Association of Regulatory Boards of Optometry

ARBO
Association of Regulatory Boards of Optometry, Inc.

3440 Toringdon Way
Suite 205 PMB #20533
Charlotte, NC 28277

Tel: (704) 970-2710
Fax: (888) 703-4848
Email: arbo@arbo.org

November 29, 2024

To: Senator Stephen West, Co-Chair Kentucky Administrative Regulation Review Subcommittee
Representative Derek Lewis, Co-Chair Kentucky Administrative Regulation Review Subcommittee
Emily Caudill, Kentucky Administrative Regulation Compiler
Ange Darnell, Kentucky Assistant Administrative Regulation Compiler
Christi LeMay, Executive Director Kentucky Board of Optometric Examiners

The Association of Regulatory Boards of Optometry submits this position statement in opposition to proposed changes to 201 KAR 5:010 (or "Bill").

The Association of Regulatory Boards of Optometry (ARBO) is a 501(c)(3) not-for-profit organization incorporated in 1919 whose membership is comprised of the state and provincial regulatory boards of optometry, including the Kentucky Board of Optometric Examiners (KBOE). ARBO's mission is "to represent and assist member licensing agencies in regulating the practice of optometry for the public welfare." An additional core component is that ARBO promotes uniformity of licensure requirements with due respect for the rights of the states and provinces to enact relevant legislation. The purpose of this submission is to express opposition to and concern over proposed rule changes to 201 KAR 5:010.

Optometry boards have a legal responsibility, similar to other healthcare boards, to recognize valid exams that measure minimum competence. Valid exams are those that can be legally defended if questioned, meaning they are developed, administered, scored, and maintained according to industry standards. These standards allow the regulatory boards to view the exams as objective measures that offer legal protections to all applicants for licensure. Passage of these exams is one requirement for eligibility to become licensed.

The proposed amendment to 201 KAR 5:010 provides for KBOE's recognition of an alternative examination, the Optometry Examining Board of Canada (OEBC) written examination, to the existing National Board of Examiners in Optometry (NBEO) Part 1 examination. It's important to note that the amendment mistakenly refers to the "Canadian OBEO," which is not an exam offered by the OEBC; this letter will refer to the OEBC examination. While ARBO respects the OEBC and its exam, there are many key differences to consider when evaluating the merits of this alternative exam pathway.

First, the NBEO Part 1 Exam and the OEBC Written Exam are not equivalent. The NBEO Part 1 consists of 350 questions that test a candidate's understanding of the basic science concepts necessary for entering clinical optometry. In contrast, the OEBC Written Exam uses case-based scenarios to evaluate how candidates respond to clinical situations. These exams assess different knowledge and skills. If candidates were to skip the NBEO Part 1 Exam, they could become licensed without demonstrating their understanding of critical scientific concepts essential for providing safe optometric care. Allowing an alternative assessment that does not test the same content undermines the goal of ensuring entry-level competence.

Second, it's important to understand the steps taken to validate the NBEO Part I Exam, which also helps defend its legality if challenged. The NBEO follows industry standards to ensure its exam is reliable and can be used by the KBOE to determine whether applicants meet legal requirements. NBEO is prepared to provide data to support their exam if necessary. These validation processes include surveys of U.S. practitioners, identifying key content areas, developing exam questions, performing statistical analyses on question

performance, and setting passing scores. ARBO suggests that the KBOE investigate whether the OEBC has followed similar validation measures and whether they can be defended in the U.S. However, as mentioned earlier, the two exams assess very different content areas.

To highlight KBOE's reasoning in the Regulatory Impact Analysis and Tiering Statement for the proposed changes, which simply states "…many individuals in Kentucky have failed the NBEO Part I Exam multiple times." This reasoning lacks substance and context, and it may imply a "lowering" of standards for licensure eligibility. Competence exams are designed to differentiate between those who meet minimum competence to enter into practice and those who do not. As a criterion-referenced exam, candidates are measured against a set standard determined by survey results. A candidate's ability to pass the exam reflects their knowledge of the subject and their overall preparation. Understanding the basic sciences in optometric care is crucial for ensuring public safety.

ARBO recognizes individual states' rights and also assists our members in promoting uniformity in licensure requirements. However, accepting alternative examinations could undermine the mobility and portability programs that exist in state statues. By allowing applicants to take alternative pathways as proposed, Kentucky may create a situation where these licensees are isolated because of the unique requirements in Kentucky that differ significantly from those of other optometry boards in the U.S. All states require passing Part I for initial licensure and may not consider the OBEC written exam as an acceptable alternative for licensure by endorsement.

Additionally, ARBO's National Board Examination Review Committee (NBERC) provides oversight of the NBEO Exams on behalf of our members. NBERC's purpose is to review the content and methodology of all of the NBEO examinations to ensure they are current and appropriate for use by ARBO member boards. The NBERC is made up of members from ARBO's member licensing Boards. NBERC provides a report to members every year after reviewing the NBEO Exams. As a member of ARBO, the KBOE has been invited to participate on the NBER Committee in the past to gain first-hand knowledge of how the NBEO Exams are developed and discuss any concerns they may have. If the KBOE were to agree to accept the OEBC written exam, it is unlikely that they would have oversight or participation in the OEBC exam program.

ARBO respects all Member Boards and the variations to licensure requirements that may best serve the public and consumers of optometric services, but is opposed to the proposed change to 201 KAR 5:010 for several reasons, including:

- Lack of Standardization: Licensing exams must meet certain standards to ensure they accurately assess the knowledge and skills required for practice. An exam that is not nationally recognized and has no oversight in the U.S. may not adhere to these standards, leading to inconsistencies in evaluating a candidate's competence.

- Incompatibility with Other States: If an alternative exam is not recognized by other states or regulatory bodies, it can create barriers for optometrists who want to practice in multiple jurisdictions. This lack of recognition can limit career mobility and hinder the ability of licensed professionals to work across state lines.

- Potential Legal Issues: If questions arise about the validity of an unrecognized exam, it may be difficult to defend the results in a legal context. This could lead to challenges in upholding licensure decisions based on that exam.

- Public Safety Concerns: Licensing authorities prioritize public safety, and the quality of care provided to patients. If an exam is allowed which is not equivalent to the nationally recognized exam, there may be concerns about whether it adequately evaluates a candidate's competency, potentially putting patients at risk.

ARBO asks that the KBOE explore these issues and the potential consequences prior to promulgating the proposed changes. ARBO and NBEO would be happy to present a more thorough analysis of the issues identified in this submission and answer questions of the KBOE. This type of information exchange and interaction is precisely the purpose of ARBO. We look forward to hearing from you and exploring opportunities to discuss this letter.

Sincerely,

Lisa Fennell, ARBO Executive Director/CEO

*On behalf of the ARBO Board of Directors:*

Jeffery Yunker, OD, President
Terri Haley, OD, Vice President
Margaret Whelan, MPA, Secretary-Treasurer
Eric Bailey, OD, Director
James Campbell, OD, Director
Luanne Chubb, OD, Director
Glenn Kawaguchi, OD, Director
Gerard Lozada, OD, Director
Patrick O'Neill, OD, Director
Linda Tharp, OD, Director

NATIONAL BOARD OF EXAMINERS IN OPTOMETRY

November 27, 2024

TO: Senator Stephen West, Co-Chair Kentucky Administrative Regulation Review Subcommittee
Representative Derek Lewis, Co-Chair Kentucky Administrative Regulation Review Subcommittee
Emily Caudill, Kentucky Administrative Regulation Compiler
Ange Darnell, Kentucky Assistant Administrative Regulation Compiler
Christi LeMay, Executive Director Kentucky Board of Optometric Examiners

FROM: Michael Ohlson, OD, FAAO, Dipl. AAO, National Board of Examiners in Optometry (NBEO) President
Jill Bryant, OD, MPH, FAAO, FSLS, FASOS, NBEO Executive Director
Brett Foley, PhD, NBEO Director of Psychometrics and Research

**RE: Proposed Regulatory Amendment to 201 KAR 5:010**

The National Board of Examiners in Optometry (NBEO®) submits the following information concerning proposed regulatory amendment 201 KAR 5:010 (attached as **Exhibit 1**). If adopted, this amendment would allow a person who wants to practice optometry in Kentucky to obtain a license without taking an exam that has questions about foundational science concepts. NBEO is not aware of any other U.S. jurisdiction that does so. NBEO feels strongly that a strong grasp of foundational science concepts is a prerequisite to providing vision and eye care to citizens of every state. If adopted, Kentucky would become the **only** U.S. state to allow for the omission of proof of fundamental understanding of basic science during the licensure process in optometry.

**Key Points From This Letter:**

- The NBEO Examination Series represents one examination delivered in three parts.
- Each examination part represents a separate and unique measurement of candidate knowledge. The sum of these examination parts represents a holistic measurement of optometric competency that regulatory boards of optometry can utilize in applications for licensure.
- The NBEO's comprehensive examination process ensures that the citizens of Kentucky receive appropriate and safe health care, and we do not believe that perceived low passage rates justify an alternative testing pathway that is less comprehensive.
- NBEO's data and research does not support excluding one or more parts of the NBEO exam licensure series.
- Pursuant to the Six-Time Limit Policy, NBEO allows an applicant to take the NBEO examination a maximum of six times. The proposed amendment invites an applicant who has exhausted their attempts under the Six-Time Limit Policy to obtain licensure, nonetheless. NBEO has serious concerns that, by doing so, an unqualified candidate could obtain licensure. We trust that the board shares NBEO's deep concerns about the consequences of unqualified candidates offering eye-care treatment.
- Critically, the OEBC exam series does <u>not</u> include basic-science testing.
    - The NBEO examination, in contrast, <u>does</u> include questions about basic sciences—in particular, anatomy, physiology, pathology, and pharmacology—because they are foundational to practicing ocular health. These sciences are prerequisites for comprehension of eye structure and function, as well as the complex

interplay between the eye and other bodily systems. While the NBEO Part I Applied Basic Science exam tests important basic science information, in all questions these topics are presented with direct optometric applications to clinical care. In other words, it is not just basic science information, but basic science in the context of real-world optometric situations.
    - As health care practitioners, optometrists have a role to play in diagnosing and identifying signs of disease in other bodily systems when caring for their patient's ocular health. This foundational knowledge also helps optometrists understand drug interactions, side effects, and contraindications—understandings that are important for ensuring patient safety. This knowledge also supports optometrists' ability to stay current with industry advancements in technology and research.
    - In sum, a strong grasp of basic sciences is needed for optometrists to effectively recognize, diagnose, and manage a wide range of ocular and systemic conditions.
- Finally, OEBC allows candidates to take their written examination through Remote Proctoring (online) whereas NBEO requires testing to be done in a Professional Testing Center. NBEO has intentionally avoided utilizing Remote Proctoring due to security and exam integrity concerns which are supported by data (see section titled "Remote Proctoring v. In-Person Testing" below).

**BACKGROUND**

The National Board of Examiners in Optometry, a 501(c)3 non-profit corporation, was founded in 1951 from the Association of Regulatory Boards of Optometry (ARBO) and the Association of Schools and Colleges of Optometry (ASCO). **NBEO's mission is to protect the public by developing, administering, scoring, and reporting results of valid examinations that assess competency in optometry.** NBEO's founders understood the value of licensure assessment as a public protection measure to guard against optometry school graduates who may not be ready for entry into the profession of optometry.

In the early days, state boards of optometry each administered their own examinations. The state boards subsequently decided to follow the path of other healthcare professions and have one national organization dedicate the entirety of its resources towards examinations that evaluate optometric competency to enter the healthcare profession. Since that time, all U.S. state optometry boards have selected NBEO to develop and administer the examinations to assess competency in optometry. NBEO follows best practices consistent with high stakes licensure examinations within the testing community (see **Exhibit 2**).

**NBEO EXAMINATIONS ARE A HOLISTIC ASSESSMENT OF OPTOMETRIC COMPETENCY**

Since the primary purpose of the collective NBEO exams is to protect the public, our goal is to ensure that individuals who are granted an optometric license possess the knowledge, clinical thinking, and skills (i.e., "expertise") necessary to practice optometry in a competent manner. This goal is accomplished by differentiating, via board examinations, those who ***do*** have the necessary expertise from those who ***do not***, at the level of the Minimally Qualified Candidate. NBEO examination results are used for licensure purposes, and must be sufficiently demanding, rigorous, and reflective of the most challenging aspects of the common scope of current optometric practice to meet the "good faith" expectations of public protection among those being licensed. NBEO acknowledges the high reliability and validity standards that state boards expect of NBEO when producing examinations.

The NBEO series of licensure exams consists of three separate parts.[1] Part I Applied Basic Science (ABS®) is a multiple-choice, computer-based exam administered at Pearson Vue testing centers that assesses candidates' mastery of the

[1] The content matrices for each examination can be easily accessed at www.optometry.org, listed under the first tab, "NBEO Exams."

underlying basic science concepts necessary for entry into optometric practice. The exam consists of 370 questions, 20 of which are unscored, pre-test items, and is administered in two sessions of 4 hours each. Every examination item has a direct correlation to optometric patient care.

Part II Patient Assessment and Management (PAM®) examination assesses clinical thinking and decision-making, along with knowledge of diagnosis and treatment. The Part II PAM exam is also a computer-based, multiple-choice exam administered at Pearson Vue testing centers. It contains 350 items and is administered over two sessions of 3.5 hours each. Part II PAM questions frequently are shown as part of an overall case wherein candidates are given clinical information, sometimes including diagnostic images. The questions for the case follow a sequence that mimics clinical thinking and decision-making; however, examinees are able to select from a list of possible answers while thinking through the case and appropriate treatment steps. The Treatment and Management of Ocular Disease (TMOD®) examination can be completed as part of Part II PAM (embedded within the exam) or may be taken as a standalone examination.

Lastly, Part III Patient Encounters and Performance Skills (PEPS®) is a performance-based exam wherein examinees are required to interact with live Standardized Patients to simulate clinical encounters or to perform optometric clinical skills that reflect practice. During the examination, which is administered at NBEO's National Center of Clinical Testing in Optometry, candidates rotate through a series of 12 stations: 10 stations are clinical encounters, and 2 stations are clinical skills. Candidates have 15 minutes at each station making the examination roughly six hours long including time for check-in, orientation, and checkout. Each station is housed within an examination room that is designed to simulate real-life optometric exam rooms. The equipment, placement of materials, and room dimensions are standardized, and the NBEO follows a multilayered protocol for quality assurance throughout the examination process.

Given that every knowledge, skill, and ability necessary for entry into the independent practice of optometry cannot be tested in the same format, the examination series provides a scaffolded path for the assessment of overall competency. Figure 1 provides a graphical representation of this holistic assessment. Each exam within the series covers an aspect of optometric competency, but it is the combination of the series of exams that represents overall competency.

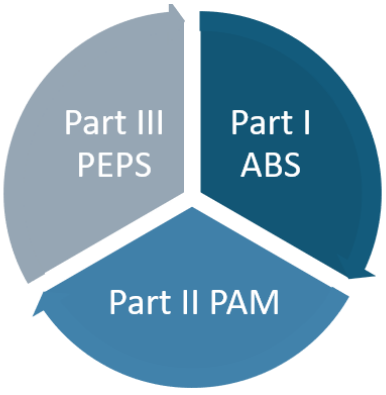


*Figure 1: NBEO exam series is a holistic assessment of optometric competency*

Each examination in the series contributes to the holistic assessment of competency. Following Miller's Pyramid of assessment (Miller, 1990), the NBEO exam series can be mapped along the pyramid by the area in which candidates demonstrate competency in each exam (see Figure 2). However, in optometry, the top of the pyramid, "Does," is truncated. The final assessment of competency in optometry is at the "Shows How" level because examinees are able to, after completing this level of assessment, apply for and receive a license to practice independently. In some other healthcare professions, examinees similarly progress through a series of licensure examinations to determine competency, but then must also

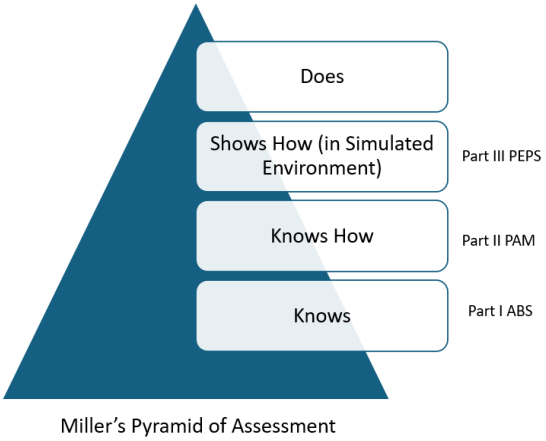


*Figure 2: NBEO examination progression through Miller's Pyramid of Assessment*

undergo a period of supervised practice outside of their graduate medical education. This period of supervised practice, or, residency, falls within the category “Does” on the pyramid. Supervised practice operates as an additional layer to the overall assessment of clinical competency.

Each piece of the exam series is important in demonstrating competency to practice the profession of optometry independently.

**WHY IS BASIC SCIENCE TESTING IMPORTANT IN OPTOMERIC LICENSURE TESTING?**

As part of its exam development process, NBEO conducts a Job Analysis (JA), during which information is gathered from optometrists throughout the United States and Canada to identify all the tasks and knowledge needed to successfully and safely perform as an optometrist. During the most recent JA, it was found that the content of the Part I Applied Basic Science examination is important in the holistic assessment of optometric competency. While Part I ABS assesses important science information related to anatomy, physiology, pharmacology, etc., in all cases questions on these topics are presented with concrete optometric applications. In other words, it is not just basic science information, but basic science in the context of real-world optometric patient care.

Professional standards require that high-stakes licensure programs, like NBEO, regularly review their exam content guidelines to ensure that they remain aligned with current best practices. Therefore, approximately every 5-7 years NBEO conducts a JA to collect input from the professional community to identify any necessary revisions.

For NBEO, the JA consists of a large-scale survey of licensed optometrists. The survey respondents are diverse with respect to demographics, geography, years of experience, training school, and modes of practice. For example, respondents include optometrists working in private practice, the VA system, and academia. The survey format helps NBEO ensure that our licensure exam content represents the full breadth of practice into which new licensees are entering.

Optometrists responding to the survey share both holistic judgements and detailed ratings related to the frequency and criticality of different elements including conditions, diseases, and procedures. Frequency refers to how often an optometrist encounters an element during practice. Criticality refers to how important a given element is to patient health, safety, and welfare. Elements with higher frequency and criticality ratings are given more weight in the exam content matrices. Optometrists are also given opportunities to bring attention to new developments in the profession including new treatments and emerging best practices.

The most recent JA was conducted in 2022. Emails were sent to 39,703 optometrists based on information from NBEO’s internal database, addresses provided by ARBO, and addresses provided by ASCO. There were 1,375 complete survey respondents, having met eligibility criteria and completed at least the survey questions related to the Part I ABS, Part II PAM, and TMOD exams. Respondents included practitioners, faculty, and regulators (i.e., regulatory board members). There were respondents from all 50 states, the District of Columbia, and seven Canadian provinces.

Respondents were presented with the current Part I ABS exam structure with respect to condition areas and disciplines the exam assessed. They were asked to provide recommendations about changes they would like to see in the relative balance of the exam content. Respondents of all types (i.e., practitioners/residents, faculty, regulators) recommended content weights that align with the current structure of the Part I ABS exam (see Table 1 below).

| | Current Weights | | Holistic Weights from Survey (middle 50% of responses) | | | |
|---|---|---|---|---|---|---|
| | Min. | Max. | All Respondents | Practitioners/ Residents | Faculty | Regulators |
| Condition Area | | | | | | |
| Refractive Status, Sensory Processes, and Oculomotor Processes | 35% | 35% | 35%-35% | 35%-35% | 35%-40% | 35%-40% |
| Normal Health, Disease, and Trauma | 65% | 65% | 65%-65% | 65%-65% | 60%-65% | 60%-65% |
| Discipline | | | | | | |
| Anatomy (Gross, Neuroanatomy, Histology, Development) | 12% | 18% | 15%-18% | 15%-18% | 15%-20% | 15%-15% |
| Biochemistry/ Physiology | 10% | 14% | 11%-12% | 11%-12% | 12%-15% | 12%-15% |
| Immunology/ Microbiology/ Pathology | 22% | 30% | 21%-26% | 21%-26% | 20%-26% | 20%-26% |
| Optics (Geometrical, Physical, Ophthalmic, Physiological) | 31% | 35% | 25%-33% | 25%-33% | 22%-33% | 28.5%-33% |
| Pharmacology | 11% | 17% | 14%-20% | 14%-20% | 14%-20% | 14%-20% |

*Table 1: Job Analysis Survey results regarding Part I ABS*

Additionally, respondents were asked open-ended questions about the Part I ABS examination. Specifically, they were asked if, given the purposes of the Part I ABS exam, they had suggestions for condition areas or disciplines that they feel should be either added or removed. Of the 1,375 respondents, 35 had comments regarding something to add, 98 had comments regarding something to remove, and 79 commented on both additions/removals. Only twelve of these respondents (<1%) recommended discontinuing the Part I ABS exam or making extensive/wholesale changes.

**In summary, NBEO has recently given a large sample of practitioners an opportunity to comment on and make recommendations for the Part I ABS exam. The vast majority provided responses indicating that they preferred the current content balance of the exam. Only a tiny percentage recommended discontinuing the exam or making large-scale changes. Taken together, we believe that these results indicate strong support for the current structure of the Part I ABS exam among optometric practitioners.**

**REMOTE PROCTORING VS. IN-PERSON TESTING**

NBEO requires testing for the multiple-choice exams to be done in a Professional Testing Center. NBEO has intentionally avoided utilizing Remote Proctoring due to security and exam integrity concerns. As described on the OEBC website, the OEBC Written Examination is "a case-based assessment delivered via computer using remote proctoring, consisting of multiple-choice questions based on different case scenarios."[2]

Figure 3 and 4 below show data presented by Pearson VUE in comparing incidents of suspected collusion in professional testing centers compared to remote proctored testing modalities. Collusion is a conspiracy between two or more people - one of whom is the candidate - to circumvent exam security measures in a way which allows someone other than the candidate to take the exam. Please refer to **Exhibit 5**, which includes the entire article from the journal, *Applied Psychological Measurement.*

[2] https://oebc.ca/taking-the-exam/preparing-for-the-exam/

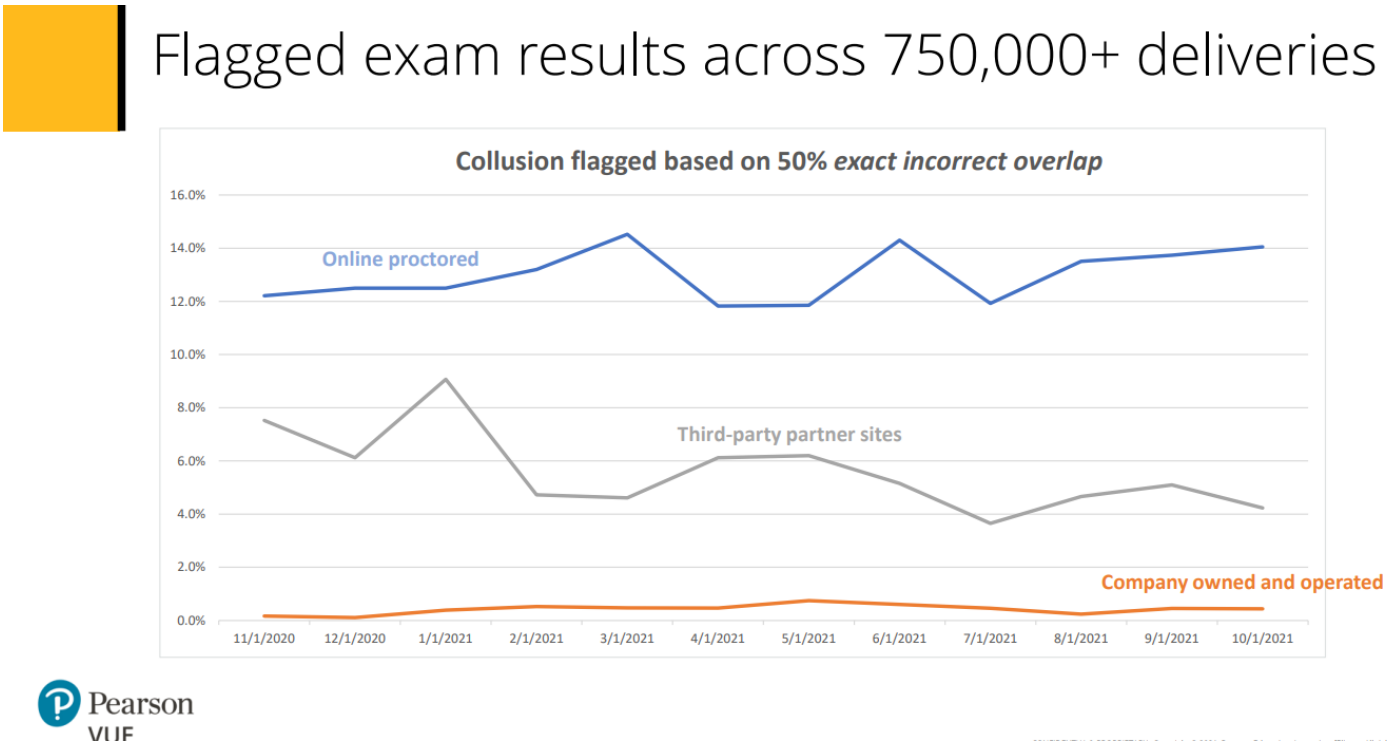


*Figure 3: Collusion flagged across exam deliveries comparing remote proctoring to testing centers*

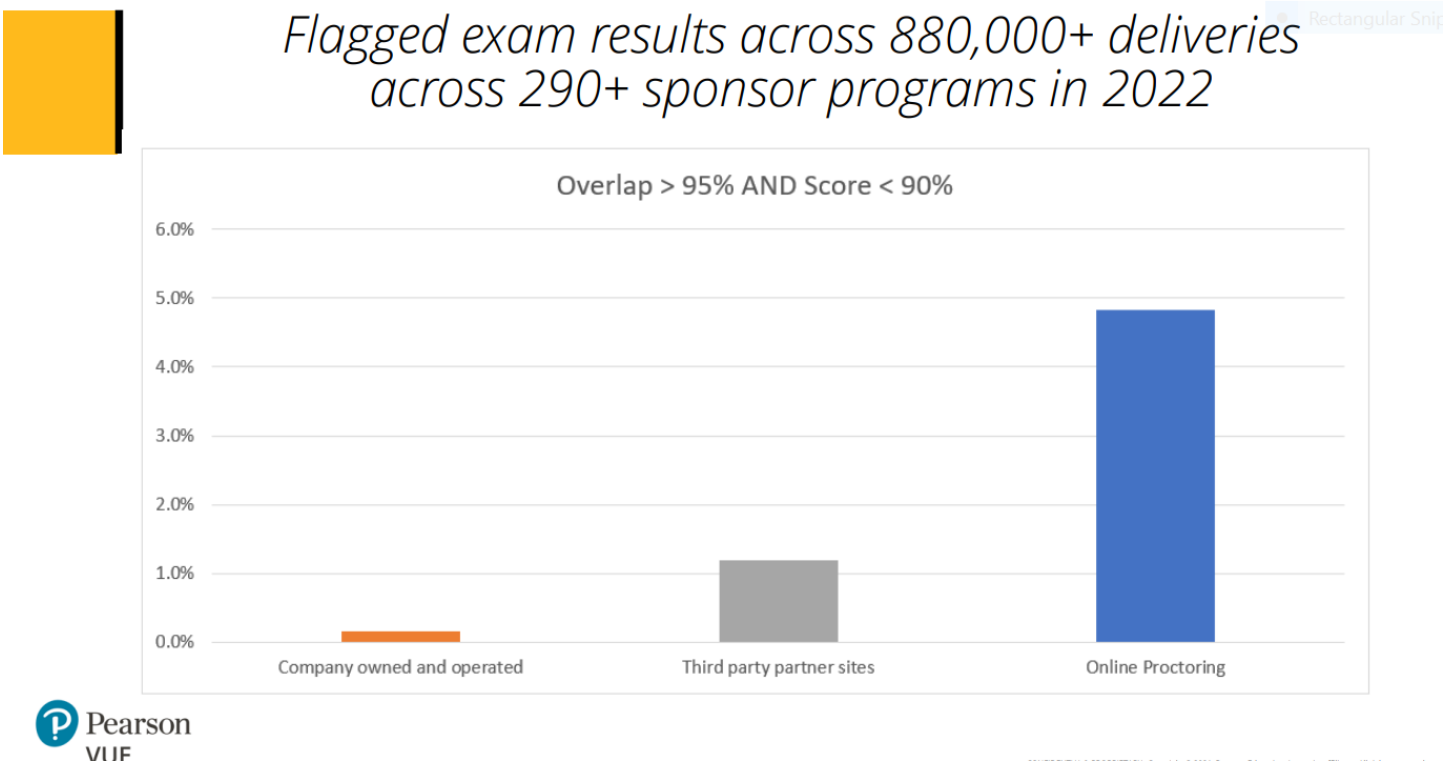


*Figure 4: Flagged exam results across exam deliveries comparing remote proctoring to testing centers*

**NBEO PASS RATES**

NBEO compiles the Institutional Yearly Performance Report to show first-time and ultimate pass rates for each school accredited by the Accreditation Council on Optometric Education (ACOE). A copy of the most recent Institutional Yearly Performance Report is attached hereto as **Exhibit 3**.

**SIX TIME LIMIT POLICY**

NBEO's Board of Directors voted to place a six-time attempt limit on the entry into the profession examination series beginning August 1, 2019. The Six-Time Limit Policy is attached hereto as **Exhibit 4**. This decision was largely based on results of a survey distributed to ARBO member boards at the 2017 ARBO Annual Meeting and thereafter by email. A majority of those respondents indicated their support of a proposed attempt limit.

Prior to 2019, NBEO had historically not placed a limit on the number of attempts that a candidate may take an examination. This was in contrast to the National Board of Medical Examiners and the National Board of Osteopathic Medical Examiners who both limit the number of examination attempts. Additionally, NBEO understands that some state medical licensing boards place a stricter limit on their applicants.

NBEO believes that implementation of our six-attempt limit is fundamental for ensuring protection of the public. In addition, there are several psychometric reasons to place a limit on retake opportunities.

1. It is beneficial from a security standpoint providing less item exposure. Exam security is better preserved when there are less attempts for candidates to visualize items.
2. Examinee scores stop significantly improving after the second or third attempt without some form of remediation (Wolkowitz, 2011 and Geving, 2005).
3. It can increase candidate motivation. If there is no retake limit, then candidates may not adequately prepare. Therefore, there is some expectation that the pass rate will increase with an examination attempt limit.
4. It may reduce the likelihood of an unqualified person passing. A false positive occurs when an incompetent candidate passes for reasons other than increased content knowledge. For example, the candidate may pass because of increased test taking strategies, decreased test anxiety, increased item exposure, etc.

The National Board of Examiners in Optometry serves to protect the public through the development, administration, scoring and reporting of valid examinations that evaluate competence. The NBEO examinations must be designed to assure state boards that those candidates which pass all Parts of the exams are competent to practice the profession of optometry as unsupervised healthcare providers. The NBEO takes its responsibility and commitment to the state boards very seriously and wants to ensure that candidates passing all Parts meet the level of a Minimally Qualified Candidate (MQC).

It is reasonable to allow candidates to have multiple opportunities to pass an examination when they have invested time and effort into acquiring the knowledge and skills measured by the examination (Millman, 1989). Further, this is logical for the barely competent candidate who failed the examination due to taking the examination on a particularly bad day, high levels of test anxiety, or another host of variables (Millman, 1989). Candidates deserve the opportunity to earn scores that are reflective of their true level of competence (Millman, 1989). However, there is a real danger that accompanies the practice of allowing indefinite opportunities to pass an examination (Millman, 1989). A candidate whose true level of functioning is below the passing standard may pass the examination on one of their unlimited reattempts resulting in a concerning false positive result (Millman, 1989). "Because the purpose of many licensing exams is to protect the public from harmful actions made by licensees, repeat testing opportunities may allow candidates to pass tests and obtain licenses without actually being highly knowledgeable" (Geving, 2005 and Millman, 1989). "In other words, retesting opportunities may increase the risk of obtaining false positives, subverting the main objective of licensure testing" (Geving, 2005).

Because the NBEO utilizes a portion of the same items repetitively for psychometric purposes, candidates may know a number of items asked before they enter the examination. Cook and Campbell (1979) suggest that "familiarity with a test can sometimes enhance performance because items and error responses are more likely to be remembered at a later testing session." According to Hausknecht (2002), test familiarity, reduced test anxiety, and increases in skill levels resulted in increased scores for the first several administrations, but test practice effects subsequently eroded or disappeared and there were no significant differences in test scores between test 3 and 4. When score gains are associated with increases in test taking abilities without growth in underlying abilities measured by the test, inferences drawn regarding the test scores from subsequent administrations will be compromised (Hausknecht, 2006). Additionally, research regarding medical examinations has shown that lower scoring candidates select more choices that would be dangerous to their patients than do candidates scoring even slightly higher (Juul, 1988).

NBEO has a responsibility towards protecting the public. Having an infinite number of exam attempts does not adequately support that mission. When candidates are allowed unlimited opportunities to take examinations with a fixed cut score, errors of measurement will increase the ratio of false positives to false negatives (Millman, 1989). Because NBEO's mission is to protect the public against incompetence, it follows that NBEO should have a process in place to minimize the risk of false positives. Millman states "for licensing and certification examinations designed to protect the public, passing incompetent candidates (false positives) is a more serious error than failing competent ones."

The proposed amendment 201 KAR 5:010 invites an applicant who has exhausted their attempts under the NBEO Six-Time Limit Policy to obtain licensure in Kentucky. NBEO has serious concerns that, by doing so, an unqualified candidate could obtain licensure.

**CONCLUSION**

Again, NBEO's mission is to protect the public by developing, administering, scoring and reporting results of valid examinations that assess competency in optometry. Utilizing basic science in optometric licensure exams is essential for producing competent, knowledgeable, and adaptable practitioners. A robust understanding of anatomy, physiology, pathology, and pharmacology enables doctors of optometry to diagnose and treat complex ocular and systemic conditions effectively, utilize new technologies, and contribute meaningfully to patient health outcomes. Basic science in optometric licensure examinations is a safeguard for public health. Basic science plays a vital role in ensuring safe, competent, and comprehensive optometric care while emphasizing patient safety and public trust.

***About NBEO*** *— Established in 1951, NBEO is an independent, non-governmental, non-profit organization whose examinations are universally accepted for optometric licensure in the United States and accepted internationally. NBEO's mission is to protect the public by developing, administering, scoring, and reporting results of valid examinations that assess competency in optometry. For more information, please visit our website at www.optometry.org.*

**References**


American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). Standards for Educational and Psychological Testing. American Educational Research Association.

Clauser, A. L., & Raymond, M. (2017). Specifying the content of credentialing examinations. In S. Davis-Becker, & C. W. Buckendahl (Eds.), *Testing in the professions: Credentialing policies and practice.* Routledge.

Cook TD, Campbell DT. *Quasi-experimentation: Design and analysis issues for field settings.* Boston: Houghton Mifflin. 1979.

Geving AM, Webb S, and Davis B*. Opportunities for Repeat Testing: Practice doesn't always make perfect*. Applied H.R.M. Research. 10(2):47-56. 2005.

Hausknecht JP, Halpert JA, DiPaolo NT, Moriarty-Gerrard MO*. Retesting in Selection: A Meta-Analysis of Practice Effects for Tests of Cognitive Ability.* Cornell University School of Industrial and Labor Relations. 2007.

Hausknecht JP, Trevor CO, Farr JL. *Retaking Ability Tests in a Selection Setting: Implications for practice effects, training performance, and turnover*. Journal of Applied Psychology. 87(2): 243-254. 2002.

**https://oebc.ca**

Juul D, Loewy EH. *The selection of critical errors on a medical school certifying examination*. Presented at the annual meeting of the National Council for Measurement in Education, New Orleans, LA. 1988.

Miller, G. E. (1990). The assessment of clinical skills / competency / performance. Academic Medicine, 65(9), S63-67.

Millman J*. If At First You Don't Succeed: Setting passing scores when more than one attempt is permitted.* Educational Researcher. 18(6): 5-9. 1989.

Myers, T., Morral, E., & Sares, T. (2019). Job analysis. In J. Henderson (Ed.), *Certification: The ICE handbook (3rd ed.).* Institute for Credentialing Excellence.

Wolkowitz A. *The effects multiple attempts on a nursing admissions examination have on the total score.* Journal of Nursing Education. 50(9): 493-501. 2011.

**BOARDS AND COMMISSIONS**
**Board of Optometric Examiners**
**(Amendment)**

**201 KAR 5:010. Application for licensure; endorsement.**

RELATES TO: KRS 218A.205(3)(g), 320.220, 320.250, 320.270
STATUTORY AUTHORITY: KRS 218A.205(3)(g), 320.240(7), 320.270(4)
NECESSITY, FUNCTION, AND CONFORMITY: KRS 320.220 requires all persons who practice optometry in this state to be licensed by the Kentucky Board of Optometric Examiners. KRS 320.250 establishes criteria for an applicant to apply for a license. KRS 320.270 grants the board the discretion to admit to practice in Kentucky persons licensed to practice optometry in other states. KRS 218A.205(3)(g) requires fingerprint-supported criminal record checks and queries to the National Practitioner Data Bank on applicants. This administrative regulation prescribes the procedures to be followed in making application to the board for a license.

Section 1.
(1) A person wishing to apply for a license to practice optometry shall submit to the board, within fifteen (15) days of board review, the following items:
(a) A completed Application for License to Practice Optometry;
(b) Birth certificate;
(c) A certified copy of college transcripts received directly from the registrar's office;
(d) A certified copy of optometry school ~~[transcripts ]~~received directly from the registrar's office;
(e) National board "NBEO" or Canadian OBEO written examination in lieu of Part 1 NBEO results;
(f) Therapeutic Management of Ocular Disease, "TMOD" results;
(g) Two (2) letters of recommendation, one (1) of which shall be from a licensed optometrist;
(h) Proof of successful completion of State Law Exam results;
(i) A ~~[passport-sized, ]~~recent photograph of head and shoulders, front view;
(j) Payment ~~[A money order or cashier's check payable to the Kentucky State Treasurer]~~ in the amount of $500 representing the non-refundable application fee; and
(k) Payment ~~[A money order or cashier's check]~~ in the amount of twenty-five (25) dollars ~~[made payable to the Kentucky State Treasurer ]~~for the purpose of submitting a query on the applicant to the National Practitioner Data Bank of the United States Department of Health and Human Services.
(2)
(a) Prior to approval for examination, the board shall review~~[receive]~~ and consider:
1. A national and state, fingerprint-supported criminal record check conducted by the:
a. Federal Bureau of Investigation; or
b. Kentucky State Police; and
2. A query for any relevant data from the National Practitioner Data Bank of the U.S. Department of Health and Human Services.
(b) Both of the items required to be furnished by this subsection shall be less than sixty (60) days old when reviewed by the board.

Section 2.
(1) A person wishing to apply for a license to practice optometry by endorsement shall submit to the board, within fifteen (15) days of board review, the following items:
(a) A completed Application for License by Endorsement to Practice Optometry;

(b) Verification that the applicant has been licensed in optometry and in active practice the past five (5) years;

(c) Information regarding any resolved, pending, or unresolved board action or malpractice suit in any state or territory;

(d) A certified copy of college transcripts received directly from the registrar's office;

(e) A certified copy of optometry school transcripts received directly from the registrar's office;

(f) A certificate of good standing from the board where the applicant is currently licensed and from all state boards where the applicant has held a license in the past;

(g) A copy of the credential that proves the applicant is therapeutically licensed;

(h) Two (2) letters of recommendation, one (1) of which shall be from a licensed optometrist;

(i) Proof of successful completion of Kentucky State Law Exam;

(j) A ~~[passport-sized, ]~~recent photograph of head and shoulders, front view;

(k) Payment ~~[A certified check or money order made payable to the Kentucky State Treasurer]~~ in the amount of $700 representing the non-refundable application fee;

(l) A notarized statement explaining why the applicant wishes to be admitted to practice in Kentucky; and

(m) Payment ~~[A money order or cashier's check]~~ in the amount of twenty-five (25) dollars ~~[made payable to the Kentucky State Treasurer ]~~for the purpose of submitting a query on the applicant to the National Practitioner Data Bank of the United States Department of Health and Human Services to retrieve any relevant data on the applicant.

(2)

(a) Prior to approval for licensure, the board shall receive and consider:

1. A national and state, fingerprint-supported criminal record check conducted by the:

a. Federal Bureau of Investigation; or

b. Kentucky State Police; and

2. A query for any relevant data from the National Practitioner Data Bank of the U.S. Department of Health and Human Services.

(b) Both of the items required to be furnished by this subsection shall be less than sixty (60) days old when reviewed by the board.

Section 3.

(1) A person whose license has been revoked pursuant to KRS 320.280(3) may apply for reinstatement of his or her license.

(2) Except as provided in subsection (3) of this section, a person applying for reinstatement shall submit to the board:

(a) Evidence of completion of the continuing education requirements established in 201 KAR 5:030; and

(b) Payment of the annual renewal fee established in 201 KAR 5:090, Section 2.

(3) To apply for reinstatement, an optometrist whose license has been revoked pursuant to KRS 320.280(3) shall submit to the board:

(a) Evidence of completion of the annual continuing education requirement for each year, or any portion of a year, that the license was not renewed ~~[up ]~~to a maximum of sixty (60) hours; and

(b) Payment of the renewal fee established in 201 KAR 5:090, Section 2~~[of $200]~~ for each year, or any portion of a year, that the license was not renewed.

Section 4. Incorporation by Reference.

(1) The following material is incorporated by reference:

(a) "Application for License to Practice Optometry", June 2024~~[August 2012]~~; and

(b) "Application for License by Endorsement to Practice Optometry", June 2024[August 2012].

(2) This material may be inspected, copied, or obtained, subject to applicable copyright law, at the Kentucky Board of Optometric Examiners, 2365 Harrodsburg Road, Suite A240, Lexington, Kentucky 40504-3333[2624 Research Park Drive, Suite 305, Lexington, Kentucky 40511], phone (859) 246-2744, Monday through Friday, 9:00 a.m. to 4:30 p.m.[8:30 a.m. to 5 p.m.].

*JOE ELLIS, OD, President*

APPROVED BY AGENCY: August 15, 2024

FILED WITH LRC: August 30, 2024 at 10:30 a.m.

PUBLIC HEARING AND COMMENT PERIOD: A public hearing on this administrative regulation shall be held on November 21, 2024, at 2:00 p.m. EST, at 2365 Harrodsburg Road, Suite A240, Lexington, Kentucky 40504. Individuals interested in being heard at this hearing shall notify this agency in writing by five workdays prior to the hearing, of their intent to attend. If no notification of intent to attend the hearing was received by that date, the hearing may be cancelled. A transcript of the public hearing will not be made unless a written request for a transcript is made. If you do not wish to be heard at the public hearing, you may submit written comments on the proposed administrative regulation. Written comments shall be accepted through November 30, 2024. Send written notification of intent to be heard at the public hearing or written comments on the proposed administrative regulation to the contact person.

CONTACT PERSON: Christi LeMay, Executive Director, 2365 Harrodsburg Road, Lexington Kentucky 40504, (859) 246-2744, email christi.lemay@ky.gov.

# REGULATORY IMPACT ANALYSIS AND TIERING STATEMENT

**Contact Person:Christi LeMay**

**(1) Provide a brief summary of:**

**(a) What this administrative regulation does:**
This amendment will allow an additional path to secure a Kentucky license.

**(b) The necessity of this administrative regulation:**
Post pandemic, there were a number of optometry school students that could not pass the National boards as required in current Kentucky regulation.

**(c) How this administrative regulation conforms to the content of the authorizing statutes:**
KRS authorizes the Board to determine license requirements.

**(d) How this administrative regulation currently assists or will assist in the effective administration of the statutes:**
This amendment adds an additional path to Kentucky licensure.

**(2) If this is an amendment to an existing administrative regulation, provide a brief summary of:**

**(a) How the amendment will change this existing administrative regulation:**
This will allow an additional path to secure a Kentucky license.

**(b) The necessity of the amendment to this administrative regulation:**
Post pandemic, there were a number of optometry school students that could not pass the National boards as required in current Kentucky regulation.

**(c) How the amendment conforms to the content of the authorizing statutes:**
KRS authorizes the Board to determine license requirements.

**(d) How the amendment will assist in the effective administration of the statutes:**
This amendment adds an additional path to Kentucky licensure.

**(3) List the type and number of individuals, businesses, organizations, or state and local governments affected by this administrative regulation:**
This amendment will affect potential applicants for a Kentucky optometrist license.

**(4) Provide an analysis of how the entities identified in question (3) will be impacted by either the implementation of this administrative regulation, if new, or by the change, if it is an amendment, including:**

**(a) List the actions that each of the regulated entities identified in question (3) will have to take to comply with this administrative regulation or amendment:**
The potential applicant will need to sit for and pass the OBOE written examination that is done online.

**(b) In complying with this administrative regulation or amendment, how much will it cost each of the entities identified in question (3):**
This amendment will give applicants a choice, not require them to take the OBOE written exam and does not incur a cost more than the accepted NBEO Part 1 exam.

**(c) As a result of compliance, what benefits will accrue to the entities identified in question (3):**
This amendment will give applicants an additional path to secure a Kentucky optometrist license.

**(5) Provide an estimate of how much it will cost the administrative body to implement this administrative regulation:**

**(a) Initially:**
$0

**(b) On a continuing basis:**
$0

**(6) What is the source of the funding to be used for the implementation and enforcement of this administrative regulation:**
N/A

**(7) Provide an assessment of whether an increase in fees or funding will be necessary to implement this administrative regulation, if new, or by the change if it is an amendment:**
There will be no increase in fees necessary.

**(8) State whether or not this administrative regulation establishes any fees or directly or indirectly increases any fees:**
This amendment does not establish any fees directly or indirectly.

**(9) TIERING: Is tiering applied?**
There are no fees associated with this amendment, therefore there is no tiering.

**FISCAL IMPACT STATEMENT**

**(1) Identify each state statute, federal statute, or federal regulation that requires or authorizes the action taken by the administrative regulation.**
KRS 218A.205(3)(g), 320.240(7), 320.270(4)

**(2) Identify the promulgating agency and any other affected state units, parts, or divisions:**
Kentucky Board of Optometric Examiners

**(a) Estimate the following for the first year:**

**Expenditures:$0**

**Revenues:$0**

**Cost Savings:$0**

**(b) How will expenditures, revenues, or cost savings differ in subsequent years?**
N/A

**(3) Identify affected local entities (for example: cities, counties, fire departments, school districts):**
This will not impact any local entities.

**(a) Estimate the following for the first year:**

**Expenditures:$0**

**Revenues:$0**

**Cost Savings:$0**

**(b) How will expenditures, revenues, or cost savings differ in subsequent years?**
There are no fees so nothing will change.

**(4) Identify additional regulated entities not listed in questions (2) or (3):**
None

**(a) Estimate the following for the first year:**

**Expenditures:$0**

**Revenues:$0**

**Cost Savings:$0**

**(b) How will expenditures, revenues, or cost savings differ in subsequent years?**
There are no fees associated so there will be no difference.

**(5) Provide a narrative to explain the:**

**(a) Fiscal impact of this administrative regulation:**
There is no fiscal impact resulting from this regulation.

**(b) Methodology and resources used to determine the fiscal impact:**
N/A

**(6) Explain:**

**(a) Whether this administrative regulation will have an overall negative or adverse major economic impact to the entities identified in questions (2) - (4). ($500,000 or more, in aggregate)**
: N/A

**(b) The methodology and resources used to reach this conclusion:**

N/A

**Exhibit 2**

**TESTING BEST PRACTICES**

NBEO follows the test development process set forth in the Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014). Exam development refers to a process whereby a measurement of an individual's knowledge, skills, and abilities (KSAs) is collected through the use of a test formed "according to a specified plan" (American Educational Research Association et al., 2014, p. 75). The development process steps are compiled in a test design plan, and test design begins with an evaluation of the intended uses of the test scores and expectations for how scores will be interpreted. That is, before beginning to define competencies to be measured, an initial step of determining the purpose of the test and what inferences from the test score are needed. Once this is known, the four phases of test development can begin. The Standards state,

> "Test design and development procedures must support the validity of the interpretations of test scores for their intended uses" (p. 75).

The validity of the intended inferences drawn from test scores must be upheld in all stages of test development. That is, the path for the creation of an examination should run as shown in Figure 1, and all steps should support validity.


*Figure 1*

The steps between determining the purpose of the exam and producing an operational exam are shown in Figure 2, and represent the guidance set forth in the Standards (American Educational Research Association et al., 2014).


*Figure 2*

# Exhibit 3

**NATIONAL BOARD OF EXAMINERS IN OPTOMETRY**

**October 1, 2022 - September 30, 2023 Institutional Yearly Performance Report**

| Schools | Number of Candidates | First Timer Pass Rate | | | Ultimate Pass Rate |
|---|---|---|---|---|---|
| | | Part I ABS* | Part II PAM* | Part III CSE* | |
| Arizona College of Optometry Midwestern University[6] | 56 | 80.36% | 89.29% | 85.71% | 98.21% |
| Chicago College of Optometry Midwestern University[6] | 51 | 49.02% | 72.55% | 86.27% | 100.00% |
| Illinois College of Optometry | 116 | 71.55% | 88.79% | 81.90% | 91.38% |
| Indiana University School of Optometry | 95 | 55.79% | 85.26% | 72.63% | 78.95% |
| Inter American University of Puerto Rico - School of Optometry[1] | 29 | 20.69% | 34.48% | 31.03% | 72.41% |
| Massachusetts College of Pharmacy and Health Sciences School of Optometry | 51 | 29.41% | 45.10% | 62.75% | 47.06% |
| Michigan College of Optometry at Ferris State University[5] | 36 | 88.89% | 97.22% | 83.33% | 100.00% |
| New England College of Optometry | 126 | 57.14% | 84.13% | 73.02% | 84.13% |
| Northeastern State University - Oklahoma College of Optometry[4] | 27 | 81.48% | 96.30% | 37.04% | 77.78% |
| Nova Southeastern University College of Optometry[7] | 126 | 48.41% | 71.43% | 70.63% | 76.19% |
| Pacific University - College of Optometry | 92 | 77.17% | 94.57% | 64.13% | 83.70% |
| Pennsylvania College of Optometry at Salus | 148 | 55.41% | 71.62% | 70.95% | 75.68% |
| Southern California College of Optometry at Marshall B. Ketchum University | 89 | 71.91% | 95.51% | 93.26% | 92.13% |
| Southern College of Optometry[2] | 137 | 75.18% | 91.24% | 87.59% | 99.27% |
| State University of New York College of Optometry | 96 | 86.46% | 93.75% | 81.25% | 94.79% |
| The Ohio State University College of Optometry | 61 | 91.80% | 96.72% | 85.25% | 98.36% |
| University of Alabama at Birmingham School of Optometry | 40 | 62.50% | 82.50% | 87.50% | 80.00% |
| University of California, Berkeley School of Optometry | 69 | 66.67% | 94.20% | 76.81% | 86.96% |
| University of Houston - College of Optometry | 89 | 62.92% | 91.01% | 77.53% | 85.39% |
| University of Missouri at St. Louis College of Optometry[5] | 44 | 77.27% | 86.36% | 61.36% | 86.36% |
| University of Pikeville - Kentucky College of Optometry | 52 | 46.15% | 71.15% | 46.15% | 61.54% |
| University of the Incarnate Word - Rosenberg School of Optometry | 57 | 40.35% | 73.68% | 68.42% | 64.91% |
| Western University of Health Sciences College of Optometry | 57 | 19.30% | 61.40% | 66.67% | 57.89% |
| **United States** | **1743** | **62.65%** | **82.79%** | **74.18%** | **83.59%** |
| Université de Montréal School of Optometry | 1 | 100.00% | 100.00% | 100.00% | 100.00% |
| University of Waterloo School of Optometry and Vision Science | 20 | 100.00% | 100.00% | 75.00% | 100.00% |
| **Canada** | **21** | **100.00%** | **100.00%** | **76.19%** | **100.00%** |
| **ASCO Institutional** | **1764** | **63.10%** | **82.99%** | **74.21%** | **83.79%** |

**Note, these universities require candidates to complete the following for graduation:**

[1] pass Part I for graduation

[2] pass Part I and Part II

[3] pass Part I, Part II, and Part III

[4] attempt but not necessarily pass Part I and Part II

[5] attempt but not necessarily pass Part I, Part II, and Part III

[6] pass Part I and attempt but not necessarily pass Part II

[7] This college previously had a requirement that candidates pass Part I in order to graduate. Within the timeframe this report covers, the college revised the policy and removed the graduation requirement, which has resulted in the report reflecting more graduates than the 2023 class size.

**General Notes:**

* All pass rates are calculated using only candidates who have graduated, as indicated by the university, during the listed time frame and have attempted all three parts at least once.

* First Timer pass rates are the percentage of the above referenced candidate group who passed the examination on their first attempt.

* Ulitmate pass rate is the percentage of the above referenced candidate group who passed all three exams.

Exhibit 4

NBEO

NATIONAL BOARD OF EXAMINERS IN OPTOMETRY®

ADVANCING THE ASSESSMENT OF COMPETENCE®

# Six-Time Limit Policy

The National Board of Examiners in Optometry (NBEO®), in furtherance of its mission to protect the public by developing, administering, scoring and reporting results of valid examinations that assess competency in optometry, must preserve the integrity of the NBEO entry-into-the-profession examinations. The entry-into-the-profession examinations are: (1) the three-part licensure exam (consisting of Part I ABS®, Part II PAM®/ TMOD®, Part III, which includes both Part III CSE® and Part III PEPS) and (2) ISE®. To effectuate its mission, the following policy provides certain limits on the number of times a candidate is eligible to take certain of NBEO's exams.

Candidates seeking optometric licensure must receive a passing score on each of the NBEO entry-into-the-profession examinations in the first six attempts of each exam.

Once a candidate has reached the six-attempt limit, the candidate must utilize the following process and be approved by a committee of NBEO's Board of Directors before the candidate is eligible to register for the exam again.

Under this policy, an "attempt" includes any instance where an examinee begins to take any portion of an exam. For the sake of clarity, an "attempt" on TMOD shall include both an attempt on the embedded TMOD exam and an attempt on the stand-alone version of the TMOD exam. Further, candidates are not allowed six separate attempts on each of the Part III CSE and the Part III PEPS examinations; rather, an attempt on either the Part III CSE or the Part III PEPS will be counted toward the combined number of six total attempts allowed on the Part III examination.

1. The candidate must have a sponsor who will assist the candidate in creating a remediation plan and support the candidate in the candidate's next attempt if granted. The sponsor must be a Chief Academic Officer or Chief Executive Officer of an ACOE-accredited optometric institution. Members of NBEO's Board of Directors will not serve as sponsors. The remediation plan must demonstrate the additional steps the candidate will undertake to prepare to take the exam an additional time and show how such steps will make it likely that the candidate will successfully pass the exam. The remediation plan must be appropriately robust based on the candidate's exam performance history and particular circumstances.

2. The candidate will submit a request via NBEO's online platform with the proposed remediation plan, a letter explaining the circumstances that led to the six-attempt limit being reached, why the candidate should be allowed an additional attempt, and a letter from the candidate's sponsor explaining why the sponsor supports the candidate's additional attempt. Candidates submitting a request under this policy will be charged a non-refundable fee of $300. Candidates who wish to submit a request for an attempt under this policy should click here.

3. The Judicial Committee of NBEO's Board of Directors or an ad hoc committee of the Board appointed by the President (the "Committee"), will review the request and all related documentation. In appointing members of the Committee, NBEO will endeavor to identify, manage and avoid any potential or perceived conflicts of interest.

4. The Committee will consider all relevant written documentation submitted in connection with the request, then deliberate and decide whether the candidate's request should be granted or denied. NBEO will send a letter to both the candidate and candidate's sponsor notifying them of the decision of the Committee. That decision will constitute NBEO's final determination of the matter.

5. If the Committee agrees to allow an additional attempt on the exam, the candidate will have the opportunity to register for the next examination following the completion of the agreed-upon remediation. The completion of remediation must be evidenced by a written confirmation from the sponsor and emailed to NBEORequests@optometry.org. Once NBEO receives that confirmation, the candidate will be eligible to register for the examination.

6. If the Committee grants a candidate an additional attempt on an entry-into-the-profession exam and the candidate is ultimately unsuccessful on the additional attempt, the candidate may not request another attempt. Any such requests made will be denied.

The National Board of Examiners in Optometry, Inc. reserves the right to amend, modify or change the terms and conditions of this policy at any time for any reason.

*Manuscripts*

# Test Security and the Pandemic: Comparison of Test Center and Online Proctor Delivery Modalities

Applied Psychological Measurement
2024, Vol. 0(0) 1–13

DOI: 10.1177/01466216241248826
journals.sagepub.com/home/apm
Sage

**Kirk A. Becker[1], Jinghua Liu, and Paul E. Jones[1]**

**Abstract**

Published information is limited regarding the security of testing programs, and even less on the relative security of different testing modalities: in-person at test centers (TC) versus remote online proctored (OP) testing. This article begins by examining indicators of test security violations across a wide range of programs in professional, admissions, and IT fields. We look at high levels of response overlap as a potential indicator of collusion to cheat on the exam and compare rates by modality and between test center types. Next, we scrutinize indicators of potential test security violations for a single large testing program over the course of 14 months, during which the program went from exclusively in-person TC testing to a mix of OP and TC testing. Test security indicators include high response overlap, large numbers of fast correct responses, large numbers of slow correct responses, large test-retest score gains, unusually fast response times for passing candidates, and measures of differential person functioning. These indicators are examined and compared prior to and after the introduction of OP testing. In addition, test-retest modality is examined for candidates who fail and retest subsequent to the introduction of OP testing, with special attention paid to test takers who change modality between the initial attempt and the retest. These data allow us to understand whether indications of content exposure increase with the introduction of OP testing, and whether testing modalities affect potential score increase in a similar way.

**Keywords**

test security, online proctoring, remote proctoring, risks to validity, high-stakes assessment

## Introduction

The COVID-19 pandemic has disrupted testing programs and changed the outlook of the testing industry. One of the most notable changes is the rapid increase in remote online proctored testing.

[1]Pearson VUE, Chicago IL, USA

**Corresponding Author:**
Kirk A. Becker, Psychometric Services, Pearson VUE, 4627 N Leclaire Ave, Chicago, IL 60630, USA.
Emails: Kirk.becker@pearson.com; Kirk.becker@yahoo.com

Online proctored (OP) testing was developed to make test taking more available and convenient, while increasing the security of heretofore proctor-less online tests in employment and educational settings. By contrast, test center (TC) delivery involves dedicated locations where test takers provide physical ID and verify their identity in person and use computers provided by the test center in a controlled environment while being surveilled by an in-person proctor. TC environments may show variation based on the policies and technology requirements of the test delivery provider and variability in proctor quality across locations. OP environments, on the other hand, add more varieties, such as differences in the individual test taker's environment, computer hardware, and connectivity.

While not proven, it is easy to imagine a possible interaction between testing in public and private environments and one's inclination to try to cheat, one's attitude or motivation toward the test session, and other personal factors that could affect performance for good or bad. Given the limited published empirical research comparing the security of the two delivery methods, testing organizations must fall back on professional judgment to decide whether OP test delivery "eliminate[s] opportunities for test takers to attain scores by fraudulent means" (AERA et al., 2014, p. 116) to the same extent as TC delivery. It is reasonable to argue that a testing vendor or organization can eliminate more opportunities to cheat if they control the physical environment. Reviews of OP vulnerability to cheating behaviors (Woodley, personal communication, June 19, 2020) have sounded the need for caution here. This worry about greater vulnerability in the OP environment may explain why, until the pandemic, relatively few high-stakes programs had considered OP delivery given the technology and resources available at the time.

Of course, the goal of completely "controlling the testing room" is aspirational. Room control means that the proctor controls what and who is in the testing room, and limits the materials (cell phones, cheat sheets, etc.) that can be taken into or out of the room. The first practical objective for any proctor is to make it very difficult for a proxy to take the exam: either take the exam for the nominal test taker or communicate the answers to the nominal test taker in real time. While cheating doubtless occurs in testing centers where a proctor is surveilling several candidates in real time, a remote proctor surveilling several candidates with less room control and a narrower field of vision would likely be less successful at inhibiting or detecting cheating, even with the help of artificial intelligence.

Like individual cheating, exam theft through covert information gathering is logically more likely in a physical environment controlled by the test taker than in a testing center. A thief could easily hide many kinds of surveillance equipment in the remote test site and quickly steal all the items on an exam. For example, Foster and Marder (2020, April) showed that it is possible for hidden cameras to "outrange" a single computer web camera and, through Wi-Fi or Bluetooth connections to external computers, automatically process item text off the computer screen without being in the field of view of a proctor confined to a one-camera view of the room. Of course, every test presents the risk of item compromise through harvesting. Cell phones, miniature cameras (even contact lens cameras), and other spy technology also present a risk for recording content in the testing center. Lower-tech efforts to harvest through candidate memory will always exist. Whatever the method, the effects of item theft (typically on hard-to-contain social media) are even more serious than individual cheating, potentially undermining the interpretation of many test scores and rendering worthless item banks that may have cost hundreds of thousands to millions of dollars to produce.

The rapid rollout of online proctoring with the onset of the pandemic created an inflection point in the attitudes of test sponsors towards testing at home. It also created a sudden change to the security environments of many exams. So far a very limited amount of data is available regarding the relative security of OP versus TC environments, and the purpose of this study is to begin to provide that quantitative data. Hurtz and Weiner (2022) compared OP and TC test takers from five

different credentialing exams on exam performance and forensic indicators, including occurrence of high similarity responses, irregular response patterns, and response speed. While some small differences in the proportion similar responses by OP versus TC candidates and slightly higher percentage of irregular cases among OP candidates, two having slightly higher percentages among TC candidates, almost all cases the effect sizes in these comparisons were negligible. We hope the results we are reporting will add to the conversation about the appropriate use of OP.

## Overview of Types of Test Security Analysis

Test security analyses involve the detection of behaviors or test results that are unusual in the general population and potentially indicative of dishonest behaviors or compromised content. Any such behavior that departs from legitimate test taking obviously threatens the validity of the test result interpretation and the downstream statistics used to construct subsequent exams. Examples of dishonest behaviors include illicitly accessing live test content for study, access to unauthorized information (e.g., cheat sheets) during testing, working with a proxy test taker, working with a dishonest testing location, and taking tests specifically to acquire test content.

Within this umbrella of illicit behaviors, "test collusion" refers to cheating in cooperation with another person, such as by answer copying, sharing test content, communicating during an exam, proxy test taking, and receiving content from instructors (Maynes, 2017). Forensic analyses can be conducted with the aim of identifying individuals, locations, or groups involved in collusion, as well as items and tests that have been compromised. These analyses are important aspects of test validity arguments in both TC and OP environments. Beginning at least as far back as the 1920s (Bird, 1927), detecting cheating on multiple-choice tests has involved finding unusual levels of agreement between test taker responses (response overlap), typically incorrect responses. When test takers are all taking the same test in the same location, copying answers is a relatively easy way to cheat on the test, either by stealing answers or collusion. The introduction of computer-based testing and randomized item order, the use of multiple different test forms, and the use of continuous rather than event-based testing has virtually removed the opportunity for answer copying in person. However, other methods of collusion such as preknowledge through access to illicit study guides and using proxy test taking, are susceptible to detection using response overlap.

The probabilistic nature of the testing situation means that even when two test takers have the same level of knowledge, they won't necessarily know the answers to all the same items, although people who have studied in the same class might have more similar response patterns than those from very different teaching environments. As scores increase to near 100% correct, the sensitivity of response overlap decreases as pairs of high-scoring test takers will have all or nearly all item responses in common. Tests covering a wide range of content, or adaptive tests, may show different patterns of agreement than fixed form tests or those covering a narrow range of content. Multiple-choice items with three plausible distractors will have less agreement than those with only one plausible distractor. So as Saupe quipped in 1960, while perfect agreement of all correct and incorrect responses seems an obvious case of cheating, "there would exist a need for estimating the likelihood of the observed correspondence under the condition of no collaboration." (Saupe, 1960, p.475). Numerous approaches have been used for determining significant response similarity for the purpose of identifying collusion (Becker & Meng, 2022; Maynes, 2012; Smith, 2019; Zopluoglu, 2017).

Other approaches to identifying cheating or security breaches include response time analyses, unusual response patterns, large retake score changes, item parameter drift, discovery of stolen content, pretest versus operational item performance, and a variety of other approaches (see Cizek & Wollack, 2017; Kingston & Clark, 2014; Wollack & Fremer, 2013). These analyses may

supplement collusion analyses, identify situations outside of collusion, or apply to situations such as CAT or other pool-based testing that limit or preclude response overlap analysis.

This article reports on the results of two studies of potential test taker fraud. In both studies, we used collusion analysis to compare security between OP and TC settings. Study 1 examines high degrees of answer overlap across several hundred testing programs in 2022, while Study 2 examines high overlap and other security indicators in a detailed analysis of a single testing program prior to and following the introduction of OP in 2020.

## Study 1

### Data

The first dataset included a convenience sample of aggregated data from over 3 million test takers across 326 testing programs from January to July of 2022. These data were compiled by an automated process and were not subject to individual data cleaning procedures prior to aggregation. These programs included professional credentialing, admissions, and IT certification programs. Programs offering CAT, LOFT, or other pool-based exam administration, and programs with short tests (<50 items) were excluded.

These data include a breakdown by testing location type and week but are otherwise aggregated by testing organization. Test location types included online proctoring (OP) and three types of TCs: (a) test centers owned and run by the testing vendor (company-owned), (b) test centers not owned by the testing vendor, but which met hardware and security standards (3rd party select), and (c) third-party test centers which had not indicated that they meet those standards (3rd party). Company-owned and 3rd party select test centers in this sample require direct line of sight for proctors, walk-throughs, and overhead cameras, while third-party TCs require at least one of those. For proctor to candidate ratios, the company-owned test centers required a maximum ratio of 15:1, while the other TCs are not specified. The OP exams made use of live real-time proctors with a single camera and up to 15 test takers assigned to each proctor using two computer monitors. Some testing programs included in these data tested in all test location types, while others tested in only a subset. There are large differences in sample size, as well as differences between the types of organizations making use of OP versus different types of testing centers. A subset of 33 organizations was selected, containing results from programs with $N > 100$ in all four test location types.

### Methods

For this dataset, the total number of correct and incorrect responses in common was calculated for all test takers each week, compared with all other test takers from the current week and the previous time period (time period was typically 2 months, but could be shorter depending on exam volume). Because of the large number of programs and volume of test takers in this initial study, only extreme levels of overlap were flagged. Test taker pairs where each candidate's number-correct score was <90% correct and where response overlap was >95%, were flagged. Items that were not answered were not counted as matching, and high overlap for retake test takers with previous attempts were not flagged. These collusion criteria could be described as very stringent, even given the large number of comparisons being made, and are more likely to detect proxy test takers than preknowledge. The number of unique flagged test takers were then counted within each test location type by week. These counts, combined with the total number of test takers, were used to evaluate likely collusion by testing location. For the subsample including only testing programs with all four testing modalities, the percent of flagged test takers were calculated for

each organization, and averages and standard deviations were computed across the 35 organizations for each test location type.

### Results

Table 1 shows the total number of test takers and the percentage of test takers with a response overlap >95% with at least one other person and a percent correct <90%. The company-owned test centers had the lowest rate of collusion, with 3rd party select test centers showing a slightly increased rate, and 3rd party test centers having the highest incidence rate among test centers. OP test takers were flagged at a higher rate than any of the three types of test centers. These values are obviously influenced by the size and characteristics of the programs included, with large and highly valued programs potentially having an outsized impact. Many programs had no test takers with 95% overlap in a given test location type.

Figure 1 shows percentage of flagged candidates by test modality and type by week over the course of 25 weeks. While these weekly flagging rates show a great deal of variability, especially for the OP modality, the relative incidence of overlap by mode and test center type is consistent: OP shows the most overlap, followed by non-select 3rd Party test centers. The Company-Owned and 3rd Party Select test centers have the lowest incidence of high overlap.

Table 2 provides the mean and standard deviation of flagging rates for the 33 testing programs using all four test location types. The percent of candidates flagged (>95% response overlap and percent correct <90%) was calculated by organization and by location type, and these percentages were then used to calculate the mean percent flagged (Mean High Overlap %) and variance of percent flagged (SD %) reported in table 2. This breakdown partials out the effect of sample size and differences in the organizational makeup of the data provided in Table 1. While this sample includes only 33 organizations, over 1.9 million test taker results are represented. These results show a similar pattern to that in Table 1, with company-owned TC showing the lowest percentage of flag, whereas 3rd party TC having the highest percentage of flag. The percentage of OP flag is higher than any of the TC flag. The variance of flags is also higher in third-party and OP settings.

## Study 2

### Data

The second dataset included data from 69,353 testing instances from November 2019 through January 2021 for a single test program. The total data set includes incomplete tests and repeat test takers, although different analyses may include or exclude certain results. Between November 2019 and March 2020, six non-overlapping test forms were administered in company-owned and third-party select test centers. Test takers retesting after failing were administered a different form

**Table 1.** Percent of Flagged Candidates by Mode and by Test Location Type in 2022.

| Modality | *N* | High overlap (%)[a] |
|---|---|---|
| TC: Company owned | >800k | 0.5 |
| TC: 3rd party – select | >200k | 0.6 |
| TC: 3rd party | >1.6 M | 1.0 |
| OP | >600k | 3.5 |

[a]Percent of test takers with a 95% response overlap with at least one other test taker, and number-correct score <90% correct.

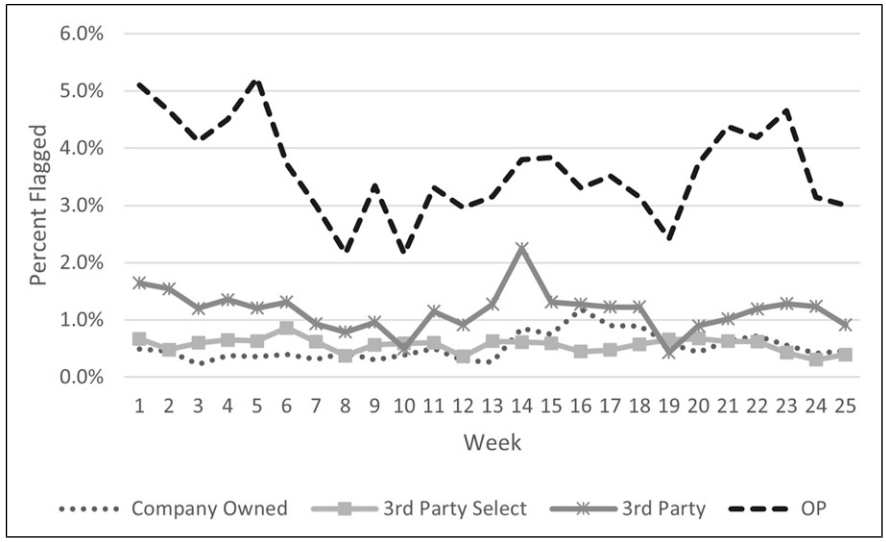


**Figure 1.** Percent of flagged candidates by mode and by test location across week of 2022.

**Table 2.** Flagged Candidates for 33 Programs Tested in all Location Types.

| Location type | Candidates included | Mean high overlap % | SD% |
|---|---|---|---|
| TC: Company owned | >100k | 0.7 | 0.02 |
| TC: 3rd party – select | >1m | 0.8 | 0.02 |
| TC: 3rd party | >50k | 1.4 | 0.04 |
| OP | >600k | 1.8 | 0.04 |

than previously administered when possible, and always a different form than the last one taken. This program began offering OP (live real-time proctors with a single camera and multiple test takers assigned to each proctor) as an option in April of 2020 during the pandemic shut down. For the first four months of OP administration, two of the six test forms that were previously administered in test centers were available in both OP and TC administration. Due to the shutdown only, OP exams were administered in April. During this period, an illicit study guide containing operational test items was discovered online. New test forms were then published in August of 2020, along with additional security protocols for OP.

## Methods

While Study 1 looked only at test taker pairs with >95% response overlap and score <90% correct, Study 2 made use of a simulated null distribution (Becker & Jones, 2022; Becker & Meng, 2022) to determine flagging criteria. The use of a null distribution will avoid interpreting high levels of overlap as statistically normal if widespread cheating is present and empirical distributions are used to interpret overlap. The distribution of overlapping responses is determined by the number of items in common, the distribution of item parameters on the test, the probability of selecting each incorrect response on items, and the relative scores of a pair of test takers.

For Study 2, flagging values for significant response overlap were determined by simulating large numbers of responses with different scores for several test lengths and identifying overlap with $p < .00001$. While Study 2 involves a single testing program, we used a general method for flagging significant overlap that did not depend on the specific characteristics of this test. This is in line with our goal to have a general flagging rule that can be calculated ahead of time for the identification of potential misconduct. To calculate the significant overlap percentages, random normal distributions ($M = 0$, $SD = 1$) of IRT item difficulties (b values) were generated for test lengths of 25, 50, and 75. Three incorrect options were simulated with probabilities of 70%, 20%, and 10%, coinciding with the observation that real distractor options tend to be unequal in strength. Overlap criteria for these discrete test lengths were applied to pairs of response vectors of similar length, accommodating the fact that test takers did not necessarily take all items on the test, and it would be difficult to simulate every possible number of overlapping items. Thus, the flagging criteria, in terms of percent of overlapping responses, for 25-item exams were used to evaluate the overlap between response vectors containing 25 to 49 items, the flagging criteria for 50-item-exams were applied to vectors of between 50 and 74 items, and criteria for 75-item exams were applied to vectors of 75 items. We refer to the percentage of cases flagged for high overlap based on the null distribution as the *Response Similarity Index* (RSI).

Like the programs in Study 1, the detailed analysis of collusion for the single program involved calculating the overlap of correct and incorrect responses for all pairs of test takers. Instead of weekly analysis, this was done using monthly data, comparing test takers who tested within a given month with all other test takers in that month, plus those who tested during the previous month. For each test taker pair, the percent correct for each test taker was also calculated. The observed response overlap (both correct and incorrect) was then compared to the expected overlap for similar test taker pairs based on the simulated distribution of response overlap in the null condition. For example, on a 75-item test, when one test taker has a score of 80% and another test taker has a score of 90%, a total response overlap greater than 89.3% (67 responses in common) would appear with probability <.00001. A test taker was flagged if they had response overlap with at least one other test taker, with $p < .00001$, and that other test taker did not share the same person ID. Response overlaps of 100% for flagged pairs (all correct and all incorrect responses identical) were specifically called out.

Test security flags were also calculated based on test and item time anomalies, as well as indicators of unusual responses. While these flags are not conclusive evidence of security violations, they are suggestive, especially when they are used collectively. These flags include:

- Slow correct responses: Number of correct responses in slowest 5%. Flagged if number is < 3 SD below mean.
- Fast correct responses: Number of correct response with time <15 seconds. Flagged if number is > 3 SDs above mean fast correct count.
- Fast Pass: Candidates passing the test in fewer than 15 minutes.
- Negative Response Time Correlation: Correlation between response time and mean response time <0.
- Low Easy versus Hard Performance Delta: Performance difference between percent correct on easy and hard items (items above or below the median item p-value). Flagged if difference <3 SD below the mean difference.
- High Operational versus Pretest Performance Delta: Performance difference between percent correct on pretest and operational items. Flagged if difference >3 SD above mean pretest/operational percent correct difference.

In addition to the response overlap and security flags mentioned, these data also included interesting within-candidate evidence bearing on exam security. Within the timeframe studied, there were 7,259 test takers who failed their first attempt and then retook the exam a second time. The results for these test takers allow us to compare the security environments for OP versus TC test delivery using differential number-correct score gains and pass rates between the first and second attempt. In analyzing these cases, we distinguished between exams that were "fully delivered" or "not fully delivered." A test that was "fully delivered" meant that either the test taker used all available time or answered all questions on the exam. A test that was not "fully delivered" in OP could have been stopped by the proctor, stopped due to technical issues, or stopped by the test taker, although that information is not available to the authors.

## Results

Figure 2 shows the RSI by month in both OP and TC environments. The average monthly volume for the period was around 3,000 test takers. We divided the time frame into three periods: Pre-OP, OP begins, and new forms published. In the five months prior to the introduction of OP, four candidates out of 17,000 tested were flagged for significant overlap. This flagging rate for over 17,000 test takers (~200 billion pairs compared) offers strong evidence that the flagging criteria are not detecting chance levels of overlap. With the introduction of OP, we see not only a higher number of flagged candidates in that modality, but also the appearance of flagged candidates in TCs. The general pattern of results was the same in both modalities, albeit to different degrees: the flagging rate started to climb with the introduction of OP, reached its peak sometime in the summer, and then declined when new forms and security protocols were introduced in August. After August, the OP channel still shows greater numbers of flags than TC, and TC shows small but non-zero numbers of flags, when prior to OP introduction there were effectively no flags.

Table 3 shows the monthly test taker counts and RSIs. The OP test takers are further broken down to show the percentage of cases equal to 100% response overlap and the percentage less than 100% overlap. TC flags are not broken down this way because a total of only 16 TC test takers had 100% overlap with at least one other test taker (where the other test taker was either TC or OP).

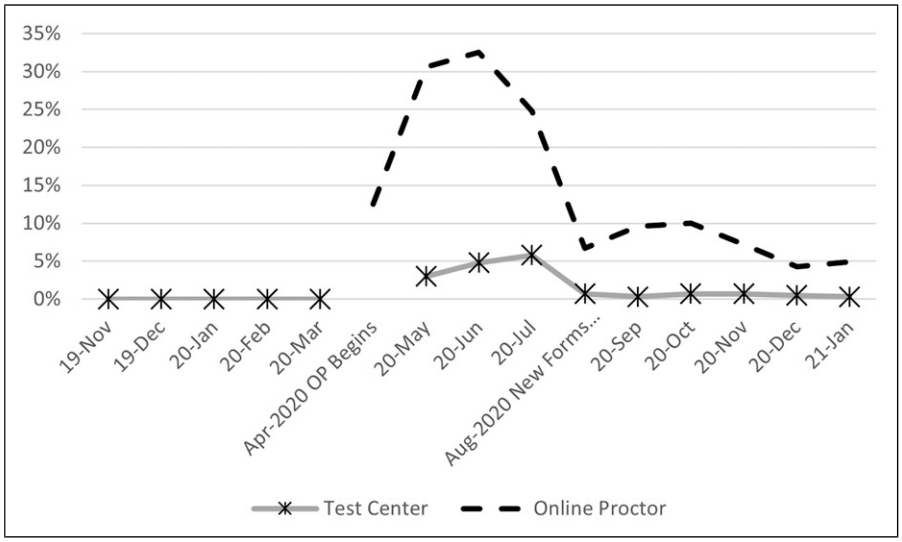


**Figure 2.** Monthly RSI during 3 time periods.

Overlap of 100% is likely a result of proxy test taking (test takers are remarkably consistent in their item responses over time) which, while not impossible in a test center, appears to be very difficult in that modality. Approximately 47% of the flagged cases in OP had 100% response overlap with at least one other test taker.

Table 4 reports the additional security flags for OP versus TC during the three time periods, as well as the test pass rates. There is a greater incidence of fast correct responses and low response time correlations among OP candidates in the initial OP and later OP time periods. Pass rates do show differences, however, there is no control over the comparability of the populations in OP versus TC. The 15 second correct and response time correlation flags also show larger differences between test takers flagged for possible collusion (i.e., significant overlap) versus those not flagged, which is shown in Table 5 along with the pass rates for these groups. It is important to note that while the data does include repeat test takers, RSI flags exclude comparisons between a retest test taker and their own previous results.

Next, we look at subsets of repeat test takers, focusing on differences between Time 1 and Time 2. These test takers failed the Time 1 attempt and then either passed or failed on the Time 2 attempt. The data we have reviewed so far suggests that collusion is more frequent for this exam in the OP mode, and perfect overlap between responses is exceedingly more frequent in OP. Thus, we might expect TC- > OP Time 2 scores and pass rates to be elevated for test takers with a propensity to cheat, but who couldn't do so under TC restrictions. We might expect less elevation among OP- > OP test takers who failed Time 1 either because they had less propensity to cheat or less ability. These persons might still be elevated more than OP- > TC or TC- > TC test takers who had more restrictions during Time 2. Obviously, we would expect greater score gains among test takers who were flagged for collusion than among those having no evidence of collusion.

Table 6 provides the Time 1 and Time 2 modality based on whether Time 1 was complete (excluding those for whom Time 2 was incomplete), as well as the score change and pass rates for

**Table 3.** Monthly Response Similarity and 100% Overlap by Testing Modality.

| | TC | | OP | | OP RSI breakdown | |
|---|---|---|---|---|---|---|
| Month | *N* | RSI (%) | *N* | RSI (%) | <100 (%) Overlap | 100 (%) Overlap |
| Pre-OP | | | | | | |
| Nov-19 | 3798 | 0.0 | | | | |
| Dec-19 | 3389 | 0.0 | | | | |
| Jan-20 | 3481 | 0.0 | | | | |
| Feb-20 | 4130 | 0.0 | | | | |
| Mar-20 | 2382 | 0.0 | | | | |
| OP begins | | | | | | |
| Apr-2020 | 7 | 0.0 | 1327 | 12.5 | 8.7 | 3.8 |
| May-20 | 1167 | 3.0 | 5141 | 30.6 | 15.0 | 15.6 |
| Jun-20 | 1747 | 4.8 | 5203 | 32.5 | 16.2 | 16.3 |
| Jul-20 | 1616 | 5.8 | 3948 | 24.8 | 12.6 | 12.3 |
| New forms published | | | | | | |
| Aug-2020 | 1543 | 0.7 | 2980 | 6.7 | 4.7 | 2.0 |
| Sep-20 | 2123 | 0.3 | 4224 | 9.6 | 6.0 | 3.6 |
| Oct-20 | 2306 | 0.7 | 4379 | 10.0 | 5.4 | 4.6 |
| Nov-20 | 1615 | 0.7 | 3203 | 7.2 | 4.1 | 3.1 |
| Dec-20 | 1823 | 0.5 | 2846 | 4.3 | 2.9 | 1.3 |
| Jan-21 | 1778 | 0.3 | 3479 | 4.9 | 3.3 | 1.6 |

**Table 4.** Security Flags by Modality and Time Period.

| Mode | *N* | Pass rate (%) | Slow correct resp. (%) | Fast correct resp. (%) | Fast pass (%) | Neg RT correl (%)[a] | Low easy versus Hard perf. Delta (%) | High oper. versus Pre. Perf. Delta (%) |
|---|---|---|---|---|---|---|---|---|
| Pre-OP | | | | | | | | |
| TC | 17176 | 62.0 | 0.8 | 0.2 | 0.0 | 0.1 | 0.1 | 0.4 |
| OP begins | | | | | | | | |
| OP | 18599 | 75.4 | 0.9 | 5.4 | 0.3 | 4.0 | 0.1 | 0.4 |
| TC | 5856 | 59.5 | 0.7 | 0.6 | 0.0 | 0.4 | 0.0 | 0.7 |
| New forms published | | | | | | | | |
| OP | 18131 | 76.7 | 1.0 | 2.4 | 0.1 | 1.6 | 0.0 | 0.1 |
| TC | 9596 | 68.5 | 0.9 | 0.2 | 0.0 | 0.1 | 0.1 | 0.2 |

[a]Negative correlation between item response times and average response times.

**Table 5.** Select Security Flags and Collusion Flags.

| | | No collusion | | | Collusion | | | |
|---|---|---|---|---|---|---|---|---|
| Mode | *N* | Fast correct resp. (%) | Neg RT correl (%) | Pass rate (%) | *N* | Fast correct resp. (%) | Neg RT correl (%) | Pass rate (%) |
| Pre-OP | | | | | | | | |
| TC | 17172 | 0.2 | 0.1 | 62 | 4 | | | |
| OP begins | | | | | | | | |
| OP | 13988 | 0.9 | 1.5 | 67.8 | 4611 | 19.2 | 11.5 | 98.2 |
| TC | 5632 | 0.4 | 0.3 | 58.2 | 224 | 6.3 | 4.5 | 92.4 |
| New forms published | | | | | | | | |
| OP | 16767 | 0.7 | 0.7 | 74.9 | 1364 | 23.0 | 12.9 | 98.5 |
| TC | 9547 | 0.2 | 0.1 | 68.3 | 49 | 8.2 | 0.0 | 89.8 |

the different sequences of modalities. The first thing to point out is that very few repeaters changed modality between Time 1 and Time 2. Test takers who began in OP but switched to TC had a much higher incomplete rate for Time 1 than those in the OP- > OP group (26% vs 7%), suggesting that they may have had technical problems and switched to test centers when able. The table also shows that the TC- > TC test takers had the lowest Time 2 pass rate (39%) and the smallest score gain (5.3), while the TC- > OP test takers had the highest pass rate (53%) and largest score gain (11.6), as expected. Finally, those with incomplete Time 1 results have a similar Time 2 pass rate to the overall first-time pass rate on the test (81%), strongly indicating that they should be excluded from further retest analyses, as the score gains are likely due to legitimately finishing the test.

Table 7 incorporates the collusion flags and excludes test takers with an incomplete Time 1 or Time 2 test. Prior to OP there were no repeat candidates flagged for collusion, the retest pass rate was 40%, and the average number-correct score change was 4.9 points. In both initial OP and later OP, retest pass rates are much higher for test takers flagged for collusion than they are for those not flagged and were even higher than the first-time complete test taker population of 81%. Score changes are largest for the TC- > OP group, and retest pass rates are also highest for that group, along with the handful of OP- > TC cases. This is true in both initial OP and later OP samples.

When comparing just the No Collusion OP- > OP and TC- > TC groups, the OP group shows about a 1-point higher score change, but a 10% higher pass rate than the TC group. While changes in the population, or changes to training and preparation may have resulted in the initial months of the pandemic, such changes would not seem to explain the differences between modalities.

## Discussion

Two salient findings emerge from these two studies. Study 1 shows that the magnitude of high answer overlap across many programs corresponds to the relative level of security presumed in each delivery mode; however, the magnitude of this difference is related to the composition of the samples evaluated. Study 2 shows that, within the context of a single program, the onset of the OP delivery option resulted in increased high answer overlap and that high answer overlap is accompanied by corroborating indicators of cheating, namely, higher incidence of rapid response times and lower correlation between observed and expected response times. Retest score gains in Study 2 also indicate that Time 1/Time 2 modalities are associated with differences in repeat performance and pass rates, as are test takers flagged for possible collusion. Results such as these should be taken seriously when evaluating the suitability of delivery mode for a particular testing application.

**Table 6.** Retest Performance by Mode and Completion Status.

| | Time 1 complete | | | Time 1 incomplete | | |
|---|---|---|---|---|---|---|
| Time 1 modality/Time 2 modality | *N* | Pass % | Score change | *N* | Pass % | Score change |
| OP- > OP | 2949 | 52 | 7.6 (10.6) | 218 | 81 | 38.2 (15.2) |
| OP- > TC | 212 | 50 | 7.8 (7.8) | 75 | 80 | 39.7 (17.8) |
| TC- > OP | 748 | 53 | 11.6 (14.7) | 1 | | |
| TC- > TC | 3056 | 39 | 5.3 (6.7) | | | |

**Table 7.** Retest Changes by Mode and Time Period for Complete Tests.

| | No collusion | | | Collusion | | |
|---|---|---|---|---|---|---|
| Seq | *N* | Pass rate (%) | Mean score change (SD) | *N* | Pass rate (%) | Mean score change (SD) |
| Pre-OP | | | | | | |
| TC- > TC | 1385 | 40 | 4.9 (6.0) | 0 | | |
| OP begins | | | | | | |
| OP- > OP | 1126 | 46 | 6.8 (8.6) | 192 | 94 | 21.2 (10.9) |
| OP- > TC | 76 | 47 | 8.1 (8.1) | 6 | | |
| TC- > OP | 393 | 41 | 8.3 (11.5) | 108 | 100 | 29.1 (10.5) |
| TC- > TC | 663 | 36 | 5.6 (7.7) | 27 | 93 | 18.1 (8.6) |
| New forms published | | | | | | |
| OP- > OP | 1499 | 49 | 6.7 (8.2) | 85 | 94 | 20.2 (12.4) |
| OP- > TC | 128 | 48 | 6.8 (6.8) | 2 | | |
| TC- > OP | 188 | 47 | 8.7 (11.8) | 36 | 100 | 27.5 (10.1) |
| TC- > TC | 974 | 37 | 5.4 (6.4) | 7 | | |

An analysis of collusion flags from the various test sponsors represented in Figure 1 (not reported for confidentiality reasons) shows different testing programs are subject to different levels of cheating for different reasons. A self-assessment, for example, is unlikely to elicit cheating as there are no stakes attached to the results and the test takers are interested in learning from the valid results of the test. If grades or other rewards were attached to the results, it is possible that some level of cheating might begin to occur. Testing programs with financial rewards attached to passing status (promotion or salary incentives), especially those with course-sized prerequisites, as opposed to curriculum-sized prerequisites, might inspire high levels of cheating. This is especially true when, as is often the case, such programs test in diverse locations in the interest of market penetration. While the results reported in this article cover 326 different testing programs and over 3 million test takers, the generalizability of these results to a specific testing program will depend on the characteristics of that program. These results do indicate that collusion is more likely in OP settings than TC overall. This pattern is maintained when controls for sample size and comparability of test centers offered is controlled for, although the magnitude changes. Moreover, as noted above, while all TC results showed lower levels of collusion than OP, test centers with fewer controls over hardware and security showed more flags than those with greater controls.

The analysis of a single test program over time in Study 2 also showed higher levels of collusion in OP than in TC. The baseline information provided for the program indicates that test content was likely not widely available prior to OP. Following the onset of OP, there was an increase in indications of preknowledge in test centers. The increase in RSI combined with the lack of 100% overlap suggests that some TC test takers were memorizing the exposed content. The high RSI along with high levels of 100% overlap suggest that OP colluders feature a combination of those memorizing exposed content and proxy test takers. It is possible, but unlikely, that numerous test takers would memorize the exposed content so well that their responses would be identical. It is more likely that individuals testing for others would respond in the same way over multiple tests (Becker & Makransky, 2011).

The additional test security flags also showed an informative interaction with collusion flags. Pass rates were much higher for flagged test takers in the overall population and the first-time retest population. Additionally, fast correct responses and negative response time correlations were also much more common in the collusion-flagged group. The authors suggest that extreme levels of response overlap combined with non-extreme scores may be indicative of proxy test takers. Research in this area is ongoing. Relatively short tests, relatively shallow content domains, and relatively exposed content will lead to high levels of correct overlap, but not necessarily exact incorrect overlap. As organizations identify verified test takers who are memorizing material versus verified proxy test takers, we will hopefully increase our knowledge of their characteristics.

This report is intended to show how one may usefully analyze operational data to evaluate the security of online proctored test administrations relative to test center test administrations. We do not think that the instances of high overlap or the test security flags reported here constitute the "true rate of misconduct." Not every cheating instance may trigger a flag. While chance overlaps to the degree reported here would be extremely unlikely, a few could happen, while other more idiosyncratic efforts to cheat might go undetected. Continued research on cheating detection, as well as effective responses to cheating, continue to develop. Many factors contributing to the likelihood of cheating and the ability to detect it could not be included here, such as the value or stakes of programs, geographic considerations, exposure, test length, or content breadth. Given the concentration of interest in these issues by test delivery providers and sponsors, it is also likely that OP technology and program actions regarding security issues will continue to change the test security environment going forward. The authors encourage individual programs using OP, and testing vendors offering this service, to evaluate the security of their testing modalities.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iD

Kirk A. Becker https://orcid.org/0000-0002-4141-8740

## Supplemental Material

Supplemental material for this article is available online.

## References

American Educational Research Association. (2014). American psychological association, national council on measurement in education. *Standards for educational and psychological testing*. AERA. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf

Becker, K. A., & Jones, P. (2022). Varying item parameters and collusion detection. *Paper presented at the national council on measurement in education annual meeting*. ERIC

Becker, K. A., & Makransky, G. (2011). Verifying candidate identity over time: Candidate response consistency for repeated test items. *Presented at the association of test publishers annual conference*. Association of Test

Becker, K. A., & Meng, H. (2022). Identifying statistically actionable collusion in remote proctored exams. *Journal of Applied Testing Technology*, *23*(Special Issue), 54–61.

Bird, C. (1927). The detection of cheating in objective examinations. *School and Society*, *25*(635), 261–262.

Cizek, G. J., & Wollack, J. A. (2017). *Handbook of quantitative methods for detecting cheating on tests*. Routledge.

Foster, C., & Marder, A. (2020). *Stealing test content with hidden cameras*. Presentation at the Association of Test Publishers annual meeting, Online.

Hurtz, G. M., & Weiner, J. A. (2022). Comparability and integrity of online remote vs. onsite proctored credentialing exams. *Journal of Applied Testing Technologies*, *23*(Special Issue), 36–45.

Kingston, N. M., & Clark, A. K. (2014). *Test fraud*. Routledge.

Maynes, D. (2012). Detection of non-independent test taking by similarity analysis. *Paper presented at the statistical detection of potential test fraud conference*. University of Wisconsin–Madison

Maynes, D. (2017). Detecting potential collusion among individual examinees using similarity analysis. In G. J. Cizek, & J. A. Wollack (Eds.), *Handbook of quantitative methods for detecting cheating on tests*. Taylor & Francis

Saupe, J. L. (1960). An empirical model for the corroboration of suspected cheating on multiple-choice tests. *Educational and Psychological Measurement*, *20*(3), 475–489. https://doi.org/10.1177/001316446002000304

Smith, R. (2019). Comparing B3 to answer similarity Index for detecting collusion. *Paper presented at the annual meeting of the conference on test security*. University of Miami

Wollack, J. A., & Fremer, J. J. (2013). *Handbook of test security*. Routledge.

Zopluoglu, C. (2017). Similarity, answer copying, and aberrance: Understanding the status quo. In G. J. Cizek, & J. A. Wollack (Eds.), *Handbook of quantitative methods for detecting cheating on tests*. Taylor & Francis