

**Advances in Nurse Hiring Processes:
Pre-employment Assessments that Predict Job Performance**

(SUBMITTED FOR REVIEW)

Primary Author:

Jim Ostmann, RN, MBA

Additional Authors:

Dan Biddle, PhD

Heather Patchell, MA

Ron Gonzales, RN, MBA

One of the largest and most comprehensive pre-employment-based research projects in the history of the nursing profession was conducted between 2008 and 2010. This project combined the work of three pre-employment assessment firms who joined together with two large medical facilities to create a pre-employment assessment process with an exceptionally high operational validity (job-relatedness) of $r = .67$. Because this testing process has such high level of validity, this assessment can report nurse applicant scores in terms of their likely job success (e.g., 75% likely to be rated as “average/above average”), which in turn has significant ramifications for those healthcare facilities who are looking for a way to screen and place qualified nursing applicants.

This project brought together over 900 professionals who participated in the development and validation (both inferential and statistical job-relatedness) of a pre-employment assessment process, which included: 13 test development professionals and industrial-organizational psychology consultants, over 30 film professionals, 492 nurse staff from a Mid-West Healthcare Facility (470 nurses and 22 nurse supervisors), and 384 nurse staff from North-East Healthcare Facility (367 nurses and 17 nurse supervisors).

The Assessments

There were three types of assessments included in this study: clinical, situational and behavioral. The Clinical Assessments were written, job knowledge tests (delivered online) that were designed to measure clinical job knowledge in the specified clinical practice area (e.g., Labor & Delivery, Medical-Surgical, etc.). Each test typically included between 30 and 70 multiple-choice items designed to measure the relative practice area in a balanced way (i.e., test content followed professional practice standards set by many nursing certification organizations, such as, the American Association of

Critical-Care Nurses (AACN), the Emergency Nurses Association (ENA), and the American Physical Therapy Association (APTA)). A total of 12 clinical practice area tests were included in the study. These tests were developed and validated by a healthcare pre-employment testing and assessment company using a content validation strategy and validation tools (software and surveys) developed by one of the nation's leading pre-employment assessment, EEO, and validation consulting firms.

The Situational Assessments were video-based, situational judgment tests designed to measure interpersonal competence in hospital-related situations. The final version of this assessment included 21 video scenarios (30-120 seconds in length) that presented complex interpersonal situations that occur in hospital environments (interacting with patients, physicians, and other co-workers) to the nurse applicant. Each video was followed by a set of written response options from which the applicant was asked to select the “most effective” and “least effective” way of handling the situation.

The Behavioral Assessment included 70 personality/behaviorally-based test items (scored using a Likert-type scale) targeted toward measuring Conscientiousness, Tough-Mindedness, Conventional, Extroversion, Stability, Teamwork, and Good Impression.

Study Results

The study was completed in five phases¹ and took over two years to complete. Each phase is briefly summarized below.

The first phase involved the development of a Job Performance Rating Survey (JPRS) that included 19 observable aspects of nursing job performance (e.g., Problem Solving, Report Transitioning, Following Clinician Instructions, Patient Care Plan

¹ While there were in fact five major elements to the study, the research was completed as an iterative process, with many steps being completed concurrently.

Management, etc.). These domains were used for the remaining phases of the study to rate job performance metrics on the nurses participating and criterion validate the use of the three assessments in the nursing profession.

The second phase included the development and validation of the Situational Assessment using job analysis and validation information from nursing staff at the Mid-West Healthcare Facility. After the original development and calibration of this test, it was subsequently cross-validated and refined through a study conducted at the North-East Healthcare Facility. The resulting final test battery (consisting of 21 videos) demonstrated substantial levels of both content- and criterion-related (inferentially job-related and statistically job-related, respectively) validity, with significant correlations to job performance ranging between $r = .20$ and $.33$ (with the highest correlations observed in the areas of Developing Patient Relationships and Patient Customer Service).²

The third phase involved evaluating the criterion-related (statistical) validity of the Clinical Assessments. This set of assessments included 12 unique, written, multiple-choice tests designed to measure mastery of the job knowledge domains associated with each relevant practice area. While these tests are based on a content-validation methodology (where they are rationally developed and connected to critical aspects of the job using subject-matter expert opinions), they also demonstrated varying degrees of significant correlations to job performance (as high as $r = .26$), with the highest relationships in areas that would be expected by using a written test (e.g., Critical Thinking, Problem Solving, and Report Transitioning).

² Based on the Mid-West Healthcare Facility study, with similar results observed in the study conducted at North-East Healthcare Facility.

The fourth phase evaluated the validity of the Behavioral Assessment. By correlating the various personality scales to the job performance metrics, several scales showed significant correlations. A final “Nurse Effectiveness Scale” was developed as a derivative of the more than 70 Behavioral Assessment items by evaluating the correlations with job performance at the test item-level using a split-half/hold-out validation study design. The resulting correlations exhibited strong relationships with job performance in the $r = .20$ s to $.30$ s in areas such as Honesty and Conscientiousness (work practices), Accountability, and Delegating/Managing Patient Care.

The fifth and final phase involved combining the datasets from all four study phases and conducting Structural Equation Modeling (SEM) to empirically derive a set of weights that could be applied in the scoring process when all three assessments are administered in hiring situations. This research revealed optimal weights of 26.4%, 32.9%, and 40.7% for the Clinical, Situational, and Behavioral Assessments, respectively. This research also revealed that the combined set of the three tests produces a correlation to job performance of $r = .38$, which explains about 15% of job performance variance. Because personnel tests typically provide only limited correlations to job performance (often ranging in the $r = .15$ to $.25$ range for single tests), this level of validity ($r = .38$) explaining 15% of job performance (derived by squaring this value) can be viewed as *outstanding*. In actual practice, only rarely do researchers encounter tests with correlations in this range. In addition, the study revealed that the three assessments measure *distinct aspects* of job performance, with a limited degree of overlap. This was encouraging to see because, in an ideal setting, tests that are combined into an overall measure should be designed to target unique (rather than overlapping) aspects of job

performance. Table 1 reveals the job performance dimensions that had the highest correlation with each assessment.

Table 1. Top Six Ranked Performance Dimensions Predicted by Each Assessment.

Written Test	Situational Assessment	Personality Scale
Administering Medications	Patient Care (Calm & Competent)	Accountability
Problem Solving	Developing Patient Relationships	Honest & Conscientious
Report Transitioning	Verbal Communication	Conflict Resolution
Critical Thinking	Patient Customer Service	New Technique Application
Multitasking	Patient Care Plan Management	Delegating/Managing Patient Care
Assertiveness	Conflict Resolution	Change Adaptation
<i>Note: The values in this table are based on the total amount of criteria variance predicted (explained) by the set of three assessments (summed), not the total variance available.</i>		

One of the compelling findings of this study (as revealed by Table 1) was the fact that each of the assessments seemed to target unique aspects of job performance. In other words, the 19 job performance dimensions evaluated in the study seemed to correlate with each of the respective assessments in a way that seemed intuitive (e.g., with the clinical test correlated with hard skills, etc.). Only one of the top six job performance dimensions predicted by each assessment overlapped with the top six predicted by the other assessments (see Table 1).

When all three Assessments are combined into an overall score (weighted in the optimal proportions derived from this study), the observed validity coefficient is very strong ($r = .38$). After making corrections for the unreliability of the raters (nurse supervisors) and indirect range restriction (because the nurses in the study were selected using other selection devices), the resulting operational validity is quite strong ($r = .67$) (Hunter, Schmidt, & Le, 2006).

Figure 1. Likelihood of Job Success Based on Combined Score Using Different Base Rate Assumptions.

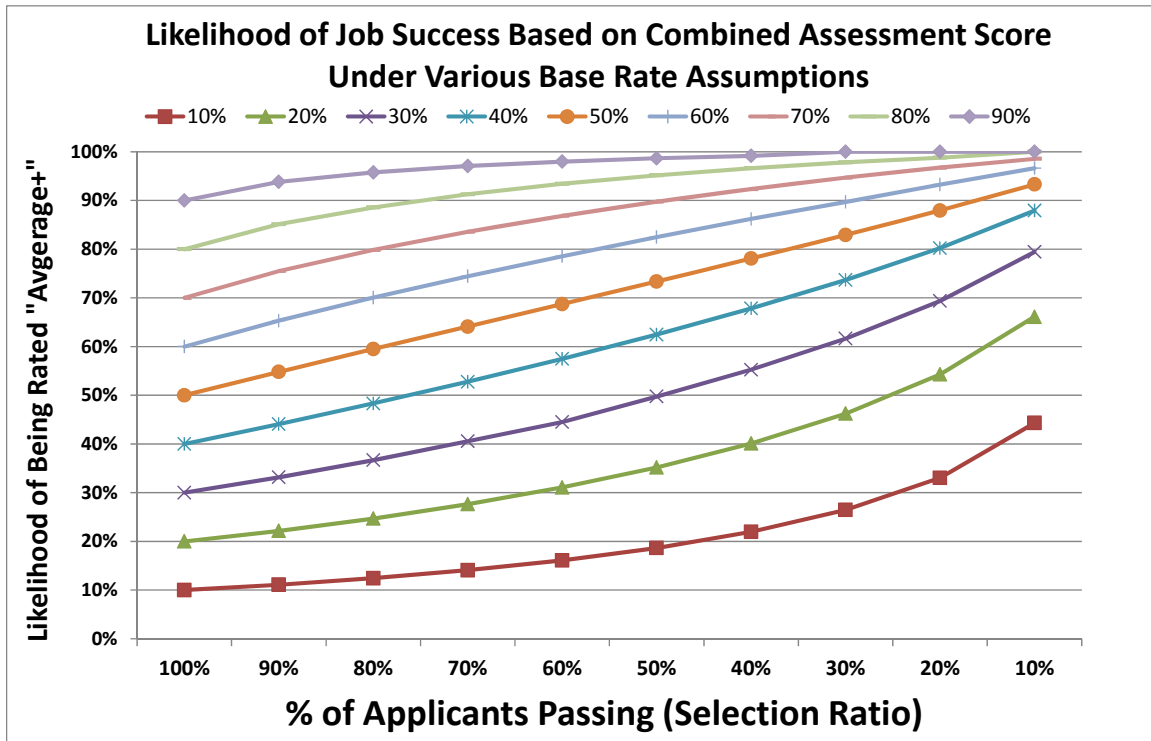


Figure 1 illustrates the *practical benefits* that will be realized by employers using the Assessments in nurse hiring processes using the tables provided by Taylor and Russell (1939), which are tables frequently used by personnel psychologists for evaluating the practical benefits from using validated assessments. Figure 1 uses the “Base Rate” (i.e., the percentage of incoming nurses who are qualified at the desired level) and the facility’s “Selection Ratio” (the percentage of applicants who will pass the assessment battery at a specified t-score) to show the percentage of tested applicants who will perform at or above the Base Rate based on the utility and effectiveness of the test. This Figure shows that using a test that has a significantly high correlation to job performance nearly ensures an improved workforce, especially when higher selection

ratios are used. For example, assuming a 60% Base Rate and using a 40% Selection Ratio, about 86% of the incoming nursing staff will be at or above the desired qualification levels—a 26% improvement over using no test (or an invalid test).

In addition to improving the job performance levels of the incoming nursing staff, substantial financial gains can also be realized by using validated assessments. For example, using an assessment battery with strong validity (e.g., $r = .67$) and assuming a 70% Base Rate and 40% Selection Ratio, a medical facility that screens 500 applicants (and hires the top 200) will realize \$8,842,365 in value over the expected tenure of the newly hired staff (assuming 5.3 average tenure, a financial value of \$10,000 associated with one standard deviation of job performance, \$70,000 annual salary, and testing cost of \$150 per applicant) (Cascio & Boudreau, 2010). This equates to about \$8,342 added value for each new hire.

In summary, the research presented indicates that a multi-faceted selection system that includes measures of job knowledge, situational judgment, and behavioral/personality traits can combine to forge a powerful selection battery. Using such tools to help predict an employee's success before investing in their acquisition, aids in the long-term reduction of hiring expenditures and improves the overall workforce by raising the quality of staff.

###

Biography

Jim Ostmann RN, MBA Chief Nursing Officer, brings over 25 years of healthcare experience to the Prophecy organization. Ostmann worked as an RN in environments such as, ER/Shock Trauma, Cardiac Catheterization Lab and ICU. He entered the corporate sector in 1997 with Marquette Medical Systems (GE Healthcare). Jim achieved his Six Sigma Green Belt and spent nine years in various roles such as e-Learning Product Manager, Subject Matter Expert, and Instructional Design.

Over the last five years, Jim has facilitated the development of dozens of pre-employment job knowledge competency assessments as well as conducting evidence-based research of behavioral and interpersonal competencies using content and criterion validity methodologies.

References

Biddle, D. A. Adverse impact and test validation: A practitioner's guide to valid and defensible employment testing. West Conshohocken, PA: Infinity Publishing, 2011.

Cascio, W., Boudreau, J. Investing in People: Financial Impact of Human Resource Initiatives (2nd ed.). Upper Saddle River, NJ: FT Press: 2010.

Hunter J, Schmidt F, & Le, H. Implications of direct and indirect range restriction for meta-analysis methods and findings. Journal of Applied Psychology, 91: 594-612, 2006.

Taylor, H., Russell, J. The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. Journal of Applied Psychology, 23, 565-578, 1939.

U.S. Department of Labor: Employment and Training Administration. Testing and assessment: an employer's guide to good practices. Washington, D.C.: Department of Labor Employment and Training Administration, 2000.