ABSTRACT

Background: Standardized time allotments are typically imposed for administration of nursing tests. There is little evidence to guide the determination of time allotment. When time allotted for tests is too limited, construct irrelevant variance in test scores may be introduced and the reliability of tests may be negatively impacted. Method: For test administration, we establish a standard time allotment and offer all students the option of extended time. Results: Many of the students use extended time, reporting that extended time reduces stress. Program outcomes have not been negatively affected. Data are provided to guide calculation of time allotment. Conclusion: Extended time may help relieve test anxiety and facilitate success for students with undiagnosed learning disabilities or non-native English speakers. Time allotment should be based on the item composition of tests using published mean item response times. Further research is needed. [J Nurs Educ. 2018;57(3):166-169.]

The use of testing is widespread in nursing education to measure student achievement (Birkhead, Kelman, Zittel, & Jatulis, 2018; Killingsworth, Kimble, & Sudia, 2015; Oermann, Saewert, Charasika, & Yarbrough, 2009). However, little is known about how much time should be allotted for nursing tests. Anecdotal evidence suggests that many nurse educators impose time limits for testing based on a 1 minute per question rule of thumb. At this hospital-based nursing education program, we used the 1 minute convention for calculating time to be allotted for testing for many years, but approximately 10 years ago, we began offering all students the option of extra time for testing.

In a power test (i.e., a test that measures the examinee’s knowledge and its application), it is important to allot a period of time sufficient so that all examinees have time to complete the test (Attali, 2004; Bridgeman, Laitusis, & Cline, 2007; Goldhammer, 2015). On the other hand, a speeded test is designed to measure the examinee’s ability to do something quickly, such as a typing test where proficiency is measured in words per minute (Attali, 2004; Goldhammer, 2015). Insufficient time in a power test may introduce a construct irrelevant variance; the test becomes a test of the examinee’s ability to quickly read and answer questions, an irrelevant difficulty and a possible basis of invalidity (Attali, 2004; Goldhammer, 2015; Halkitis & Jones, 1996; Sireci, Scarpati, & Li, 2005). If the examinee is unable to complete the test and the unanswered questions are marked wrong, or if the examinee rushes to complete the examination by guessing, this detracts from the validity evidence supporting the interpretation of the test results and may negatively influence the reliability of the test results (Attali, 2004; Goldhammer, 2015). Macfarlane (2004), in a discussion of teaching with integrity, spoke to the concept of fairness in the assessment. He posits that fairness is a teaching virtue; vices related to fairness in testing are arbitrariness and inflexibility. The period of time allotted for a test should not be arbitrary, but rather be based on best evidence, and flexibility with respect to time may enhance fairness. The purpose of this article is to briefly explore what is known about allotting time for testing and to describe an innovative practice in this aspect of test administration.

Brief Literature Review

Time for Testing

Brothen (2012) provided a thorough historical review of what few recommendations there are as to how to allot time for testing. Brothen researched publications as far back as 1924 to...
explore this question without finding any statistical substantiation of recommendations for test time allotment. He noted that in 1999, an instructional textbook for teachers of psychology recommended one minute per multiple choice question without citing evidence for this recommendation, and that in 2002, McKeachie, in a teaching handbook for college teachers, suggested a rule of thumb of one minute per question without citing evidence for that rule. In fact, the rule of thumb remains in the most recent edition of McKeachie’s book, still without citing evidence for this advice (Svinicki & McKeachie, 2014).

Several textbooks for nurse educators refer to the allotted time for tests but do not specify how long the testing period should be. However, one resource for nurse educators does state that “each multiple-choice item should take about 1-minute to read and answer” (Quinn Griffin & Novotny, 2012, p. 155) but provides no evidence supporting this statement.

Halkitis and Jones (1996) examined the relationship of three independent variables (item difficulty, item discrimination, and item word count) to the dependent variable item response time in a computerized testing situation. They found a statistically significant increase in item response time as item length increased and item discrimination increased. In other words, as one might expect, longer, more challenging test questions require more time to answer. The authors implied that given this finding, additional time should be allotted for these types of items.

Attali (2004) provided a theoretical discussion of the effects of speediness on the reliability of the results of multiple choice tests in a mathematical simulation model. He concluded that reliability is reduced when examinees run out of time because they guess at answers to finish the test, which causes adverse effects on the psychometric properties of the test.

**Non-Native English Speakers**

Nursing education programs in the United States are charged with increasing diversity in student populations to reflect the diversity of the U.S. population (American Association of Colleges of Nursing, 2015). This includes enrolling students whose first language is not English, or non-native English speakers. Bosher and Bowles (2008) stated that, “For [non-native English speakers], every test becomes a test of language proficiency” (p. 166). Non-native English speakers may read more slowly (Abedi, 2002). The National Council of State Boards of Nursing (NCSBN) has analyzed data from the National Council Licensure Examination (NCLEX). Those data show that candidates whose primary language was not English had a mean item response time of 75.10 seconds, compared with candidates whose first language was English, who had a mean item response time of 60.16 seconds (NCSBN, 2005).

**Students With Learning Disabilities**

The Americans with Disabilities Act (ADA) requires that accommodations be made for students with disabilities. According to one meta-analysis of 59 studies that focused on the effects of test accommodations, the provision of extra time for testing for students with disabilities was one of the most common accommodations (Sireci et al., 2005). Those authors found that studies show that extra time does help students with disabilities, but they also reported that extra time helps students without disabilities as well. They specifically recommend that “accommodations for [students with disabilities] apply in full force to accommodations for English language learners” (Sireci et al., 2005, p. 486).

One conclusion they reach is that additional research is needed to determine the appropriate length of time for any test period, either standard or accommodated.

With respect to NCLEX, research suggests that there is not an unfair advantage for those candidates with disabilities who are permitted extra time (Woo, Hagge, & Dickison, 2013). The authors concluded that “NCLEX-RN items performed largely the same for candidates with and without extended time accommodations” (p. 14), and in contrast to other researchers they found that extended time was of no benefit to candidates without disabilities.

Most of the research on accommodations for students with disabilities assumes that the diagnosis of disability has been made and is silent about undiagnosed disabilities. However, Williams and Ceci (1999), in a thought-provoking discussion of accommodations, suggested that students with undiagnosed disabilities are unfairly penalized because they do not receive accommodations. The authors also suggested that there are many reasons why students may need extra time for testing, including test anxiety, compulsiveness, fatigue, or physiologic factors.

**Test Anxiety**

Some authors have suggested that nursing students are more likely to have test anxiety than students in other disciplines and that test anxiety negatively affects student success in nursing programs (Gibson, 2014; Shapiro, 2014). Those authors recommend that nurse educators and students apply known interventions or develop new interventions to ameliorate this problem. However, much of the research on test anxiety in nursing students focuses on changing student behaviors and responses, not on changing the conditions that provoke test anxiety, such as restrictive time limits. Brodersen (2017) provided an integrated review of the literature disseminating suggested interventions for test anxiety for undergraduate nursing students. She identified only five types of interventions related to changing the conditions for testing (collaborative testing, crib sheets, humorous examination items, music therapy, and aromatherapy), whereas she identified 14 types of interventions aimed at changing student’s behaviors and responses (e.g., hypnotherapy, systematic desensitization, guided reflection, test-taking skills workshop).

McDonald (2014) wrote about the psychological environment in nursing education programs for the administration of tests. She stated that excessive anxiety in a testing situation negatively affects student performance and recommended that teachers avoid certain behaviors that may intensify test anxiety. Among other things, these behaviors include emphasizing time limits and urging students to work quickly.

**NCLEX and Time for Testing**

Anecdotal evidence suggests that many nurse educators believe that the NCLEX is a timed test and therefore it is necessary to set restrictive time limits for teacher-made tests so that students will learn how to test quickly and efficiently in preparation for NCLEX. In fact, according to the NCSBN, the NCLEX is not a
timed test, and examinees may take as long as necessary to answer each question within the 6 hours allotted for the examination (Hong Qian, personal communication, January 23, 2015).

Alternate-format items were added to the NCLEX in 2003 because they are thought to test higher order thinking (McDonald, 2014; Wendt & Kenny, 2009). Certain alternate format items do take longer to answer (Wendt, 2008), but Wendt’s research shows that there has not been a rise in the number of candidates who run out of time on NCLEX since the advent of alternate-format questions, suggesting that the 6-hour time frame is sufficient for most candidates. Qian, Woo, and Kim 2015 presented the following data quantifying mean response time for various item formats on NCLEX: multiple-choice items, 68 seconds; multiple-response items, 75 seconds; ordered-response items, 113 seconds; and fill-in-the-blank dose calculation items, 242 seconds. These data suggest that the traditional 1 minute rule is probably inadequate for standard multiple choice questions and is definitely inadequate for alternate-format items.

Using data published by the NCSBN in the NCLEX Program Reports obtained by subscription to Mountain Measurement, Inc., it is possible to calculate the mean response time per question on the NCLEX for all candidates. For the three 12-month periods between April 2012 and March 2015, the mean number of questions taken by all candidates was 124 and the mean number of minutes expended for the entire examination by all candidates was 145.3. Therefore, the mean response time per question was 1.17 minutes.

In conclusion, the literature suggests that for a power test, restrictive time limits may introduce construct irrelevant variance, items written to test a higher order of cognition take longer to answer, and that there are many reasons why students may take longer to complete tests including limited facility with English, disabilities, and test anxiety.

Innovative Testing Practice
Innovation

For over 10 years, we have been allowing all students the option for extra time on examinations. This practice was initially implemented to allow students some sense of control over the anxiety-provoking nature of high-stakes testing that is so familiar to all nurse educators and nursing students. Upon further reflection, it is clear that this practice also allows non-native English-speaking students and those students with undiagnosed learning disabilities more time to demonstrate learning and the application of knowledge.

Background

Our hospital-based school of nursing offers an associate degree program. It is a commuter school and almost all students are nontraditional. Our grading policy states that 90% of a student’s grade in the didactic portion of all nursing courses must be constructed from points achieved on teacher-made tests (one point per question). The passing grade for any nursing course is 80% (B).

Method

Here is what we do: we calculate the standard test time period and add half of that again for the extended test period.

For example, using the 1 minute rule of thumb and adding extra time for alternate format questions, a 50-question test with 40 traditional multiple choice questions and 10 alternate-format questions, including several dose calculation questions, would be allotted for a standard test time period of 60 minutes and an extended test period of an additional 30 minutes. We accomplish this by establishing an optional early start time of 8:30 a.m. and a standard start time of 9:00 a.m. Students may enter the testing room only at 8:30 a.m. (after which the door is closed and secured) or 9:00 a.m., and may not straggle in at their discretion during that half-hour interval. Students are assigned seats in the testing room and those who enter at the standard time are asked to be ready to be seated quietly to minimize disruption when they enter the testing room. Students are informed of the early start option at the beginning of the semester and it is outlined in each course syllabus.

Outcomes

Faculty report that many of the students take advantage of the early start option. Anecdotally, students report that they like the early start option, citing their belief that the early start time reduces stress. A retrospective statistical analysis of improvement in test scores before and after the implementation of the early start option was not attempted because there are many variables influencing test scores. In particular, the same students take different tests as the semester progresses (so test scores cannot be compared), and different students take the same tests in subsequent semesters (so test scores cannot be compared). However, it can be said that using the early start option has not negatively affected program outcomes. In fact, the program’s graduation rate has slowly been inching upward; for the last four cohorts of students entering the program the mean graduation rate was 82%. In the 10 years since the implementation of the early start option, the mean first-time NCLEX pass rate of graduates of this nursing program has consistently remained above 90%, except for 2009, when the mean pass rate dropped to 87%.

Discussion and Recommendations

Offering all students extended time on tests may help relieve test anxiety and facilitate success on tests for students with undiagnosed learning disabilities or those for whom English is not their first language. Offering all students extended time on tests will not necessarily negatively affect outcomes such as a program’s NCLEX pass rates. Nurse educators should analyze the item type composition of their tests and consider adding time for testing using published response times for item types, especially for tests with alternate format items. However, the question remains—what is the time allotment standard that should be used for nursing tests? Since a before-and-after study design is not the strongest to show causation, a prospective study should be done where students are randomly assigned to different testing times, if the ethical implications can be overcome by careful study design. Further research is needed.

References


