ABSTRACT

Background: When a student makes multiple attempts at an admissions assessment, an institution must decide how to incorporate the resulting information into its admissions decision. However, little research exists to guide this decision within a nursing admissions context. Method: This article examines patterns in retesting scores from a nursing admissions assessment and the correlations of four retesting score treatments with scores on an early program assessment. Results: Although test scores do increase with each subsequent attempt, the average of all attempt scores is more highly correlated, in almost all instances, with future performance in a nursing program than any one test score. Conclusion: These results indicate that nursing programs presented with multiple scores from an examinee would be best served by using an average of all available scores when making admissions decisions. [J Nurs Educ. 2018;57(10):578-583.]

As nursing education programs establish criteria to admit qualified students, many use scores from standardized admissions examinations as a source of information about student preparedness. Realizing the importance of these scores to the admissions decision, prospective students may choose to take an examination multiple times, either to maximize their score to better their chances or to meet a program’s required minimum score. When a student presents multiple scores from an examination, programs may have difficulty in determining how to use all available information in making a decision about the student. Which is the best indicator of a student’s readiness to succeed in the program: The first test score? The most recent score? The highest score? The average of all available scores?

Although many educators have anecdotal impressions of how to address this practical problem, the question bears empirical investigation. This article examines several common treatments of multiple test scores and their relationship to early performance in a nursing program, with the goal of providing practical recommendations for programs making admissions decisions. For the purpose of this analysis, four treatments of scores from a nationally available nursing entrance examination, the Test of Essential Academic Skills (TEAS®) examination from Assessment Technologies Institute, LLC (ATI®) are investigated: first attempt, most recent attempt, highest attempt, and average of all attempts. Early program performance is operationalized as the score on a nationally available standardized assessment typically given during the first semester of a nursing program: ATI’s Fundamentals of Nursing examination. The research questions are as follows:

1. How do the first, most recent (last), highest (maximum), and average of all attempts (mean) scores of individuals with multiple TEAS attempts compare with scores of individuals with a single attempt?
   a. How do score patterns vary by number of TEAS attempts?
   b. How do score patterns vary for Associate’s (ADN) and Bachelor’s (BSN) degree program types?

2. What is the correlation of each of the score types (first, last, maximum, and mean) with early program performance on the Fundamentals of Nursing assessment?
a. How does the pattern of correlations vary by number of TEAS attempts?
b. How does the pattern of correlations vary by program type (ADN and BSN)?

LITERATURE REVIEW

Admissions personnel are tasked with selecting prospective students who are likely to be successful in nursing programs. Many selection criteria are often required when applying for admission to a program, as no single criterion is a perfect predictor of success. Standardized test scores’ use has grown in popularity in recent years. Standardized tests have been shown to be a predictor of nursing program success when considered alongside other factors, such as grade point averages, achievement in science courses, and admissions essays (Cunningham, Manier, Anderson, & Sarnosky, 2014; Schmidt & Mac-Williams, 2011).

However, approximately half of students retake standardized tests, which complicates the admissions decision (Patterson, Mattern, & Swerdzewski, 2012; Roszkowski & Spreat, 2016). Why would a student retake a standardized test? Although there are many reasons why an examinee might decide to retake an admissions test, the only definitive rationale is to raise a score (Lane & Feltz, 1976; Roszkowski & Spreat, 2016; Wolkowitz, 2011a; Wightman, 1990). If students are attempting to raise their scores, what does that tell us about the student, and how are admissions officers supposed to treat multiple scores on an examination?

An applicant’s decision to retake an examination may tell us something about motivation. For instance, Roszkowski and Spreat (2016) discuss Simon’s (1955) theories of satisficing and optimizing when retaking the SAT college admissions test. Examinees who are optimizers are trying to improve their score to be the best it can be, while those who are satisficers are trying to reach some minimal threshold. Other researchers consider retaking an examination to be an indicator of grit or determination (Roszkowski & Spreat, 2016; Zhang & Patterson, 2010). Zhang and Patterson (2010) examined the persistence of those taking the general education diploma (GED) and found that motivation played a role. They found a positive relationship between retesting and passing the GED when an examinee’s goal was to gain entrance into a 2-year college after earning a GED. Regardless of examinees’ motivation, the end goal of raising scores is the same.

To evaluate validity concerns related to admissions policies, many studies have described the nature of taking an admissions examination multiple times. With repeated attempts, scores typically increase and retesters tend to have lower initial scores, compared with students who test only once (Roszkowski & Spreat, 2016; Wolkowitz, 2011a). Villado, Randall, and Zimmer (2016) examined this issue and found that even though scores tend to increase over multiple attempts, single and repeat attempt scores show similar levels of correlation with the criterion. Repeat testers also tend to take their final attempt closer to program entrance, compared with those who only test once (Patterson, Mattern, & Swerdzewski, 2012). These concerns must be considered when comparing scores for students who take entrance examinations multiple times.

Consider policies that admissions offices could implement when prospective students have attempted an examination multiple times. They might choose to use the first attempt, most recent attempt, highest single attempt, average of all attempts, or some combination of the highest scores across multiple sections of an examination (Patterson et al., 2012; Roszkowski & Spreat, 2016). Patterson et al. (2012) examined this issue in relation to the SAT and found that none of the methods undermined the predictive validity of the test. Currently, no consensus exists in the literature for the proper treatment of multiple scores, although using the average of all attempts is often recommended (Dalessandro & McLeod, 1999; Roszkowski & Spreat, 2016; Wightman, 1990; Zhao, Oppler, Dunleavy, & Kroopnick, 2010).

Specialty programs, such as nursing, attract a considerably more homogeneous group of examinees than general college admissions tests. However, little exists in the nursing literature related to this topic. Therefore, policies from other specialty programs, such as medical and law schools, are useful for comparison. These specialty programs incorporate various guidelines for multiple attempts of the MCAT and LSAT, which are consistent with the aforementioned polices used for general college admission (Dalessandro et al., 1999; Wightman, 1990; Zhao et al., 2010).

The College Board (2015) also provides some insight into how nursing programs handle multiple attempts of the SAT by disclosing the practices of institutions that have chosen to share their policies. A search for institutions with “nursing” in the title yielded 13 institutions which consider the highest sections for each examinee, six that consider all submitted examinee scores, two that request that examinees contact the institution for the policy, and one that considers the highest single sitting for the SAT.

Many nursing programs use more specific admission tests, such as the TEAS V developed by ATI, for use in program entrance decisions. The TEAS V results provide a recommended set of thresholds for institutions to consider when evaluating students: Developmental, Basic, Proficient, Advanced, and Exemplary (ATI, 2010; Wolkowitz, 2010, 2011b). Nursing admissions programs choose to implement the recommended thresholds, with the Proficient often being the lowest level of acceptable performance for admission. Therefore, it is reasonable to conclude that many nursing students retaking the TEAS V may fall into the “satisficing” category in that they are trying to reach a particular level required for entrance.

Other entrance examinations, such as the SAT, ACT, MCAT, and LSAT, do not have official cut scores. For those examinations, the process of devising acceptable thresholds for admission is left up to institutions (Albanese, Farrell, & Dottl, 2005; Briggs, 2009; Kreiter, 2007). A study by Briggs (2009) surveyed institutions to assess their use of entrance examination scores and how small differences in scores increase the likelihood of acceptance. The study found that 20% to 25% of institutions set internal cut scores for the SAT and ACT, and those intuitions were more likely to agree that small increases, such as 10 to 20 points on the SAT, would increase a student’s likelihood of admission. Therefore, institutions are using examination scores similarly, regardless of the presence or absence of official cut scores.

As this literature review shows, recent research addresses the issue of multiple attempts for college entrance examina-
Table 1 shows the number of students by program type (ADN and BSN) having each number of TEAS attempt scores available.

Table 1 shows that although the majority of students (approximately 75% in each program type in this data set) take the TEAS only once, the remaining students (n = 11,076 ADN; n = 7,796 BSN) take the assessment multiple times. The large number of examinees with multiple attempts underscores the utility of this article’s focus. It should be noted that there are students who have taken the TEAS more than six times; however, due to small sample sizes in these groups, their scores are not reported in this article.

Instrumentation

Admissions Test. The admissions test of interest was the TEAS V. The TEAS V assesses objectives in the content areas of Reading, Writing, Mathematics, and Science. The test contains 110 multiple-choice items, which are divided into four sections: Test of Essential Academic Skills (TEAS) V, which assesses the ability to read, write, and calculate, as well as the ability to understand scientific concepts.

The TEAS V is a standardized test that is used to assess the academic readiness of students applying to nursing programs. The test is administered by ATI, which is the largest provider of standardized testing in the healthcare field. ATI recommends that the Adjusted Composite Score be used for admissions decisions, as this score is standardized, objective, and comparable across institutions.

Analysis

To determine how the first, last, maximum, and mean scores of individuals with multiple TEAS attempts compare with scores of individuals with a single attempt, the four score types were calculated for all examinees. From these scores, the arithmetic mean of each type was then calculated for each group, defined by number of TEAS attempts and program type.

Table 2 provides the mean first, last, maximum, and mean scores for ADN students with multiple TEAS attempts, separated by the number of TEAS attempts. The mean TEAS score for students with a single attempt is shown for comparison.

For ADN students, the mean first attempt score for students with only one attempt is 72.3% correct and decreases for groups with additional attempts. The mean first attempt score for students with a single attempt is 72.3% correct and decreases for groups with additional attempts. The mean first attempt score for students taking the TEAS six times is 55.3% correct, which is a mean difference of −17 points. However, a comparison of six-time-testers’ last and maximum scores to single-testers’ first-time scores reveals dramatically reduced mean differences of −5.2 and −3.1 points, respectively.

Table 3 provides the mean first, last, maximum, and mean scores for BSN students with multiple TEAS attempts, separated by the number of TEAS attempts. The mean TEAS score for students with a single attempt is shown for comparison.

The pattern for BSN students is similar to ADN students; the mean first attempt score for students with only one attempt is 75.5% correct and decreases for groups with additional attempts. The mean first attempt score for students taking the
TEAS six times is 59.2% correct, a mean difference of –16.3 points, comparable to that in the ADN sample. Again, a comparison of six-time-testers’ last and maximum scores to single-testers’ first-time scores reveals dramatically reduced mean differences of –3.9 and –2.4 points, respectively. As might be expected, the mean of the all scores option produces score differences that are less extreme than either the first attempt or last and maximum options for both ADN and BSN students.

To answer research question two, the correlation between each score type (first, last, maximum, and mean) and Fundamentals score was calculated for groups of examinees with each number of TEAS attempts. Tables 4-5 present these correlations for ADN and BSN students, respectively.

With one exception for six-time testers applying to ADN programs, the score type most highly correlated with Fundamentals scores is the average of all TEAS attempts (mean).

### DISCUSSION

The analyses presented here lead to several conclusions of practical import for institutional decision makers, especially as they seek to make decisions about examinees presenting scores from multiple attempts on an admissions assessment.

Question one explored the pattern of mean scores by number of TEAS attempts for first, last, maximum, and average scores. Results revealed that for both BSN and ADN students, the first attempt score is highest for those examinees who take the TEAS only once and decreases successively for groups with additional TEAS attempts. In the extreme case, individuals who ultimately take the TEAS six times have a mean first attempt score as much as 17 points (ADN; 16.3, BSN) below their single-attempt peers.

However, by the time they complete all TEAS testing, individuals with six TEAS attempts appear much more like single-testers than they did at first attempt. This finding is consistent with prior research indicating that, on average, students’ admissions test scores increase with each additional attempt (Wolkowitz, 2011a). For both ADN and BSN students, the gap between single-attempt scores and multiple-attempt last and maximum scores is meaningfully less than for first-attempt scores. Because of this, an institution might make different admissions decisions based on its approach to multiple test scores. Understanding which score treatment is most correlated with future performance is a matter of practical significance to institutional decision makers.

Question two asked which approach to multiple test scores was most highly correlated with future performance early in a nursing program, as measured by a nationally available standardized test of nursing fundamentals knowledge. The data in Tables 4-5 show that although the correlations between all TEAS score treatments and Nursing Fundamentals scores are significant for most groups, the correlation is generally greatest for the mean test score approach. This finding is consistent with literature from other fields recommending that programs use the average of all of an examinee’s test scores in making decisions (Dalessandro & McLeod, 1999; Roszkowski & Sprat, 2016; Wightman, 1990; Zhao et al., 2010). Conceptually, use of the average of examination scores seems to strike a happy medium; it provides an opportunity for examinees to remediate and improve while taking into account the full history of their examination attempts. Taken as a whole, the data presented here
suggest that the score treatment most strongly related to future performance in a nursing program is the average of all an examinee’s test attempts.

**CONCLUSION AND LIMITATIONS**

The analyses here are presented with two noted limitations. The first is that the data included may be assumed to represent a restriction of range; only students having both TEAS and Fundamentals scores were included, meaning that they had to have been admitted to a nursing program and completed at least the first semester (to the point of administration of the Fundamentals assessment). Presumably, students with very low TEAS scores would not have met these criteria. As a result, the correlations presented between TEAS and Fundamentals likely constitute an underestimate of the true relationship between scores on the two tests. However, in this case, the emphasis of this article is on the relative magnitude of correlations between each score treatment and Fundamentals, which would all be equally affected by the restriction of range.

The second limitation is that these analyses do not consider the admissions policy of the schools to which students applied, specifically whether the institution has a minimum threshold score necessary for acceptance. Prior research has discussed the concept of satisfiers and maximizers. Without knowing the admissions policies of the institutions to which individuals applied, it is not possible to examine any potential differences in the relationship between entrance examination score and future performance. Future research should incorporate such institution-level admissions policy information.

The question of how to incorporate the resulting information into the admissions decisions is a practical and important one for nursing education programs. Indeed, because examinee scores increase, on average, with each attempt, the approach an institution takes to handling multiple test scores has a real impact on applicants’ futures. This article used data from a nursing admissions examination to examine this question and found that, on balance, the score treatment most highly correlated with future performance early in a nursing program is the average of scores from all examination attempts. This finding held true regardless of the number of test attempts and for both BSN and ADN program types, with the exception of six-time-testers in ADN programs. These conclusions are of practical import for institutional decision makers and should provide evidence-based guidance for using scores from multiple attempts at an admissions assessment.

**REFERENCES**


Roszkowski, M.J., & Spreat, S. (2016). Retaking the SAT may boost scores but this doesn’t hurt validity. *Journal of the National College Testing Association, 2*, 1-16.


Zhao, X., Oppler, S., Dunleavy, D., & Kroopnick, M. (2010). Validity of four approaches of using repeaters’ MCAT scores in medical school admissions to predict USMLE Step 1 total scores [Supplemental material]. *Academic Medicine, 85*, S64-S67.