Proposed Title:

The Development and Validation of Three Pre-Employment Assessments for Nursing

(SUBMITTED FOR REVIEW)

Abstract Proposal

The process of hiring effective nurses is a difficult task that can have monumental repercussions for medical facilities. If nurses without proper skills are selected, patients can suffer from insufficient quality of care and potentially life threatening conditions. Although the knowledge an applicant receives while being educated to become a nurse is extremely important, there are untrained skills that factor into their on-the-job success. These skills can be referred to as "soft skills" and include bedside manner, personality, effectiveness of communication, decision making, and so on. In order for medical facilities to select and maintain a high-performing nurse staff, nurse hiring managers must incorporate evaluations for soft skills in their hiring process.

With the purpose of creating an assessment by which nurse applicants can be functionally evaluated for technical knowledge and soft skills, a comprehensive test validation study was performed. The study (conducted between 2008 and 2011) involved three test publishers, two large healthcare facilities, 837 registered nurses and 39 nurse supervisors.

Two types of test validation methodologies were applied in the study (content and criterion-related validity) to develop three assessments: technical knowledge (Clinical), situational judgment (Situational), and behavioral/personality (Behavioral). The Clinical Assessment is written in question and multiple-choice answer form, and was designed to measure an applicant's job knowledge relevant to their specific practice area (e.g., labor and delivery, surgical, etc.). The Situational Assessment, a situational judgment test (SJT), is video-based and was designed to measure interpersonal competency in medical facility-related situations. The behavioral test was designed to measure different aspects of personality and behavior, including: conscientiousness, tough-mindedness, conventionality, extroversion, stability, teamwork, and ability to make a good impression, all of which this study has shown to be pivotal to the success of a nurse.

This study was conducted in several different phases over two years. Throughout the phases, a Job Performance Rating Survey (JPRS) was created and analyzed; the clinical, situational and behavioral assessments were developed and validated; analyses of the assessments were done to determine the manner in which each should be applied to a composite score; and then the exams were correlated, individually and as a compilation, with job performance as measured by the JPRS. Results from the study indicate that using these three assessments together as one

assessment is *very beneficial* (U.S. Department of Labor 2000) with regard to predicting a nurse applicant's job performance and success (as based on the assessment's positive correlation with job performance). Subsequent research on the exam analyses has shown that implementing the assessment in hiring decisions (with specified cutoff rates) can increase the quality of a medical facility's nurse staff, decrease hiring costs over a few short years, and potentially decrease the risk of adverse impact (civil rights) litigation that can sometimes occur in healthcare facilities.

Hiring Practices in the Nursing Field

The hiring practices generally utilized in the nursing field include a variation of written exams and interviews. These practices are useful because they allow hiring managers to objectively evaluate an applicant's clinical knowledge, as well as take a brief look at the their verbal communication skills. Written exams are what give managers an objective view on an applicant's technical knowledge and are therefore often considered pivotal inclusions in the hiring process. In this test study, 12 different comprehensive clinical assessments (each under a different specification, e.g., labor and delivery) were developed and validated. With regard to scores on the clinical assessments, results indicate that they predict an applicant's job performance across the 19 job performance dimensions evaluated in the study. Because the assessments are designed to measure a baseline of technical knowledge as well as knowledge specific to practice areas, employers who implement the clinical assessments have the ability to explore how proficient applicants are in their specified practice areas and predict how well they will incorporate that knowledge into their on-the-job performance.

Interviews, on the other hand, are important for giving hiring managers an impression of the applicant, including their presence, demeanor and verbal skills. As many hiring managers know, interviews can be difficult to standardize, which means their results often lack the ability to objectively distinguish between applicants. In order to curtail this issue while still maintaining the capability to gain insight into an applicant's teamworking skills, work practice styles, and interpersonal skills, behavioral tests can and should be used.

The behavioral portion of the assessment evaluated in this study showed that a subject's scores correlated positively with, and were predictive of, the subject's job performance across 19 job performance dimensions evaluated in the presented study (especially in the "soft skill"

dimensions evaluated in the study). So while traditional interviewing methods can still be used to aid in the hiring process, the behavioral portion of the assessment should also be utilized to give employers an objective edge to hiring for traits otherwise overlooked. Also, if hiring managers utilize this ability to measure personality and behavioral traits, they can maintain a cohesive work environment by hiring applicants who have the characteristics necessary to excel in the field of nursing, as well as in their specific medical facility.

Nurses and On-the-job Success

Most hiring managers would agree that a nurse applicant's potential for being successful at their job is the main focus when screening prospective employees. Unlike many other professions, in the nursing field, job success is based on a complex combination of knowledge, skills, abilities, and personal characteristics (KSAPCs). Other than the previously mentioned technical knowledge and interpersonal skills, the ability to make good judgments and react properly in stressful situations is of the utmost importance. If an agency is to create and maintain a proficient workforce, decrease turn-over rates and lower hiring costs long-term, it is pivotal to assess applicants for these traits before hiring them.

Of the aforementioned traits, traditional methods of hiring nurses often lack proficiency in evaluating an applicant's propensity toward making good judgments and reacting properly to stressful situations. A method that has been adopted by agencies across the nation is the implementation of situational judgment tests (SJTs) (Weekley & Jones, 1997). SJTs can be delivered in written form or via video segments, called vignettes. While written SJTs are effective means to deliver hypothetical scenarios of hospital interactions, their effectiveness can be affected by extraneous or unrelated variables. For instance the applicant's level of reading comprehension or lack of imagination could affect their test results. While reading comprehension and other such variables can be significant, they are generally already tested for in the clinical assessments. So in order to maintain independence between the KSAPCs evaluated by each assessment and to create a more hands on and realistic experience for the nurse applicants, this study employed a video based SJT.

The creation of the SJT, as well as the other portions of the assessment, involved the collaboration of information gathered from Subject Matter Experts in the field of Nursing

(NSMEs). The NSMEs created a list of KSAPCs necessary for on-the-job success, this list was then used as a framework for developing *critical incidents* (highly critical events likely to occur in a hospital setting) that were compiled into scenarios and then into video vignettes. Once the vignettes were compiled, NSMEs created a key for each scenario, each of which contains four reactions that a nurse could take to handle the situation. Of the four reactions, one is considered the most effective response, one is the least effective response, and the other two are neither least nor most effective, however, one is more effective than the other. These responses are presented to the nurse applicant just after the vignette is played and they are asked to indicate which reaction the applicant considers to be the best as well as which reaction the applicant considers to be the worst.

Results from the SJT indicate that subjects' scores are predictive of their average job performance rating. Consequently, hiring managers can utilize this test to take an in depth look into the actual decision making skills of a nurse applicant long before investing in their acquisition.

Practical Implications of the Pre-employment Assessments

Using the average job performance criteria, statistical analyses of the assessments demonstrate corrected validity coefficients (i.e., correlations) of r = .21, .24, and .28 for the clinical, situational, and behavioral assessments, respectively. Some job performance dimensions even reached as high as r = .27, .35, and .38. In effect, there is some overlap between what each portion of the assessment predicts, which is not optimal, however, they do map onto separate job performance dimensions in complimentary ways (Table 1 and Figure 1).

[Insert Table 1 and Figure 2 about here]

To compute a corrected validity coefficient for the combined score, range restriction estimates were computed by comparing the test score standard deviations between the subjects and the applicants (after the first year of administration). This analysis revealed a ratio between the combined and weighted composite score of U = 1.19, indicating that the applicant variance was 19% larger than the study sample. Additional steps were necessary to compute the corrected validity, including estimating the reliability of the score composite (r = 0.753). An estimated operational validity coefficient was computed using these values, resulting in an elevated r = .67,

indicating that the assessment predicts about 45% of job performance. This level of prediction is considered *very beneficial* by the U.S. Department of Labor, and is extremely effective in taking the guesswork out of the hiring process. Building a selection system with validities as high as r = .67 requires using multi-faceted and (relatively) uncorrelated tests that can work in complimentary ways to explain job performance variance.

Gains in Workforce Performance by Using Validated Selection Systems

The following table and discussion elaborate on the practical effectiveness of using an assessment battery with an overall correlation to job performance of r = .67.

[Insert table 2 here]

Table 2 shows the expected gains in job performance success of the hired nursing staff from using a validated selection system with a correlation of r = .67 to job performance. The t-score column shows the combined (and weighted) score (set with a mean of 80 and standard deviation of 10 for each of the three tests). The Selection Ratio column shows the percentage of applicants that will likely pass the cutoff score at each corresponding t-score (based on normative data). The Base Rate column represents the percentage of incoming applicants who would likely receive an "average/above average" overall job performance rating. In hospital settings where the incoming applicant population is extremely well qualified the Base Rate will likely be high (e.g., 90%).

The values in the table that intersect the Selection Ratio and Base Rate represent the likelihood (i.e., expectancy) that a nurse hired using the complete assessment will subsequently receive job performance ratings in the 50th percentile or higher (i.e., average/above average) (Taylor & Russell, 1939). For example, consider a hospital with a 70% Base Rate where the cutoff score is set at a t-score of 82.53, where about 40% of the applicants will pass. In this situation, 92% of the incoming nurses will likely receive "average/above average" job performance ratings when subsequently evaluated by their supervisors, compared to the 70% base rate if no test (or an invalid test) was used. The difference between screening in 92% "average/above average" nurses versus 70% (when using no test) is 22%. This 22% hiring advantage (or "utility") can be directly attributed to using the validated tests in the hiring process.

Financial Advantages from Using Validated Assessments

Using validated assessments that are highly correlated with job success can also result in positive financial outcomes. For example, using the same set of assessments (with r = .67) and hiring scenario above (i.e., 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.

Conclusion

The results of this study provide evidence that a well-rounded nurse selection system should include (at least) three components: situational judgment tests, behavioral/personality assessments (for screening their general nursing staff), and clinical (job knowledge) assessments for making functional area placements (e.g., medical-surgical, emergency room, etc.). When using this three-pronged testing strategy, healthcare institutions—especially those agencies with high hiring volumes—are likely to benefit from hiring a well-rounded and high-performing nursing staff, substantial gains in overall workforce outcomes, financial advantages, and perhaps even reduced liability (from using validated pre-employment assessments).

References:

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

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

U.S. Department of Labor (2000). *Testing and Assessment: An Employer's Guide to Good Practices (p. 3-10).*

Weekley, J.A. & Jones, C. (1997). Video-based situational testing. *Personnel Psychology, 50, 25–49.*

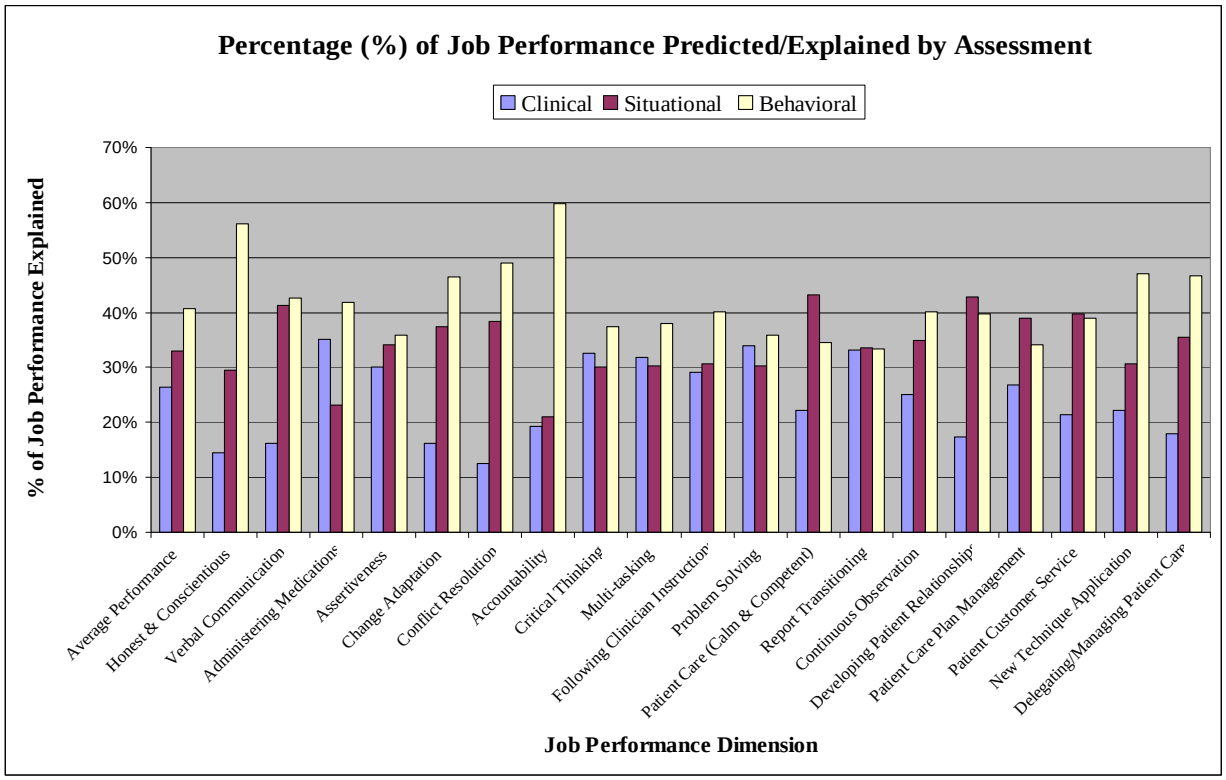
Tables:

Table 1. Percentage of Job Performance Predicted/Explained by Assessment.

Percentage (%) of Job Performance Predicted/Explained by Assessment[1,2]			
Performance Dimension	**Clinical**	**Situational**	**Behavioral**
Average Performance	26%	33%	41%
Honest & Conscientious	14%	29%	56%
Verbal Communication	16%	41%	43%
Administering Medications	35%	23%	42%
Assertiveness	30%	34%	36%
Change Adaptation	16%	37%	46%
Conflict Resolution	13%	38%	49%
Accountability	19%	21%	60%
Critical Thinking	33%	30%	37%
Multitasking	32%	30%	38%
Following Clinician Instructions	29%	31%	40%
Problem Solving	34%	30%	36%
Patient Care (Calm & Competent)	22%	43%	34%
Report Transitioning	33%	34%	33%
Continuous Observation	25%	35%	40%
Developing Patient Relationships	17%	43%	40%
Patient Care Plan Management	27%	39%	34%
Patient Customer Service	21%	40%	39%
New Technique Application	22%	31%	47%
Delegating/Managing Patient Care	18%	35%	47%
Unit Weighted Average	24%	34%	42%

Notes: (1) After controlling for the predictive variance of each of the other two assessments; (2) The values in this table are based on the total amount of criteria variance predicted (explained) by the set of three assessments, not the total variance available.

Figure 1. Percentage of Job Performance Predicted/Explained by Assessment



Note: This figure displays the percentage of the total variance explained by each of the three assessments by summing the variance explained by each test, and dividing by the total variance explained by all three tests (traditional regression methods were not completed because the dataset used for the three tests had limited overlap).

Table 2. Percentage of “Average/Above Average” Nurses Hired Based on Various Base Rate and Selection Ratio Assumptions

T-Score	Selection Ratio	Base Rate (% of Applicants who would likely receive "average/above average" performance ratings)								
<67.18%	100%	10%	20%	30%	40%	50%	60%	70%	80%	90%
67.18%	90%	11%	22%	33%	44%	55%	65%	75%	85%	94%
71.58%	80%	12%	25%	37%	48%	59%	70%	80%	89%	96%
74.76%	70%	14%	28%	41%	53%	64%	74%	84%	91%	97%
77.47%	60%	16%	31%	44%	57%	69%	79%	87%	93%	98%
80.00%	50%	19%	35%	50%	62%	73%	82%	90%	95%	99%
82.53%	40%	22%	40%	55%	68%	78%	86%	92%	97%	99%
85.24%	30%	26%	46%	62%	74%	83%	90%	95%	98%	100%
88.42%	20%	33%	54%	69%	80%	88%	93%	97%	99%	100%
92.82%	10%	44%	66%	79%	88%	93%	97%	99%	100%	100%