Abstract

Standardized examinations were created to evaluate student academic aptitude and eliminate bias from college admissions. The validity of standardized examinations in predicting student success in doctoral health professions education has been minimally explored. A scoping review was conducted to determine what is known about the predictive value of standardized tests in doctoral health professions education in the United States and Canada. A total of 323 titles and abstracts were reviewed for inclusion and exclusion criteria. Fifteen full-text articles were ultimately chosen for inclusion in this study. The results indicate conflicting data and underscore the necessity for further research in this area.

Key Words: doctoral health profession, standardized exam, academic success

Background

Standardized testing is used as a formal assessment of academic ability in order to predict student success in higher education at both the undergraduate and graduate levels. Standardized tests differ from in-class assessments of knowledge because they are administered in a controlled environment, thus allowing for comparison of student performance that is presumably independent of socioeconomic status (SES), gender and race. Two supporting arguments for standardized testing assume that the standardization process eliminates potential for bias and that standardized tests can accurately assess students’ intellectual ability.

The first large-scale standardized tests were administered in 1901 by the College Entrance Examination Board, which is known today as the College Board. One reason for creating and administering the tests was to reduce the volume and variety of pre-matriculation exams required by each undergraduate institution. Due to a national push to mandate aptitude tests for college admission, the College Board administered the first Scholastic Aptitude Test (SAT) in 1926 to thousands of students. In 1959, the American College Testing Company administered the ACT for the first time. Shortly after the first administration of the SAT, American graduate institutions followed suit. The first Medical College Admissions Test (MCAT) was given in 1928, the first Graduate Record Examination (GRE) in 1949, and the first Law School Admissions Test (LSAT) was given in 1948. The first Optometry College Admission Test (OCAT) was developed in 1971 by the Psychological Corporation under the sponsorship of the Association of Schools and Colleges of Optometry and was administered for the first time in the fall of 1972.

Despite attempts to eliminate bias from standardized examinations, many in academia argue that requiring standardized exams for admission into undergraduate and graduate programs significantly disadvantages students of female gender and lower SES and those in under-represented minority (URM) groups. Miller and Stassun argue that requiring the GRE significantly decreases the opportunities for women, URM students and lower SES students to enter the science, technology, engineering and math (STEM) professions. They point out that the Educational Testing Service, which administers the GRE, publicizes that women score 80 points lower in the physical sciences than men, and African Americans score 200 points below white test-takers on the exam.

Moneta and Koehler report that students with low SES perform worse on standardized exams, possibly due to lack of access to academic preparation and lack of funds to pay for exam retakes if the first score is low. Nankervis argues that the SAT underestimates future success of female test-takers because males average 35 points higher than females in the mathematics section. In another publication, Wilson discusses a study demonstrating that metrics-based file reviews of applicants excluded twice the number of applicants who identified as historically URM, and moving away from metrics-based admissions processes resulted in a remarkable increase in admission of URM students to a doctoral biomedical science program. Wilson also reiterates that women and URM students score lower on the GRE than white and Asian-American men; therefore, using GRE scores to stratify doctoral applicants significantly reduces the diversity of the applicant pool.

In addition to the reported bias against women, URM students and students of lower SES, there are conflicting reports about the ability of standardized exams to predict academic success at both the undergraduate and graduate level. Kuncel argues that standardized exams are effective predictors of performance in graduate school but the combination of standardized exam scores and undergraduate grade point average (uGPA) gives the most accurate prediction of academic success. He also states that student motivation and interest are crucial for continued exertion throughout graduate school and cannot be measured with standardized exams.

In contrast, Miller and Stassun argue that there is a weak correlation between the GRE and success in STEM fields. They
point out that research from the Educational Testing Service shows that the predictive validity of the GRE is limited to first-year graduate grade point average (gGPA) but academic success is much broader than first-year gGPA. Academic success encompasses first-year gGPA, gGPA, degree attainment, licensing examination performance, faculty evaluation of students, residency attainment and completion, and numerous other outcomes that cannot be predicted by GRE scores.

Similarly, Moneta-Koehler and colleagues found that GRE scores are moderate predictors of first-semester graduate grades and weak to moderate predictors of overall gGPA. They also found that the GRE does not predict other skills necessary to succeed in biomedical doctoral programs and concluded that the limited benefits of the GRE do not outweigh the expense of excluding URM students and students of lower SES from entering graduate biomedical doctoral programs.

| Table 1. Click to enlarge |

While a reasonable amount of literature about the ability of standardized exams to predict success in biomedical graduate programs exists, current studies exploring the validity of standardized exams in predicting academic success in doctoral health professions programs are relatively scant, particularly in the field of optometry. To gain insight into how many doctoral health professions programs require standardized exams as part of their admissions requirements, the authors first identified 12 health professions to explore, including optometry, dentistry, allopathic medicine, osteopathic medicine, podiatry, audiology, physical therapy, veterinary medicine, occupational therapy, pharmacy, acupuncture and chiropractic programs. Nursing was not included due to the wide range of doctoral-level nursing programs, the diverse pathways to attainment of doctoral degrees within the nursing profession, and the differing requirements for each program. Once the programs for inclusion were identified, their respective standardized test requirements were summarized (Table 1).

As suggested by Arksey and O’Malley, a scoping review methodology is well-suited for four primary contexts. The review question undertaken in this research addresses three of the four circumstances identified: 1) to examine the extent, range and nature of research activity; 2) to summarize and disseminate research findings; and 3) to identify research gaps in the existing literature. This project utilized a scoping review methodology to gain a deeper understanding of the current status of standardized testing in health professions admissions processes, along with any research evaluating the predictive power of pre-admissions standardized test results for ultimate academic and/or professional success. Furthermore, this research seeks to summarize what is currently known by gathering and assessing published works relevant to this inquiry, and upon review of the summary, to identify needs for further research on this topic.

**Methods**

The scoping review methodological framework adheres to the guidelines suggested by Arksey and O’Malley and was conducted as an iterative process following the five suggested stages: Stage 1: identifying the research question; Stage 2: identifying relevant studies; Stage 3: study selection; Stage 4: charting the data; and Stage 5: collating, summarizing and reporting the results.

The investigators determined the inclusion and exclusion criteria prior to searching the literature. The inclusion criteria focused specifically on scholarly, peer-reviewed indexed literature describing information related to health professions education, admissions to health professions education programs, and standardized testing. The time period for inclusion was limited to the past 10 years, and language was limited to English. Only articles published about health professions education in the United States and Canada were included due to potential differences in health professions education relative to governmental, regulatory, economic and cultural factors in countries outside North America.

Articles with a focus on non-health professions programs, such as biomedical sciences, and articles from gray literature sources were excluded. Gray literature was defined as a thesis, dissertation, non-peer reviewed study, conference proceeding or editorial. It was noted that one article could have multiple reasons for rejection. To establish clearly defined guidelines for rejection, each reason was enumerated. In each case where a paper was excluded, the primary reason for rejection was noted.

**Literature search**
A combination of methods was used to locate articles for this scoping review. Keywords were used to retrieve the broadest possible number of articles related to the research question, and controlled language was used to construct a narrow, defined search strategy. A discussion among the investigators, serving as content experts, and two vision science librarians, serving as technical experts, reviewed the research question in depth and resulted in a listing of the following keywords for a broad search: educational measurement, school admissions criteria, academic success, GRE admissions health professions, standardized testing graduate education. A review of the controlled language (Table 2) available in electronic databases was conducted using the keywords to locate appropriate terms to create focused search strategies. All keywords, controlled language terms and subsequent search strategies were vetted by both vision science librarians.

The databases included in the study were selected based on library subscription, availability of controlled language search option, and comprehensive coverage of the topic. The investigators and the consulting librarians determined which search filters (Table 3) would yield the best, most relevant results. The searches of the following databases were conducted in August 2019: PubMed/Medline, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Cochrane Library.

All searches for all databases were filtered for English-only articles published in the past 10 years. Embase had the following additional filters added: Embase-only articles, articles in press, and review-only articles. After a review of the keywords, the controlled language terms were chosen to best retrieve search results that were focused more on the topic being searched.

Inter-rater reliability calibration

After mechanical and manual de-duplication of all articles resulting from the full electronic search, all investigators participated in a calibration session to determine inter-rater reliability. In this session, 14 abstracts were reviewed to gauge inter-rater reliability regarding the application of the inclusion and exclusion criteria. This session took place prior to the title, abstract and full-text review. The decision-making process for each of the articles reviewed was documented on a case-by-case basis then summarized to identify major categories for the reasons to include or exclude an article. Through discussion, the investigators’ inter-rater reliability for title/abstract review was able to reach 100% agreement.

After calibration, each investigator reviewed two-thirds of the titles and abstracts, with two investigators randomly assigned to each article. If consensus was reached by two of the investigators to either include or exclude an article, that action was immediately taken. If there was no consensus, the third investigator was used as a tiebreaker to determine final inclusion or exclusion of the article. Once a full consensus was achieved regarding every article, each investigator was randomly assigned 11 full-text articles for a thorough review.

As suggested by Arksey and O’Malley, data extracted from each source were charted and entered into a data-charting form using the database program Excel. Data were charted independently by each investigator with confirmations by co-investigators when questions arose. Data charting focused on summarizing each publication’s process and methodology, predictors and outcomes. For each source included in the scoping review, the following variables were included: authors, year of publication, study location (to ensure United States or Canada), health professions studied, study design utilized, standardized admissions test evaluated, outcome measures or indicators of success and means of measuring outcomes, statistical test, results and significance level, and the publication’s main conclusions.

Results

A total of 323 articles underwent title and abstract review by two authors to determine inclusion for full article review. Of the 323 articles reviewed, the two authors agreed on 305 (94.4%) of the articles, with 18 (5.6%) articles requiring a title and abstract review by the third author. Of the 18 titles and abstracts that underwent a review by the third author, five were accepted for the full article review, resulting in a total of 33 articles accepted for full article review and 290 rejected articles. The primary reasons for rejection were summarized in Table 4. After the full article review, 18 additional articles
were excluded for the reasons summarized in Table 5. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart for this review can be found in Figure 1.

Table 4. Click to enlarge

<table>
<thead>
<tr>
<th>Primary Reason for Rejection</th>
<th>Number of Articles Rejected</th>
<th>Percentage of Articles in this Category (from total of 232)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not about admissions / off topic</td>
<td>138</td>
<td>42.1%</td>
</tr>
<tr>
<td>Not from North America</td>
<td>60</td>
<td>21.1%</td>
</tr>
<tr>
<td>No standardized test</td>
<td>31</td>
<td>9.6%</td>
</tr>
<tr>
<td>Biomedical / PhD / other non-clinical</td>
<td>16</td>
<td>5.5%</td>
</tr>
<tr>
<td>Duplicate</td>
<td>9</td>
<td>2.8%</td>
</tr>
<tr>
<td>About residency education</td>
<td>8</td>
<td>2.5%</td>
</tr>
<tr>
<td>Baccalaureate or other non-graduate level</td>
<td>5</td>
<td>1.5%</td>
</tr>
<tr>
<td>Other reasons or multiple reasons</td>
<td>15</td>
<td>4.6%</td>
</tr>
<tr>
<td>Total Rejected</td>
<td>280</td>
<td>89.4%</td>
</tr>
</tbody>
</table>

Table 5. Click to enlarge

The final scoping review was completed for a total of 15 papers (Table 6). The six different health professions represented in the 15 papers included in the scoping review were pharmacy (5 papers), dental medicine (3 papers), veterinary medicine (2 papers), physical therapy (1 paper), allopathic medicine (3 papers), and allopathic medicine combined with PhD (1 paper).

Four different standardized tests were represented in the papers included in the scoping review: GRE (four articles), Pharmacy College Admission Test (PCAT - five articles), Dental Admission Test (DAT - two articles), and MCAT (four articles).

The majority of papers included in the scoping review assessed more than one primary outcome variable. Eleven of the papers included some assessment of the association between standardized tests and gGPA, including the ability to predict the gGPA at different points in the program, such as at the time of graduation, in the first-year curriculum, multiple years in the program, or for specific courses such as basic science courses or clinical evaluations. Eleven of the papers evaluated the association between standardized tests and performance on board examinations. Three papers evaluated the predictive value of standardized exams on residency success. One paper evaluated the predictive value of the MCAT on gGPA, time to defend PhD, board scores, publication number and career outcomes.

Ten of the papers reported positive findings. The DAT, GRE, MCAT and PCAT were all found to be predictive of board examination results. The GRE, DAT and PCAT were found to be positive predictors of gGPA.

Five of the papers reported somewhat mixed results. For example, it was found that the PCAT was predictive of gGPA; however, PCAT scores were inadequate when used alone. PCAT scores were also found to be less strongly predictive of pharmacy program GPA than the uGPA. One paper found that PCAT scores were predictive of pharmacy program GPA but not predictive of board examination results. Similarly, another paper found that the DAT was less predictive than the uGPA for the gGPA. One paper found that the GRE indirectly predicted board scores via the gGPA. Thus, in essence, the gGPA was more predictive of board scores than the GRE.
Three papers did not find an association between a standardized test and the outcome of interest. In two studies, the MCAT did not predict board examination results or gGPA. In one study, the GRE did not predict specialty board results.

Several limitations are noted for this scoping review. The review did not yield any optometry-specific literature using the databases that the investigators chose; thus, optometry-specific literature was not included in the scoping review. Additionally, the review focused only on graduate, doctoral-level health professions in the United States and Canada. Other research for non-doctoral-level professions as well as research in other countries may have provided additional insight into this research question. The authors decided not to include literature about nursing programs due to the wide range of programs and varying pathways to degree attainment within the profession as it differs greatly from the traditional academic trajectory of an optometry student. Including studies from doctoral nursing programs may add additional insight about the question at hand. The time frame for inclusion in the scoping review was limited, and publications before and after the period of the review may provide additional information. The scoping review used qualitative techniques for interpreting the data and did not employ quantitative methods. Additional research on this topic might have been found by including published abstracts, conference proceedings or other sources of gray literature. Despite these limitations, this scoping review seeks to shed light on a topic of great interest for the profession of optometry, particularly due to the paucity of current research on this topic as it relates directly to optometric education.

Discussion

Standardized tests have been used to assess academic aptitude in order to determine student preparedness for higher education. Initially designed to minimize potential for bias and increase accuracy in assessing students’ intellectual ability, standardized tests have come under increasing scrutiny due to possible bias against low-income, minority and female students. In addition to this drawback, many have questioned the true ability of standardized exams to predict student success, particularly in under-represented populations. The purpose of this study was to conduct a scoping review of the ability of standardized tests to predict success in doctoral-level health professions programs in order to shed light on the role of these exams for admission into optometry school.

The results of the scoping review suggest that health professions programs are invested in evaluating the predictive value of standardized testing as a tool to be utilized in the admissions process. Most of the publications included in this scoping review assessed the relationship between standardized testing and academic achievement within a specific health professions program. Few publications carried the assessments further into correlations with ultimate success in clinical practice. The publications included in this scoping review demonstrate disagreement about the value of standardized exams in predicting success in doctoral-level health professions education.

The limited number of articles included in this scoping review suggests that there is not an abundance of solid evidence to support the value of standardized testing for admissions decision-making in the health professions. There certainly appears to be some evidence of the value of standardized exams to predict academic success, but the magnitude of the potential benefits of testing has not been compared to the potential costs of limiting access to, and perpetuating bias against, under-represented student groups. Because the predictive value of standardized exams cannot be compared for applicants who did not matriculate into a health professions program, there is no means of knowing how many of those candidates would have been successful in the programs they were denied entrance into based on their exam scores. There is also no way to measure the ultimate impact on the pipeline of healthcare providers and the public who would have been served by the individuals who were denied entrance into the various programs.

Since the first administration of the OCAT in 1972 (which was renamed Optometry Admission Test in 1987), few studies have explored the ability of the exam to predict success in optometry school. Out of eight studies that looked at the predictive value of the OAT, four studies used the first- and second-year optometry GPA as the main outcome measures. Another four studies looked at additional outcome measures, including class rank at graduation, clinic
Standardized Tests as Predictors of Success in Health Professions Education: a Scoping Review

This scoping review demonstrates the limited body of research on a critical topic and highlights the need for further exploration to fully understand the complexities of the value of standardized testing as part of the admissions process for doctoral health professions programs, particularly in regard to optometry. The paucity of studies on this topic, particularly with regard to optometry, is troubling due to the ever-increasing demand for qualified and diverse healthcare providers. Understanding the framework of doctoral health professions education and the lessons that have been learned by other doctoral health professions will help guide decision-making by optometry admissions committees until new and relevant optometry-specific research is published. The current studies that exist on this topic suggest conflicting data about the ability of standardized exams to predict student success in doctoral health professions education. This finding makes it even more critical to invest in research on this topic as standardized exams can heavily disadvantage certain student populations, thus negatively impacting the profession of optometry by decreasing the diversity of their graduates.

Conclusion

While there is certainly an increasing move away from requiring standardized examinations across many disciplines, including the health professions, due to several reasons. If the recent decision by Indiana University School of Optometry (IUSO) to move toward test-optional admissions requirements is any indication, the profession of optometry may start to follow suit. One of IUSO’s motives for the pivotal change is its finding that “test scores are becoming a weaker predictor of future academic success,” which is at least partially consistent with the results of this scoping review. In light of this, the glaring lack of optometry representation in the literature search on this topic is greatly concerning. It needs to be ameliorated with future research so that optometry admissions committees have sound and current data on which to base their admissions decisions so they may ensure student success in their programs without affecting the diversity of their graduates.
information that admissions committees could utilize when selecting students into their programs. Additionally, case studies to further elaborate on perceived barriers and biases associated with the OAT are necessary to complete the picture of why this research needs to be conducted in the first place.

Acknowledgments

1. Rudy R. Barreras, Assessment & Public Relations Librarian, Western University of Health Sciences
2. Angela Lee, Associate Professor, Health Science Librarian, Pacific University

References


Dr. Jakirlic [njakirlic@westernu.edu] is an Assistant Professor at Western University of Health Sciences College of Optometry. She teaches various topics related to ocular disease and sees patients in clinic.

Dr. Ooley is an Associate Professor at Pacific University College of Optometry. She mainly teaches systemic disease and anatomy, and she sees patients in clinic.

Dr. Hoppe is the Founding Dean of the College of Optometry at Western University of Health Sciences. She teaches evidence-based decision-making.