Evaluating the Validity of ATI Exams for Predicting MCCC Registered Nursing Course and Program Outcomes

Concurrent validity refers to the extent to which the results of a particular test or measure correspond to those of a previously established measure of the same construct. Both measures are taken at the same point in time. For example, how well does an automated measurement of blood pressure correspond with a manual cuff measurement? Predictive validity refers to the extent to which the scores on a particular measure (predictor) correspond with the actual construct they are designed to predict (criterion). The criterion is measured at a later time. For example, how well does a nursing entrance exam predict first-semester program performance?

Method

To determine the utility of using ATI exams to predict MCCC registered nursing course and program outcomes, student data were tracked on ATI subject matter exams, ATI exam retakes, the ATI Comprehensive Predictor Exam, course grades, overall GPA, and the National Council Licensure Examination (NCLEX). The ATI subject matter exams are intended to have concurrent validity in predicting course outcomes, as well as predictive validity in determining NCLEX readiness. The ATI Comprehensive Predictor Exam is intended to have predictive validity in predicting NCLEX performance. Passing the NCLEX is necessary for state licensure.

Fall 2012 (*N* = 29) and Winter 2013 (*N* = 30) nursing cohorts were examined for a period of two years. The graduation rate within this two-year period was 62.1% (*N* = 18) for the Fall 2012 cohort, 33.3% (*N* = 10) for the Winter 2013 cohort, and 47.5% for the data combined. The limited cohort sample size and program attrition warrant viewing these findings as a preliminary analysis.

To test the concurrent validity of the ATI subject tests, each one was correlated with a course(s) that covered the same topic (e.g. mental health). Overall GPA was used instead of a course grade when ATI content was identified as being covered throughout the program. ATI subject matter exams were administered at the end of each course. Students scored <1, 1, 2, or 3 (most proficient) on each ATI exam and also received a percentage score. The levels are described by ATI as being produced by nurse educator content experts. While levels correspond with slightly different percentage scores across tests, an example of the level – percentage correspondence for RN Adult Medical Surgical Test can be found in Table 1.

Table 1
RN Adult Medical Surgical 2013 Form Proficiency Correspondence

ATI Proficiency Level	NCLEX Standard	ATI Proficiency Percentage
Level 3	Likely to Exceed	81.1% - 100.0%
Level 2	Fairly Certain to Meet	68.9% - 80.0%
Level 1	Just Likely to Meet	56.7% - 67.8%
Below Level 1	Significant Risk	0.0% - 55.6%

Results

ATI Subject Tests

Table 2 illustrates the results of the correlation between course grade on a 4.0 scale and ATI level and ATI percentage score. The number of students with scores is displayed within the parentheses, which also provides some insight as to when students exited or took a break from the program. Generally in social science a correlation of plus or minus .1 to .3 is considered weak, .3 to .5 is moderate, and >.5 is strong. Findings with a p-value below .05 are considered reliable. The ATI subject matter exams with the strongest relationship with corresponding course grades were MH – N110, OB – N204, PED – N205, F – N105, and MS – N208. For instance we are 33.6% better off predicting performance in N110 upon knowing one's performance on the ATI Mental Health exam. The number of statistically significant relationships indicates that ATI level served as a slightly more useful predictor than ATI percentage score when calculating the relationship with course grades. The ATI Leadership Test was least useful in course performance and overall GPA (see Table 3) prediction.

Table 2
Relationship between ATI Subject Tests (First-Attempt) and Course Performance

ATI Test	Course	Course Grade & ATI Level Correlation (Kendall's tau-b)	Course Grade & ATI Percentage Correlation (Kendall's tau-b)
F = Fundamentals	N103	-.10 (43)	.12 (43)
F = Fundamentals	N105	.33* (45)	.28* (45)
MH = Mental Health	N110	.58*** (49)	.57*** (49)
OB = Obstetrics	N204	.44** (37)	.30* (37)
PED = Pediatrics	N205	.33* (38)	.39** (38)
MS = Medical Surgical	N105	.11 (36)	.18 (36)
MS = Medical Surgical	N208	.33* (33)	.22 (33)
L = Leadership	N210	.02 (28)	-.08 (28)
L = Leadership	N212	.34+ (28)	.23 (28)

Note: + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3
Relationship between ATI Subject Tests and GPA

ATI Test	Course	Overall GPA & ATI Level Correlation (Kendall's tau-b)	Overall GPA & ATI Percentage Correlation (Kendall's tau-b)
N = Nutrition	Throughout program	.32* (28)	.17 (28)
PHM = Pharmaceutical	Throughout program	.34* (28)	.27* (28)
L = Leadership	Throughout program	.23 (28)	.17 (28)

Note: * $p < .05$

When a student received less than level two on an ATI subject test, remediation was strongly suggested by faculty. To determine whether remediation increased ATI subject matter test scores, students' test scores pre and post-remediation were compared. When students chose to remediate, their test score increased an average of 9.7%, $p < .001$. Due to the small sample sizes of the individual tests, only the remediation total was tested for statistical significance. Table 4 illustrates the results.

Table 4
ATI Remediation Need and Effectiveness

ATI Test	*N* without Remediation Need	*N* with Remediation Need	*N* Completed ATI Subject Post-Test	Mean Score Increase (%)	Standard Deviation
F = Fundamentals	32	13	6	13.9	7.1
MH = Mental Health	37	12	7	7.4	4.2
OB = Obstetrics	34	3	2	10.3	0.8
PED = Pediatrics	28	10	7	8.1	5.8
MS = Medical Surgical	25	11	N/A	N/A	N/A
N = Nutrition	18	18	N/A	N/A	N/A
PHM = Pharmaceutical	11	26	N/A	N/A	N/A
L = Leadership	7	21	N/A	N/A	N/A
Total	192 instances	114 instances	22 instances	9.7***	5.9

Note: *** $p < .001$. One remediation Mental Health score of "0%" was deleted from the analysis.

ATI Comprehensive Predictor Exam

ATI Comprehensive Predictor Exam reports by ATI predicted that 100% of the 28 students who took the NCLEX test would pass; the scores ranged from 57.3% to 79.3% likelihood of passing. In actuality 89.3% of the students passed the first-time and 100% were successful after another try. Logistic regression was used to examine the degree to which exam performance predicted NCLEX outcome. The overall model was significant, $\chi^2(1, N = 28) = 4.1$, $p < .05$. In multiple regression the effect size, R^2, refers to the proportion of variance explained by the predictor(s); in other words, how much better we are at predicting NCLEX outcome upon knowing ATI Comprehensive Predictor Exam score. Logistic regression uses effect size measures Cox & Snell R^2 (.14) and Nagelkerke R^2 (.28) which are not equivalent in terms of interpretation.

Conclusion

The ATI Comprehensive Predictor Exam demonstrated a degree of predictive validity by enhancing the ability to predict first-time NCLEX test passage. Five out of the six ATI subject tests demonstrated concurrent validity in predicting outcomes in courses covering similar content. The sixth test, Leadership, was close to statistical significance. Pharmacy, Nutrition, and Leadership ATI subject tests were identified as being covered throughout the nursing program; performance on the Pharmacy and Nutrition tests were significantly related to overall GPA. Students who retook an ATI subject test increased their score by an average of 9.7%, suggesting that remediation was helpful. Given the small sample size of only 28 students who completed the program, it was not possible to conduct an analysis

of the contribution of each of the eight ATI subject tests (final performance level) and their combined effect in predicting ATI Comprehensive Predictor Exam performance. With increased data collection over time, this analysis could be conducted, and the analyses within this paper rerun with the new data, to extend the findings beyond “preliminary”.