliability/precision should be provided for the interpretation for each intended score use (AERA, APA, NCME, 2014, p. 42).

A strategy that is commonly used to increase reliability is to lengthen examinations. Having uniformly high-quality items also contributes to reliability.

Internal Consistency Reliability

The Joint Commission uses the alpha reliability coefficient (Cronbach, 1951) as one index of score reliability for the DLOSCE. Coefficient alpha is an index of internal consistency reliability, and can range from zero to one, with higher values indicating higher reliability. Once an estimate of score reliability has been obtained, the standard error of measurement for the examination scores can be estimated as follows:

$$SEM = s\sqrt{1 - r_{xx}} \qquad 10.1$$

where *s* is the standard deviation of the scores, and r_{xx} is the reliability estimate. Under the assumption that random errors are normally distributed, test users can construct a 95 percent confidence interval around a candidate's score by adding and subtracting 1.96 standard errors of measurement from the score. From 2020-2023, alpha reliability coefficients for the DLOSCE ranged across examination forms from .77 to .79. The standard error of measurement ranged from 3.28 to 3.39.

Classification Accuracy and Classification Consistency

When scores on an examination are used as a basis for making pass/fail decisions, it is important for the test developer to demonstrate that the pass/fail decisions are reliable as described in Standard 2.16 below (AERA, APA, NCME, 2014, p. 46-47).

Standard 2.16
When a test or combination of measures is used to make classification decisions, estimates should be provided of the percentage of test takers who would be classified in the same way on two replications of the procedure (AERA, APA, NCME, 2014, p. 46).

To evaluate reliability, testing programs typically estimate *classification accuracy* and *classification consistency*. Classification accuracy is the probability that a candidate's pass/fail result on an examination reflects the decision that would be made had their true skill level been known. Classification consistency is the probability that a candidate would receive the same pass/fail result on two hypothetical, successive administrations of an examination. The JCNDE estimates classification accuracy and consistency for the DLOSCE using a Classical Test Theory-based version of the method described in Rudner (2001), whereby a common standard error of measurement is used for each candidate (see Equation 10.1). From 2020-2023, estimates of classification accuracy for the DLOSCE ranged across examination forms from .938 to .945. Estimates of classification consistency ranged from .910 to .916.

11. Test Administration

The DLOSCE is administered during fixed administration windows throughout the year. Prometric administers the examination at its Testing Centers located throughout the United States and its territories. Once eligible, candidates can schedule an examination for any business day within the administration window, conditional on testing center availability. The administration schedule for the DLOSCE is provided in Table 11.1.

Table 11.1
DLOSCE Administration Schedule

Section	Minutes
Tutorial (optional)	25
Section 1 (37 items)	75
Scheduled break (optional)	10
Section 2 (37 items)	75
Section 3 (2 prescription tasks)	10
Scheduled break (optional)	10
Section 4 (37 items)	75
Scheduled break (optional)	10
Section 5 (37 items)	75
Post-examination survey	20
Total Time	6 hr. 45 min.

The DLOSCE Candidate Guide details DLOSCE candidate eligibility requirements and the DLOSCE application process. The guide is made available to candidates through the JCNDE website (https://jcnde.ada.org).

12. Results Reporting

This chapter describes how DLOSCE results are reported to various stakeholders, including candidates, dental boards, and dental education programs. JCNDE results reporting procedures are consistent with the best practices outlined in Standards 6.10 and 6.16.

Standard 6.10
When test score information is released, those responsible for testing programs should provide interpretations appropriate to the audience. The interpretations should describe in simple language what the test covers, what scores represent, the precision/reliability of the scores, and how scores are intended to be used (AERA, APA, NCME, 2014, p. 119).

Standard 6.16
Transmission of individually identified test scores to authorized individuals or institutions should be done in a manner that protects the confidential nature of the scores and pertinent ancillary information (AERA, APA, NCME, 2014, p. 121).

Reporting DLOSCE Results to Candidates

DLOSCE results are reported to candidates through a secure, password protected electronic portal. Results are typically made available to the candidate within four weeks of the close of the corresponding DLOSCE administration window, barring unusual circumstances (e.g., a candidate's results are being voided or withheld).

DLOSCE candidates who pass the examination receive a report indicating their result is "Pass," but do not receive numeric scores. Candidates who fail the DLOSCE receive a report indicating their result is "Fail" along with their numeric overall DLOSCE scale score. DLOSCE scale scores can range from 49 to 99 in one-point increments; candidates must earn a scale score of 75 or higher to pass the examination. For remediation purposes, candidates who fail the DLOSCE are also provided with a graphical depiction of their performance in the following areas:

- Restorative Dentistry
- Prosthodontics
- Oral Pathology, Pain Management, and Temporomandibular Dysfunction
- Periodontics
- Oral Surgery
- Endodontics
- Orthodontics
- Medical Emergencies

The numeric scores represented in the graphic are placed on a common scale so candidates can compare their relative performance in the different areas and identify areas where they are most in need of remediation. Consistent with best practices outlined in the *Standards*, the results report issued to candidates who fail the DLOSCE contains explanatory text that is intended to help candidates interpret their results accurately.

Reporting DLOSCE Results to Dental Boards

When candidates apply to take the DLOSCE, they also indicate which dental boards should receive their official results. The JCNDE reports Pass/Fail results to the requested dental boards through a secure, password protected electronic portal. A history of the candidate's Pass/Fail results is made available to each dental board requested to receive results. Numeric scores are not reported to dental boards.

Reporting DLOSCE Results to Dental Schools

A candidate's Pass/Fail results are reported to the candidate's dental school, provided that the school is accredited by the Commission on Dental Accreditation (CODA). Results are reported to the school's dean or designee through a secure, password protected electronic portal. CODA accredited dental schools receive periodic reports that describe how their students on average perform on the examination, as compared to the national student average.

13. Convergent Validity Evidence

Convergent validity evidence is established when scores on an examination are positively correlated with scores from other measures of similar constructs. The following section provides

convergent validity evidence in support of the DLOSCE. Convergent validity evidence is discussed in Standard 1.16 as follows:

Standard 1.16
When validity evidence includes empirical analyses of responses to test items together with data on other variables, the rationale for selecting the additional variables should be provided. Where appropriate and feasible, evidence concerning the constructs represented by other variables, as well as their technical properties, should be presented or cited. Attention should be drawn to any likely sources of dependence (or lack of independence) among variables other than dependencies among the construct(s) they represent (AERA, APA, NCME, 2014, p. 27-28).

Correlation with National Board Dental Examination Scores

The JCNDE used data from 2020-2023 administrations of the DLOSCE to examine the relationships between scores on the DLOSCE and scores on Parts I and II of the National Board Dental Examination (NBDE). The JCNDE also examined the relationship between scores on the DLOSCE and scores on the Integrated National Board Dental Examination (INBDE). The samples for these analyses consisted of 373 candidates who took the DLOSCE and Parts I and II, and 116 candidates who took the DLOSCE and INBDE. For all analyses, only scores from candidates' first examination attempts were used.

The observed Pearson correlations between DLOSCE and NBDE scores are provided in Table 13.1. Performance on the DLOSCE showed a moderately strong correlation with performance on NBDE Part II, which measures knowledge and cognitive skills in the clinical sciences. Performance on the DLOSCE showed a moderate correlation with performance on NBDE Part I, which measures knowledge and cognitive skills in the biomedical sciences.

Table 13.1
Observed Pearson Correlations among DLOSCE and NBDE Scores: 2020-2023
N = 373 candidates

Score	DLOSCE	NBDE Part I	NBDE Part II
DLOSCE	1.00		
NBDE Part I	.29	1.00	
NBDE Part II	.48	.60	1.00

Note. Estimates are based on data from candidates attempting the examination for the first time.

Table 13.2 shows the fully disattenuated Pearson correlations between scores on the DLOSCE and scores on NBDE Parts I and II. The fully disattenuated correlations correct for unreliability in the measured variables and represent estimates of what the correlations would be if the variables were measured with perfect reliability.

Table 13.2
Fully Disattenuated Pearson Correlations among DLOSCE and NBDE Scores: 2020-2023
N = 373 candidates

Score	DLOSCE	NBDE Part I	NBDE Part II
DLOSCE	1.00		
NBDE Part I	.33	1.00	
NBDE Part II	.56	.64	1.00

Note. Disattenuated correlations were estimated using the following reliability coefficients: DLOSCE=.79; NBDE Part I=.95; NBDE Part II=.95. Estimates are based on data from candidates attempting the examination for the first time.

Table 13.3 shows the observed and fully disattenuated Pearson correlations between scores on the DLOSCE and scores on the INBDE. These fully disattenuated correlations indicate that the relationship between performance on the DLOSCE and the NBDE Part II is strong.

Table 13.3
Observed and Fully Disattenuated Pearson Correlations between DLOSCE and INBDE Scores: 2020-2023
N = 116 candidates

Observed Correlation	.66
Disattenuated Correlation	.79

Note. Disattenuated correlations were estimated using the following reliability coefficients: DLOSCE=.79; INBDE=.90. Estimates are based on data from candidates attempting the examination for the first time.

The correlations presented in Tables 13.1 through 13.3 provide strong convergent validity evidence in support of the intended interpretation and use of DLOSCE results. As expected, DLOSCE scores correlated more strongly with INBDE and NBDE Part II scores than they did with NBDE Part I scores. This is due to the fact that DLOSCE, INBDE, and NBDE Part II all focus on the clinical sciences, while NBDE Part I focuses on the biomedical sciences. In reviewing the disattenuated correlations, it should be noted that perfect correlations between DLOSCE, INBDE, and NBDE Parts I and II scores would not be desirable, since perfect correlations would indicate that these examinations measure an identical construct. The DLOSCE is intended to measure a construct that is related to, but not identical to the constructs measured by the INBDE and NBDE Parts I and II. The DLOSCE assesses a candidate's skill with with respect to the clinical tasks that a dentist performs while providing direct, chair-side treatment to patients. The INBDE, on the other hand, assesses a broader set of dental cognitive skills integrating the biomedical, behavioral, and clinical sciences. The NBDE Parts I and II assess candidate knowledge and skills in the biomedical sciences and clinical sciences, respectively.

Correlation with Clinical Performance in Dental School

The JCNDE conducted a study to understand the relationship between candidates' scores on the DLOSCE and their clinical performance in dental school. To do so, the JCNDE collected dental school performance information for 40 DLOSCE candidates who attempted the

examination in the summer of 2020. All 40 candidates in the sample were students in the graduating class of 2020 in the same CODA-accredited dental education program in the United States. The study focused on two indicators of candidate performance in dental school: 1) performance in clinical courses in the third-year of dental school, and 2) final clinical performance, as rated by the program's senior associate dean for academic affairs. Results of the correlational analyses are described in the section that follows.

Clinical performance in third-year dental courses. Candidate clinical performance in third-year dental courses was evaluated using course performance data provided by the dental education program that participated in the study.[1] For each third-year course, students had an opportunity to earn a letter of commendation (LOC) by demonstrating superior performance. For the present study, third-year performance was defined as the total number of LOCs earned during the year (referred to hereafter as the LOC Sum Score). The third-year curriculum included 14 courses, so LOC Sum Scores could range from zero to 14. The third-year courses covered the following areas: Oral Pathology, Oral Surgery, Pediatric Dentistry, Prosthodontics, Endodontics, Orthodontics, Medical Emergencies, Management of Medically Compromised Patients, Health Promotion, and Dental Practice Operations.

Table 13.3 presents the observed and disattenuated correlations between DLOSCE scale scores, and the LOC Sum Scores for third-year dental courses. As shown in the table, the observed Pearson correlation between the two outcomes was .37 ($p < .05$), which represents a moderate positive relationship. Visual inspection of the scatter plot of the two variables, however, revealed that a single outlier exerted a strong influence on the correlation. For this reason, Table 13.3 also presents the Pearson correlation calculated with the aforementioned outlier removed. After removing the outlier, the observed Pearson correlation was .57 ($p < .05$). The partially disattenuated correlations presented in Table 13.3 correct for unreliability in the LOC Sum Scores, and represent estimates of what the correlations between the two variables would be if the LOC Sum Scores were perfectly reliable. The fully disattenuated correlations correct for unreliability in both the LOC Sum Scores and the DLOSCE scores, and represent estimates of what the correlations would be if the LOC sum scores and the DLOSCE scores were both perfectly reliable.

Table 13.3
Correlation between DLOSCE Scale Scores and Clinical Performance in Third-Year Dental Courses: 2020

	Observed Pearson Correlation	Partially Disattenuated Pearson Correlation	Fully Disattenuated Pearson Correlation
Full Study Sample (N=40)	.37*	.46*	.51*
Study Sample with Outlier Removed (N=39)	.57*	.70*	.76*

Note. The partially disattenuated correlations were estimated assuming a reliability coefficient of .66 for the letter of commendation (LOC) sum scores. The fully disattenuated correlations were estimated assuming reliability coefficients of .66 for the LOC sum scores and .83 for the DLOSCE scores.
*$p < .05$

[1]Performance information for fourth-year courses was not available.

Final clinical performance in dental school. The concurrent validity study also examined the relationship between candidates' performance on the DLOSCE, and their clinical performance in dental school, as rated by the program's senior associate dean for academic affairs. For this portion of the study, the senior associate dean was asked to categorize the 40 students in the analytic sample into one of three groups based on their final clinical performance in dental school, relative to their 4th year peers.[2] The question posed to the dean appeared as follows:

> *Please place each student into one of the following categories, with regard to their final clinical performance in dental school, relative to their 4th year peers.*
>
> *Top 20%* *Middle 60%* *Bottom 20%*

The 40 DLOSCE candidates from the participating school were classified by the dean into one of the three performance groups. Table 13.4 presents a summary of the DLOSCE scale scores for the candidates in each group. Prior to calculating the means, the DLOSCE scale scores for the 40 candidates in the sample were standardized to have a mean of zero and a standard deviation of one.[3]

Table 13.4
Descriptive Statistics for Standardized DLOSCE Scale Scores, by Group
(N=40)

Academic Dean's Rating of the Candidate's Final Clinical Performance in Dental School†	Candidates	Mean of Standardized DLOSCE Scale Scores	Standard Deviation of Standardized DLOSCE Scale Scores
Top 20% of Class	8	0.59	0.90
Middle 60% of Class	24	0.03	0.95
Bottom 20% of Class	8	-0.69	0.92

†Each candidate was classified into a performance category by their senior associate dean for academic affairs.

On average, candidates rated as being in the top 20% of their class in terms of their final clinical performance scored .59 standard deviations above the mean on the DLOSCE, while candidates rated as being in the bottom 20% of their class scored .69 standard deviations below the mean on the DLOSCE. This amounts to a difference of 1.28 standard deviations between the lowest and highest performing groups. Stated slightly differently, the preceding indicates that candidates demonstrating the strongest clinical performance in dental school achieved

---

[2]The student performance ratings were collected after the pass/fail DLOSCE results had been released to the dental education program that participated in the study. However, the senior associate dean did not have access to candidates' numeric DLOSCE scores at any time.

[3]A data sharing agreement was established with the dental school that participated in the validity study. The agreement stipulated that the study results reported in this technical report would not contain specific information about the school's performance on the examination. The DLOSCE scale scores were standardized as a means of abiding by the terms of the agreement. The standardized scores can only be interpreted normatively, and therefore provide no information about the participating school's examination performance relative to the DLOSCE performance standard.

DLOSCE scores that were on average 1.28 standard deviation units higher than those demonstrating the weakest clinical performance in dental school. This is a sizable performance difference on the DLOSCE, and is interpreted as representing a very large effect size by commonly applied interpretive guidelines within the field of psychological measurement. The polyserial correlation between the clinical performance ratings and the DLOSCE scale scores was .47 ($p < .05$), which is also understood as representing a strong effect size.[4]

Overall, the results presented in Tables 13.3 through 13.4 indicate that performance on the DLOSCE is positively correlated with clinical performance in dental school. Again, these findings provide strong convergent validity evidence in support of the intended interpretation and use of DLOSCE results. Notably, the correlations observed here are substantially stronger than those observed in similar convergent validity studies involving clinical licensure examinations that include patient-based and manikin components (Chambers, 2011; Formicola et al., 1998; Gadbury-Amyot et al., 2014; Hangorsky, 1981; Ranney et al., 2004).

14. Test Security

General Principles

Effective examination security procedures are critical to the success of any examination program. Responsibilities for examination security are clearly defined for test developers, test administrators, and examination users. Examination security is maintained through test development and test administration procedures in a variety of ways. DTS policies address issues related to examination security and are reviewed periodically by the Joint Commission and its staff.

Security Audit

In 2008, Caveon Test Security, an independent organization, conducted a security audit of DTS. This audit was conducted to identify potential security risks, propose specific measures to ameliorate or diminish any potential risks, and provide recommendations to support security planning. The findings of the audit supported the overall security measures implemented within DTS.

Identification of Secure Materials

The *Standards* highlight the importance of maintaining appropriate data security, including protections for candidate score information and sensitive ancillary information (AERA, APA, NCME, 2014, p. 121).

---

[4]As noted in discussions involving Table 13.3, an outlier was present in the analytic sample that exerted a strong influence on the correlation between LOC Sum Scores and DLOSCE scale scores. While this particular candidate was an outlier in the LOC analysis, this candidate did not have a large impact on results of the analysis involving faculty ratings of candidate performance. For completeness, the analysis involving faculty ratings was also performed after removing this individual from the data set. After removing the individual, the mean difference between DLOSCE scale scores for the Top 20% and Bottom 20% clinical performance groups increased from 1.28 to 1.34 standard deviations, and the polyserial correlation between clinical performance ratings and DLOSCE scale scores increased from .47 to .49.

Standard 6.16
Transmission of individually identified test scores to authorized individuals or institutions should be done in a manner that protects the confidential nature of the scores and pertinent ancillary information (AERA, APA, NCME, 2014, p. 121).

Accordingly, the Joint Commission has identified certain materials as secure. These include the following:

1. individual items and item materials (e.g. radiographs, clinical photographs);
2. scoring materials (e.g., item analyses, answer keys, and statistical analyses);
3. computer scoring software;
4. standard setting materials and meeting notes;
5. item banks; and
6. candidate personal information.

Departmental Procedures

The Joint Commission and DTS have a number of procedures in place that are designed to increase examination security. Relevant procedures are described in the section below.

Policies and legal issues. All items and examinations are copyrighted to establish ownership and restrict their use or dissemination through unauthorized means. Policies and procedures for handling secure materials require continuous secure custody of materials and a chain of evidence attesting to the status and location of secure materials.

Personnel. The team that maintains the security of examination materials includes Joint Commission staff, vendors, and volunteers. Personnel who handle examination materials must be screened at the time of hire or selection for assignment to disqualify individuals who could represent an unacceptable risk. All staff members are trained in procedures for handling secure materials and are required to comply with policies on confidentiality and conflict of interest. The examination development staff maintain security on examination materials during the development process.

All vendors are responsible for maintaining security of examination materials. Joint Commission staff review vendors' operations to ensure compliance with security policy. Service agreements with vendors must reasonably adhere to the Joint Commission's security procedures.

Volunteers who assist in the development of items and editions of the examination must complete agreements regarding confidentiality, copyright assignment, and conflicts of interest. Volunteers are prohibited from releasing information about examination content.

Facilities. Access to the offices of the Joint Commission is restricted and secure.

Security of Test Materials in Electronic Format. Departmental and vendor computers are protected with firewalls, login identifications, passwords, and other current forms of security. Access to electronic files is limited to authorized individuals.

Testing Procedures. Examinations are administered by Prometric at its nationwide testing centers, unless additional test facilities are authorized by the JCNDE. The DLOSCE Candidate Guide describe procedures for identification of candidates, including requirements for positive

identification through biometrics. Candidate conduct is closely monitored during the testing appointment. Examination regulations and testing center policies are designed to deter policy violations and prevent security breaches.

Policies and Procedures for Dealing with Breaches in Security

The Joint Commission provides specific procedures for observing and reporting security breaches and communicates them to test administrators. The Joint Commission promptly investigates reports of security breaches and responds appropriately given the nature and severity of the breach. When the source of a security breach is identified, the Joint Commission takes legal action or imposes appropriate sanctions.

15. Rights and Responsibilities of Test Takers

Documentation Provided to Candidates

The *Standards* indicate that information about an examination should be provided to all test takers, free of charge and in accessible formats (AERA, APA, NCME, 2014, p. 133-134).

Standard 8.1
Information about test content and purposes that is available to any test taker prior to testing should be available to all test takers. Shared information should be available free of charge and in accessible formats (AERA, APA, NCME, 2014, p. 133).

Standard 8.2
Test takers should be provided in advance with as much information about the test, the testing process, the intended test use, test scoring criteria, testing policy, availability of accommodations, and confidentiality protection as is consistent with obtaining valid responses and making appropriate interpretations of test scores (AERA, APA, NCME, 2014, p. 134).

Accordingly, the Joint Commission annually publishes the DLOSCE Guide. This document provides detailed information related to Joint Commission policy, rules and conduct, the format and content of the examination, eligibility requirements, examination regulations, the appeal process, examination scoring, and examples of item formats. The JCNDE also makes publicly available a set of DLOSCE practice questions that is provided free of charge. Each year the DLOSCE Guide is updated and amended as necessary. The guide and DLOSCE practice questions are available through the Joint Commission's website atjcnde.ada.org. This technical report also serves as a source of documentation that is freely available to all DLOSCE candidates through the JCNDE website.

Fair Treatment and Recourse

According to the *Standards*, candidates are entitled to fair treatment. This includes the right to information regarding available means of recourse pertaining to irregularities and appeals (AERA, APA, NCME, 2014, p. 137).

Standard 8.11
In educational and credentialing testing programs, when it is deemed necessary to cancel or withhold a test taker's score because of possible testing irregularities, including suspected misconduct, the type of evidence and the general procedures to be used to investigate the irregularity should be explained to all test takers whose scores are directly affected by the decision. Test takers should be given a timely opportunity to provide evidence that the score should not be canceled or withheld. Evidence considered in deciding on the final action should be made available to the test taker on request (AERA, APA, NCME, 2014, p. 137).

Standard 8.12
In educational and credentialing testing programs, a test taker is entitled to fair treatment and a reasonable resolution process, appropriate to the particular circumstances, regarding charges associated with testing irregularities, or challenges issued by the test taker regarding accuracies of the scoring or scoring key. Test takers are entitled to be informed of any available means of recourse (AERA, APA, NCME, 2014, p. 137).

For the DLOSCE, candidates whose results are subject to being voided are notified by written correspondence and provided with a copy of the Limited Right of Appeal for Examination Candidates. Candidates are notified of the appeal decision approximately 60 days after receipt of the appeal. When considering an appeal, the JCNDE strives to ensure that examination results accurately reflect candidates' skills, and that the appealing candidate has an opportunity to gain DLOSCE certification equal to, but not greater than, the opportunity provided to other candidates. The JCNDE strives to handle irregularities and their investigation in a professional, fair, objective, and confidential manner.

16. Candidate Performance

Table 16.1 provides DLOSCE administration volumes and failure rates, by candidate group and year.

Table 16.1
DLOSCE Administration Volumes and Failure Rates, by Candidate Group and Year

	Accredited*				Non-Accredited**				Total	
	First Attempt		Retake		First Attempt		Retake		First Attempts and Retakes	
Year	Number	% Failing	Number	% Failing	Number	% Failing	Number	% Failing	Number	% Failing
2020	231	9.5	2	0.0	14	57.1	1	100.0	248	12.5
2021	97	5.2	10	30.0	9	11.1	6	66.7	122	10.7
2022	52	21.2	5	40.0	6	0.0	2	50.0	65	21.5
2023	71	9.9	2	0.0	11	45.5	0	0.0	84	14.3

* Indicates candidates educated by dental education programs accredited by CODA.
** Indicates candidates educated by dental education programs not accredited by CODA. Failure rates for this group should be interpreted with caution due to the small sample size present.

References

American Dental Association, American Dental Education Association, American Student Dental Association. (September, 2018). *Report of the Task Force on Assessment of Readiness for Practice.*

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing.* Washington, DC: Author.

Angoff, W. H. (1971). Scales, norms and equivalent Scores, In R.L. Thorndike (Ed.), *Educational Measurement* (2nd ed., pp. 508-600). Washington, DC: American Council on Education.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 3*, 296-322.

Buckendahl, C. W., Smith, R. W., Impara, J. C., & Plake, B. S. (2006). A Comparison of Angoff and Bookmark standard setting methods. *Journal of Educational Measurement. 39*, 253-263.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chambers, D.W. (2011). Board-to-board consistency in dental licensure examinations. *Journal of Dental Education, 75*, 1310-1315.

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

Formicola, A.J., Lichtenthal, R., Schmidt, H.J., & Myers, R. (1998). Elevating clinical licensing examinations to professional testing standards. *New York State Dental Journal, 64*, 38-44.

Friedrichsen, S.W. (2016). Moving toward 21st-century clinical licensure examinations in dentistry. *Journal of Dental Education, 80,* 639-640.

Gadbury-Amyot, C.C., McCracken, M.S., Woldt, J.L., & Brennan, R.L. (2014). Validity and reliability of portfolio assessment of student competence in two dental school populations: A four-year study. *Journal of Dental Education, 78,* 657-667.

Gerrow, J.D., Murphy, H.J., Boyd, M.A., & Scott, D.A. (2003). Concurrent validity of written and OSCE components of the Canadian dental certification Examinations, *Journal of Dental Education, 67,* 896-901.

Gerrow, J.D., Murphy, H.J., Boyd, M.A., & Scott, D.A. (2006). An analysis of the contribution of a patient-based component to a clinical licensure examination. *Journal of the American Dental Association, 137,* 1434-1439.

Hangorsky, U. (1981). Clinical competency levels of fourth-year dental students as determined by board examiners and faculty members. *Journal of the American Dental Association, 102,* 35-37.

Horn, J. (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Impara, J. C., & Plake, B. S. (2005). Standard setting: An alternative approach. *Journal of Educational Measurement. 34*, 353-366.

Kaiser, H. E. (1960). The application of electronic computers to factor analysis. *Education & Psychological Measurement, 20*, 141-151.

Kane, M.T. (2016). Validation strategies: Delineating and validating proposed interpretations and uses of test scores. In L.S. Land, M.R. Raymond & T.M. Haladyna, *Handbook of Test Development* (pp. 64-80). Routledge.

Kane, M. T., Kingsbury C., Colton, D., Estes, C. (1989). Combining data on criticality and frequency in developing test plans for licensure and certification examination. *Journal of Educational Measurement 26*, 17-27.

Kolen, M. J. (1986). Standard errors of Tucker equating. *Applied Psychological Measurement, 9*, 209-223.

Kolen, M. J., & Brennan, R. J. (1995). *Test equating: Methods and practices*. New York: Springer.

Kuder, G.F., and Richardson, M.W. (1937). The theory of estimation of test reliability. *Psychometrika, 2*, 151-160.

Lewis, D. M., Mitzel, H. C., Mercado, R., Schulz, E. M. (2012) The Bookmark standard setting procedure. In G. J. Cizek (Ed.) *Setting performance standards: Foundations, methods, and innovations* (pp. 225-254). New York: Taylor and Francis.

National Dental Examining Board of Canada (2019). *Technical report: Objective structured clinical examination*. Ottawa, ON.

Ranney, R.R., Gunsolley, J.C., Miller, L.S., & Wood, M. (2004). The relationship between performance in dental school and performance on a clinical examination for licensure: A nine-year study. *Journal of the American Dental Association, 135,* 1146-1153.

Rudner, L. M. (2001). Computing the expected proportions of misclassified examinees. *Practical Assessment, Research & Evaluation, 7*.

Wilcox, R.R. (2017). *Introduction to Robust Estimation & Hypothesis Testing. 4th edition*. Elsevier, Amsterdam, The Netherlands.

Ziebert, A.J. & Waldschmidt, D.M. (2020). The Dental Licensure OSCE: A modern licensure examination for dentistry. *Journal of the California Dental Association, 48(7), 331-338.*

Appendix A

Clinical Content Areas for General Dentistry

#	Diagnosis and Treatment Planning
1	Interpret patient information and medical data to assess and manage patients.
2	Identify the chief complaint and understand the contributing factors.
3	Perform head and neck and intraoral examinations, interpreting and evaluating the clinical findings.
4	Use clinical and epidemiological data to diagnose and establish a prognosis for dental abnormalities and pathology.
5	Recognize the normal range of clinical findings and distinguish significant deviations that require monitoring, treatment, or management.
6	Predict the most likely diagnostic result given available patient information.
7	Interpret diagnostic results to inform understanding of the patient's condition.
8	Recognize the manifestations of systemic disease and how the disease and its management may affect the delivery of dental care.
9	Recognize the interrelationship between oral health and systemic disease, and implement strategies for improving overall health.
10	Select the diagnostic tools most likely to establish or confirm the diagnosis
11	Collect information from diverse sources (patient, guardian, patient records, allied staff, and other healthcare professionals) to make informed decisions.
12	Formulate a comprehensive diagnosis and treatment plan for patient management.
13	Discuss etiologies, treatment alternatives, and prognoses with patients so they are educated and can make informed decisions concerning the management of their care.
14	Understand how patient attributes (e.g., gender, age, race, ethnicity, and special needs), social background and values influence the provision of oral health care at all stages of life.
15	Interact and communicate with patients using psychological, social, and behavioral principles.

#	Oral Health Management
16	Prevent, recognize and manage medical emergencies (e.g., cardiac arrest).
17	Prevent, recognize and manage dental emergencies.
18	Recognize and manage acute pain, hemorrhage, trauma, and infection of the orofacial complex.
19	Prevent, diagnose and manage pain during treatment.
20	Prevent, diagnose and manage pulpal and periradicular diseases.
21	Prevent, diagnose and manage caries.
22	Prevent, diagnose and manage periodontal diseases.
23	Prevent, diagnose and manage oral mucosal and osseous diseases.
24	Recognize, manage and report patient abuse and neglect.
25	Recognize and manage substance abuse.
26	Select and administer or prescribe pharmacological agents in the treatment of dental patients.
27	Anticipate, prevent, and manage complications arising from the use of therapeutic and pharmacological agents in patient care.
28	Diagnose endodontic conditions and perform endodontic procedures.
29	Diagnose and manage the restorative needs of the partially or completely edentulous patient.
30	Restore tooth function, structure, and esthetics by replacing missing and defective tooth structure, while promoting soft and hard tissue health.
31	Perform prosthetic restorations (fixed or removable) and implant procedures for the edentulous and partially edentulous patient.
32	Diagnose and manage oral surgical treatment needs.
33	Perform oral surgical procedures.
34	Prevent, diagnose and manage developmental or acquired occlusal problems.

#	Oral Health Management
35	Prevent, diagnose and manage temporomandibular disorders.
36	Diagnose and manage patients requiring modification of oral tissues to optimize form, function and esthetics.
37	Evaluate outcomes of comprehensive dental care.
38	Manage the oral esthetic needs of patients.

#	Practice and Profession
39	Evaluate and integrate emerging trends in health care.
40	Evaluate social and economic trends and adapt to accommodate their impact on oral health care.
41	Evaluate scientific literature and integrate new knowledge and best research outcomes with patient values and other sources of information to make decisions about treatment.
42	Practice within the general dentist’s scope of competence and consult with or refer to professional colleagues when indicated.
43	Evaluate and utilize available and emerging resources (e.g., laboratory and clinical resources, information technology) to facilitate patient care, practice management, and professional development.
44	Conduct practice activities in a manner that manages risk and is consistent with jurisprudence and ethical requirements in dentistry and healthcare.
45	Recognize and respond to situations involving ethical and jurisprudence considerations.
46	Maintain patient records in accordance with jurisprudence and ethical requirements.
47	Conduct practice related business activities and financial operations in accordance with sound business practices and jurisprudence (e.g., OSHA and HIPAA).
48	Develop a catastrophe preparedness plan for the dental practice.
49	Manage, coordinate and supervise the activity of allied dental health personnel.
50	Assess one's personal level of skills and knowledge relative to dental practice.
51	Adhere to standard precautions for infection control for all clinical procedures.
52	Use prevention, intervention, and patient education strategies to maximize oral health.
53	Collaborate with dental team members and other health care professionals to promote health and manage disease in communities.
54	Evaluate and implement systems of oral health care management and delivery that will address the needs of patient populations served.
55	Apply quality assurance, assessment and improvement concepts to improve outcomes.
56	Communicate case design to laboratory technicians and evaluate the resultant restoration or prosthesis.

Appendix B

DLOSCE Test Specifications

Topic	Percent

Restorative Dentistry
- Diagnosis
- Preparations
- Restorations
  - Direct
  - Indirect
	24%

Prosthodontics
  - Removable
  - Fixed
  - Implants
	19%

Oral Pathology, Pain Management, and Temporomandibular Dysfunction - Oral pathology/Oral medicine - Orofacial pain and temporomandibular dysfunction	13%
Periodontics - Diagnosis - Treatment planning - Etiology	10%
Oral Surgery - Diagnosis - Treatment planning - Extractions	9%
Endodontics - Diagnosis - Treatment planning - Emergency management - Post-treatment evaluation	8%
Orthodontics - Treatment screening - Space management	6%
Medical Emergencies - Diagnosis - Management	6%
Prescriptions - Antibiotic - Analgesic	5%
	100%
Additional Notes: Diagnosis and Treatment Planning—as well as Occlusion—are covered across the topics listed above. The DLOSCE includes questions involving patients of various types and backgrounds, including pediatric, geriatric, special needs, and medically complex patients.	

Joint Commission on National Dental Examinations
Department of Testing Services
211 East Chicago Avenue
Chicago, Illinois 60611-2637