EDUCATION
RESEARCH SERVICE

COMMENTARY

Revised methodology for the examinations of the National Board of Chiropractic Examiners: *Impact on institutions, faculty, and students*

Michael R. Wiles, DC, MEd, MS, EdD, Craig S. Little, DC, MEd, EdD, and John P. Mrozek, DC, MEd, EdD

For over 50 years, the National Board of Chiropractic Examiners (NBCE) has administered pre-licensure examinations to chiropractic students and graduates. During this time, the testing process has been continually refined and improved, consistent with the evolving science and practices of standardized testing. NBCE test results are provided to chiropractic program leaders who use these data to improve their curricula as part of their own ongoing efforts to refine and improve the academic programs. Finally, the Council on Chiropractic Education (CCE) requires accredited chiropractic programs to report their NBCE scores to ensure that benchmarks set by the CCE are met. With this symbiotic relationship between the NBCE, CCE, and chiropractic programs (as well as state licensing authorities), it is very important that these groups collaborate and communicate with transparency and diplomacy. In particular, the chiropractic program leaders—and their students as the end users—are vitally interested in monitoring changes at the NBCE and CCE levels that may impact their programs. Recent changes in testing methodology for the NBCE examinations need to be understood and monitored to ensure that they result in their intended outcome, which is greater validity of the testing process. This commentary reflects the views and concerns of 3 chiropractic educational leaders and is intended to facilitate further discussion among chiropractic program leaders toward strengthening the aforementioned symbiotic relationship.

Key Indexing Terms: Chiropractic; Education; Certification/Standards

J Chiropr Educ 2020;34(1):68–70 DOI 10.7899/JCE-19-28

The expansion of standardized testing into virtually all areas of the American educational system has provided policymakers, accreditation agencies, and school administrators with renewed interest in including test scores into institutional performance evaluation criteria.[1] The Council on Chiropractic Education (CCE), a programmatic accreditation agency, mandates that, "Doctor of Chiropractic Programs (DCPs) must disclose up-to-date results of student performance on national board examinations and completion rates on the program website." In the United States, the organization responsible for administrating the national board examinations is the National Board of Chiropractic Examiners (NBCE). Additionally, CCE requires that "the overall weighted average for the 4 recent years' NBCE Parts I, II, III, and IV exam success rates must not be less than 80%."[2] Therefore, there is a need for a symbiotic relationship between the chiropractic programs and the NBCE with reciprocal elements of mutual dependency. Recently, there have been several changes on the NBCE side of this relationship that are being looked at carefully by chiropractic educators, including testing methodology and analysis, reporting of results to students and chiropractic programs, the registration process for students and programs, and the retake policy. There are many factors to consider in the evaluation and design of assessment instruments, the most important of which are reliability, validity, educational impact, acceptability, and examination costs.[3] The extent to which these changes impact the programmatic side of this symbiotic relationship and the views and concerns of the authors as stakeholders are the subject of this commentary.

Programs must be able to utilize NBCE examination data to report student academic achievement to their accreditors and to make constructive and appropriate changes to their curricula. Programs must first rely on the validity and reliability of the data and properly interpret them in making curricular decisions. This point is important, as programs constantly refine curricula to reflect best (and current) practices, including anticipating future trends in order to decrease the lag between curricular change and its practical outcome in chiropractic practice. At the same time, programs must be responsive to the content of NBCE examinations, which may also lag current practices.

The validity of test score interpretation is the focus of Standard 9 in the Standards for Educational and

Psychological Testing.[4] Standard 9.0 states, "The users are responsible for knowing the validity evidence in support of the intended interpretations of scores on test that they use, from test selection through the use of scores, as well as common positive and negative consequences of test use." Standard 9.3 continues, "The test user should have a clear rationale for the intended uses of a test or evaluation procedure in terms of the validity of interpretations based on the scores and the contribution the scores make to the assessment and decision-making process." Test users who interpret and use scores are responsible for ascertaining that there is appropriate validity evidence supporting their interpretations of test results.

One important change in NBCE examination analyses has been the introduction of item response theory (IRT) methodology, replacing (some might say supplementing) classical test theory (CTT) methodology, which is described in this issue by Himelfarb et al.[5] IRT is a measurement framework that has become the focal point in large-scale assessment, surpassing CTT.[6] Measurement models under IRT specify the probability of a correct response to an item, which is both a condition of a test taker's ability and a specific item's difficulty. CTT uses simple definitions and weak statistical assumptions. Moreover, CTT focuses on results for groups or test takers rather than for individual examinees; therefore, individual ability is not estimated, and the probability of the correct response is unconditional. This type of estimation is weaker when compared to IRT. CTT works with item statistics, such as item difficulty and item-test correlation, which are completely dependent on the particular sample, and CTT does not provide ready means for result generalization from one group of examinees to another.[7] IRT, on the other hand, relates a test taker's ability to the probability of item response. The predictions made under IRT are more precise and of a wider range. Similar to CTT, IRT produces unconditional estimates (for groups). In addition, with IRT, conditional estimates (for individuals) are also available.

Shifting from CTT to IRT is a positive step, as IRT appears to have greater utility in score production for the purpose of large-scale assessment practice. This also means that programs now receive new data reports that offer opportunities as well as challenges in using the data for program improvement(s).

Many (perhaps all) chiropractic programs have used NBCE data for institutional evaluation and accountability. Overall pass rates have been useful to fulfill the reporting requirements of the CCE. Pass rates of the various domains of Parts I–III have been helpful, especially when compared to national averages, to identify possible areas of programmatic quality improvement. Some programs have established benchmarks or thresholds to identify specific domains and subdomains where trends on weak scores indicate a need for curricular attention.

The new NBCE program-level reports utilize scaled domain means in place of arithmetic means but without the domain pass rates previously provided. Scaled domain means can vary from one cohort to the next and create somewhat of a moving target for programs (compared to the previous long-term trend data). Program officials may need to review their current practices in light of these new data to determine thresholds for program improvement actions. The NBCE currently provides programs with data regarding first-time fail rates, domain pass rates (including number and percentage of failing students), and specific domain scores per individual student for Parts I and II, a practice that the authors of this commentary endorse and appreciate.

Students were also affected considerably by the reporting and retesting practices of the NBCE, which were amended in early 2019 and reversed back to prior practice within 6 months. Currently, students receive domain scores regardless of passing or failing the domains. Failing 1 or 2 domains requires the retaking of only those domains, whereas students must retake the entire exam only if they fail more than 2 domains. Failing students receive feedback on their performance from the NBCE. These feedback reports provide an overall board score and a percentile rank for each domain and category that illustrates where the student fell in relation to the sitting cohort. Percentile rankings do not indicate if a student's performance in any domain was at a passing level; rather, they communicate to students only how they performed relative to the sitting cohort. Remediation efforts are considerably more effective if students can focus on failed domains rather than on domains passed but with lower percentile rankings.

The application of IRT methodology is an important step in creating examinations that take into consideration individual examinee "ability" regardless of the cohort characteristics and, like CTT methodology, are also useful for cohort analysis. Additionally, the NBCE has transitioned to digital radiology images in the Part IV examination. This appears congruent with the shifting landscape in diagnostic imaging. Many chiropractic educational programs include digital imaging in their curriculum. Himelfarb and et al[8] address the minimal difference between the 2 platforms (plain film and digital) regarding assessment and describe the many positive attributes of digital representation. In addition, to decrease bias and content underrepresentation, the NBCE now features 20 image stations in Part IV, a valuable improvement to the exam experience.

In summary, while there have been some bumps in the road related to process and communication between the NBCE and the colleges (including the administrative changes related to student registration for the examinations), overall the NBCE has made some positive changes to enhance the reliability and validity of its examination results. This is of vital importance to programs, which also impacts their regional and programmatic accreditation. The CCE includes NBCE examination data as a measure of program effectiveness and relies on such data as an important third-party verification of student academic achievement.

Continued communication and transparency within the symbiotic relationship between the colleges and the NBCE will be of extreme importance in the future. Educators are

beginning to take note of the "stasis in global chiropractic education"[9] and questioning its response to a rapidly changing educational environment. Among several key factors in "reimagining chiropractic education"[9] is the role of standardized examinations, such as the NBCE. Ultimately, stakeholders rely on independent third-party testing agencies such as the NBCE to confirm practitioner competence to governments and the profession and, most important, to reassure the public of the competence of licensees.

We are indeed in for some interesting times as our profession leaves its adolescence and grapples with best practices in standardized testing, and the use of these data that will require transparency, diplomacy, and, above all, communication.

FUNDING AND CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare relevant to this work.

About the Authors

Michael Wiles is the dean of the College of Chiropractic Medicine at Keiser University (2085 Vista Parkway, West Palm Beach, FL 33411; mwiles@keiseruniversity.edu). Craig Little is the president of the Council on Chiropractic Education (8049 North 85th Way, Scottsdale, AZ 85258; little@cce-usa.org). John Mrozek is the vice president for Academic Affairs at Texas Chiropractic College (5912 Spencer Highway, Pasadena, TX 77505; jmrozek@txchiro.edu). Address correspondence to Michael Wiles (2085 Vista Parkway, West Palm Beach, FL 33411; mwiles@keiseruniversity.edu). This article was received May 21, 2019; revised June 4 and November 1, 2019; and accepted November 1, 2019.

Author Contributions

Concept development: MRW, JPM, CSL. Design: MRW, JPM, CSL. Supervision: MW. Data collection/processing: n/a. Analysis/interpretation: n/a. Literature search: MRW, JPM, CSL. Writing: MRW, JPM, CSL. Critical review: MRW, JPM, CSL.

REFERENCES

1. Grissom J, Kalogrides D, Loeb S. Using student test scores to measure principal performance. *Educ Eval Policy Anal*. 2014;20(10):1–26.
2. Council on Chiropractic Education. *Manual of Policies of the Council on Chiropractic Inc*. Scottsdale, AZ: The Council; 2019. http://www.cce-usa.org/uploads/1/0/6/5/106500339/2019-01_cce_manual_of_policies.pdf.
3. Chandratilake M, Davis M, Ponnamperuma G. Evaluating and designing assessments for medical education. *Internet J Med Educ*. 2009;1(1):1–7.
4. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association, American Psychological Association, National Council on Measurement in Education; 2014.
5. Himelfarb I, Shotts BL, Tang N, Smith M. Score production and objective methods used by the NBCE for post-exam analyses. *J Chiropr Educ*. 2020;34(1):35–42.
6. Embretson SE, Reise SP. *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
7. Yen W, Fitzpatrick AR. Item response theory. In: *Educational Measurement*. 4th ed. Westport, CT: Praeger Publishers; 2006:111–153.
8. Himelfarb I, Seron MA, Hyland JK, et al. The transition to digital presentation of the diagnostic imaging domain of the part IV examination of the National Board of Chiropractic Examiners. *J Chiropr Educ*. 2020;34(1):52–67.
9. Ebrall P. The first chiropractic hackathon: reimagining chiropractic education. *Chiropr J Aust*. 2018; 46:151–161.