White Paper

Retrieval Practice: Enhancing Application Skills
Nurse educators require educational tools that assist them in supporting students to be successful in the program and on the licensing and certification exams. Cumulative testing is well supported in the educational research literature, both as a learning tool for elementary and college students and as a tool to obtain data for remediation and program revision. Research with college students has supported the use of ‘retrieval learning in addition to study as a method of learning (Roediger & Karpicke, 2006, 2008; Kang, McDermott & Roediger, 2007; Karpicke & Roediger, 2007; Swanson, Holtzman, & Butler, 2010; Karpicke & Blunt, 2011; Agarwal, Bain, & Chamberlain, 2011; Karpicke, 2012; Soderstrom, & Bjork, 2014; and Weimer, 2014). Schools preparing nurses have become aware that cumulative testing throughout the program supports success in the program and on the licensing exam (Yoho, Young, Adamson, & Britt, 2007; Murray, Merriman, & Adamson, 2008; Parrone, Sredl, Miller, Phillips, & Donaubauer, 2008; Hinderer, Dibartolo, & Walsh, 2014). Standardized cumulative exams are available from Elsevier HESI for use throughout the curriculum. Elsevier/HESI provides a standardized Admission Assessment (A2), a Registered Nurse Exit (E2), Practical Nurse Exit (E2), Specialty exams, a Custom Mid-Curriculum exam, as well as, Concept-Based Curriculum Exams and other Custom exams. Elsevier/HESI also provides APRN End of Program exams for Family, Adult Gerontology, and Nurse Executive as well as APRN Pathophysiology, Pharmacotherapeutics, and Physical Assessment examinations that are standardized. The exit exams for the PN, RN and APRN programs are based on the blueprints for the NCLEX-PN and NCLEX-RN licensing exams and the APRN certification exams respectively. These exams allow faculty to consistently and authoritatively evaluate their students’ learning, give direction for remediation, and evaluate the strength of their curricula.

Investigators have found Elsevier HESI examinations, including the Admission Assessment, Exit Exam, Mid-Curriculum Exam, Specialty, and APRN End of Program to be successful measures for benchmarking program outcomes, measuring student achievement, guiding remediation prior to licensure candidacy, and facilitating effective hospital orientation.

**Validity and Reliability**

Content validity for the HESI Exit Exams (E2) and the APRN End of Program exams are achieved through use of the NCLEX-RN, NCLEX-PN, and APRN Certification Exam blueprints to determine content, types of questions, and reading level. Content validity for the Specialty Exams, Mid-Curriculum, and other custom exams is supported by basing them on the Elsevier test book content used in the classroom setting and by using masters and doctorally prepared expert nurses to write the items. Reliability is determined for each edition and version by conducting item analyses on each exam and statistically calculating reliability. Elsevier end-of-program testing and Exit Exams range in the highest categories for estimated reliability coefficients using the Kuder Richardson Formula 20 (KR-20) and range from 0.90 to 0.94.

**Predictive Validity**

Numerous studies have been conducted with the A2 entrance exam and the specialty exams to address predictive validity in relation to success in the first year and success in courses (Murray, Merriman, & Adamson, 2008; Knauss & Willson, 2013; Underwood, Williams, Lee, & Brunnert, 2013; Hinderer, Dibartolo, & Walsh, 2014; Manieri, E., De Lima, M., & Ghosal, N., 2015). Research has also been completed with the mid-curriculum exam and the need for remediation (Harding, 2010). Research with the specialty exams was incorporated into the ninth and eleventh predictive validity research with the exit exam (Zweighaft, 2011; 2016). In these studies, E2 scores were significantly higher for those schools using the specialty exams in their courses (t=12.42; p=.0001) (Zweighaft, 2011; 2016).

Regularly, research is conducted to assess the continued predictive validity of the E2 exit exam. For the eleven completed validity studies, the E2 was found to have 94.8% to 99.2% accuracy in predicting NCLEX-RN success for students who achieved the recommended score of 900 or greater on the E2. Table 1 outlines RN studies conducted, dates, number of participants, and study outcomes. The RN predictive accuracy ranges from 94.8% to 99%. Table 2 outlines the PN studies, also highly accurate at predicting NCLEX-PN success (98.4% to 100%).
Table 1: Predictive Accuracy of Registered Nurse (RN) HESI™ Exit Exam

<table>
<thead>
<tr>
<th>RN Study</th>
<th>Dates</th>
<th>Participants</th>
<th>Predictive Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>1996-1997</td>
<td>2,555</td>
<td>97.3</td>
</tr>
<tr>
<td>Two</td>
<td>1997-1998</td>
<td>3,296</td>
<td>96.5</td>
</tr>
<tr>
<td>Three</td>
<td>1998-1999</td>
<td>5,588</td>
<td>97.6</td>
</tr>
<tr>
<td>Four</td>
<td>1999-2000</td>
<td>5,903</td>
<td>98.3</td>
</tr>
<tr>
<td>Five</td>
<td>2001-2002</td>
<td>9,695</td>
<td>97.8</td>
</tr>
<tr>
<td>Six</td>
<td>2004</td>
<td>10,147</td>
<td>96.4</td>
</tr>
<tr>
<td>Seven</td>
<td>2006-2007</td>
<td>4,383</td>
<td>99.2</td>
</tr>
<tr>
<td>Eight</td>
<td>2007-2008</td>
<td>4,134</td>
<td>98.3</td>
</tr>
<tr>
<td>Nine</td>
<td>2008-2009</td>
<td>3,790</td>
<td>96.6</td>
</tr>
<tr>
<td>Ten</td>
<td>2009-2010</td>
<td>5,038</td>
<td>98.3</td>
</tr>
<tr>
<td>Eleven</td>
<td>2010-2011</td>
<td>9,215</td>
<td>94.8</td>
</tr>
</tbody>
</table>

Table 2: Predictive Accuracy of the Practical Nurse (PN) HESI Exit Exam

<table>
<thead>
<tr>
<th>Study</th>
<th>Dates</th>
<th>Participants</th>
<th>Predictive Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>1996-1997</td>
<td>170</td>
<td>100.0</td>
</tr>
<tr>
<td>Two</td>
<td>1997-1998</td>
<td>456</td>
<td>97.9</td>
</tr>
<tr>
<td>Three</td>
<td>1998-1999</td>
<td>689</td>
<td>99.1</td>
</tr>
<tr>
<td>Four</td>
<td>1999-2000</td>
<td>897</td>
<td>99.4</td>
</tr>
<tr>
<td>Five</td>
<td>2006-2008</td>
<td>1,149</td>
<td>99.5</td>
</tr>
<tr>
<td>Six</td>
<td>2009-2011</td>
<td>962</td>
<td>98.4</td>
</tr>
</tbody>
</table>

Three investigations have been conducted to assess predictive accuracy of the APRN End of Program HESI exam (Binder & Nibert, 2007; Binder, Jones, & Fuentes, 2008; Willson & Goodman, 2013). The investigators found that Advanced Practice Registered Nurse (APRN) students scoring 800 and higher had a 100% certification pass rate on the certification exam regardless of certifying body - American Academy of Nurse Practitioners (AANP) or American Nurses Credentialing Center (ANCC). Further research with a larger sample size is planned for 2016. Table 3 outlines APRN studies conducted, dates, number of participants, and study outcomes.
### Table 3

<table>
<thead>
<tr>
<th>APRN Study</th>
<th>Dates</th>
<th>Participants</th>
<th>Predictive Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Nurse Practitioner</td>
<td>2005-2007</td>
<td>49</td>
<td>100.0</td>
</tr>
<tr>
<td>Family and Adult NP</td>
<td>2008-2009</td>
<td>170</td>
<td>100.0</td>
</tr>
<tr>
<td>Family and Adult NP</td>
<td>2011</td>
<td>120</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Implementation Science**

The Implementation of HESI™ Exams includes student remediation strategies, policy setting, and curriculum evaluation that is based on the evidence.

**HESI Admission Exam (A2)**

Murray, Merriman, & Adamson (2008) used a longitudinal descriptive design to evaluate the A2 in predicting student success in an ADN and a BSN program. A2 scores and grades throughout the curriculum for 68 ADN and 69 BSN students were used. The A2 was administered after admission to the program for placement and remediation. A bivariate regression analysis was used to assess the accuracy of the A2 composite in predicting course grades. In the ADN program, composite scores were significantly positively correlated with eight of the nine nursing course grades from four levels of the program ($r=0.253-0.442$; $p=.05-.01$). In the BSN program, the A2 composite score was significantly, positively correlated with 10 of the 20 nursing course grades ($r=0.241-0.374$; $p=.05-.01$). Four of the ten courses were in the sophomore year, five were in the junior year, and one was in the senior year. All 69 students in the BSN program completed the program within two years of admission. Although course grades were available for only 68 of the ADN students, 217 students took the A2 exam. Of these students, 80 did not complete the program within two years of admission. A t-test was used to assess the difference in A2 scores for those who completed the program and those who did not. Scores on the A2 were significantly higher ($p<.001$) for those who completed the program within two years.

Knauss and Willson (2013) completed a retrospective study with 157 associate degree nursing students. They used a composite score of 75% including basic math, reading comprehension, vocabulary/general knowledge, and grammar. The investigators found a moderate significant positive correlation between the A2 composite score and grades in Nursing I ($r=.532$; $p<.01$) and grades in Nursing II ($r=.455$; $p<.01$). There was also a positive correlation between each of the exam components and course grades with vocabulary/general knowledge scores having the highest correlation (Nursing I: $r=.371$; $p<.01$ and Nursing II: $r=.359$; $p<.01$).

Underwood, Williams, Lee, and Brunnert (2013) evaluated correlations between A2 subscores in reading comprehension, vocabulary and general knowledge, math, and anatomy and physiology. They combined the reading comprehension and the vocabulary and general knowledge scores to form one English composite score. The investigators used A2 scores from 184 baccalaureate students and correlated them with the final grades in the first three nursing courses. The highest correlations were between the English composite and the three courses ($r=.503, .581, and .414; p<.01$). Correlations between A&P and the three course grades were the next highest ($r=.350, .402, and .404; p<.01$). The lowest, but still statistically significant correlations were between Math scores and the three final course grades ($r=.297, .239, and .253; p<.01$).
**Mid-Curriculum Exam**

Harding (2010) completed a longitudinal, descriptive study with 52 second year associate degree students who completed the program to assess the usefulness of the Mid-Curriculum exam in identifying at-risk students. The original sample of 68 students was reduced because 16 students did not complete the program. The MC-HESI was customized to their curriculum, included 105 items, and was administered during the final week of the second semester of a four-semester nursing sequence. The HESI E2 was administered six weeks prior to completion of the program. Program variables included grades from the two Capstone courses. There were significant correlations between the MC-HESI scores and students’ GPAs on admission and at the time of the test (r=0.381; p<.01 and r= 0.445; p<.01). The MC-HESI scores were significantly correlated with the two Capstone course scores (r=0.522 and 0.513; p<.01). The 16 students who were not successful in the program all had scores less than 800 on the MC-HESI exam. A t-test was used to assess the difference in scores for the successful versus the unsuccessful group. There was a significant difference in scores with those who were unsuccessful scoring significantly lower (=823.3 and 743.4; t=2.532; p=.017).

Yoho, Young, Adamson, and Britt (2007) used a descriptive correlational design to determine the predictive accuracy of the HESI A2, Mid-Curriculum (MC) and E2 with the NCLEX-RN in a sample of 77 ADN students. Pearson correlations were used to assess relationships between the A2 Math and Reading Comprehension scores and the Mid-Curriculum exam scores. A2 Math scores were not significantly correlated with the MC scores (r=.129; p>.05). A2 Reading Comprehension scores significantly positively correlated with MC scores (r=.412; p=.01). MC scores were positively correlated with E2 scores (r=.617; p=.01). The first version of the E2 was 94.83% accurate in predicting NCLEX-RN success.

**Specialty Exams Improve NCLEX-RN Pass Rates**

Zwieghaft (2013) in the ninth validity study examined the use of specialty exams in relation to E2 scores and NCLEX success in a sample of 63 participating schools (baccalaureate, ADN, and diploma), representing a total of 3,790 students. Students had taken one or more HESI Specialty Exams during their course work. There was a significant difference in mean Exit Exam scores for schools (n=43) that used Specialty Exams during their programs (= 865.7) and for non-users (n=20) (= 837.3). Using the benchmark score of 850 and above, all eight Specialty Exams were significantly predictive (p ≤ .0001 to .0034) of NCLEX-RN success. The top three most predictive specialty exams were Critical Care (r=11.16), Pediatrics (r= 5.51), and Medical-Surgical (r= 5.50). Schools of nursing used Specialty Exams as course tests, course final exams, remediation, and for curriculum evaluation.

Zwieghaft (2014), in the eleventh validity study, again examined the impact of specialty exams with a sample of 138 schools including 8,464 students. Ninety nine (71.74%; 5900 students) schools used the specialty exams and 39 (28.26%; 2564 students) did not. There was a significant difference in E2 scores between schools that used the specialty exams and those that did not (t=12.42; p=.0001). This difference held across types of schools: baccalaureate (t=3.53; p=.0004), ADN (t=6.14; p=.0001) except diploma (t=1.68; p=.0943). It is possible that the result for diploma schools was affected by the sample size, 175 users of the exams and only 23 nonusers.

**Repeat Testing After Remediation**

Adamson and Britt (2009) in the sixth validity addressed predictive accuracy of the exam for repeat testings of one to three times. On V-1, 4,715 students scored 900 and above, and 4,547 (96.44%) passed the NCLEX-RN® examination on their first attempt. On V-2, 822 students scored 900 and above, and 764 (92.94%) passed the NCLEX-RN® examination on their first attempt. On V-3, 200 students scored 900 and above, and 165 (82.50%) passed the NCLEX-RN® examination on their first attempt. A one-way ANOVA was used to determine if differences existed in the predictive accuracy of the E2 among the three versions administered: V-1, V-2, and V-3. Findings indicated that there was no significant difference in the predictive accuracy between V-1 and V-2, but V-3 was significantly (P < .001) less accurate in predicting NCLEX-RN® examination success than both V-1 and V-2.
Langford and Young (2010), in the ninth validity study, asked deans and directors at 154 RN schools (66 responded; 43%) and 92 PN schools (26 responded; 28%) if they mandated retakes of the E2 exam prior to graduation. Twenty-four (36.4%) of the schools did not require repeat testing; 9 schools each permitted 1-2 re-takes (versions 1, 2, and 3); 24 schools allowed 4-8 re-takes. Analysis for predictive validity indicated that the exams retained predictive accuracy over three attempts on the exam with accuracy for the RN group at 98.3%, 94.9%, and 95.9% and for the PN group 98.4%, 94.4%, and 100%.

Young and Willson (2012), in the seventh validity study, assessed the predictive validity of the E2 for students required to take the exam one to three times to achieve success. Of the 4383 total student sample, 1075 students scored 900 or above on V1 of the E2 and 1066 (99.16%) passed the NCLEX-RN on the first attempt. Of the 730 (16.66%) students who were required to take V-2 of the E2, 271 scored 900 and above on the second version of the E2 and 259 (95.57%) passed the NCLEX-RN on their first attempt. Three hundred sixty seven students were required to take V3 of the E2. Of the 148 students who scored 900 and above on V-3, 138 (93.24%) students passed the NCLEX-RN on their first attempt. Significantly more students who scored 900 and above on V-1 of the E2 successfully completed the NCLEX-RN on their first attempt than those who scored 900 and above on V-2 and V-3 of the E2 ($\chi^2=31.4156$, P < .000). Accuracy of the E2 in predicting NCLEX-RN success was 97.93%, regardless of whether the student was required to take the examination up to three times before achieving the faculty-designated E2 benchmark score.

Barton, Willson, Langford, and Schreiner (2014) assessed the predictive validity of the E2 at different required scoring levels 900 and above, 850-899, 800-849, 700-799, and 699 or less. Although HESI recommends a minimum score of 900, many schools have lower benchmarks. Predictive accuracy decreased with each descending scoring level: 900 or greater=98.26%, 850-899=95.13%, 800-849=92.43%, 700-799=86%, and 699 or less= 71.30%.

**Policies Related to the HESI Exit Exam**

Barton, Willson, Langford, and Schreiner (2014) surveyed schools about school policies related to standardized testing. Fifty-four (84.4%) schools responded representing 5,438 students with 3,084 from ADN programs and 2,0354 from baccalaureate programs. Sixty-eight percent (37) of the schools set a benchmark of 850 and 18.5% (10) at 900. Thirteen percent (7) set other minimum scores between 700 and 950. Forty-four percent (24) required mandatory achievement of the benchmark with retesting one or more times (22; 41%) or required remediation (35; 64%). Consequences for not meeting the benchmark were set by 30 (56%) schools and included course failure (27; 50%), delayed NCLEX candidacy (13; 24%), and delayed graduation (9; 17%). The authors also assessed differences in HESI E2 scores based on the presence or absence of a policy requiring passage of the E2 benchmark set by the school. Schools with the policy had a mean score of 907.2 and those without the policy had a mean score of 855 (t=13.365; p< .0001).

**Test Preparation**

Barton, Willson, Langford, and Schreiner (2014) received surveys from 99 schools with 64 completing the section related to policies. Sixty-one percent (39) indicated that they provided students with a preparation plan and 85% (33) of those schools mandated that the student participate in the plan. Policies included self-guided review (e.g., case studies, test items, study guide) (33; 85%), faculty guided group review (22; 56%), faculty guided individual review (18; 46%), formal review (e.g., HESI, ATI, Kaplan) (17; 44%), peer/mentor tutoring (13; 33%), and self-guided formal review online course (1; 13%). Schools with this requirement had a mean score of 905.47 on the E2 and those without a policy had a mean score of 853.13 (t=13.677; p<.0001).

**Remediation Requirements**

Langford and Young (2010) surveyed deans and directors at 154 RN schools (66 responded; 43%) and 92 PN schools (26 responded; 28%) about remediation policies. Forty-seven (71%) of the RN schools required remediation with most requiring two to six weeks. All thirteen (50%) of the responding PN schools required two to six weeks of remediation. Eighteen of the 66 RN schools required Evolve online, seven the HESI Live Review, 11 online case studies, 12 a NCLEX preparation book, 20 live tutoring, 17 computer-based tutoring, and 17 other courses. Six of the PN schools required
Evolve online; none required Hesi Live Review, two online case studies, three a NCLEX preparation book, four live tutoring, five computer-based tutoring, and six required other courses.

Barton, Willson, Langford, and Schreiner (2014) surveyed 99 schools about policies and 64 (65%) completed the survey. The authors used a t-test to determine if there was a significant difference in HESI Exit Exam scores based on whether or not remediation was required after students failed to meet the benchmark. Those with the requirement had a mean of 885.18 and those without had a mean of 849.91 (t = 9.742 and p<.0001). Those schools with required remediation had significantly higher mean scores on the E2 than those without required remediation.

Use of Case Studies

Dufrene, Hodges, and Vandenberg (2016) evaluated the use of case studies and the impact on specialty exam scores. The investigators used the case studies included with the HESI specialty exam packages for Fundamentals and Health Assessment. One hundred thirty-one students were divided into four groups: Group A, junior students had no case studies, but took the specialty exams, Group B, junior students had no case studies and no specialty exam, but three Fundamentals case studies and the Fundamentals specialty exam as seniors, Group C, juniors completed five case studies in Fundamentals and seven in Health Assessment plus five fundamental case studies as seniors, Group D, juniors completed ten case studies in Fundamentals and five in Health Assessment. Students in the three groups with case studies scored higher on the Fundamentals Specialty Exam than the group that did not complete case studies (B: 523-1031, C: 513-1178, D: 374-1122 compared to A: 386-849). Groups C and D scored higher on the Health Assessment exam than did group A (C: 661-1222, D: 491-1053 compared to A: 426-970). Group B did not take the Fundamentals exam. The groups were primarily Hispanic (32%, 49%, 41%, & 47%), followed by Caucasian (29%, 24%, 35%, & 11%), African American (21%, 18%, 9%, & 17%), and Asian (14%, 9%. 9%, & 25%).

Advanced Practice RN End of Program Exams

Binder, Jones, and Fuentes (2008) correlated FNP exit exam scores for three groups of students from three campuses taking the exit exam with success on the FNP certification exam. All students (49) who scored at the recommended level of 800 or above on the APRN End of Program Exam passed either the AANP or ANCC certification exam.

Binder and Nibert (2009) evaluated the predictive validity of the HESI scores for students achieving scores >800 on the APRN End of Program exams with their ANCC or AANP certification exam outcomes. The data analysis indicated acceptable and recommended scoring levels. The standardized FNP and ANP exams were effective in assessing students’ preparedness for the specialty accreditation exams and also provided evidenced-based measures of curricular outcomes. One hundred percent of the nurses who scored greater than 800 on the exam were successful on the certification exam.

Willson and Goodman (2015) surveyed 35 graduate nursing schools that administered the HESI APRN exam for Family and/or Adult-Gerontology about their APRN national certification exam outcomes, implementation strategies, and testing policies. Outcomes were reported for 141 APRN test takers, 96 Family APRN test takers and 45 Adult-Gerontology test takers. They found that scores of 700 or greater (113 students; 80.1%) accurately predicted certification pass rates 100% of the time. Fifty-three percent of the students scored 800 or greater. Only one school had set a mandatory testing policy and few had established bench mark scores, or included the test as part of a graded experience.

Focused Nurse Orientations with Identified Remediation and Mentoring Topics

A HESI custom exam was evaluated as a measure of how much orientation is necessary to bring RN graduates to the level of clinical competence required for clinical practice. Ryan, & Tatum (2012) explored the relationship between general knowledge of pediatric nursing, critical thinking (CT) ability of newly hired nurses, as determined by a custom Health Education System Inc. (HESI E2) exam, and the length of time new graduates had to stay in orientation before meeting the required clinical competencies. A descriptive correlational study design was used. The customized exam was administered to 98 nurses who were beginning employment at a large pediatric institution. The RNs’ critical thinking ability significantly correlated (r=.325; p=.004) with length of time in orientation, indicating that those with a higher
critical thinking score required fewer days to meet orientation objectives. Orientation time was decreased by 50%. The custom examination provided an objective assessment of critical thinking ability in the nurse’s specific domain of knowledge of pediatric care.

References


Resources:

Ask your Elsevier Education Solutions Consultant for more information including:

- HESI Assessment RN brochure and HESI Assessment PN brochure
- HESI Assessment Testing & Remediation Research Bibliography
- State of the Science: HESI™ Remediation and NCLEX® Preparation Bibliography
- Eighth Validity RN Exit Exam Study Summary Report
- Ninth Validity RN Exit Exam Study Summary Report
- Tenth Validity RN Exit Exam Study Summary Report
- Tenth Validity PN Exit Exam Study Summary Report

Future Research:

- There are multiple research studies currently underway with Nurse Educators and Health Professionals investigating testing with HESI products and remediation, use of case studies, use of practice test questions and student outcomes. For more details, please contact Dr. Terry Throckmorton, Principal Researcher, HESI Review and Testing at 713-346-6927 or t.throckmorton@elsevier.com.