Dear Reader:

Please join me in paying tribute to Steve Nettles and Jim Zukowski as they step down from the editorial board of CLEAR Exam Review: A Journal. We thank them for their extraordinary service and dedication over the last 26+ years. Steve joined the editorial board with Volume 3, No. 1 in Winter 1992, and Jim joined in Winter 1994. In volunteering countless hours of their time and considerable expertise, they have ensured that CLEAR members and subscribers stay current on the latest news, resources, and thinking from the testing and measurement field.

Over the years, this extraordinary publication has influenced the work of occupational regulators across CLEAR’s broad membership. We thank Steve and Jim for their role in bringing this relevant and timely information to the readership, helping our organizations to better protect the public we all ultimately serve.

We are also indebted to and thank Adrienne Cadle and Sarah Wennik, who have agreed to join the editorial board. Recognizing that they will bring new perspectives to this important publication, we look forward to future issues and the continued success of CLEAR Exam Review: A Journal.

Like me, I know you will enjoy the contents of this latest issue. We applaud the work of all those involved in the production of this much-valued publication.

Sincerely,

Ginny Hanrahan
CLEAR President
CLEAR Exam Review is a journal, published twice a year, reviewing issues affecting testing and credentialing. CER is published by the Council on Licensure, Enforcement, and Regulation, 108 Wind Haven Drive, Suite A, Nicholasville, KY 40356.

Design and composition of this journal have been underwritten by Prometric, which specializes in the design, development, and full-service operation of high-quality licensing, certification and other adult examination programs.

Subscriptions to CER are sent free of charge to all CLEAR members and are available for $30 per year to others. Contact CLEAR at (859) 269-1289 or cer@clearhq.org for membership and subscription information.

Advertisements and Classified (e.g., position vacancies) for CER may be reserved by contacting CLEAR at the address or phone number noted above. Ads are limited in size to 1/4 or 1/2 page, and cost $100 or $200, respectively, per issue.

Editorial Board
Steven Nettles
Retired, Applied Measurement Professionals, Inc.

Jim Zukowski
Independent Consultant

Coeditor
Elizabeth Witt, Ph.D.
Witt Measurement Consulting
Laingsburg, MI
WittMeasure@aol.com

Coeditor
Sandra Greenberg, Ph.D.
ACT
New York, NY
Sandy.Greenberg@act.org

CONTENTS

From the Editors ........................................... 1
Sandra Greenberg, Ph.D.
Elizabeth Witt, Ph.D.

Columns

Abstracts and Updates ................................. 3
George T. Gray, Ed.D.

Legal Beat ...................................................... 9
Dale J. Atkinson, Esq.

Recent CLEAR Quick Poll Results .................. 12
Carla M. Caro, M.A.

Articles

Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams ..................................................... 15
Lisa M. Abrams, Ph.D., Katherine A. Reynolds, Ph.D., and Michael O’Leary, Ph.D.

One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On ..........24
Margaret Bent, Ph.D., Sarah Carroll, Ph.D., and Paul Grace, M.S.
From the Editors

SANDRA GREENBERG, Ph.D.
ELIZABETH WITT, Ph.D.

As we near the end of 2020, we are reflective of the difficult times that we have all endured—the losses suffered around the world; the extent of the changes we have witnessed, from the most profound to the most mundane; and the way we have endured. CLEAR witnessed its first-ever virtual CLEAR Annual Educational Conference (and while we didn’t have to travel, we did invite each other into our homes for Hollywood Squares mornings, afternoons, and evenings). And, yes, this is the second time we have edited CLEAR Exam Review (CER) without the enjoyment of seeing each other. We offer our sincerest appreciation to our guest and column authors, and to you our readers. Read carefully; we hope you enjoy CER and take with you new learnings and appreciation for those in our credentialing community.

Now that CER is distributed in electronic form, we are attempting to include more links to online resources in articles’ reference lists. Although we do our best to keep these up to date, websites tend to change over time. We try to ensure that sufficient information is provided so that, should a link not lead to the article, the reader may still locate the article online through a simple search.

Before describing the contents of this issue, we want to take a moment to remember Dr. Rose McCallin, who passed away on November 12, 2020. Rose was a psychometrician who worked with the Colorado Department of Regulatory Agencies (DORA) as director of examination and statistics. She was involved in the National Council on Measurement in Education (NCME), serving on the NCME Board of Directors. She was active in CLEAR, serving on the publications committee, ERAC, and the testing and examinations program subcommittee. She also served briefly as a reviewer for CER. Rose will be remembered for her wit, intelligence, and personality. Her contribution to the regulatory community will be missed.

This issue of CER contains Abstract and Updates, Legal Beat, and Quick Poll results as well as two feature articles on topics as diverse as implementing an innovative competency assessment platform and aligning arguments in support of scoring interpretation and use claims of credentialing exams.

In Abstracts and Updates, George Gray describes a new edition of a 20-chapter volume covering assessment topics in health professions education, a set of articles on standard setting published in Educational Measurement: Issues and Practice, and articles on an enhancement to a job analysis for a medical maintenance of certification examination and proctoring for internet-based testing. Taken together, George Gray points us to much valuable information for those that are new to credentialing as well as for those who are experienced credentialing hands.
In *Legal Beat*, Dale Atkinson presents a legal take on a question that has been debated for years—for how long do test results count? Consider the following: Once candidates have passed the licensure examination and meet other statutory criteria to licensure eligibility, they generally do not have to reestablish entry-level competence through a testing process. Continued competence is demonstrated through ongoing practice and continuing education requirements as part of licensure renewal. That said, if an applicant has never practiced and has not completed continuing education because they are not licensed, for how long should their entry-level competence examination results be considered as valid?

Carla Caro summarizes the timely results of a *Quick Poll* survey conducted by CLEAR in July 2020 in the midst of the COVID-19 pandemic. When it became evident that the pandemic was going to be disrupting the work of regulatory bodies, licensure/registration authorities, and certification bodies for more than the short term, the *Quick Poll* was implemented to get a snapshot of responses to the situation. The poll was designed to assess how regulators were responding to the pandemic, what changes they had implemented, and what they anticipated for the future. Carla Caro summarizes the types of testing changes implemented, including postponing and modifying examination formats, transitioning between computer-based testing and paper-and-pencil testing, and extending testing windows and testing locations. In sum, regulatory bodies have shifted focus, changed policies and procedures, and met the challenges in several ways, some of which may permanently alter the landscape.

In a feature article, Margaret Bent and colleagues present a saga about one organization’s near decade-long development and implementation of an innovative continuing competence assessment platform built on gaming technology. Their organization was confronted with the everlasting need—designing a recredentialing program that engages practitioners in a continuous cycle of learning and assessment. The example they provide, including numerous related challenges and solutions, may provide a basis for others to consider. The authors go on to state that building such an assessment platform requires a major initial and ongoing investment of time and other resources, but the benefits may outweigh the costs.

In a feature article, Lisa Abrams and her co-authors present an approach to capitalizing on a powerful source of content validity evidence for credentialing examinations—one borrowed from educational achievement testing. The authors provide two different case studies to illustrate how content alignment studies might increase public confidence in the appropriateness of credentialing exam content. Moreover, the authors believe that there is much to be learned from educational K-12 testing, where content standards are more clearly explicated and technical documentation and reporting occur in ways that support the interpretation and use validity argument at all stages of the testing program.

As always, we hope you find this collection of articles useful and thought provoking. We look forward to a new year and wish you successful adventures in credentialing.
This issue’s column describes a new edition of a volume covering assessment topics in health professions education, a set of articles on standard setting published in *Educational Measurement: Issues and Practice*, plus articles on an enhancement to a job analysis for a medical maintenance of certification examination, and an article on proctoring for internet-based testing.

**Health Professions Assessment**


This book is a wide-ranging introduction to assessment in health professions education. The rationale is that it meets an unfulfilled need as it targets education in health professions rather than training in psychometrics or licensure and certification testing. The book has 20 relatively short chapters covering a wide range of topics. As introductory material covering breadth rather than depth, there is still considerable information that is immediately useful, such as conceptual frameworks and calculations for reliability of test scores and item analysis. Given the limitations of space, there is some reliance on “rules of thumb” rather than more detail, but the last chapter, “Engaging With Your Statistician,” points the direction toward developing a firm foundation for collaboration between the reader and a measurement consultant. Based on my own experience working as an assessment consultant in a university medical center and in health program licensure and certification testing, I can envision a number of situations where chapters of this book could be used as a springboard for discussion as well as a resource for health programs educators. For additional reading, each chapter contains two or more pages of references at the end.

The chapters of the book are

1. Introduction to Assessment in the Health Professions
2. Validity and Quality
3. Reliability
4. Generalizability Theory
5. Statistics of Testing
6. Standard Setting
7. Written Tests: Writing High-Quality Constructed-Response and Selected-Response Items
8. Oral Examinations
9. Performance Tests
10. Workplace-Based Assessment
11. Narrative Assessment
12. Assessment Portfolios
13. Key Features Approach
14. Simulations in Assessment
15. Situational Judgment Tests
16. Programmatic Assessment: An Avenue to a Different Assessment Culture
17. Assessment Affecting Learning
18. Assessment in Mastery Learning Settings
19. Item Response Theory
20. Engaging With Your Statistician
Abstracts and Updates

Here is a brief look at the content of the chapters. Chapter 1 introduces George Miller’s pyramid, a model for describing the “levels of knowledge and skills assessed in health professions education” (p. 4), as well as a description of written tests, oral examinations, performance tests, workplace-based assessment, narrative assessments, and portfolios. The Miller model hierarchy (knows, knows how, shows how, does) is used throughout the book to illustrate the targets of the various assessments discussed.

Chapter 2 focuses on purpose-driven assessment and introduces validity frameworks by Kane and Messick. A detailed table lists threats to assessment validity and quality. Chapter 3 (Reliability) introduces measurement fundamentals and includes the calculation of Cronbach’s alpha for reliability. Chapter 4, Generalizability Theory, provides a brief introduction to this topic. Chapter 5 is an introduction to test statistics: reported scores, reliability of scores, and item analysis. Several formulas are produced to illustrate hand calculation of results.

Chapter 6 addresses standard setting, first conceptually and then with specific methods: Angoff, Ebel, Borderline Group Method, the Contrasting Group Method, the Body of Work Method, and the Patient Safety Method. Chapter 7 describes guidelines for writing test items for written tests: selected response (e.g., multiple choice) and constructed response (short-answer and long-answer formats). The latter presentation includes a discussion of scoring rubrics, a topic often neglected in classroom settings. Chapter 8 discusses oral examinations, a topic that has a long history in medical licensure, and includes examples as well as a discussion of related issues. Chapter 9 (Performance Tests) describes standardized patients and the multi-station Objective Structured Clinical Examination (OSCE).

Chapter 10, Workplace-Based Assessment, addresses issues and methods in competency-based education. Chapter 11 (Narrative Assessment) includes documents that are completely narrative and those that involve a rating scale as well as written comments. Chapter 12 discusses assessment portfolios.

The key features approach to assessment (Chapter 13) using medical cases focuses on critical, case-specific decisions or actions (p. 199). Chapter 14 covers a range of simulations in assessment, including responsive mannequin simulators. Chapter 15 focuses on situational judgment tests. “Situational judgment tests have three parts—a situational presentation, questions asked, and response required” (p. 229).

Chapter 16 covers programmatic assessment, a model based on an “overarching competency framework.” Guidelines are described, including a mix of assessment methods, continuous dialogue with the learner, and decisions made by a committee of assessors (p. 247). Chapter 17 provides something of a reversal of perspective: how assessment affects learning. Chapter 18 covers mastery learning: instruction and assessment of learning using a criterion to indicate mastery of the content.

Chapter 19 introduces item response theory (IRT). This is a difficult topic to introduce in 16 pages, but the guiding concept is that test scores can be calculated on an interval scale if the assumptions of the IRT model are met. The conversion from a number correct or percent correct program to IRT is not usually easy, but sometimes you get lucky. I remember one certification program that wanted to move to IRT and happened to have 12 test forms, with individual test forms having links of common items to other test forms in the set. Each form was calibrated, and the common items were used to put the items on the same scale. The ability to effect this transition quickly was a very pleasant surprise.
Abstracts and Updates

The final chapter is on working with your statistician. A book containing so many topics cannot include enough
detail to make even an enthusiastic learner an expert; however, the end-of-chapter references are an excellent
source for follow-up. Finally, a shout-out goes to the authors for emphasizing threats to the validity of assessments
and how to minimize them. This is an area where increased awareness is always beneficial.

Standard Setting Articles

The first issue of *Educational Measurement: Issues and Practice* for 2020 (v. 39, 1) contains a special section:
“Issues and Advances in Standard Setting Methods”; seven articles are included.

A personal recollection provides a little context for the discussion of passing score studies. The Angoff method was
introduced in 1971 in a chapter that Angoff wrote on scales, norms, and equivalent scores. There were actually two
Angoff methods, one occupying a short section in the text and the other presented in a footnote. Historically, the
footnote method (four sentences long and rating the probability that the minimally acceptable person would answer
correctly) won out and became quite popular, although the other method (the 0/1 rating for whether the hypothetical
minimally acceptable person would answer the item correctly) has also been used. Almost 50 years later, the Angoff
method is likely the most popular standard setting method, although I remember that a conference I attended in
1994 included a strong argument to the effect that the concept of a hypothetical minimally knowledgeable person
was problematic for the subject matter experts who were asked to perform the ratings. This resulted in some
pushback at the time, but it did not endure. Here we are 26 years later with more methods, but most involve some
aspect of rater judgment of minimum examination candidate competency.

The introduction to the group of standard setting articles by Everson and Forte states that the articles are written
by ‘a diverse group of scholars who offer a range [of] concerns with some of the more ‘established’ or ‘widely
accepted’ procedures . . . a story line that starts with ‘let’s blow it all up’ and proceeds through several versions and
iterations of why traditional methods, many of which are based on expert panelists’ judgments of the ‘proficient’
examinee, are not as solid as we may think. The unifying thread . . . of these articles is the focus on ways to
improve standard setting by reducing the subjectivity of panelists’ judgments and enhancing the validity and utility
of the resulting performance classifications” (p. 7).


This article suggests that Achievement Level Descriptors (ALDs), commonly used in educational testing, might be
used in place of traditional standard setting procedures if incorporated into test design. “When test items are written
to target specific ALDs, especially with explicitly identified task complexity components and expected item difficulty,
and when the items are subject to expert verification that they include the essential features for providing evidence
that the ALDs are met, there should be no further need for evaluation. Instead, empirical data can and should be
used to identify which items do or do not perform consistently with . . . the ALDs the items are targeting” (p. 9).

“If the relationship between item targets and empirical data are [sic] sufficiently strong, then the traditional standard-
setting workshop is either unnecessary or altered in nature and significantly reduced in scope” (p. 8).
Abstracts and Updates


In the abstract of this article, the authors state, “This article presents results from three experiments where panelists were asked to give Bookmark-type ratings to separate items into groups based on item difficulty data. Results of the experiments showed, consistent with results often observed with the Angoff method, that panelists typically and paradoxically perceived hard items to be too easy and easy items to be too hard. These perceptions were reflected in panelists often placing their Bookmarks too early for hard items and often placing their Bookmarks too late for easy items” (p. 22).


As passing score decisions are based on subjective judgments, there are consequences for examinees when a false negative or false positive decision results. The authors state, “When determining a standard, standard setting panelists implicitly consider the negative consequences of the decisions made from test use. We propose the conscious weight method and subconscious weight method to bring more objectivity to the standard setting process” (p. 30).

The two methods mentioned in the title of the article are described as follows: “To help validate the use of the cut score, we propose two new methods to quantify the judgments of the participants. First, we suggest using the implicit weight of the consensus cut score as consequential validity evidence. Thus, the subconscious weight method (SWM) uses the implied weight of harm due to a false positive misclassification compared to harm due to a false negative misclassification. We use this weight to understand the mindset of the panelists and/or policymakers involved in establishing the standard. The second method, the CWM [Conscious Weight Method], incorporates an additional round of impact data into the standard setting process. Individual standard setting panelists are provided the degree to which their selected cut score weighs the harm of a false positive misclassification compared to harm of a false negative misclassification. Panelists are then afforded the opportunity to adjust their proposed cut score based on this weight” (pp. 31-32).


In the Bookmark method, judges are asked to place a bookmark in a book of items ordered by level of difficulty to set the passing score for the examination. Instructions to the judges include considering a response probability for the examinee answering correctly when they place the bookmark in the book of items. This study compared the results when groups of judges were required to consider one of two different response probabilities, .67 or .80. The results of the experiment indicated that the median bookmark position was the same for both conditions, but this outcome resulted in different scores on the proficiency scale.

The authors indicate that conceptual problems for judges performing ratings using the Angoff method are not avoided when the bookmark method is used. “Practitioners who choose to use the bookmark method should do so with the awareness that content experts are likely to have considerable difficulty incorporating the response probability into their judgments. Careful training around this aspect of the procedure certainly is important” (p. 43).
Abstracts and Updates


The authors state, “In this study, we build upon a standard-setting method that was proposed to be used with a test modeled with a cognitive diagnosis model (CDM) . . . We adapt part of that methodology to (a) define the borderline profiles that are conceptualized by standard setting panelists, (b) predict how examinees with these different profiles will perform on a test, and (c) propose modifications of current standard-setting methods to address differences in borderline profiles” (p. 46).


Like the Hofstee method, the Beuk method is a widely used compromise method in standard setting. Rather than rate individual test items, judges are asked to judge the ideal pass rate for a set of items and the ideal percent of items correct to pass. The article provides a thorough discussion of the Beuk method. Seventeen studies are cited, and a table is shown illustrating Beuk outcome and importance weights. Wyse indicates, “This article illustrates some of the challenges with the assumptions of the Beuk standard-setting method, and it offers a simple way to test these assumptions by asking panelists to assign importance weights. In my view, the importance weights provide a simple mechanism to evaluate whether the slope of the Beuk line represents the intentions of the panelists” (p. 59).


Certification and licensure examinations routinely report a total pass rate for an examination date or over a period of time (e.g., annual). This article emphasizes four types of pass rates: (1) terminal pass rate, “the number of test takers who passed over the number of test takers who took an attempt at the exam,” (2) total attempts pass rate, “the number of passes over the number of test attempts,” (3) first attempt pass rate, “what percent of examinees pass on their first attempt,” and (4) repeat attempts pass rate, “what percentage of examinees who repeat the exam passed” (p. 62). A table is included in the article listing all four pass rates for 14 credentialing programs. The authors state that “the key take-home message of this study is that there is a need for credentialing and educational programs to explicitly state which pass rates they are reporting and make sure that the meaning and use of the pass rates they choose to report are well aligned with how the statistics should be used” (p. 67).

Enhancing a Maintenance of Certification Examination


This article describes an innovative procedure to include practicing physician input in the American Board of Internal Medicine’s Maintenance of Certification (MOC) Examination. A survey was conducted that included blueprint topic descriptions and patient-related tasks. The survey for the cardiology section is discussed in the article. The survey included ratings of importance of 188 blueprint topic descriptions, crossed with five tasks such as diagnosis and treatment. The selected items for new test forms are updated based on the importance ratings from this survey.
Internet-Based Proctoring


This article is a review of practices and technical features associated with internet-based testing, with implications considered for certification and licensure testing as well as testing by schools. Both un-proctored and proctored internet-based assessments are considered, with implications not only for security of examinations but accessibility (opportunity to test) and assessment design to minimize unauthorized behaviors.

Conclusions include the following remarks: “General findings currently support the use of live and AI [artificial intelligence] proctoring in that they minimize cheating, secure test content, and provide comparable score distributions . . . Top-tier security levels are costly, and testing organizations need to seek security levels commensurate with the use and consequences stemming from test scores.” (p. 26). The last point is that with differing internet access, “equity or opportunity to test is a concern that organizations need to address” (p. 26).
Legal Beat

Do Examination Results Expire?
DALE J. ATKINSON, ESQ.

Dale Atkinson is the managing member of the law firm of Atkinson & Atkinson, LLC
http://www.atkinsonfirm.com/home

Through statutory enactments, legislatures require that applicants for governmentally issued licenses meet certain criteria. State-based licensure of the professions is premised upon consumer protection and requires applicants to meet statutory requirements. Regulatory boards are legislatively created and are delegated with the authority to, among other things, issue and deny applications for licensure. Boards are populated with a mix of persons from the profession and public members. Professional representation on the boards adds needed expertise and promotes efficiencies in carrying out the intent of the statutes. This expertise assists in interpreting statutes, promulgating regulations, assessing complaints, and promoting overall efficiencies in the administrative processes. Public members provide balance to the process and ensure consumer representation. In the end, both professional and consumer members fulfill the legislative intent by enforcing the practice act in the interest of public protection.

During the ongoing COVID-19 pandemic and into the foreseeable future, statutory criteria for licensure of the professions will continue to be scrutinized. In fact, countless Executive Orders have modified, reduced, and eliminated some criteria for licensure, albeit temporarily. Long prior to the recent COVID pandemic, such eligibility criteria have been referred to as barriers to licensure and characterized as inhibiting the pursuit of one’s chosen profession. This added attention is caused by a multitude of factors, including the political climate, a push by economists, and the racial issues brought to the forefront by recent events. Licensure eligibility criteria for regulated professions generally includes components such as education, experience, examination, good moral character assessments, and, in some jurisdictions, formal criminal background checks. Each of these requirements serves its purpose, with some professions placing added emphasis on specific criteria.

Successful completion of an entry-level competence examination is a prerequisite to licensure in most professions. These examinations are developed, scored, and maintained following industry standards that assure their validity and reliability for use by governmental entities. They are built on a blueprint established through a job analysis whereby knowledge, skills, and abilities essential to practice from day one are identified and assessed. Psychometric principles allow for statistical analyses to substantiate performance and pass-fail results.

Education plays a very different role from entry-level competence examinations. Education involves academic freedom and the ongoing interaction between professors and students. The experience spans a significant period of time, and students are tested on an ongoing basis. During an academic experience, students are not tested on entry-level competence but rather experience an educational setting and develop life-long learning opportunities. Yet, when arguments arise addressing barriers to licensure, many seek to eliminate the entry-level competence determiner, the licensing examinations.

Once candidates have passed the licensure examination and meet other statutory criteria to licensure eligibility, they generally do not have to reestablish entry-level competence through a testing process. Continued competence is demonstrated through ongoing practice and continuing education requirements as part of
licensure renewal. If an applicant has not been practicing and has not completed continuing education because they are not licensed, one might ask, how long should the entry-level competence examination results be valid? Consider the following.

An applicant for a law license (Applicant) issued by the Michigan Supreme Court passed the bar examination in 2001. However, he was not issued a license because he was unable to obtain a favorable character and fitness report. In 2016, the Applicant received a favorable character and fitness report and sought the issuance of a law license. The Michigan Board of Law Examiners (Board), however, denied his request to use his 2001 passing score on the bar exam, citing a policy that invalidates passing scores on the bar exam after a three-year period. The Applicant filed suit in Federal District Court against the Board and its president, arguing a violation of his substantive due process rights.

The Applicant argued that he has a constitutional liberty or property interest in his bar examination score and that failure to recognize his 2001 passing result was an arbitrary exercise of the powers of government. He also argued that there exists no duly enacted statute or promulgated rule that invalidates an exam score after three years. The Applicant sought partial summary judgment on the issue of substantive due process and the recognition of his exam score.

By statute, the Michigan Supreme Court may adopt rules and regulations concerning the examination of applicants for admission to the bar. Further, the Board may adopt regulations concerning its functions and duties, subject to approval of the Supreme Court. The Michigan Supreme Court recognizes that the Board has substantial discretion in its rulemaking to govern the details of the bar examinations. Specifically, the Board’s Rule 7 invalidates a bar exam score after three years if a candidate has not been granted licensure.

The Applicant previously requested an extension of the certification of his examination score. One-year extensions were granted in 2004 and 2005, and a two-year extension was granted in 2007. After finally receiving his favorable character and fitness report in 2016, the Applicant sought waiver of the bar exam. This waiver request was denied.

Before addressing the merits of the complaint, the Court addressed some procedural arguments propounded by the Board. First, the Board argued that the Applicant’s claims were moot because the Board subsequently revoked the 2016 approval of the character and fitness report. The Court disagreed, finding that the issue to be determined was whether the three-year limit of certifying the bar exam results satisfied due process.

The Board also argued that it was entitled to judicial immunity under the 11th Amendment, but the Court rejected the argument, citing previous rulings in this case addressing motions to dismiss. The Board finally urged the Court to exercise its discretion in resolving the matter under the Declaratory Judgment Act. The Court held that the factors necessary for declining to enter the declarative relief were not present. Thus, the Court proceeded to address the substance of the Applicant’s complaint.

In order for an applicant to substantiate a substantive due process claim, the Court first addressed whether there must be a protected property interest at stake. In this case, the Court held that it would address the due process considerations even without a constitutional liberty or property interest. The core of any due process claim is protection of the individual against arbitrary government action, a clause that significantly restricts government
Legal Beat

from abusing its powers. It guarantees procedural fairness and, in the case of substantive rights, protects individual liberties.

As determined by the United States Supreme Court, two issues must be addressed when analyzing substantive due process claims. First, the court must carefully describe the asserted liberty interest. Second, the court must assess whether the asserted liberty interest is fundamental such that it is deeply rooted in history and tradition. The Applicant contended that he has a liberty or property interest in the 2001 passing score on the bar examination and such was arbitrarily interfered with through application of a rule not approved by the Michigan Supreme Court.

The Court assessed the Applicant’s arguments under both a broad interpretation analysis and a narrow interpretation. On a broad basis, the Court declined to recognize that the Applicant’s desire to pursue a legal career constitutes a fundamental liberty interest. The Court noted that the three-year rule did not preclude him from pursuing his career as an attorney but merely required him to take and pass the bar examination again. The Court also cited previous case law that supports the application of high standards as a prerequisite to issuance of a law license without offending the due process provisions, so long as the qualification(s) have a rational connection to the fitness to practice.

Under a narrow analysis, the Court identified that 14th Amendment due process rights are created through federal recognition of property and liberty interests. State-created rights give rise to procedural due process but typically do not constitute fundamental rights protected by substantive due process. Thus, the Court concluded that the Applicant does not have a fundamental right to his passing bar examination result. Any liberty or property right interest he has in the examination score was created by the state in the form of criteria for becoming a member of the bar. Accordingly, the Court held that the Applicant does not have a fundamental right to substantive due process related to the expiration of his passing score on the bar exam.

In spite of determining that the Applicant is not entitled to substantive due process rights based on a lack of a fundamental right, the Court addressed whether the three-year expiration period shocks the conscience or is arbitrary and capricious. In situations where government action either shocks the conscience or is arbitrary and capricious, substantive due process can apply, even without a demonstration of a fundamental right or interest. The Court held that the three-year rule does not shock the conscience nor is it arbitrary and capricious.

In the end, the Court found no fundamental interest at stake and that the three-year rule constituted a right of the state to determine the criteria for eligibility for admission to the bar. How long an examination score is valid creates many interesting legal and practical questions. In the case of someone who has not ever been licensed, the expiration of a passing score after a specified period of time appears to be left to the individual states to determine. Without the recognition of a fundamental right or interest, the right of the states to require an applicant to re-take and pass the bar examination as a condition of licensure eligibility will withstand challenges under substantive due process analysis. One might ask if licensees can be required to re-establish their entry-level competence based upon the passage of time.

**Lawrence v. Pelton**, 2020 U.S. Dist. LEXIS 197848
Clear periodically issues Quick Poll surveys asking members about their credentialing practices, policies, and issues. These Quick Polls are not designed as scientific studies but allow us to gather snapshot data regarding current practices in credentialing. Clear administered a Quick Poll in July of 2020 when it became evident that the COVID-19 pandemic was going to be disrupting the work of regulatory bodies, licensure/registration authorities, and certification bodies for more than the short term. The poll was designed to assess how regulators were responding to the pandemic, what changes they had implemented, and what they anticipated for the future.

**Questions:**

As a result of the emergence of COVID-19, did your organization make any changes to the way you assess your candidates for licensure/registration?

- Yes
- No

Are you planning on implementing these changes permanently, regardless of the status of the pandemic?

- Yes
- No
- Not sure

If you answered yes or no, please describe the rationale for the decision. If you responded not sure, what factors will affect your decision?

**Number of responses:** 85

**Results:** Of the 85 respondents who answered the poll, two-thirds (66%, N=56) indicated that they had made changes in response to the pandemic, while one-third (34%, N=29) said they had not. Despite the fact that the poll questions were specifically about the way organizations had changed their assessments for licensure or registration, respondents took the opportunity not only to discuss changes in assessments but also to describe other organizational adjustments they had made in response to the pandemic.

Respondent organizations listed a number of changes to their assessment processes in response to the pandemic. In some cases, they had been required to make changes to be in compliance with their jurisdiction’s COVID 19-related rules. Respondents mentioned the following types of changes specifically related to assessment or examination:
Recent CLEAR Quick Poll Results

- cancelling or postponing of exams (especially for clinical, oral, or skills demonstrations);
- modifying clinical testing environments to account for distancing requirements and providing PPE as needed;
- implementing computer-based testing (CBT) or accelerating previous plans to transition from paper-and-pencil to CBT;
- moving to a paper-and-pencil administration due to limited seating at computer-based locations;
- increasing the number of locations for testing and extending testing windows;
- extending deadlines for testing;
- waiving clinical skills testing for essential healthcare workers in response to jurisdictional needs; and
- using online live remote or virtual proctoring, including for clinical observations.

Respondents also described a range of changes not related to assessment per se. Many organizations described changes in how they handled administrative responsibilities, particularly those related to documentation. Some moved to e-submission of documentation (including such documents as transcripts and employment verifications). Organizations waived reciprocity to quickly enable qualified healthcare workers from other jurisdictions into the field, did not require criminal background checks because they could not obtain these in a timely manner, and changed the way they confirmed applicants’ identity by using virtual communications rather than live interviews.

Respondents also indicated whether they planned to make the changes permanent, regardless of the course of the pandemic. Of those who had indicated they made changes, 25% (N=14) said Yes, 20% said No (N=11), and 55% (N=31) said they were not sure.

Similar to the responses to whether they had made changes, organizations discussed both their assessment processes and other organizational changes. Among the reasons for making changes permanent were increased efficiencies, particularly in relation to administrative activities such as record-keeping. As one respondent stated, “Eventually, we aim to be completely paper-free with registration and registration renewal.” In addition, for organizations that had already intended to make changes (for example, from paper-and-pencil testing to CBT, reduction in exam length, or moving to remote proctoring), the pandemic merely accelerated the pace of the change. One respondent noted that changes were not only saving the organization money but also saving candidates money and travel time and travel costs.
Recent CLEAR Quick Poll Results

On the other hand, many organizations indicated that they would be reverting to previous administrative processes and procedures once the pandemic eased. Many respondents indicated that changes were driven by state of emergency declarations and they would return to previous processes for both testing and administrative functions once those had lifted. The resumption of clinical testing was mentioned in regard to assessment, and the resumption of criminal background checks and fingerprinting were among the processes mentioned in the context of public protection. As one respondent stated, “A background check and examinations to assess basic professional competence are necessary, and we will return to requiring successful completion of those requirements prior to issuance of a license when those services are more readily available.”

Most respondents who had made changes were unsure if the changes would become permanent. Almost half of those who were unsure did not provide a reason for their response. Of those who provided an explanation, in many cases their “Not sure” response was because the person completing the Quick Poll was not the decision-maker, and the ultimate determination would be made by a higher-up, the organization’s board, or the jurisdictional regulatory authority. One respondent mentioned that seat time issues might ease, lessening the need for online testing, and another discussed how live remote proctoring might be viewed by accrediting bodies as a factor in their decision-making. Some respondents expressed concerns that safety might be compromised due to the changes that had been made.

Clearly, the COVID-19 pandemic has caused tremendous disruption in how regulatory organizations and bodies conduct their essential functions. Unfortunately, the optimism of the poll’s July and August respondents that things might go back to normal sooner rather than later may have been premature. Some of the “temporary” changes may become more entrenched as the pandemic continues to impact the work of regulatory bodies. Adjustment by testing companies to provide safer onsite testing may cause some organizations to return to using those facilities, and improvements in online testing and live remote proctoring may permit other organizations to continue pivoting to this testing modality. In sum, the regulatory environment has been profoundly affected by the pandemic, and regulatory bodies have shifted focus, changed policies and procedures, and met the challenges in several ways, some of which may permanently alter the landscape.

We will have to see how things roll out . . . remote proctoring hasn't been super smooth
Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams

LISA M. ABRAMS, Ph.D., associate professor of research, assessment and evaluation, Virginia Commonwealth University
KATHERINE A. REYNOLDS, Ph.D., assistant research director, questionnaire development and policy research, TIMSS & PIRLS International Study Center, Boston College
MICHAEL O’LEARY, Ph.D., Prometric chair in assessment, Centre for Assessment Research, Policy and Practice in Education, Dublin City University

Introduction
Determining the inferences that can be drawn from test scores and ascertaining appropriate uses for those scores are essential considerations in test design and specification. Articulation of these ideas is commonly referred to as the interpretation and use argument (IUA) and involves providing documentary evidence to demonstrate the validity of the intended inferences, conclusions, and/or decisions made based on test scores. The Standards for Educational and Psychological Testing (AERA et al., 2014), hereafter referred to as the Standards, are especially relevant: “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” (p. 125). Ferrara and Lai (2016) extend this expectation by including evidence required over time and across all stages of test specification, design, development, and implementation. Credentialing exam developers can draw on alignment models dominant in educational (K-12 in particular) testing to meet increased evidentiary expectations at the test specification stage in ways that support the industry’s Standards for the Accreditation of Certification Programs (NCCA, 2014).

The initial test specification stage requires “. . . documentation of the purpose and intended uses of the test, as well as detailed decisions about test content, format, test length, psychometric characteristics of the items and test, delivery mode, administration, scoring and score reporting” (AERA et al., 2014, p. 76). Test specifications identify the content domains measured in relation to the test’s purpose. Credentialing tests are intended to measure the extent to which examinees have acquired the necessary knowledge, skills, and dispositions to engage in safe and effective practice in the workplace. In contrast, achievement tests aim to measure mastery of specific content and skills and are most commonly used in educational settings. Credentialing tests focus on job or occupation responsibilities, while achievement tests center on academic content and skills. In both cases, considerable emphasis is placed initially on defining test content with precision and clarity so that items and tasks can be developed and subsequently used to make valid inferences about examinees (Webb, 2006). However, as Ferrara and Lai (2016) argue, procedures for routine documentation of test content validity evidence are more established in achievement testing than they are in credentialing examinations.

Practices implemented in educational achievement testing can benefit those working in the field of credentialing. Even though the specifics of the test development procedures may differ across these two broad fields, the fundamental principles are similar. Buckendahl (2017) explains that the lines distinguishing
Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams

Educational assessments, credentialing exams, and tests for employment have become increasingly blurred as the use of test results has expanded to serve a range of purposes and stakeholder needs, thus affording an opportunity for credentialing exam developers to consider the applicability and benefits of well-established practices in achievement testing. With this in mind, we first consider current test content specification practices in the credentialing field and then discuss how these can be strengthened by drawing on best practices in achievement testing.

Content Specification in Credentialing Exams

Credentialing exams, including licensure and certification tests, are administered to assess readiness for a professional role or to ensure that workers can carry out job-related responsibilities safely and in accordance with professional guidelines and standards. Specifying the test content, usually in the form of professional knowledge, skills, and judgments (KSJs), prior to item and task development is fundamental to test quality. The IUA for credentialing exams suggests that passing scores demonstrate a level of mastery of the KSJs that denote a degree of competence or readiness to perform a professional practice or activity. A key underlying assumption for the validity argument is that test content and tasks measure the KSJ domains necessary for effective practice. Examining the alignment between test content, tasks, and the KSJ domains is essential to demonstrating that several of these assumptions have been met and the IUA is credible, thus constituting an important source of validity evidence. As described in the Standards (AERA et al., 2014), “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” (p. 178). Accordingly, content specifications for credentialing exams are based on empirical job or practice analysis to determine the knowledge, skills, and attitudes needed to perform workplace responsibilities effectively (Raymond, 2016; Raymond & Neustel, 2006).

Practice analysis requires systematic and thorough methods of questionnaire design, data collection, and analysis in order to arrive at accurate indicators of job duties and successful performance. Approaches to practice analyses vary. They include developing a working theory of the profession, identifying behaviors that enable or constrain performance, and conducting a task inventory (Raymond & Neustel, 2006). Completed practice analysis studies result in process- and content-based specifications, which are similar to blueprints or frameworks that specify the broader content domains measured by achievement tests. A content-process matrix is used to create a framework for item development for credentialing exams. The content-process matrix is one of the more common ways to integrate the work-related knowledge needed to apply this information to authentic tasks or processes (Raymond & Neustel, 2006). Documentation of the content specification process is one source of evidence to support validity and use arguments.

Emerging Practices in Test Documentation

As noted earlier, the Standards (AERA et al., 2014) provide guidelines for best practices in assessment documentation. In addition, the National Commission for Certifying Agencies’ Standards for the Accreditation of Certification Programs (NCCA, 2014) sets out the practices, policies, and procedures programs should have with respect to (1) purpose, governance, and resources; (2) responsibilities to stakeholders; (3) assessment instruments; (4) recertification; and (5) maintaining accreditation (p. 6). However, reviews of documentation and
Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams

technical reporting for individual credentialing exams have found several persistent limitations, including a lack of comprehensive and publicly available documentation for certification and licensure testing programs (Ferrara & Lai, 2016). Ferrara and Lai found that reports based on different types of validity evidence were published at different times and for different audiences. Although general information for test-takers and candidates was often readily accessible, a single comprehensive program document was rarely available and often difficult to obtain. Large-scale and national testing programs had more complete information available (in technical and lay forms), compared to smaller and regional programs. They concluded that most documentation presented technical evidence and information for individual programs in discrete or disconnected ways; a coherent argument or chain of evidence to support claims about interpretations and uses of test scores was often missing.

In an effort to advance the science of test design, Ferrara and Lai (2016) recommend a new framework for test documentation based on the interpretation/use argument, which they call the IUA Report. They describe the report as “a collection of technical, procedural, and other evidence that, taken together, provides a comprehensive and coherent collection of evidence to support claims about how test scores can be interpreted and used” (p. 613). They argue that focusing on capturing evidence at each stage of the test design, development, and implementation process ensures a systematic focus on validity. The IUA Report expands current practices, especially in the credentialing field, and offers new ways of thinking about collecting and presenting evidence in a single document that evaluates purported interpretation and use claims at every stage of the process. Ferrara and Lai’s exemplar IUA Report includes developing claims and providing evidence for the following steps: “(1) determination of testing program policies and articulation of intended interpretations and use of test scores; (2) test design and development; (3) test implementation; (4) response scoring; (5) technical analyses; (6) delivery of scores and other feedback to examinees, candidates, and other test users . . . ; and (7) interpretation of score reports to guide decisions and take other actions” (p. 615-616). They suggest claims at the test design and development stage demonstrate that “item development procedures produce items that elicit evidence of targeted content knowledge and skills” and that evidence for this claim may include “. . . research evidence that the items elicit targeted knowledge and skills” (p. 615). They go on to explain that this claim could be supported by independent reviews as well as depth-of-knowledge judgments by subject matter experts who can determine the extent to which the cognitive process needed to respond to a test item aligns with the required knowledge, skill, or ability (p. 619). The Webb model of alignment, commonly used in educational testing, provides an example of one method to document claims as part of test design and development.

Alignment Studies in Achievement Testing

A series of federal education reform efforts in the United States in early 2000 substantially expanded educational testing programs and required a peer review process to provide different kinds of evidence to support the use of test scores. Alignment between state-developed content standards and their corresponding assessments was required; as a result, empirical studies of alignment became common. Among several models, Webb’s model emerged as the dominant one and was regarded as providing the strongest quantitative information to evaluate alignment based on several criteria (Martone & Sireci, 2009).

According to Webb (1997; 2006), alignment is defined as the degree to which assessment and learning objectives agree. Alignment can be measured using four criteria: (1) categorical concurrence—the
sufficiency of the item sample from the content domains measured by the test, (2) *depth-of-knowledge consistency*—the extent to which test items meet/exceed the cognitive demand expressed in the associated content standard, (3) *range-of-knowledge correspondence*—the diversity of content and extent to which different content standards are represented in the assessment, and (4) *balance of representation*—an index representing the relative emphasis placed on individual standards measured by the assessment (Webb, 2007). Each criterion has different thresholds that can be used to indicate the extent or degree of alignment. Tests are not in alignment or out of alignment—rather, a series of criteria can be used to collectively determine the degree of coherence among standards, assessments, and instruction (Resnick, Rothman, Slattery & Vranek, 2004; Roach, Elliot & Webb, 2003). Webb’s model is consistent with several of the NCCA standards guiding instrumentation development, reporting, and interpretation. For example, the alignment criteria speak directly to one of the essential elements of NCCA’s *Standard 15: Examination Specifications*, which states that “the plan for weighting sections of an examination must be based on a job analysis; the plan must provide precise direction regarding the weighting structure for each section” (p. 22). Drawing on Webb’s approach allows for clear documentation and evidence of the NCCA standards.

The Webb Model of Alignment—Implications for Test Documentation

Implementation of Webb’s model requires two main phases. The first involves making determinations about the level of cognitive demand or depth-of-knowledge (DOK) evident in the content standard.1 This phase requires that content area experts work as a review committee to determine the DOK levels for the content standards by group consensus. Webb developed distinct DOK categories to reflect different levels of cognition: *Level 1*—recall and reproduction, *Level 2*—skills and concepts, *Level 3*—strategic thinking, and *Level 4*—extended thinking. These categories are similar to the concepts in Bloom’s Taxonomy—recall knowledge, comprehension, application, analysis, synthesis, and evaluation. The second phase of the alignment study requires the content experts to closely examine each test item to determine (1) the appropriate content standard/domain the test item is designed to measure and (2) the DOK level of the corresponding item. The reviewers independently code each test item. The reliability of reviewers’ ratings is an important factor in this phase of the alignment review process. The information obtained during phases 1 and 2 can be used to produce and examine the degree of alignment according to the four criteria:

1. **Categorical Concurrence**: This criterion addresses questions about the extent to which the content measured on the test is the same as the content expressed in the content standards or learning objectives. Webb’s evaluation criteria suggest that at least six test items measuring content from a reporting category are needed for a reasonably reliable estimate of students’ content mastery on a subscale or standard (Webb, Alt & Ely, 2005, p. 110). This alignment criterion is closely related to NCCA’s Standards 15 and 20 that address construct specification and score reliability.

2. **Depth-of-Knowledge Consistency**: This criterion is focused on the match or agreement between the cognitive process or DOK expressed in the learning objectives and the aligned test item. Webb et al. (2005) indicates that for depth-of-knowledge consistency to exist between the assessment and the reporting category, at least 50% of targeted objectives should be “hit” by items of the appropriate complexity. Webb’s cut-point is based on the “assumption that a minimal passing score for any one objective/domain of greater than 50% would require the student to successfully answer at least some

---

1 This is a very general description of the Webb alignment model process. For more about Webb’s alignment methodology, see the following 2009 publication: https://www.nagb.gov/content/nagb/assets/documents/publications/design-document-final.pdf
Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams

items at or above the depth-of-knowledge level of the corresponding standard” (Webb et al., 2005, p. 111). Here, this criterion is consistent with essential elements of NCCA’s Standard 15, where the cognitive or performance task required to respond to the test item is clearly associated with “what the examination is intended to measure” such as knowledge, skills, and competency (p. 22).

3. **Range-of-Knowledge Correspondence**: This criterion addresses item sampling and the sufficiency of content coverage. According to Webb et al., 50% of the standards should have at least one related test item in order to determine the range-of-correspondence criterion “acceptable.” Similarly, alignment on this criterion is determined “weakly” met if 41%-49% of the objectives for a reporting category had a corresponding test item and “not met” if less than 41% of the standards had at least one corresponding test item. The ideas of this alignment criterion support essential elements of NCCA’s *Standard 16: Examination Development*, where “the sampling plan for the examination must correspond to the examination specifications” (p. 23).

4. **Balance of Representation**: This criterion relates to range-of-knowledge and addresses the extent to which content is emphasized or distributed on a test. A BOR index is “computed by considering the difference in the proportion of objectives and the proportion of related assessment items for that objective. An index of 1 indicates perfect balance and is obtained if the corresponding items related to a standard are equally distributed among the objectives for a given standard” (Webb et al., 2005, p. 112). Webb suggests that a BOR index of .70 or greater suggests that the criterion has been met.

Table 1 provides a hypothetical example of how the criteria might be applied and could be included as part of an IUA report.

### Table 1. Alignment Criteria Evaluation Categories

<table>
<thead>
<tr>
<th>Alignment Criteria</th>
<th>Strength of Alignment Evidence: Were the Alignment Criteria Met?</th>
<th>YES</th>
<th>WEAK</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorical Concurrency</td>
<td></td>
<td>6 or more test items</td>
<td>4-5 test items</td>
<td>less than 4 test items</td>
</tr>
<tr>
<td>Depth-of-Knowledge Consistency</td>
<td></td>
<td>greater than 50% agreement</td>
<td>41%-50% agreement</td>
<td>agreement of 40% or less</td>
</tr>
<tr>
<td>Range-of-Knowledge Correspondence</td>
<td></td>
<td>50% or greater</td>
<td>41%-49%</td>
<td>40% or less</td>
</tr>
<tr>
<td>Balance of Representation</td>
<td></td>
<td>Values of .70 or greater</td>
<td>Values between .60 and .69</td>
<td>Values less than .60</td>
</tr>
</tbody>
</table>

The following formula is used to compute the BOR index: $\text{BOR} = \frac{1}{\text{H}}\left(\frac{1}{\text{O}} - \frac{1}{\text{H}}\right)^{2}$

- $\text{O}$ – the total number of standards with corresponding test items within a reporting category
- $\text{H}$ – the total number of test items corresponding to a reporting category
Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams

Alignment in Professional Exams

Formal alignment studies appear to be less common in the areas of licensure and certification, perhaps because job analyses and other content domain definition activities are not necessarily explicated through formal standards or policy in the same way as K-12 education. It is also possible that such studies are simply not available to the public. However, the Webb model could be adapted for use in licensure or certification examinations. This section highlights themes from two content alignment studies conducted for credentialing exams for teachers and principals, with particular attention to how the procedures used compare to the Webb model. A summary of these comparisons is presented in Table 2.

In the first of these studies, Reese, Tannenbaum, and Kuku (2015) examined the correspondence between the Praxis Performance Assessment for Teachers (PPAT) and the Interstate Teacher Assessment and Support Consortium (InTASC) Model Core Teaching Standards. The PPAT was explicitly designed to measure the nine InTASC Standards applicable to teacher-candidates that “could be demonstrated during the candidate’s preservice teaching assignment, and could be effectively assessed with a structured performance assessment” (p. 2). In the second study, Swigget (2019) presents the results of a distance-based alignment study concerning the Performance Assessment for School Leaders (PASL) and the Professional Standards for Educational Leaders (PSEL). It is worth noting that both studies concern performance assessments, where assessment tasks differ considerably from the multiple-choice or constructed response items typically used in educational achievement tests.

Both Reese et al. (2015) and Swigget (2019) solicited alignment-related judgments from relevant experts—similar to Webb’s model. Specifically, both studies involved expert judgments to identify the professional standards most associated with each performance task. Reese et al. (2015) had panelists use a 5-point scale (1 = not measured, 5 = directly measured) to indicate whether or not each InTASC Standard was captured by a particular step within a given PPAT assessment task. Standards receiving a 4 or 5 from at least seven panelists were considered aligned to a task step. These results were compared to the test developers’ framework and intended association of the InTASC standard measures by each task step. These procedures are most comparable to the categorical concurrence criterion in the Webb model.

Swigget (2019) prompted review panelists to indicate whether each of three broad PASL tasks were aligned to each of the PSEL standards. PASL tasks include the compilation of “written responses, supporting instructional materials, and artifacts (e.g., student work)” (Swigget, 2019, p. 2). If an overarching PASL task was identified as aligned to a standard, panelists further indicated to which PSEL supporting elements the task was aligned. This approach allowed for calculation of how many PSEL standards were represented by the assessment, as well as the percentage of supporting elements assessed. Again, this might be considered comparable to the Webb model’s categorical concurrence criterion.

Reese et al. (2015) also asked panelists to evaluate the applicability of performance indicators for completion of assessment tasks; scoring rubric adequacy; and the relevance, importance, and authenticity of each PPAT task. Scoring rubric adequacy might be considered comparable to Webb’s DOK consistency criterion, as panelists were prompted to render dichotomous judgments of whether performance descriptions at the highest level for each task step reflected the performance indicators accompanying the standards. However, panelists did not formally classify these indicators along any sort of DOK continuum. These additional judgments imply a
Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams

broader view of “alignment” than the Webb model, invoking not only alignment among assessment tasks and standards but also scoring materials and tasks, as well as assessment content and the tasks routinely performed in the teaching profession. These are examples of additional validity evidence that can be provided in an IUA Report, but they are distinct from content alignment as conceptualized by Webb. Both alignment studies could be strengthened through incorporation of additional Webb model indices, which would yield evidence not only of task-standard alignment but also of cognitive level alignment and overall representativeness of the assessments with respect to the standards.

The lack of publicly available content alignment studies in professional testing represents an opportunity for those working in credentialing and licensure to capitalize on a powerful source of content validity evidence. Normalizing the publication of these kinds of studies might serve to increase public confidence in the appropriateness of credentialing exam content, especially considering the relatively straightforward results obtained from content alignment studies. Use of well-established alignment study models, such as the Webb model, also lends credence and procedural validity to the process.

| Table 2. Summary Comparison of the Webb Model, Reese et al. (2015), and Swigget (2019) |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Alignment Procedures           | Categorical Concurrence         | Depth-of-Knowledge Consistency  | Range-of-Knowledge Correspondence |
| Webb Model                     | Panelists classify items to standards | Panelists classify standards and items by DOK levels | Calculation of the percentage of standards represented by assessment items according to panelist classification | Examines the balance of standard representation across an assessment |
| Reese et al. (2015)            | Panelists use a Likert scale to indicate if a standard is measured by a task | Not directly addressed—tangentially addressed by panelists’ judgments of scoring rubrics compared to performance indicators | Calculation of total number of InTASC standards represented across the PPAT tasks | Not directly addressed |
| Swigget (2019)                 | Panelists provide yes/no judgments about task alignment to each standard | Not directly addressed | Calculation of the percentage of standards/supporting elements represented by assessment items according to panelist dichotomous judgments | Not directly addressed |

Conclusions
This paper describes how credentialing test developers can respond to increased calls to enhance technical documentation and reporting in ways that support an interpretation and use validity argument at all stages of
Advancing Alignment Arguments in Supporting Scoring Interpretation and Use Claims of Credentialing Exams

the testing program. Specific attention to the content specification stage is essential for designing high quality assessments. The use of systematic methods, such as implementing a performance or job analysis, heightens the validity of the inferences made based on credentialing exam scores. Especially in cases where these inferences can lead to high-stakes consequences for examinees and require strong assurances of validity evidence, Webb’s model provides an approach to bolster technical documentation. We have argued that while alignment studies are not currently a common feature of the credentialing exam development process, there is much to be learned from such studies in educational K-12 testing, where content domains and standards are more clearly explicated. Webb’s model provides test developers with a clear set of quantitative indicators that can be used to inform the initial stages of test development and demonstrate validity evidence for test content. In helping to mitigate validity challenges, the model has implications for how legal defensibility in testing might be conceptualized and for how quality assurance during important stages in the test development process can be actioned.
Reference List


Abstract

Five years have passed since the National Board for Certification in Occupational Therapy (NBCOT®) launched the NBCOT Navigator™, an innovative virtual continuing competency assessment platform built on gaming technology. Developing the Navigator required a significant investment of time and resources that continued once it was launched. In this article, we will review how we have maintained the Navigator over the past five years. We will share information on its use and adoption by key stakeholder groups, its tool development, and the critical components we put in place to safeguard the platform's continued validity and relevance. We hope sharing this review of the key elements that made the platform successful will assist other certification organizations considering the development of a virtual competency assessment platform.

Introduction

Certification programs typically have some form of recertification program for which certificants must meet requirements for maintaining or enhancing their competence before renewing their certification. In addition, major accreditation standards for certifying bodies (e.g., NCCA Standards for the Accreditation of Certification Programs, 2014; ISO/IEC 17024:2012; ABNS Accreditation Standards for Examination-Based Certification Programs, 2016) have standards that link recertification to the assessment or promotion of continuing competence. Although the concepts of recertification and continuing competence are now mainstream in the field of credentialing, the question of how to design a recertification program that engages certificants in a continuous cycle of learning and assessment remains a widely discussed and debated topic (Granatir & Robinson, 2019).

Stakeholders, including the public and employers, expect certified individuals to engage in ongoing education and development to stay updated on new advances and technology in their field, as well as to maintain...
One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On

a foundation of core knowledge, skills, and abilities throughout their career. In recent years, however, certificants and the public alike have begun to question the validity and fairness of the long-standing practice of using a traditional high-stakes exam to demonstrate continuing competence (Granatir & Robinson, 2019). As a result, several certification programs have begun to explore innovative approaches to their recertification programs (e.g., Leslie et al., 2018; Sun et al., 2016).

In 2012, the National Board for Certification in Occupational Therapy (NBCOT®) took a strategic approach to creating a new product for its certification renewal program. The goal was to develop innovative tools for certificants to assess their current practice knowledge and experience new avenues for professional growth. We wanted the tools to be fun, engaging, continuously accessible on a digital platform, and free of charge for current certificants to use. With these goals in mind, NBCOT partnered with a serious gaming company, BreakAway Games, to help create the NBCOT Navigator™. The Navigator uses gaming technology to present a variety of assessment tools that allow certificants in NBCOT’s Occupational Therapist Registered (OTR®) and Certified Occupational Therapy Assistant (COTA®) certification programs to assess their knowledge across numerous practice areas.

The NBCOT Navigator was launched in 2015. Like other successful certification products, we were cognizant of the need to have a robust program in place to support the Navigator’s continued relevance and validity, as well as a strategy for supporting its growth and enhancement. In this article, we review the steps we took to implement, maintain, update, and expand our innovative competency assessment platform. Many of the considerations we share are applicable to other virtual competency assessment platforms, regardless of the format. We hope that by sharing information about our platform’s use and adoption by key stakeholder groups, its tool development, and the critical components put in place to preserve the integrity of the product, we will accurately depict the time and investment required to implement and maintain this type of platform.

Background of the NBCOT Navigator

The Navigator initiative was partly a response to two influential reports by the Institute of Medicine (IOM, 2003, 2010), whose vision for continuing competence for healthcare professionals was centered around the competencies required to meet patients’ needs, including delivering “patient-centered care as members of an interdisciplinary team, [and] emphasizing evidence-based practice, quality improvement approaches, and informatics” (IOM, 2003, p. 3). These competencies were again recommended in IOM’s 2010 report, Redesigning Continuing Education in the Health Professions, as an appropriate basis for building a continuing education system in healthcare. In 2012, NBCOT conducted its first practice analysis focused on occupational therapy (OT) practice at the point of certification renewal. This practice analysis resulted in six domains that drew on the IOM competencies and a validated task list that was applicable to both the OTR and COTA programs. This certification renewal content outline formed the basis for building content for the NBCOT Navigator.

The Navigator incorporates a web-based assessment delivery engine, a certificant dashboard, and an interface that supports a suite of competency assessment products (for a detailed description of the Navigator’s tools and features, see McNamara et al., 2019). Certificants access the Navigator by first completing a self-

1 A short video describing the features of the Navigator tools may be viewed here: https://vimeo.com/304878367. See also the description on the NBCOT website: https://www.nbcot.org/Certificants/Navigator#overview. Note that the actual tools are only available to certificants via their MyNBCOT account.
reflection activity that focuses on their current and future professional practice goals. This activity results in a customized list of tools for certificants to address their competency assessment needs. Each tool provides virtual opportunities for certificants to engage in practice-related assessment experiences that reflect the real-world demands of current healthcare environments. Upon completion of the tools, certificants receive feedback on their performance, along with links to recommended evidenced-based articles and free resources that are accessible from certificants’ accounts. For successfully completing the tools, certificants earn units, called competency assessment units (CAUs), that can be used toward certification renewal requirements.²

Growth Over Time

When it launched, the Navigator contained four distinct competency assessment products: case simulations, the PICO (Patient/Problem, Intervention, Comparison, Outcome) Tool, mini practice quizzes, and the OT (Occupational Therapy) Knowledge Library. Since 2015, NBCOT has added three additional products to the platform: Orthotic Builder, Physical Agent Modalities, and Management Challenge (see Table 1). Each of these products contains separate assessment tools.

In addition to developing more competency assessment products for the Navigator, NBCOT has increased the number of assessment tools each product contains (see Table 2). For example, the case simulations have more than doubled from 14 in 2015 to over 30 in 2020. Additional case simulations and mini practice quiz tools are scheduled for release in early 2021, with further expansion of tools and products planned for the future. This expansion is in response to trends in OT practice, feedback from subject matter experts, and survey results from certificants in the field.

Table 1. NBCOT Navigator Products

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Simulation</td>
<td>An animated tool that allows certificants to provide OT services to virtual clients in a variety of practice settings—acute care, rehabilitation facilities, community, home, and work environments. Certificants progress through modules that encompass the full spectrum of the OT process: chart reviews, assessments, intervention planning, providing OT services, and discharge planning.</td>
</tr>
<tr>
<td>PICO Tool</td>
<td>A product that simulates the process of completing an evidence-based literature search to answer patient-related health questions. Certificants work through modules to form an effective search question, complete an evidence-based literature search, and identify the best evidence to answer specific health questions posed by a community of aliens living in outer space.</td>
</tr>
<tr>
<td>Mini Practice Quizzes</td>
<td>Short multiple-choice quizzes grounded in evidence-based literature that assess certificants’ knowledge of contemporary OT practice. The quizzes cover major OT practice areas including pediatrics, school systems, administration/management, acute care, rehabilitation, education/research, work/industry, wellness, and home health.</td>
</tr>
</tbody>
</table>

² In addition to the Navigator, NBCOT certificants may choose from over 30 different options for earning units towards their renewal (e.g., workshops/courses, conference presentations, publishing, professional service, fieldwork supervision). There are limits to how many units may be accrued via certain renewal activities, including the Navigator, which is capped at 14 out of the required 36 units in the three year renewal cycle.
## One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On

### Table 1. NBCOT Navigator Products cont.

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OT Knowledge Library</td>
<td>A series of tools in which a clinical definition, symptom, intervention, or other prompt must be correctly categorized. Certificants can use these tools to assess their ability to apply appropriate practice-based OT knowledge, identify areas for further learning, and access evidence-based resources.</td>
</tr>
<tr>
<td>Orthotic Builder</td>
<td>An animated tool set in a mock village. Certificants work through modules to fabricate an orthosis to address the specific needs of the village workers, who are experiencing an upper extremity injury or condition. The village regains its productivity as each worker receives the correct orthosis.</td>
</tr>
<tr>
<td>Physical Agent Modalities</td>
<td>A tool for which certificants must make informed decisions to select and administer the correct physical agent modality to clients in an outpatient rehabilitation facility who have a variety of conditions.</td>
</tr>
<tr>
<td>Management Challenge</td>
<td>A tool that allows certificants to assess their skills in managing the day-to-day operations of a busy outpatient rehabilitation facility. Certificants are presented with challenges that OT managers frequently encounter, such as organizing staff schedules, responding to unexpected events, and staying updated on facility policies and procedures while maintaining a successful work environment.</td>
</tr>
</tbody>
</table>

### Table 2. Growth in Navigator Products, 2015–2020

<table>
<thead>
<tr>
<th>Product</th>
<th>Number of Tools in 2015</th>
<th>Number of Tools in 2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Simulations</td>
<td>14</td>
<td>33*</td>
</tr>
<tr>
<td>PICO Tool</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Mini Practice Quizzes</td>
<td>28</td>
<td>28**</td>
</tr>
<tr>
<td>OT Knowledge Library</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Orthotic Builder</td>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>Physical Agent Modalities</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>Management Challenge</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>67</td>
<td>104</td>
</tr>
</tbody>
</table>

*6 new case simulations were developed in 2020 for launch in 2021.
**17 new mini practice quizzes were developed in 2020 for launch in 2021.
One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On

Increase in Tool Use Over Time

The mini practice quizzes, case simulations, and OT Knowledge Library are the most popular products to date (see Figure 1). These products were some of the first to launch in the Navigator.

![Figure 1. Increasing Use of Top Navigator Product, 2015–2019](image)

The total number of tools certificants interact with rose as familiarity with the products increased. From 2016—the first full year the Navigator was available to certificants—to 2019, there was a 70% increase in the number of tools used. There was also a steady increase in the number of units certificants earned toward certification renewal. As shown in Figure 2, certificants accrued just over 11,000 CAUs (Competency Assessment Units) in 2016; but by 2019, the number of units accrued increased to over 26,000. In addition to growing familiarity with Navigator products, this increase in use is also due to marketing efforts, more states adopting the Navigator tools as an option for licensure renewal, and an expansion of tools available to certificants.
 Adoption by the Regulatory Community

Since launching the Navigator, NBCOT has been conducting outreach work with state regulatory boards to discuss the possibility of accepting the Navigator as an option for licensure renewal. We have a dedicated team focused on communicating with various stakeholder groups, including occupational therapy state regulatory boards, to build strong outreach relationships. Through these efforts, the boards were asked to consider recognition and acceptance of the NBCOT Navigator as an optional activity for licensure renewal purposes. In 2015, West Virginia became the first state to accept units accrued from the Navigator as an option for licensure renewal. Nineteen states quickly followed suit that year, and by 2016, 14 more states adopted the Navigator. In 2017, four states joined in, and five additional states agreed to accept the Navigator by mid-2020. To date, a total of 43 states, plus the District of Columbia, recognize the Navigator as an option for partial completion of licensure renewal requirements. Of the remaining states, four do not have continuing education requirements in their regulations, and we have been communicating with the other three on the potential for future adoption.
One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On

Technology Growth

When the Navigator launched, it was only made available on computers. Although certificants have been requesting access to the Navigator via tablets since then, we were initially hesitant to do so due to security concerns. However, with advances in security technology and market trends, including the ubiquitous use of tablets as an extension of—or replacement for—personal computers, NBCOT moved forward in 2018 to make the Navigator available on tablets.

Once this objective was established, NBCOT and our partner, BreakAway Games, collaborated on the design for application interface functionality options, graphical user interface mockups, and icon-based player progress maps. Beginning in April 2019, all current and in-development content for Navigator tools was reviewed, edited, and enhanced to prepare for the tablet version. Early content reviews revealed several issues that needed to be resolved for optimal playability on the tablet platform. The most significant of these were increasing the text size and contrast for readability and the need for closed captioning when using the mute function to enhance the play experience for all users. We completed the Navigator application in 2020, and it is now available to download on tablets through Apple or Google Play.

Maintaining the Navigator

Maintaining a platform like the Navigator is an ongoing process that requires a dedicated team and ongoing investment of time and resources. One of the main goals of a continuing competency program is ensuring that professionals remain current on the latest developments in their field. Therefore, in addition to staying updated on changing technology, the Navigator content must also be current. A major part of confirming that Navigator tools are aligned with current occupational therapy practice is making sure that they are based on a current practice analysis.

In 2018, NBCOT convened a panel of subject matter experts to review key aspects of NBCOT’s certification renewal program. The panel reaffirmed the applicability of the 2012 domains and tasks for both the OTR and COTA certification programs; however, the panel suggested minor modifications, which indicated the need for an updated practice analysis.

In 2019, we conducted a second practice analysis focusing on ongoing professional practice. Guided by feedback from the 2018 certification renewal review panel and a thorough review of the existing certification renewal content outline, the practice analysis panel upheld the six domains of ongoing professional practice but clarified the domain names and definitions and updated the tasks within each domain to better reflect occupational therapy practice in 2019. We then created a national validation survey and sent an invitation to all certificants who had been certified for three or more years (53,037 COTA and 112,287 OTR certificants). Of those invited, 11,848 OTR and 3,552 COTA certificants accepted the invitation and agreed to participate in the survey in exchange for one Professional Development Unit (PDU). A total of 5,995 OTR and 1,560 COTA certificants who received the survey submitted valid responses, resulting in final effective response rates of 51% and 44%, respectively. The results of the validation survey supported the domains and tasks identified by the ongoing professional practice analysis panel and provided valuable information about current practice trends and demographics.
One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On

The 2019 practice analysis documented the link between current professional practice and NBCOT’s certification renewal program, and the results are used to inform updates to the Navigator’s content. In addition to the practice analysis study, we regularly seek input from stakeholders on practice trends that guide content discussions for new tools to add to the platform.

Because all the Navigator tools are evidence-based with relevant peer-reviewed journal articles provided to users after completing the tools, the Navigator team needs to stay up to date on the latest empirical research. NBCOT has an in-house team of occupational therapists who work exclusively as Navigator content developers and who collaborate closely with panels of subject matter experts to link every tool to evidence-based, contemporary occupational therapy practice.

We routinely use surveys to help identify topics for new tools. By gathering input from our Navigator content development team, subject matter experts, and survey respondents, we have a sound basis for developing new tools that are contemporary and relevant to certificants. Recent topics stakeholder groups suggested that were converted to new tools include telehealth, health literacy, and practice ethics.

Challenges We Encountered

The NBCOT Navigator offers many benefits to certificants, such as allowing them to assess their knowledge of contemporary occupational therapy practice across a wide variety of topics and settings and earning units that they can use toward certification renewal. However, we faced a few challenges before achieving these successes. Some of the earliest hurdles included raising awareness about the platform among stakeholders, including the regulatory community; showing how certificants can benefit from the Navigator; and demonstrating how to use it. Communicating effectively about these points required ongoing marketing and public relations efforts ranging from creating regular social media content to presenting and exhibiting at professional conferences.

Another hurdle we faced was the time it takes to develop new tools, particularly for more technologically advanced products like the case simulations. This issue is especially relevant for a continuing competency assessment platform, given the focus on ensuring certificants are up to date on the latest practice knowledge. The time it takes to go from draft to design to delivery means that we have to be proactive and adhere to tight production schedules if we are to make tools available to certificants in a timely fashion. Additionally, we have to be flexible and able to adjust production schedules to accommodate for any new practice updates or unexpected changes to regulations.

We also need to have the capacity to respond to questions about the platform and how to use it. Not all certificants are technologically savvy, nor do all certificants possess the appropriate technology to take advantage of gamified tools. In short, part of maintaining any innovative virtual competency assessment platform entails having staff who can respond to questions about it.

Where We Are Today

Since its launch in 2015, we have seen a steady increase each year in the number of certificants using the Navigator to accrue units towards their certification renewal. Indeed, as of early September 2020, we have seen a 23% increase in the number of certificants using the Navigator to earn units towards their certification renewal this year compared to the whole of 2019. As more states accept Navigator tools as an option for licensure renewal
One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On

and as awareness of the tools increases, we anticipate that usage of the Navigator tools for renewal purposes will increase further.

A recent survey of 1,595 Navigator users indicated that 52% of respondents agreed and 39% somewhat agreed that the NBCOT Navigator helped them identify areas of interest for their continued competency development. These results are encouraging, as are the results of questions designed to gauge certificants’ satisfaction with the Navigator (see Figure 3) and its effectiveness at improving the way certificants engage in self-directed learning and apply new knowledge to their practice (see Figure 4). We use these survey results to help target future improvements to the Navigator.

The findings of our recent internal market research surveys are consistent with the results of a study that explored U.S. occupational therapists’ perceptions of online competence assessment and associated evidence-based resources (Myers, 2019). This survey involved 266 occupational therapists who were certifying or recertifying in March of 2016. A majority of participants agreed or strongly agreed that the Navigator’s tools helped them identify clinical knowledge (80%) and skill (61%) needs. Less experienced occupational therapists (≤ 5 years of experience) were significantly more likely to agree that the Navigator helped them to identify clinical skill areas that needed improvement than more experienced occupational therapists (≥ 6 years of experience). Of the 46% who obtained copies of the evidence-based resources, 73% read at least one. And, of the 58% of participants who said they maintained a professional development plan, slightly more than half (52%) were likely or extremely likely to modify their plan based on their participation in Navigator activities.

Figure 3. NBCOT Navigator 2019-20 User Survey Results, Navigator Characteristics

To what extent do each of the following characteristics describe the NBCOT Navigator?

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Describes very well</th>
<th>Somewhat</th>
<th>Does not describe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fun to complete</td>
<td>41%</td>
<td>42%</td>
<td>17%</td>
</tr>
<tr>
<td>Dynamic, interactive</td>
<td>56%</td>
<td>34%</td>
<td>10%</td>
</tr>
<tr>
<td>Short, quick to complete</td>
<td>45%</td>
<td>33%</td>
<td>21%</td>
</tr>
<tr>
<td>Up-to-date/Based on the latest research</td>
<td>72%</td>
<td>26%</td>
<td>2%</td>
</tr>
<tr>
<td>Easy to access from anywhere</td>
<td>64%</td>
<td>27%</td>
<td>9%</td>
</tr>
<tr>
<td>Useful to set competency goals</td>
<td>48%</td>
<td>41%</td>
<td>11%</td>
</tr>
<tr>
<td>Available 24/7</td>
<td>90%</td>
<td>8%</td>
<td>2%</td>
</tr>
<tr>
<td>Relevant to my OT practice</td>
<td>56%</td>
<td>39%</td>
<td>4%</td>
</tr>
</tbody>
</table>
One Organization’s Journey to Implement an Innovative Competency Assessment Platform: The NBCOT Navigator Five Years On

Figure 4. NBCOT Navigator 2019-20 User Survey Results, Practice Improvements

To what extent did use of the NBCOT Navigator encourage you to . . .

- Seek out additional information on certain topics: 33% (To a great extent), 44% (To a moderate extent), 23% (To a small extent)
- Make changes to the way you practice: 16% (To a great extent), 40% (To a moderate extent), 43% (To a small extent)
- Apply evidence supplied at the end of play to your practice: 34% (To a great extent), 41% (To a moderate extent), 26% (To a small extent)
- Engage in self-directed learning: 42% (To a great extent), 40% (To a moderate extent), 18% (To a small extent)
- Reflect on ways to change or improve your current practice: 23% (To a great extent), 44% (To a moderate extent), 33% (To a small extent)

Conclusion

Many certification organizations are exploring innovative approaches to developing a recertification program that promotes continuous learning and assessment of competencies. We hope that by sharing our experiences developing and maintaining the NBCOT Navigator over the past five years, certification organizations exploring a similar approach to continuing competency will gain an understanding of what resources are required to manage such a program.

Once a virtual competency assessment platform has been implemented, certification programs should prepare for the possibility that the platform’s growth and widespread adoption by certificants and other stakeholders may be gradual and may require a significant amount of outreach and marketing. Organizations that are considering implementing this type of platform should also be aware that ensuring that the platform remains current—both the content and the technology—is a significant commitment that may require a dedicated group of staff and volunteer subject matter experts. It is also prudent to mention that developing gamified tools that are creative, engaging, and fun typically requires a partnership with a vendor experienced with serious game development. Building a continuing competency assessment platform such as the NBCOT Navigator requires a big investment of time and other resources up front and on an ongoing basis, but the benefits to certificants and other stakeholders are clear.


