

Integrated National Board Dental Examination (INBDE)

Technical Report

Integrated National Board Dental Examination

Technical Report

2024

Executive Summary

The Technical Report for the Integrated National Board Dental Examinations (INBDE) is the main source of validity evidence available to state licensing boards and other users of INBDE results. Validity is the most important consideration for any examination program. For the INBDE, 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). Successful completion of a credentialing examination indicates candidates have achieved an acceptable level of performance in an area of knowledge. Some of the principal information presented in the Technical Report is summarized below.

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

Content: Content specifications for National Board Examinations (NBEs) are based on validity studies involving practice analyses conducted roughly every five years. Test constructors are responsible for recommending minor modifications during the interim period. 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 INBDE using Joint Commission guidelines for writing high-quality items.

Standard Setting and Scoring: The INBDE is criterion-referenced and not norm-referenced. This means examination results and pass/fail points 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 is ultimately determined by the JCNDE. The standards are 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 exam results on a common metric regardless of which examination form was administered.

Administration: The ADA maintains a high level of security on all examination materials. Strict precautions in place at testing centers help ensure test content remains secure. The Joint Commission offer the INBDE via computer at Prometric professional level testing centers throughout the United States, US territories, and Canada. 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, jcnde.ada.org.

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	5
2. INBDE Overview	6
3. Historical Perspective	6
4. Professional Test Standards	10
5. Overview of Validity	11
6. Content Basis for the Examination	12
7. Test Design and Development	18
8. Scoring and Equating Methods	25
9. Standard Setting	31
10. Score Reliability	32
11. Test Administration	35
13. Comparisons with NBDE Parts I and II	37
14. Test Security	39
15. Rights and Responsibilities of Test Takers	41
16. Trends in Candidate Performance	42
References	43
INBDE Clinical Content Areas	45
INBDE Foundation Knowledge Areas	49
INBDE Test Specifications	50

1. Introduction

Purpose of the INBDE 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 state 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 Joint Commission endeavors to provide the highest quality examination programs possible.

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 from Chapter 7 is provided below:

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 the INBDE validation efforts that have occurred to date. Some of the validity evidence presented in this technical report is based on data obtained through the INBDE field testing program – particularly the 2017-2018 INBDE Field Test conducted by the Joint Commission. The 2017-2018 Field Test involved a large-scale administration of the INBDE to dental students from across the United States (see Chapter 3 for details).

The Joint Commission on National Dental Examinations

The Joint Commission is the agency that oversees 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.

2. INBDE 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 INBDE program is to measure whether a candidate possesses the knowledge and cognitive skills required for the safe, independent practice of entry-level general dentistry. The intended examinee population for the INBDE 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 INBDE results concerns the candidate's ability to apply knowledge of biomedical, clinical, and behavioral sciences along with cognitive skills to understand and solve problems in clinical and professional contexts. A passing score on the INBDE indicates that a candidate is able to apply the aforementioned knowledge and skills at the level required for the safe, independent practice of entry-level general dentistry. INBDE 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 INBDE.

3. Historical Perspective

The National Board of Dental Examiners was established in 1928 as a standing committee of the ADA to provide and conduct written examinations to be used by state boards of dentistry for licensing dentists. These examinations were to provide a national standard for the biomedical and clinical dental sciences knowledge necessary for the competent practice of dentistry. The practical demonstrations of clinical skills were reserved for individual states to administer. The National Board's responsibilities included developing and administering National Board examinations and formulating the rules and regulations pertaining to them.

Current National Board Examinations bear little similarity to the first editions, which were administered in 1933 and 1934. Advances in examination methodology caused the most dramatic changes. The examination format was changed in the early 1950s from essay questions to multiple-choice questions. This led to the adoption of norm-referenced scoring procedures, and the National Board delegated examination construction to committees of dentists and dental hygienists who were subject-matter specialists. In the 1960s, the Council on National Board Examinations, which succeeded the National Board of Dental Examiners, was among the earliest testing agencies to employ computer scoring and use statistical techniques to identify candidates suspected of rule violations.

In the early 1980s, the JCNDE, which succeeded the Council on National Board Examinations, instituted the procedure of equating examinations by means of anchor items which would appear on more than one examination form. This was done to implement a consistent standard for minimally acceptable performance across examination forms, and it ended the era of norm-referenced scoring. The pass rate on the examinations thereafter fluctuated only to the degree

that candidates' abilities changed or the standard itself was changed. In 1992, a comprehensive, case-based NBDE Part II replaced the NBDE Part II battery consisting of seven individual examinations. Also, at that time, a criterion-referenced method of setting the performance standard based on Rasch psychometric theory was instituted for NBDE Part II. In 2007, a comprehensive NBDE Part I examination replaced the traditional battery of four individual examinations. In 2012, the Joint Commission moved to pass/fail reporting of results for candidates who passed the NBDE.

State boards of dentistry and candidates came to accept the NBDE over time. The first candidates completed National Board examinations in 1934. For the five-year period from 1934 through 1938, an average of 70 candidates per year received National Board certificates. By 1938, 11 states accepted National Board results. State board participation remained low until the mid-1950s. By 1960, 33 state boards and the District of Columbia accepted National Board results, and by 1976, 48 states plus the District of Columbia, Puerto Rico, and the Virgin Islands accepted them. By 1990, all U.S. licensing jurisdictions accepted the National Board Examinations as fulfillment of the written examination requirement for licensure.

INBDE Development Timeline

The decision to discontinue the Part I and Part II sequence and pursue development of an integrated examination emerged through careful deliberation and consideration of the needs of patients, the state of the profession of dentistry, and the context and environment within which it operated. In May 2005, the American Dental Education Association's (ADEA) Commission on Change and Innovation (CCI) examined dental education and recommended changes in curriculum design, instructional methods, assessment techniques and the evaluation of outcomes. In response, in 2007 the Chair of the Joint Commission recommended review of the structure and composition of National Board Dental Examination (NBDE) Test Construction Committees. The Joint Commission subsequently appointed an *ad hoc* committee for this purpose. The *ad hoc* committee's examination of Test Construction Committee structure led to fundamental questions about examination content. The Joint Commission reviewed the committee's report, decided more study was needed, and in 2008 appointed an *ad hoc* Committee on Strategic Planning. In considering the future of the Joint Commission, the Committee on Strategic Planning performed various environmental scans, developed a mission statement for the Joint Commission, and recommended merging NBDE Parts I and II. In alignment with the Mission Statement and as result of the environmental analysis, the Joint Commission adopted the following resolution:

“ . . . that the Joint Commission appoints a Committee for an Integrated Examination (CIE) to develop and validate a new examination instrument for dentistry that integrates basic, behavioral and clinical sciences to assess entry level competency in dental practice to supplant Part I and Part II.”

In 2009 the Joint Commission appointed the *ad hoc* Committee for an Integrated Examination (CIE) to develop and validate a new examination instrument for dentistry that integrates biomedical, behavioral and clinical sciences. The INBDE, as the examination came to be known, would be designed to assess entry level competency in dental practice, and would ultimately supplant National Board Dental Examination Parts I and II.

From 2009 through 2015, the CIE worked to establish preliminary policies and procedures for the INBDE program. This work included the establishment of the initial content domain for the examination, along with the preliminary test specifications (see Chapter 6), the development of preliminary operational recommendations, which outlined policies and procedures related to INBDE development and administration, the development of model INBDE items for use by INBDE test construction committees, and the completion of a number of psychometric research studies, including an investigation into potential scoring models for the examination. Each year, the CIE delivered a report to the Joint Commission, outlining its progress and its recommendations for the examination program. From 2015 through 2018, the JCNDE conducted a comprehensive field testing program in support of the INBDE. The field testing program had three major phases:

1. The 2015 INBDE Sample Item Survey, and the 2016 Follow-up Study
2. The 2016-2017 INBDE Short Form Field Test
3. The 2017-2018 INBDE Field Test

A brief summary of the purpose and results of each phase of the field testing program is provided in the section below. The results of the 2017-2018 Field Test are described in greater detail throughout the various sections of this technical report.

2015 INBDE Sample Item Survey. The purpose of the INBDE sample item survey was twofold: 1) to understand how dental students apply knowledge of the biomedical, clinical, and behavioral sciences in responding to INBDE items, and 2) to collect feedback from dental students regarding the presentation of examination content. Three forms of the sample item survey were developed, each containing five items that were selected to be broadly representative of the Foundation Knowledge areas and clinical content areas covered on the INBDE. Each form was administered online to candidates who attempted or were scheduled to attempt the NBDE Part II on or after July 1, 2015. Participation was voluntary, and eligible participants received a survey link in their NBDE Part II candidate eligibility notification email. A total of 170 candidates responded to the survey; 119 (70%) completed the entire survey; the remaining 51 (30%) completed only part of the survey. The overall response rate was 3.8%.

Staff collected and analyzed participant responses to the sample items, along with their feedback regarding item content and design. Overall, the students' feedback on the sample items was very positive. Students indicated that – in responding to the sample items – they were able to apply their knowledge and clinical experiences in a practical manner, and that the sample items were straightforward, fair and clinically relevant.

Due to the low response rate in the initial study, the CIE conducted a follow-up study involving a simplified, selected-response version of the sample item survey. The simplified survey was made available to all NBDE Part II candidates beginning in December 2015. Participation was voluntary, and eligible participants received a survey link in their NBDE Part II candidate eligibility notification email. The survey was closed in early February 2016. A total of 31 candidates completed all or part of the survey. None of the participants in the follow-up study had also completed the original sample item survey. Feedback obtained from the simplified survey was positive and mirrored the feedback obtained from the original study. The majority of participants had a positive overall impression of the sample INBDE items.

Overall, the results of the two sample item surveys indicated no major adjustments were needed to the INBDE item development process in place at the time, and no major changes should be

immediately implemented to the INBDE format or item writing approach. In light of the positive results, the Joint Commission proceeded to the next stage of the field testing program: The INBDE Short Form Field Test.

2016-2017 INBDE Short Form Field Test. The purpose of the INBDE Short Form Field Test was to provide an empirical evaluation of item development, test administration, and scoring methods for an abbreviated version of the INBDE, referred to as the INBDE Short Form. The Short Form Field Test took place from October 2016 through January 2017. A total of 840 candidates participated. All participants were eligible NBDE Part II candidates enrolled in accredited dental schools. Participants were randomly assigned to complete one of two 120-item INBDE forms. Each form met the preliminary INBDE test specifications established in 2011 (see Chapter 6). Approximately 40 percent of the items on each form were part of cases. Participants were removed from the final analytic sample if there was clear evidence that they gave little effort during the examination. The final sample comprised 704 participants from 59 dental schools. The candidates in the final sample were largely representative of the NBDE population with respect to performance on the NBDE Parts I and II.

Overall, the results of the Short Form Field Test were positive. Key findings were as follows:

1. Participant scores on the 120-item INBDE forms were positively correlated with scores on NBDE Parts I and II. The disattenuated correlation between the INBDE and NBDE Part I scores ranged across the two examination forms from .48 to .62. The disattenuated correlation between the INBDE and NBDE Part II scores ranged from .74 to .75.
2. Participant scores on the 120-item INBDE forms showed reliabilities in the expected range. Projections based on the Spearman-Brown prophecy formula indicated that an INBDE form containing 400 scored items would provide highly reliable scores.
3. Almost 90 percent of participants indicated that they were satisfied with their experience on the Field Test.

Due to the positive findings from the Short Form Field Test, the Joint Commission decided to move forward with implementation of the 2017-2018 INBDE Field Test.

2017-2018 INBDE Field Test. The purpose of the 2017-2018 INBDE Field Test was to examine the psychometric properties of the INBDE, and identify any limitations that needed to be addressed before the examination could be made available to dental boards for use in licensure decision making. The field test took place from November 2017 through January 2018. Participating students were administered a 300-item version of the INBDE. Students received between \$300 and \$500 for participating, depending on how they performed. The Field Test was administered at secure Prometric testing centers in the US and Canada.

The target population for the field test involved individuals from CODA-accredited dental schools who were in the dental class of 2018. This population consisted of nearly 6,000 candidates from more than 60 dental schools in the US and Canada. Staff invited 5,006 candidates from the target population to take part in the Field Test. Of those, 1,304 (26%) participated. Of the 1,304 participants who completed the Field Test, 124 (9.5%) were removed from the sample because they appeared to show little effort during the examination. The remaining 1,180 participants comprised the final analytic sample for the study. Sixty-five dental schools were represented in the final sample. Schools differed with respect to their total number of students who were

eligible and their field test participation rates. Schools with very low participation rates tended to be slightly underrepresented, but not in a way that was likely to distort the overall study findings.

As a whole, the results of the 2017-2018 INBDE Field Test were positive and provided strong validity evidence in support of the INBDE. Key findings were as follows:

- 1) Performance on the field test was highly correlated with performance on NBDE Parts I and II. The disattenuated correlation with NBDE Part I was .65, and the disattenuated correlation with NBDE Part II was .84. Both values were slightly higher than the correlations observed for the 2016-2017 INBDE Short Form Field Test.
- 2) Data from the field test suggested that an INBDE form containing an increased number of scored items would provide high rates of classification accuracy and classification consistency.
- 3) Participants indicated that the INBDE items on the field test form were clinically relevant and integrated: 92.7% agreed with the statement “*The questions on this test were clinically relevant,*” and 94.5% agreed with the statement “*The questions required me to integrate my biomedical and discipline-based knowledge with my clinical training.*”

Based on the positive findings from the INBDE field testing program, the Joint Commission determined in 2018 that the CIE had successfully fulfilled its charge, and that all remaining INBDE development work should be transitioned to the JCNDE. The JCNDE released formal notice of INBDE implementation in July 2019. The INBDE was made available for administration 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 INBDE results concerns the candidate’s ability to apply knowledge of biomedical, clinical, and behavioral sciences along with cognitive skills to understand and solve problems in clinical and professional contexts. A passing score on the INBDE indicates that a candidate is able to do so at the level required for the safe, independent practice of entry-level general dentistry. INBDE results are used by state 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 INBDE. This technical report presents validity evidence and additional references that support the intended interpretation and use of INBDE results, as suggested in 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. Score Reliability
11. Test Administration
12. Results Reporting and Interpretation
13. Comparisons with NBDE Parts I and II
14. Test Security
15. Rights and Responsibilities of Test Takers

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 validity evidence. This document organizes, describes, and presents a large amount of

validity evidence. In the process, it provides 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 INBDE 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 INBDE content domain and describes the theoretical rationale and empirical evidence that support it.

The Domain of Dentistry

The INBDE is designed to evaluate dental candidate cognitive skills based on the JCNDE's Domain of Dentistry. The Domain of Dentistry represents the clinical content areas and Foundation Knowledge Areas required for the safe, independent, general practice of dentistry by entry-level practitioners. It contains 56 clinical content areas and 10 Foundation Knowledge Areas. The clinical content areas are grouped into three component sections: 1) Diagnosis and Treatment Planning; 2) Oral Health Management; and 3) Practice and Profession. The 56 clinical content areas and ten Foundation Knowledge Areas that comprise the Domain of Dentistry are presented in Appendices A and B, respectively.

The Domain of Dentistry is structured as a matrix of clinical content areas (rows) and Foundation Knowledge Areas (columns). Each cell of the matrix represents the contribution of a given Foundation Knowledge Area to the performance of a particular clinical content area. Each cell is weighted according to its importance to entry-level dentistry, based on findings from the Joint Commission's most recent comprehensive practice analysis. This weight helps determine the number of examination items allocated to each cell.

The INBDE test specifications are provided in Appendix C. The specifications provide the overall percentage of examination items that are allocated to each clinical component section and each Foundation Knowledge Area, as well as the percentage of items allocated to the intersection of each clinical component section and Foundation Knowledge Area. Each form of the examination is built to meet the INBDE test specifications.

The current Domain of Dentistry was developed over a period of nearly ten years. A summary of the key development activities is provided in Table 6.1.

Table 6.1
Development of the Domain of Dentistry

Year	Activity
2010	The JCNDE developed 65 clinical content areas, describing the tasks and activities required for safe, independent, general practice of dentistry by entry-level practitioners.
2011	The JCNDE developed and administered a practice analysis survey to gather feedback on the 65 clinical content areas from a nationally representative sample of active, full-time dentists who had been in practice for five years or less.
	The JCNDE developed the 10 Foundation Knowledge Areas which describe the biomedical and clinical science knowledge that required for successful performance in the 65 clinical content areas.
	JCNDE developed an initial draft of the Domain of Dentistry, by integrating the 65 clinical content areas and the 10 Foundation Knowledge areas.
	The CIE convened two separate panels of subject matter experts to review the Domain of Dentistry. Panelists rated the strength of the perceived relationship between each Foundation Knowledge Area and each clinical content area. The results of this rating activity provided support for the appropriateness of the draft model.
2012	The JCNDE developed preliminary test specifications for the INBDE based on the results of the practice analysis survey, and the findings of the subject matter expert review panels.
2015	The JCNDE scrutinized the 65 clinical content area statements and revised them to be better suited for practice analysis purposes. From this review process emerged an updated set of 56 clinical content areas.
2016	The JCNDE developed and administered a practice analysis survey to gather feedback on the 56 clinical content areas from active, full-time dentists who had been in practice for ten years or less.
2017	The JCNDE convened a panel of subject matter experts to review the results of the 2016 practice analysis and recommend updates to the preliminary INBDE test specifications.
	The JCNDE approved the updates to the INBDE test specifications recommended by the aforementioned subject matter expert panel.

The section below provides a more detailed description of the process the JCNDE employed to develop the Domain of Dentistry and the INBDE test specifications.

Establishing the Domain of Dentistry

In 2010, the JCNDE developed 65 statements describing the tasks and activities required for the safe, independent, general practice of dentistry by entry level practitioners. These clinical content areas were adapted from competencies and standards contained in the following frameworks:

- The 2008 American Dental Education Association (ADEA) Competencies for the New General Dentist
- The 2007 Commission on Dental Accreditation's (CODA) Accreditation Standards for Dental Education Programs

The 65 clinical content areas were organized into the following three component sections: 1) Diagnosis and Treatment Planning, 2) Oral Health Management, and 3) Practice and Profession.

In 2011, the JCNDE administered a practice analysis survey to gather feedback on the 65 clinical content areas from a nationally representative sample of active, full-time dentists who had been in practice for five years or less. Over 700 respondents rated the frequency with which each clinical content area was utilized in entry level practice, as well as the criticality of each area to patient care. A statistical model was used to derive an overall "importance" rating for each content area, based on the content area's frequency and criticality ratings. The survey results provided support for the appropriateness of the clinical content areas. All 65 areas were judged on average to be at least "moderately important to patient care."

Later in 2011, the JCNDE identified ten areas of biomedical and clinical science knowledge that were required for successful performance in the clinical content areas. In developing these ten Foundation Knowledge Areas, the JCNDE relied on the following frameworks from dentistry and medicine:

- *American Dental Education Association (ADEA) Foundation Knowledge and Skills for the New General Dentist (as approved by the 2011 ADEA House of Delegates)*
- *Scientific Foundations for Future Physicians*, a 2009 report of the Association of American Medical Colleges (AAMC) and Howard Hughes Medical Institute (HHMI)

The JCNDE subsequently drafted a document entitled, *Foundation Knowledge for the General Dentist*, to accompany the ten Foundation Knowledge Areas. The document provides descriptions of each Foundation Knowledge Area, including subcategories, annotations and examples.

In 2011, the JCNDE established an initial draft of the Domain of Dentistry by integrating the 65 clinical content areas with the ten Foundation Knowledge Areas. The Domain presented the tasks and activities performed by general dentists, as well as the knowledge areas required to perform those tasks and activities. Subsequently, the CIE convened two panels of subject matter experts to permit empirical evaluation of the relationship between the clinical content and Foundation Knowledge Areas within the Domain of Dentistry. The first panel was composed of seven practicing dentists with strong science backgrounds who represented the major geographic regions of the U.S. The second panel was composed of nine male and nine female dentists in practice for five years or less who represented the ADA's 17 trustee districts. Panelists rated the strength of the perceived relationship between each Foundation Knowledge Area and each clinical content area. The results of this rating activity provided support for the appropriateness of the draft model; the ten Foundation Knowledge Areas were perceived as being related to performance on one or more of the 65 clinical content areas. Subsequently, the JCNDE posted the draft model of the Domain of Dentistry and the Foundation Knowledge for the General Dentist to the Joint Commission's INBDE website to solicit feedback from communities of interest. The feedback period was roughly six months in duration. The JCNDE reviewed comments received, and incorporated them into the two documents as appropriate.

Establishing Preliminary INBDE Test Specifications

In 2012, the JCNDE developed preliminary test specifications for the INBDE, based on findings from the aforementioned practice analysis and science review panels. The specifications were hypothetical in nature, outlining the number and percentage of items devoted to the 65 clinical content areas and ten Foundation Knowledge Areas under two plausible test lengths: 300 items and 450 items. The test specifications were derived using a methodology JCNDE staff created specifically for the INBDE. The importance ratings from the practice analysis were used to determine the number of items assigned to each clinical content area on the examination. Within a given content area, items were allocated to Foundation Knowledge Areas based on the perceived strength of the relationship between each Foundation Knowledge Area and the content area as identified through the science review panels. A paper describing the methodology was presented at the annual conference of the National Council on Measurement in Education (NCME) (Waldschmidt, Yang, and Christensen, 2013).

Refining the Domain of Dentistry

In 2015, the JCNDE reviewed the 65 clinical content areas comprising the model of the Domain of Dentistry. The JCNDE scrutinized and revised the statements so they were better suited for practice analysis purposes. From this review process emerged an updated set of 56 clinical content areas that represent the tasks and activities required for the safe, entry-level practice of dentistry. Among these clinical content areas, 13 involved Diagnosis and Treatment Planning, 23 involved Oral Health Management, and 20 involved Practice and Profession. Shortly thereafter, the JCNDE conducted a survey to collect feedback from practicing dentists on the relevance and comprehensiveness of the revised set of 56 clinical content areas. A total of 166 respondents completed the survey. Respondents consisted of practitioners from across the US who had various levels of practice experience. Ninety-three percent of the respondents either agreed or strongly agreed that the 56 clinical content areas were comprehensive in covering the tasks and activities that underlie dental practice.

In 2016, the JCNDE administered a practice analysis survey to gather feedback on the 56 clinical content areas from a sample of active, full-time dentists who had been in practice for ten years or less. The first section of the survey gathered information about the responding dentist and his or her practice environment. The second section consisted of a list of the 56 clinical content areas. 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 Joint Commission 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 ranged across clinical content areas 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.

Refining the INBDE Test Specifications

In 2017, the JCNDE convened a panel of subject matter experts to review the results of the 2016 practice analysis and recommend updates to the preliminary INBDE test specifications. The panel was made up of ten dentists, including three members of the Joint Commission and three members of the CIE. In establishing their recommendation, the panelists completed the following activities:

1. The panelists completed a linking exercise in which they individually rated the strength of the relationship between each clinical content area and Foundational Knowledge area. The following rating scale was used:
 4. Very strong relationship
 3. Strong relationship
 2. Moderate relationship
 1. Slight relationship
 0. No relationship
2. The panelists then reviewed the results of the 2016 practice analysis, which provided a starting place for the number of items that should be allocated to each clinical content area. Panelists then scrutinized the clinical content areas, identified clinical content areas that overlapped, and made adjustments to the number of items per area to account for the overlap. This established the recommended number of items per clinical content area.
3. The panelists were asked to think at an overall conceptual level regarding the percentage of items on an entry-level dental licensure examination that would be appropriate to assign to each Foundation Knowledge Area, based on their general knowledge of dentistry, the size of the area, the depth of knowledge required within each area, and the importance of each area to patient care. Each panelist then made a corresponding initial recommendation regarding the percentage of items they felt should be allotted to each Foundation Knowledge Area. These recommendations were then averaged to represent a group conceptual recommendation.
4. Ratings from the linking exercise were averaged across panelists and used to allocate the items in each clinical content area across the ten Foundation Knowledge Areas. The above thus yielded two separate recommendations regarding the overall percentage of

items to allocate to Foundation Knowledge Areas; one based on a conceptual analysis, and the other based on the linking exercise. Members reviewed the total number of items allocated to each Foundation Knowledge Area, as determined by the linking exercise, and compared these totals to their initial conceptual recommendations. Additionally, members were made aware of the Foundation Knowledge Area allocations associated with the preliminary INBDE test specifications. In making decisions for 2017 and beyond, members felt that both the conceptual and linking estimates provided useful information. The linking exercise required careful analysis and consideration of the relationship between each Foundation Knowledge Area and each clinical content area, while the conceptual approach was more global but perhaps more broadly sensitive to the size and complexity of each area of knowledge. In recognition of the contributions of each approach, members felt that it would be appropriate to weight results of the linking exercise by 75 percent and the conceptual exercise by 25 percent. Members also worked as a group to identify overlap among the Foundation Knowledge Areas, and made adjustments to the number of items per area to account for the overlap. The preceding process resulted in the establishment of a recommended overall number of items to be allocated to each of the ten Foundation Knowledge Areas.

5. Having established the overall number of items per clinical content area and per Foundation Knowledge Area, the panelists worked via an iterative process to determine the number of items that should be allocated to each cell of the 56 clinical content area by ten Foundation Knowledge Area Domain of Dentistry. The process ended when the total number of items for each clinical content area and Foundation Knowledge Area was within one item of the recommended number determined in the previous steps, and the overall number of items allocated was 500. This established the final recommendation for the INBDE test specifications.

In 2017, the JCNDE reviewed and approved the INBDE test specifications recommended by the subject matter expert panel described above. The specifications are provided in Appendix C. Each INBDE form is built to meet the specifications.

The JCNDE conducts comprehensive practice analyses on a periodic basis, and uses practice analysis results in combination with subject matter expert judgments to ensure that the INBDE test specifications continue to reflect dental practice. In the time period between practice analyses, INBDE test constructors evaluate the specifications and – accompanied by appropriate justification – recommend minor changes as needed, for consideration by the JCNDE.

Clinical Relevance and Integration

Clinical Relevance and Integration are two key concepts that inform the development of items within the INBDE Domain of Dentistry. The Joint Commission defines these concepts as follows:

Clinical Relevance refers to factors that impact patient outcomes in clinical and professional contexts. This includes all aspects of patient care and also encompasses considerations involving how dentists approach the practice of dentistry (Practice Relevance), and keep up with advances that impact the profession (Professional Relevance). Broadly speaking, Clinical Relevance involves the actual experiences of entry-level, general dentists, practicing independently, as they work to improve patient

outcomes. Clinical relevance is maximized in the INBDE when there is a strong degree of fidelity between the content of examination items, the knowledge and cognitive skills required to answer those items, and the actual experiences of entry-level, general dentists.

Integration brings to bear knowledge of biomedical, clinical, and behavioral sciences along with cognitive skills to understand and solve problems in clinical and professional contexts.

As mentioned previously, the purpose of the INBDE program is to measure whether a candidate possesses the knowledge and cognitive skills required for the safe, independent practice of entry-level general dentistry. The JCNDE views the development of clinically relevant examination content as a strong means of fulfilling that purpose. Integration is seen as a pathway to achieving clinical relevance. The results of the 2017-2018 INBDE Field Test suggest that candidates view INBDE items as both clinically relevant and integrated. More than 92 percent of the Field Test participants indicated that the INBDE items on the Field Test were clinically relevant, and more than 94 percent indicated that the items required them to integrate their biomedical and discipline-based knowledge with their clinical training.

7. Test Design and Development

The overall design of an examination is a crucial step in test development. The INBDE is designed with the full participation of content expert teams and supervised by staff specialists from the Joint Commission's 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 provide technical support and guidance to help ensure the desired technical qualities are achieved during the examination design phase. 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 (AERA, APA, NCME, 2014, p. 85).

Examination Format

The INBDE is a comprehensive examination consisting of 500 items. Sixty percent of the items on each examination form are standalone; the remaining 40 percent are part of cases. A case consists of two or more items that reference a common stimulus or scenario. There are two types of cases on the INBDE:

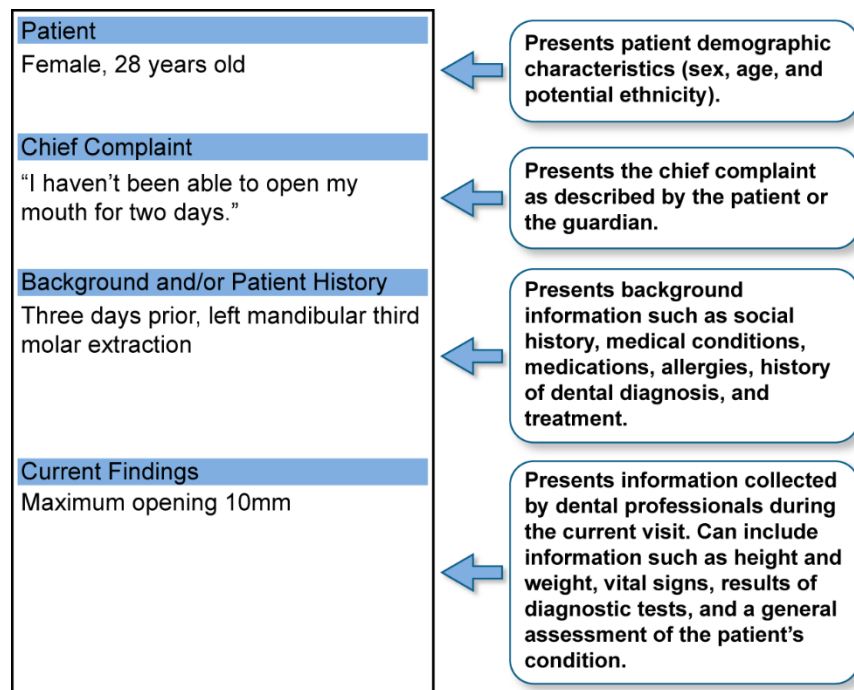
- Single-stimulus (image or patient box) cases
- Comprehensive, multiple-image cases, including but not limited to the Patient Box, dental charting, radiographs, and photographs.

Each case contains 2 to 5 items about various aspects of the patient's dental care. These items might derive from any of the examination's content areas, including Practice and

Profession. The proportion stemming from any particular clinical content area and Foundational Knowledge area depends upon the nature of the case itself. In responding to cases, the candidate must:

- interpret the findings and information provided;
- identify problems and make diagnoses;
- select materials, technique, and armamentarium;
- apply treatment;
- evaluate progress and complications; and
- establish procedures for prevention and maintenance.

Some INBDE items involve 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 several benefits associated with using the Patient Box format to present patient information. Specifically, the Patient Box

- permits the candidate to focus on the content of the item, as opposed to how items are worded;
- simplifies the item writing process for test constructors, allowing them to focus on the concept being tested;
- reduces construct-irrelevant variance by lowering the level of language skills required to understand the items; 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, development of the Patient Box was intended to maximize construct-relevant variance and minimize construct-irrelevant variance.

The Patient Box can be used for both standalone items and for cases. 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 INBDE Examination Guide) also includes information concerning the Patient Box. More than 91 percent of the participants in the 2017-2018 INBDE Field Test indicated that the Patient Box was an effective way to present patient information.

INBDE Test Constructors

The Joint Commission relies on subject matter experts – called test constructors – to develop and review INBDE items and forms. The role of test constructors is fundamental to the examination's validity argument. Test constructors are responsible for constructing a clear, precise, and cohesive group of items for each examination and for providing content-related validity evidence. 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). The section below presents this information, as it pertains to the INBDE.

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).

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

- Submit new items, according to JCNDE guidelines, within the designated time frame.
The number of new items 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, specifications, and content outlines, within the designated time frame.
- Construct additional items for the item banks when 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 the standard curricula, and suggest modifications to the test specifications as needed.
- Consider special issues and make recommendations at the request of the JCNDE.

- Safeguard the security and confidentiality of the examination by declining any arrangement to assist with review courses or reviewing books pertaining to the examination while serving as a test constructor and for at least one year following the final term of appointment.
- Comply with the ADA's policy on professional conduct. The policy includes prohibitions against sexual harassment and other forms of unlawful conduct.

The INBDE Item Development Guide, and the INBDE test specifications provide basic information to new test constructors. New test constructors also attend an orientation session where they receive information about the INBDE program, and the item development and review process.

Teams and Requirements. INBDE test constructors work in teams, referred to as Test Construction Teams (TCTs), to develop INBDE items. Most TCTs meet multiple times per year, and most meetings are 2½ days in duration. The four main categories of INBDE 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. Each Item Writing and Review team consists of at least sixteen 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 accuracy, currency, validity, adherence to the test specifications, and item guidelines outlined by the Joint Commission. Each team is composed of experts associated with the Clinical Content Sections of the Joint Commission's Domain of Dentistry. The item development TCTs include the following:

- **Diagnosis and Treatment Planning (at least 5 TCs)**
 - This includes all aspects of diagnosis, assessment, interpretation of laboratory results, physical evaluation of structure, appearance and function, as well as background, characteristics, circumstances, underlying conditions and psychological, social or behavioral influences to derive an accurate diagnosis and treatment plan, as well as reasonable and safe treatment alternatives for any patient.
- **Oral Health Management Section (at least 6 TCs)**
 - This includes all aspects of patient care delivery, including provision and modification of treatment, handling of complications, intervention outcomes assessment, and all developments that arise in the course of managing a patient's oral health care.
- **Practice and Profession Section (at least 5 TCs)**
 - This includes the understanding, implementation and monitoring of effective systems, procedures and policies that support professional practice integrity, quality control and self or organizational performance assessment. It also includes the systems, procedures and policies that enable skillful integration of new knowledge, products and technology, and that minimize risk and enhance the quality of care provided or improve community oral health.

The following definition of **General Dentist** applies to INBDE TCs:

- a) Graduate of a dental program accredited by either the Commission on Dental Accreditation (CODA) or the Commission on Dental Accreditation of Canada (CDAC).
- b) Full-time or part-time practitioner or clinician/scientist with at least five years of experience in the United States or Canada.
- c) Currently licensed as a dentist in the United States or Canada.

Case Development Teams. The Case Development teams are responsible for reviewing images and case materials to identify clinically relevant scenarios and concepts that could serve as the basis for item development. These teams also write items to the identified scenarios and concepts. Each team consists of five or more General Dentists who have expertise across the three clinical component sections.

Clinical Relevance Review Teams. Each Clinical Relevance Review team consists of five or more General Dentists, each of whom must have attended at least one meeting in Item Writing and Review, or Case Development. The Clinical Relevance Review teams confirm the appropriateness of examination items in terms of clinical relevance. The teams are also responsible for final categorization of items relative to the Domain of Dentistry and test specifications, and in support of the general needs of the INBDE program.

Form Review Teams. Form Review teams provide a final review of INBDE 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 diverse, meets the test specifications, and that there is no unintended overlap among the items included on each form. These teams also determine the order of cases when appropriate. Each Form Review team consists of one or more General Dentists, one or more experts from Diagnosis and Treatment Planning, one or more experts from Oral Health Management, one or more experts from Practice and Profession, and one or more additional members of an existing team with expertise in areas determined by staff.

Test Constructor Selection. The Joint Commission annually advertises its need for test constructors. All applications are processed by staff and forwarded to the Joint Commission's Dental Test Constructor Recommendation Committee, which is responsible for recommending individuals for the INBDE Test Constructor Pool.

The JCNDE annually approves and reapproves test constructors into the INBDE 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 JCNDE 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. Individuals are invited to attend a given meeting. Should they accept, they are considered part of the team for that calendar 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 will always have a sufficient number of volunteers with the required expertise, so that meeting goals can be accomplished efficiently and effectively. Each test

constructor invited to attend an INBDE TCT meeting receives a copy of the INBDE Item Development Guide, along with the INBDE test specifications. Membership in the ADA is required for those in the clinical sciences.

Item Validation

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

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).

INBDE test constructors are expected to draft new items and submit them to DTS staff in advance of test construction meetings. Once new items are received by staff, they are entered into the JCNDE item bank so they can be reviewed during test construction meetings. Several types of item reviews take place during test construction meetings. Reviews may be conducted by either the original Item Writing and Review team, or other TCTs and/or individuals. The categories of review are described below.

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 item writing guidelines, as detailed in the INBDE Item Development Guide. Similarly, item stimulus materials are reviewed to ensure they are legible and in accordance with modern dental practice. Feedback and comments generated during the editorial review process are returned to the original Item Writing and Review team for consideration.

Fairness and sensitivity review. Items are reviewed based on fairness and sensitivity considerations, in alignment with the item writing guidelines. A fairness and sensitivity review takes place as part of the original item writing 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. After a successful pilot, this updated review process was approved for full and ongoing implementation across all JCNDE examinations in 2024. Through the updated review process, Fairness and Sensitivity Reviewers evaluate examination content to help ensure that test questions adhere to fairness considerations as outlined in the *Standards*. In pursuing this charge, Fairness and Sensitivity Reviewers facilitate the accurate measurement of candidate knowledge skills, and abilities, helping to ensure that examinations avoid sources of construct irrelevant variance that might interfere with accurate skill measurement. Fairness and Sensitivity Reviewers are asked

to focus specifically on the aforementioned 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.

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

Case material quality review. During this process, Case Development teams review case material submissions to determine if those materials are of sufficient quality to serve as the basis for item writing. The team makes a recommendation to accept or reject the materials, and then prepares accepted materials for use by Item Writing and Review teams. An editorial/graphics staff member is available to support this team. An intellectual property review is included as part of this team's responsibility.

Clinical relevance review. The INBDE is designed for licensure purposes, to help state boards understand whether a candidate possesses the necessary cognitive skills to enter the profession and safely practice dentistry. The general dentist is thus of focal importance to this examination program. During the clinical relevance review, Clinical Relevance Review teams review 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.

Item performance review. Items that survive the reviews described above are eligible to be placed on examination forms as unscored pretest items. The purpose of including pretest items on an examination is to gather data that can be used to evaluate the items' psychometric performance, in anticipation of using the item on a production examination. Pretest items are administered randomly throughout the examination and are presented in the same manner as scored items so they cannot be identified by candidates. After pretest items have been administered to a sufficient number of candidates, DTS calculates the following statistics for each item:

- 1) The p-value, defined as the proportion of candidates who answered the item correctly.
- 2) The point-biserial correlation, defined as the Pearson correlation between performance on the item and performance on the examination.

The p-value is an indicator of an item's difficulty, and the point-biserial correlation is an indicator of an item's ability to discriminate among candidates of different ability levels. If during pretesting, an item shows a p-value below .30 or above .95, or a point-biserial correlation below .03, it becomes a candidate for elimination or revision. In some cases, test constructors discuss these non-performing items and decide whether to retain, revise, or remove them. In other cases, the non-performing items are simply retired from the item bank and never administered again. When pursued, the revision process typically involves 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 as a pretest item. If test constructors determine that an item cannot be improved through revision, the item is designated as unusable. Items that

survive the validation process described above are eligible to become scored items on future forms of the examination.

Examination Form Assembly

INBDE forms are built according to strict content and statistical specifications. Each examination form is reviewed by test constructors before it is made available for administration. The steps in the INBDE form assembly process are as follows:

- 1) Staff determine the number of examination forms that will be created.
- 2) Staff review the INBDE test specifications, which indicate the number of scored items each form should contain from each area of the Domain of Dentistry.
- 3) Staff review the bank of items that are eligible to be included as scored items on an examination form. To be eligible, items must have survived the item content and performance review steps described in this chapter, and must have been placed onto the established INBDE scale of measurement via the equating process (see Chapter 8).
- 4) Staff use a computer program to select the scored items for each examination form. The computer program is designed to create examination forms that meet the INBDE test specifications and have similar psychometric properties (e.g., difficulty, reliability).
- 5) Staff review the bank of items that are eligible to be included as pretest items on an examination form. To be eligible, items must have survived the initial item content review steps described in Chapter 7.
- 6) Staff select the pretest items for each examination form.
- 7) A Form Review team consisting of subject matter experts reviews each examination form and replaces any items that should not be included on the form. The test constructors on the Form Review team do not make content edits to examination items during this step. Rather, they focus on the basic principles of form assembly (e.g., ensuring that items that evaluate identical concepts do not appear on the same form).
- 8) Staff conduct a final check of the forms to ensure that they meet the INBDE test specifications and have similar psychometric properties.

Once examination forms have been finalized, staff deliver them to the test administration vendor for publication. During the publication process, staff conduct a series of additional quality assurance reviews to ensure that items are delivered correctly during the examination. When the quality assurance reviews have been completed, the examination forms can be administered to INBDE candidates.

8. Scoring and Equating Methods

Scoring Methods

There are two common approaches to scoring licensure examinations. Under the first approach, referred to as *compensatory scoring*, 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 scoring approach, referred to as *conjunctive scoring*, an examination is divided into separately scored sections, and the candidate must pass each section in order to pass the examination. Conjunctive scoring is often used when the topic areas on an examination are 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, compensatory scoring 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 the 2017-2018 INBDE Field Test data indicates that candidate performance on the INBDE topic areas is highly correlated. Disattenuated correlations among scores on the INBDE's three clinical component sections ranged from .94 to 1.02.¹ Disattenuated correlations among scores on the INBDE's ten Foundation Knowledge Areas, ranged from .70 to 1.03. Additionally, exploratory factor analysis of candidate scores on the ten Foundation Knowledge Areas suggested they were indicators of a common underlying ability that could be well represented with a single score; the ratio of the first to second eigenvalue from the factor analysis was 5.4, 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 compensatory scoring approach for the INBDE, whereby a candidate's result (i.e., pass or fail) is based on a single score that is determined from their performance on the entire examination. The section below describes how the overall score and corresponding pass/fail result is determined for each INBDE candidate.

Scoring Model. The INBDE is scored using an item response theory (IRT) model called the three-parameter logistic (3PL) model. The 3PL model can be expressed mathematically as follows:

$$P(x_{ij} = 1|\theta_i, a_j, b_j, c_j) = c_j + (1 - c_j) \frac{e^{a_j(\theta_i - b_j)}}{1 + e^{a_j(\theta_i - b_j)}} \quad 8.1$$

where $P(x_{ij} = 1|\theta_i, a_j, b_j, c_j)$ is the probability that candidate i will respond correctly to item j , θ_i is the ability level of candidate i , a_j is the discrimination parameter for item j , b_j is the difficulty parameter for item j , c_j is the pseudo-guessing parameter for item j which represents the probability that the item will be answered correctly by a candidate with a very low level of ability. The value e is a mathematical constant that is approximately equal to 2.71828.

Under the 3PL model, a candidate's ability (θ) represents his or her standing on the construct measured by the examination. When the 3PL model is used, a candidate's ability estimate is not based simply on the number of items they answer correctly. Rather, the ability estimate is based on the candidate's entire item response pattern involving a set of items with known characteristics relative to the three parameters of interest. When the 3PL model shows adequate fit to the response data, it provides more accurate estimates of candidate abilities than simple number correct scores (Kolen & Brennan, 1995). A strength of the 3PL model is that it allows each test item to have a unique discrimination, difficulty, and pseudo-guessing parameter. In contrast, the two parameter IRT model permits items to have unique discrimination and difficulty parameters, while assuming that the pseudo-guessing parameter is zero for all items. Lastly the one parameter model permits items to have unique difficulty parameters, while assuming that every item is equally discriminating and that the pseudo-guessing parameter is zero for all items. Analysis of INBDE response data has consistently

¹A fully disattenuated correlation is an estimate of what the observed correlation between two variables would be if both were measured with perfect reliability (i.e., the true correlation between the variables). Disattenuated correlations are useful for determining if the relationship between two variables is actually weak, or if it just appears weak because it is being heavily attenuated by measurement error. Estimates of disattenuated correlations can exceed 1.0.

demonstrated that INBDE items differ considerably in terms of their ability to discriminate among candidates of different ability levels. Additionally, because INBDE items are multiple choice, even very low ability candidates have some chance of selecting a correct response. These considerations make the 3PL the preferred scoring model for the INBDE. Additional information about the model is available in de Ayala (2009).

The JCNDE uses Bilog software to estimate the parameters of the 3PL model (Zimowski, Muraki, Mislevy, & Bock, 2003). JCNDE psychometricians seek to estimate INBDE item parameters using samples of at least 300 candidates, per the recommendation of Chuah, Drasgow, and Luecht (2006). Item parameters (a , b , c) are estimated in a Bayesian framework using Bilog's default prior distributions. Candidate abilities are determined using maximum likelihood estimation. Candidate ability estimates are transformed into scale scores which range from 49 to 99 in one-point increments. A scale score of 75 represents the minimum ability level required for the safe, independent practice of entry-level general dentistry, as determined through standard setting activities (see Chapter 9). A candidate must achieve a scale score of 75 or higher in order to pass the INBDE.

The application of the 3PL model relies on two key assumptions about the data to which the model is applied: 1) unidimensionality, and 2) local independence (Hambleton et al, 1991). The unidimensionality assumption means that only one ability is measured by the items in the examination. Evidence of meeting the unidimensionality assumption involves empirical findings indicating the presence of a dominant factor influencing examination performance. This dominant factor reflects the ability measured by the examination. An example of a violation of the unidimensionality assumption would be a situation where performance on a written examination of mechanical skills is influenced by candidate language skills in addition to candidate mechanical skills.

The JCNDE uses a statistical technique called factor analysis to assess whether or not the unidimensionality assumption is feasible for INBDE data. As mentioned previously, an exploratory factor analysis of data from the 2017-2018 INBDE Field Test provided support for the idea that a single dominant factor underlies candidate performance on the INBDE's ten Foundation Knowledge Areas; the ratio of the first to second eigenvalue from the factor analysis was 5.4, 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. Exploratory factor analysis of item-level data from the Field Test using the program NOHARM (Fraser & McDonald, 2003) also indicated that no clear content-related multidimensionality was present. Accordingly, the JCNDE is confident that INBDE data are sufficiently unidimensional to permit the use of the 3PL model for determining candidate results. It is also worth noting that the results of several studies indicate that candidate scores derived from an IRT analysis are highly robust to moderate violations of the unidimensionality assumption (Dorans & Kingston, 1985; Drasgow & Parsons, 1983).

The local independence assumption means that the ability specified in the model is the only factor responsible for the relationships among candidate responses to test items. Per this assumption, candidate responses to any pair of items should be statistically independent, when the ability influencing examination performance is statistically held constant. A violation of the local independence assumption might occur for cases, where multiple items rely on a common stimulus. Items appearing within a case involving case materials may be more highly correlated with each other than with items appearing elsewhere in the examination, due to the common

stimulus materials. This suggests the presence of an additional factor influencing responses on an item, and the presence of this factor would constitute the violation of local independence.

The JCNDE has typically used Yen's (1984) Q_3 index to detect violations of the local independence assumption. The Q_3 index is the Pearson correlation between response residuals for a pair of test items. The response residual for candidate i on item j is calculated as follows:

$$r_{ij} = x_{ij} - P(x_{ij} = 1 | \theta_i, a_j, b_j, c_j) \quad 8.2$$

where x_{ij} is the observed response for candidate i on item j (either 1 or 0 in the case of dichotomously scored multiple choice items), and $P(x_{ij} = 1 | \theta_i, a_j, b_j, c_j)$ is the probability that candidate i will respond correctly to item j . The Q_3 index for items j_1 and j_2 is then the correlation between the residuals for items j_1 and j_2 , across all candidates. Per Yen's (1984) recommendation, the JCNDE flags item pairs as conditionally dependent if the absolute value of the Q_3 index exceeds 0.2. In these instances, the issue can be addressed by simply eliminating one of the conditionally dependent items from the examination form. Analysis of data from the 2017-2018 INBDE Field Test indicates that violations of the local independence assumption are rare for the INBDE. Of the hundreds of items pairs analyzed, less than one-tenth of 1 percent showed Q_3 indexes whose absolute values exceeded the cutoff value of 0.2.

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, 2015, p. 118).

Accordingly, the JCNDE has established strict quality control measures to facilitate accurate scoring of the INBDE. Each week the roster of candidates scheduled to complete the INBDE is compared with the candidates appearing in result files, to ensure no result files are missing. Each examination is independently scored by two analysts from DTS, and the resulting scores are compared against one another to ensure they are identical before results are released to candidates. DTS staff maintain extensive documentation related to the examination scoring process and its various quality assurance steps.

Equating Methods

Multiple forms of the INBDE are available for administration. The JCNDE takes care to ensure that all forms of the examination meet the INBDE test specifications and are similar in terms of their psychometric properties (e.g., difficulty, reliability). 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 account for these differences. The equating process ensures that all INBDE candidates are held to the same performance standard, regardless of which examination form they attempt. The *Standards* indicate that test developers should provide evidence for the claim that results from different forms of an examination may be used

interchangeably (AERA, APA, NCME, 2014, p. 105). The methods used to equate INBDE results across examination forms are described below.

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 an equating method that is based on IRT. The basic steps in the equating process are as follows. First, items from separate 3PL model calibrations are placed on a common scale of measurement using a common-item nonequivalent groups equating design. A description of this process is provided below. Next, multiple INBDE forms, which are similar in terms of content, difficulty and reliability, are assembled using the items that have been placed on the established scale. When a form is administered to a candidate, the candidate's ability (θ) is then estimated based on his or her responses to the already-calibrated items. Candidate abilities determined in this way are on the same scale of measurement and can be directly compared to one another, even though candidates may have attempted different forms of the examination.

In a common-item nonequivalent groups equating 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 is used to place the items from the two forms on the same scale. The items from the first examination form (Form X) are calibrated using data from one sample of candidates (Sample 1), and the items from the second examination form (Form Y) are calibrated using data from another sample of candidates (Sample 2). The two calibrations are conducted separately. Next, the slope (A) and intercept (B) of a linear equation are determined, that can be used to place item parameters from Form Y on the same scale as item parameters from Form X. The slope of the linear equation is calculated as follows:

$$A = \frac{s(b_X)}{s(b_Y)} \quad 8.3$$

where $s(b_X)$ is the standard deviation of the item difficulty parameter estimates for the anchor items calibrated using data from Sample 1, and $s(b_Y)$ is the standard deviation of the item difficulty parameter estimates for the anchor items calibrated using data from Sample 2. The intercept of the linear equation is then calculated as follows:

$$B = \bar{x}(b_X) - A\bar{x}(b_Y) \quad 8.4$$

where $\bar{x}(b_X)$ is the mean of the item difficulty parameter estimates for the anchor items calibrated using data from Sample 1, A is the slope determined via equation 8.3, and $\bar{x}(b_Y)$ is

the mean of the item difficulty parameter estimates for the anchor items calibrated using data from Sample 2.

Once the slope and intercept of the linear equation have been determined, they can be used to place item parameters from Form Y on the same scale as item parameters from Form X. The discrimination parameter for item j on Form Y can be transformed as follows:

$$a_{Xj} = \frac{a_{Yj}}{A} \quad 8.5$$

where a_{Xj} is the discrimination parameter for item j on the Form X scale, a_{Yj} is the discrimination parameter for item j on the Form Y scale, and A is the slope determined via equation 8.3. The difficulty parameter for item j on Form Y can be transformed as follows:

$$b_{Xj} = Ab_{Yj} + B \quad 8.6$$

where b_{Xj} is the difficulty parameter for item j on the Form X scale, b_{Yj} is the difficulty parameter for item j on the Form Y scale, and A is the slope determined via equation 8.3, and B is the intercept determined via equation 8.4. No transformation is required for the pseudo-guessing parameter (Kolen & Brennan, 1995).

When the above described method is applied to the INBDE, the items that comprise the anchor test 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). Furthermore, anchor test items that perform differently in the two samples are not used in the calculation of the slope (A) and intercept (B) of the linear transformation equation that is used to place item parameters onto a common scale (see Equations 8.3 and 8.4). The JCNDE uses the item parameter plots described in Kolen and Brennan (1995) to identify items that perform differently in the two samples so they can be excluded from the data set prior to the calculation of the slope and intercept.

The equating approach described above was used to place INBDE items pretested in 2019 onto the same scale of measurement as INBDE items pretested in 2018. This established an initial calibrated item bank for the INBDE. The first INBDE forms were assembled from this calibrated item bank. Because the items had been placed on the same scale of measurement, and the examination forms were built to the same detailed content specifications, candidate ability estimates derived from the administration of these forms were equated and could be directly compared to one another, even though candidates may have attempted different forms.

As mentioned previously, each INBDE form contains scored items, which are used to estimate candidate abilities, and unscored items which are not. Scored items are drawn from the calibrated INBDE item bank and have therefore been placed on a common scale of measurement. Unscored items, referred to as *pretest* items, are new items that are being administered to candidates for the first time. Pretest items are administered randomly throughout the examination and are presented in the same manner as scored items so they cannot be identified by candidates. If items show adequate psychometric performance during the pretesting phase, they are eligible to become scored items on future forms of the

examination. Before this can occur, however, the items must be placed onto the established INBDE scale of measurement so they can become part of the calibrated item bank. This is accomplished by fixing the parameters for the scored items to their already-estimated values, while simultaneously estimating the parameters of the pretest items. The JCNDE uses the aforementioned pretesting approach to continually expand the calibrated INBDE item bank.

9. Standard Setting

A critical step in the development of any licensure examination is the setting of the performance standard that candidates must meet in order to pass the examination (AERA, APA, NCME, 2014, p. 100-101). Standard 5.21 indicates that subject matter experts involved in setting performance standards should be qualified, and that the process for setting the standard should be well described and documented (AERA, APA, NCME, 2014, p. 107-108). The section below describes the process that was used to determine the performance standard for the INBDE.

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 for the INBDE

In January 2024, the JCNDE convened a standard setting panel to recommend a new performance standard for the INBDE. The standard setting panel established a recommendation using the Bookmark standard setting method (Lewis, Mitzel, Mercado, & Schulz, 2012). The standard setting activities took place at the ADA headquarters over a two-day period and involved the following steps:

1. A standard setting panel was convened. The standard setting panel was composed of ten 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 INBDE. This included a description of the test specifications, test construction methods, scoring, equating, and results reporting methods. Panel members were also provided with historical information about candidate performance on the NBDEs. Finally, panel members completed an abbreviated version of the INBDE which was representative of a full version with respect to content, difficulty level, timing, and item formats.
3. The panel members engaged in a complete and thorough discussion of the characteristics and behaviors of the “just qualified” (i.e., minimally competent) candidate and of the importance of individual content elements on the examination.
4. Following the discussion phase, panel members were trained in the Bookmark standard setting method and were given an opportunity to practice the method using provided practice materials.
5. Panel members reviewed a large set of examination items that had been placed into an Ordered Item Booklet (OIB) assembled as follows:

- a. Each page of the OIB contained one item.
 - b. 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.
 - c. The items included in the OIB spanned a representative range of difficulty levels.
 - d. After reviewing the OIB, each panel member was asked to independently “bookmark” the page number in the OIB of the last item that a minimally competent candidate would have at least a two-thirds chance of answering correctly. A cut score for the examination was derived from the median of the committee members’ bookmark placements using the method described by Lewis et al. (2012).
6. 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 performance standard associated with the median bookmark placement (i.e., the percent of candidates who would fail if the standard was implemented).
 7. Steps 4 and 5 as described above were repeated three times. After each replication of the process, panel members were provided an opportunity to ask questions, express any concerns, and engage in group discussion. The final recommended performance standard for the examination was based on the standard derived in the third and final round of the process.
 8. At the conclusion of this process, panel members were asked to complete an evaluation questionnaire regarding their impressions of the process. Of the nine panel members who completed the evaluation, eight (89%) indicated they were either “very confident” or “confident” that the panel had recommended an appropriate passing standard. The remaining panel member indicating they were “somewhat confident.”

The recommended performance standard resulting from the 2024 INBDE standard setting activities was reviewed and approved by the Joint Commission in May 2024. The standard was implemented on June 1, 2024.

10. Score 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. These factors include fatigue, background noise, and candidate motivation. Reliability indices assess the degree to which random error affects scores. Low score reliability indicates the strong presence of random sources of measurement error, whereas high score reliability indicates the absence of such sources of error. Chapter 2 of the *Standards* highlights the importance of reporting the reliability of test-based decisions for high stakes licensing examinations (AERA, APA, NCME, 2014, p. 46-47). The overarching standard for Chapter 2 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.

Classification Accuracy and Classification Consistency

When scores on an examination are used as a basis for making pass/fail decisions, it is critical to ensure 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. For the INBDE, the Joint Commission estimates classification accuracy and consistency using the methods described in Rudner (2001), which are based on IRT. For this method, estimates of classification accuracy and classification consistency can be determined using IRT candidate ability estimates and their standard errors, which are produced during the 3PL model estimation process (see Chapter 8).

Estimates of classification accuracy and classification consistency for the INBDE are presented in Table 10.1. The estimates are calculated using data from first-time INBDE candidates who were students in dental education programs accredited by CODA.

Table 10.1
Classification Accuracy and Classification Consistency

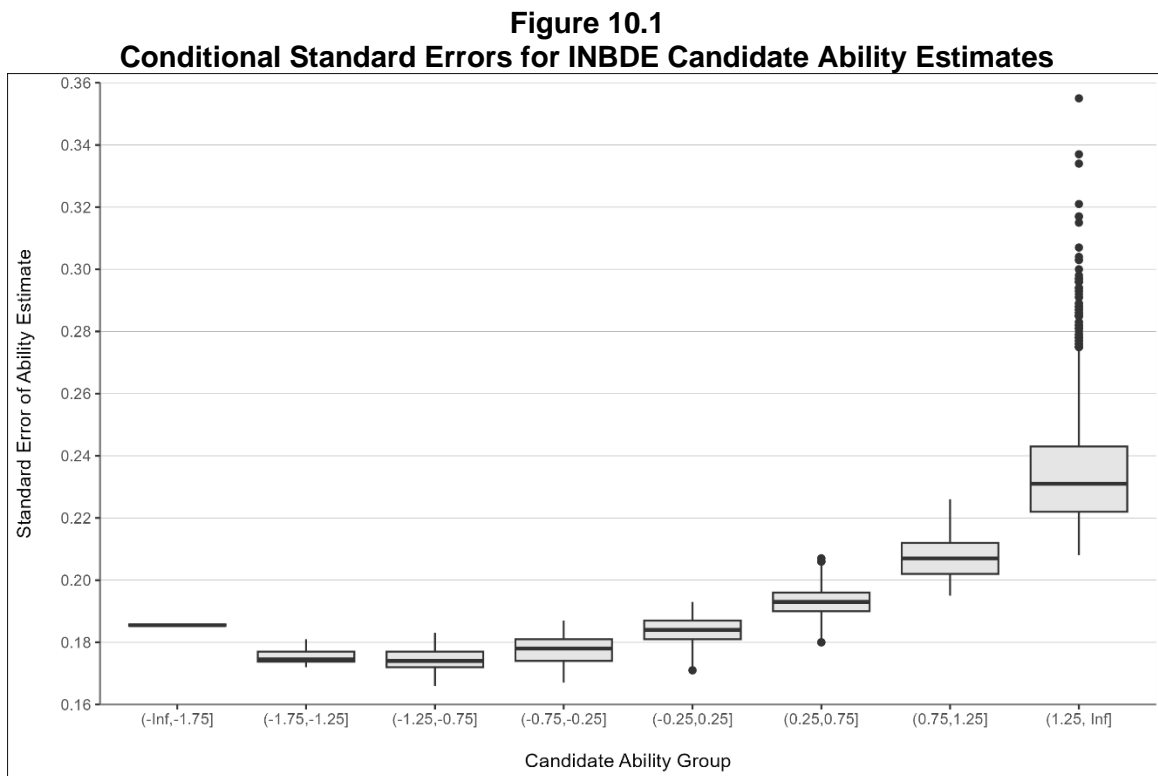
Year	Candidates	Classification Accuracy	Classification Consistency
2024	6,546	.96 to .98	.94 to .97

Note. The values presented represent the range of estimates calculated across examination forms.

Conditional Standard Errors

As described in Chapter 8, the INBDE is scored using the 3PL IRT model. In an IRT framework, measurement error is conditional on candidate ability. JCNDE staff strive to assemble INBDE forms that enable low levels of measurement error in the area of the candidate ability continuum that surrounds the INBDE performance standard. Doing so helps to ensure high levels of

classification accuracy and consistency. In order for this to occur, the difficulty of the items that comprise each examination form must be appropriate with respect to the examination purpose. Figure 10.1 shows conditional standard errors for INBDE candidate ability estimates based on data from 6,546 INBDE candidates trained in dental education programs accredited by CODA, who were attempting the examination for the first time in 2024. The standard errors are presented on the 3PL candidate ability metric as opposed to the INBDE scale score metric (see Chapter 8 for details). To create the figure, INBDE candidates were categorized into eight groups based on their 3PL ability estimates, and the standard errors of the ability estimates for the candidates in each group were summarized via Box plots. Inspection of Figure 10.1 reveals two notable findings: 1) standard errors for INBDE candidate ability estimates are relatively low across the entire ability continuum, with few instances exceeding .33; and 2) standard errors tend to be lowest in the area of the candidate ability continuum that surrounds the INBDE performance standard which is located in the interval between -1.75 and -.75. Overall, the conditional standard errors summarized in Figure 10.1 indicate that INBDE forms are being assembled in a way that is optimal given the examination's purpose.



Kuder Richardson Formula 20 (KR-20) Reliability

The Joint Commission also uses Kuder Richardson Formula 20 (KR-20; Kuder & Richardson, 1937) as an index of score reliability for the INBDE. The KR-20 reliability coefficient is an index of internal consistency reliability for examinations with items scored dichotomously (e.g., right or wrong). The coefficient can range from zero to one, with higher values indicating higher reliability. Table 10.2 presents KR-20 reliability coefficients for recent administrations of the

INBDE. The reliability estimates are calculated using data from first-time INBDE candidates who were students in dental education programs accredited by CODA.

Table 10.2
KR-20 Reliability Estimates for INBDE Scores

Year	Candidates	KR-20 Reliability
2024	6,546	.88 to .90

Note. The values presented represent the range of estimates calculated across examination forms.

11. Test Administration

The INBDE is administered year-round. Prometric administers the examination at its testing centers located throughout the United States, its territories, and Canada. Once eligible, candidates can schedule an examination for any business day, conditional on testing center availability. The administration schedule for the INBDE is provided in Table 11.1.

Table 11.1
INBDE Administration Schedule

Day 1		
Section	Content	Minutes
Tutorial (optional)		15
Section 1	100 standalone items	105
Schedule Break 1 (optional)		15
Section 2	100 standalone items	105
Scheduled Break 2 (optional)		30
Section 3	100 standalone items	105
Scheduled Break 3 (optional)		15
Section 4	60 items (cases)	105
8 hours and 15 minutes		
Day 2		
Tutorial (optional)		15
Section 5	70 items (cases)	105
Scheduled Break 4 (optional)		15
Section 6	70 items (cases)	105
Post-exam Survey		15
4 hours and 15 minutes		

The INBDE Examination Guide details INBDE candidate eligibility requirements and the INBDE application process. The guide is made available to candidates through the JCNDE website, jcnde.ada.org.

12. Results Reporting and Interpretation

This chapter describes results reporting to the various stakeholders, such as candidates, dental boards, and dental education programs. The Joint Commission's results reporting procedures are consistent with the standards indicated below.

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 INBDE Results to Candidates

INBDE 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 testing appointment date, unless the candidate's results were voided or withheld.

As noted previously, in 2012 the JCNDE moved to the pass/fail reporting of results for the National Board Examinations. Therefore, INBDE candidates who pass the examination receive a report indicating their result is "Pass" but do not receive numeric scores. Candidates who fail the INBDE receive a report indicating their result is "Fail" along with their numeric overall INBDE scale score. INBDE 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 INBDE are also provided with a graphical depiction of their performance in the following areas:

Clinical Component Sections

- Diagnosis and Treatment Planning
- Oral Health Management
- Practice and Profession

Foundation Knowledge Areas

- Molecular, biochemical, cellular, and systems-level development, structure and function
- Physics and chemistry to explain normal biology and pathobiology
- Physics and chemistry to explain the characteristics and use of technologies and materials
- Principles of genetic, congenital and developmental diseases and conditions and their clinical features to understand patient risk
- Cellular and molecular bases of immune and non-immune host defense mechanisms
- General and disease-specific pathology to assess patient risk
- Biology of microorganisms in physiology and pathology
- Pharmacology
- Sociology, psychology, ethics and other behavioral sciences
- Research methodology and analysis, and informatics tools

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 INBDE contains explanatory text that is intended to help candidates interpret their results.

Reporting INBDE Results to Dental Boards

When a candidate applies to take the INBDE, they indicate which dental boards their results should be reported to. The JCNDE reports results to the requested dental boards through a secure, password protected electronic portal. Results are typically made available to the requested dental boards within four weeks of the testing appointment date, unless the candidate's results were voided or withheld. A history of the candidate's NBDE results is made available to each dental board requested to receive results.

Reporting INBDE Results to Dental Education Programs

INBDE results are reported to dental education programs that are accredited by the Commission on Dental Accreditation (CODA). Results are reported to the program's dean or designee through a secure, password protected electronic portal. Each month, dental education programs receive a report containing examination results for their current students and recent graduates who attempted the examination during the preceding month. If three or more students or recent graduates from the program attempted the INBDE in the preceding 12-month period, the report also contains information about how the program performed relative to other programs during that period. This normative information is provided at the overall level, by clinical component section, and by Foundation Knowledge Area.

At the beginning of each calendar year, dental education programs receive an annual report that describes how their program performed relative to other programs during the previous calendar year. The annual report also provides historical data, going back five years, so the program can monitor performance trends over time.

13. Comparisons with NBDE Parts I and II

The INBDE is intended to measure a construct that is similar but not identical to those measured by NBDE Parts I and II. The INBDE relies less on knowledge and information recall than did NBDE Parts I and II, and instead emphasizes the relevance of the biomedical sciences through items that link science and clinical application. Items on the INBDE require candidates to draw from their knowledge of biomedical science, behavioral science, and clinical science, in order solve patient problems. The INBDE Item Development Guide provides a detailed comparison of scientific areas covered by the INBDE Foundation Knowledge Areas and those covered by NBDE Parts I and II. The Item Development Guide is made publicly available through the JCNDE website, jcnde.ada.org.

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 INBDE. Convergent validity evidence is discussed as follows within the *Standards*.

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).

To provide convergent validity evidence in support of the INBDE, the JCNDE used data from the 2017-2018 INBDE Field Test to examine the relationship between scores on the INBDE and scores on NBDE Parts I and II. This was possible because each candidate who participated in the Field Test also attempted NBDE Parts I and II. The observed and disattenuated correlations among INBDE scores and NBDE scores are provided in Table 13.1 and 13.2, respectively. As shown in the tables, performance on the INBDE was highly correlated with performance on NBDE Parts I and II, although INBDE scores showed a stronger correlation with scores from NBDE Part II.

Table 13.1
Observed correlations among INBDE and NBDE scores (N = 1,180 participants)

Score	INBDE	NBDE Part I	NBDE Part II
INBDE	1.00		
NBDE Part I	.58	1.00	
NBDE Part II	.74	.65	1.00

Table 13.2
Disattenuated correlations among INBDE and NBDE scores (N = 1,180 participants)

Score	INBDE	NBDE Part I	NBDE Part II
INBDE	1.00		
NBDE Part I	.65	1.00	
NBDE Part II	.84	.71	1.00

Note. Disattenuated correlations were estimated using the following reliability coefficients: INBDE=.87; NBDE Part I=.94; NBDE Part II=.91.

The moderate to strong correlations presented in Tables 13.1 and 13.2 provide sound convergent validity evidence in support of the intended interpretation and use of INBDE results. In reviewing the disattenuated correlations, it should be noted that a perfect correlation between the INBDE and NBDE Part II scores would not be desirable, since it would indicate that the INBDE and NBDE Part II measure an identical construct. As noted previously, the INBDE is intended to measure a construct that is similar but not identical to those measured by NBDE Parts I and II. The domains of interest for Parts I and II target knowledge and skills in the biomedical science and clinical sciences, respectively. The INBDE, on the other hand, targets

the integration of knowledge and skills in complex ways to deal with patient problems from a more holistic perspective. This is exemplified in the two key concepts that inform INBDE development: Clinical Relevance and Integration. Clinical Relevance refers to factors that impact patient outcomes in clinical/professional contexts. This includes all aspects of patient care and encompasses considerations involving how dentists approach the practice of dentistry (Practice Relevance) and keep up with the profession and advances that impact the profession (Professional Relevance). Integration brings to bear knowledge of basic, clinical, and/or behavioral sciences along with cognitive skills to understand and solve problems in clinical and professional contexts.

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, which is the department within the ADA that conducts examination programs for the Joint Commission. The 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).

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. The following are examples:

1. individual items and case materials (e.g. radiographs, clinical photographs, and dental charts in development, in camera-ready copy, and in electronic files for transmission to administration sites);
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 the security of examination materials. Joint Commission staff reviews vendors' operations to ensure compliance with security policy. All service agreements with vendors require adherence 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.

Security of Test Materials in Electronic Format. Departmental and vendor computers are protected with firewalls, login identifications, passwords and other 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 INBDE Guide describes procedures for identification of candidates, including requirements for positive identification through biometrics. Candidates' conduct is closely monitored during the testing appointment. Examination regulations and testing center policies are designed to deter policy violations and breaches of security.

Policies and Procedures for Dealing with Breaches in Security

The Joint Commission provides specific procedures for observing and reporting breaches in security and communicates them to test administrators. It promptly investigates reports of security breaches and ensures examination items are removed from use when it determines security has been breached. 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 INBDE Guide. This document provides detailed information related to the Joint Commission's examination policies, 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 INBDE practice questions that are provided free of charge. Each year the INBDE Guide is updated and amended as necessary. The guide and INBDE practice questions are available through the Joint Commission's website. This technical report also serves as a source of documentation that is freely available to all INBDE 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 INBDE, 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 INBDE 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. Trends in Candidate Performance

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

Table 16.1
INBDE Failure Rate, by Candidate Group and Year

	Accredited [†]						Non-Accredited [‡]						Total	
	First Attempt ^a		Mixed Attempt ^b		Retake ^c		First Attempt ^a		Mixed Attempt ^b		Retake ^c		All Attempts	
	N	% Fail	N	% Fail	N	% Fail	N	% Fail	N	% Fail	N	% Fail	N	% Fail
2020*	204	1.0	1	0.0	69	7.3	147	38.8	0	-	117	58.1	538	24.5
2021	2,018	1.3	3	0.0	245	16.0	1,340	33.1	1	0.0	971	55.8	4,578	22.3
2022	5,837	0.8	15	0.0	234	13.2	2,144	25.3	10	20.0	1,145	44.7	9,385	12.1
2023	6,648	0.4	11	0.0	136	14.0	3,277	16.0	24	8.3	1,212	33.8	11,308	8.7
2024*	6,636	4.8	-	-	168	32.1	4,137	25.3	-	-	919	52.8	11,860	16.1

^a Indicates candidates who had never previously attempted the INBDE, NBDE Part I, or NBDE Part II.

^b Indicates candidates who passed NBDE Part I on their first attempt and subsequently elected to attempt the INBDE instead of NBDE Part II.

^c Indicates candidates who had previously attempted and failed the INBDE, NBDE Part I, or NBDE Part II.

[†] Indicates candidates trained by dental education programs accredited by CODA.

[‡] Indicates candidates trained by dental education programs not accredited by CODA.

*A new standard was introduced this year, based on updated standard setting activities.

References

- Alpern, R. J., et al. (2009). *Scientific Foundations for Future Physicians*. Report of the AAMC-HHMI Committee. Association of American Medical Colleges. Washington, DC.
- American Dental Education Association (2013). Foundation knowledge and skills for the new general dentist (as approved by the 2011 ADEA house of delegates). *Journal of Dental Education*, 77, 903-907.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: Author.
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296-322.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Chuah, S.C., Drasgow, F., & Luecht, R. (2006). How big is big enough? Sample size requirements for CAST item parameter estimation. *Applied Measurement in Education*, 19, 241-255.
- de Ayala, R.J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford Press.
- Dorans N. J., Kingston N. M. (1985). The effects of violations of unidimensionality on the estimation of item and ability parameters and on item response theory equating of the GRE verbal scale. *Journal of Educational Measurement*, 22, 249-262.
- Drasgow F., Parsons C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, 7, 189-199.
- Fraser, C., & McDonald, R.P. (2003). *NOHARM: A Windows program for fitting both unidimensional and multidimensional normal ogive models of latent trait theory* [Computer program]. Wellend, ON: Niagara College.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, Inc.
- Horn, J. (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Joint Commission on National Dental Examinations, American Dental Association (2019). *INBDE Item Development Guide*. Chicago: Author.
- 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., & 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.
- Rudner, L.M. (2001). Computing the expected proportions of misclassified examinees. *Practical Assessment, Research & Evaluation*, 7.
- Waldschmidt, D.M., Yang, C.L., & Christensen, M. (2013, April). Developing Test Specifications for an Integrated Examination Based on Practice Analysis and Linkage Data. Paper presented at the meeting of the National Council on Measurement in Education, San Francisco, CA.
- Yen, W.M. (1984). Effects of local item dependence on the fit an equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8, 125-145.
- Zimowski, M.F., Muraki, E., Mislevy, R.J., & Bock, R.D. (2003). *BILOG-MG 3*. Chicago, IL: Scientific Software International.

Appendix A

INBDE Clinical Content Areas

#	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.
35	Prevent, diagnose and manage temporomandibular disorders.

#	Oral Health Management
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

INBDE Foundation Knowledge Areas

FK1	Molecular, biochemical, cellular, and systems-level development, structure and function
FK2	Physics and chemistry to explain normal biology and pathobiology
FK3	Physics and chemistry to explain the characteristics and use of technologies and materials
FK4	Principles of genetic, congenital and developmental diseases and conditions and their clinical features to understand patient risk
FK5	Cellular and molecular bases of immune and non-immune host defense mechanisms
FK6	General and disease-specific pathology to assess patient risk
FK7	Biology of microorganisms in physiology and pathology
FK8	Pharmacology
FK9	Behavioral sciences, ethics, and jurisprudence
FK10	Research methodology and analysis, and informatics tools

Appendix C

INBDE Test Specifications

	FK 1	FK 2	FK 3	FK 4	FK 5	FK 6	FK 7	FK 8	FK 9	FK 10	Total
Diagnosis and Treatment Planning	5.0%	2.2%	1.8%	4.6%	3.6%	5.2%	4.2%	3.2%	3.0%	3.4%	36.2%
Oral Health Management	6.8%	4.4%	4.4%	4.2%	4.2%	3.8%	4.2%	4.4%	2.8%	2.8%	42.0%
Practice and Profession	0.4%	0.2%	1.8%	1.8%	1.2%	2.8%	2.2%	3.0%	4.8%	3.6%	21.8%
Total	12.2%	6.8%	8.0%	10.6%	9.0%	11.8%	10.6%	10.6%	10.6%	9.8%	100.0%

- FK 1. Molecular, biochemical, cellular, and systems-level development, structure and function
- FK 2. Physics and chemistry to explain normal biology and pathobiology
- FK 3. Physics and chemistry to explain the characteristics and use of technologies and materials
- FK 4. Principles of genetic, congenital and developmental diseases and conditions and their clinical features to understand patient risk
- FK 5. Cellular and molecular bases of immune and non-immune host defense mechanisms
- FK 6. General and disease-specific pathology to assess patient risk
- FK 7. Biology of microorganisms in physiology and pathology
- FK 8. Pharmacology
- FK 9. Behavioral sciences, ethics, and jurisprudence
- FK 10. Research methodology and analysis, and informatics tools



Joint Commission on National Dental Examinations
Department of Testing Services
401 North Michigan Avenue, Suite 3300
Chicago, Illinois, 60611-4250