How to Build the Ideal Nurse Hiring Assessment

Friday, March 23, 2012, 9:00 AM & 10:00 AM
Session Number: CS-3

Presenters:
James Ostmann Sr, RN, MBA, Chief Nursing Officer,
Prophecy Healthcare, High Point, NC
Dan Biddle, PhD, President CEO,
Biddle Consulting Group, Inc., Folsom, CA

Prophecy Healthcare, Inc.
A leading provider of healthcare assessments: over 1.5 million exams administered
Standardized competency testing focuses on screening skill levels pre- and post-hire
Rigorous and comprehensive validation program
Over 100 nursing, allied, and home healthcare job knowledge competency exams
Over 50 self-assessment clinical-skills checklists

Biddle Consulting Group, Inc.
Leading U.S. EEO consulting firm
Nationally-recognized experts in test development and validation
Experience in 200+ state and federal cases involving EEO compliance & testing

PeopleClues
International test publishing firm that provides modern & validated behavioral assessments
Online platform specifically designed for the commercial market
Assessments used by thousands of companies to hire, train, and promote
The Climate…

• In June 2011, Wanted Analytics reported that employers and staffing agencies posted more than 121,000 new job ads for Registered Nurses.
• Typical annual turnover rates for RNs average 14% (KPMG survey).
• Bureau of Labor Statistics reports “the health care industry added 428,000 jobs throughout the 18-month recession from December 2007 until June 2009, and has continued to grow at a steady rate since the end of the recession.”
• Hospitals, long-term care facilities, and other ambulatory care settings added 37,000 new jobs in March 2011, the biggest monthly increase recorded by any employment sector.
• The BLS confirmed that 283,000 jobs have been added in the healthcare sector within the last year.

The Question…

• With such high growth rates expected, how do healthcare institutions hire the most qualified “best fitting” nurses for their openings?
• What knowledges, skills, abilities, and personal characteristics (KSAPCs) should hiring authorities pay most attention to when building a robust hiring system for nurses?
• Finally, how can these “key KSAPCs” be accurately measured in the shortest testing session possible?

The Presentation

• We’ll review the answers we found to these questions while working with over 900 professionals, including:
  – 492 nurse staff from Saint Francis Medical Center (470 nurses and 22 nurse supervisors)
  – 384 nurse staff from Frederick Memorial Hospital (367 nurses and 17 nurse supervisors)
  – 13 test development professionals and industrial-organizational psychology consultants
  – 30 film professionals
While Reviewing our Process...

- Keep in mind... we want you to replicate, borrow, steal, or just learn from our process
- We don’t claim to have the “best” process, but we have uncovered some of the key traits to target in your testing program
- We’ve also learned how your hiring program might benefit from using different tests in different ways for hiring nurses

Our testing program resulted in a 3-pronged assessment strategy called “Prophecy”

- Our goal was to create a stronger alignment between the nurse job requirements and nurse testing programs

- A typical nurse hiring process measures only limited clinical background (sometimes through a written test) and interpersonal competence (through interviews)

- **Prophecy** measures cognitive, interpersonal, and behavioral
  - Intentionally
  - Thoroughly
  - With limited overlap

What Makes Up Nurse Job Performance?

- But how should test skills be tested?
- How should they be weighted?
- Which skills are overlapping?

We tried to answer these questions while developing Prophecy...
The Prophecy Development Process

- Prophecy Overview
- Project Phase 1: Job Performance
- Project Phase 2: Situational
- Project Phase 3: Clinical
- Project Phase 4: Behavioral
- Project Phase 5: Combining the Assessments

2008-2011 Validation Process

Prophecy is the aggregation of three assessment tools. The three pieces work together to form the most comprehensive employment predictor for the nursing field.

Twelve Clinical Assessments were included in the study, each having 30-70 items measuring clinical specialty job knowledge (e.g., Medical-Surgical, Emergency Care, etc.)
The Situational Assessment included 21 video-based situational judgment scenarios designed to measure the applicants' interpersonal competence using a variety of complex situations that emerge in hospital settings.

A 70+ item behavioral/personality test bank was included in the study. Twenty (20) of the most predictive items were distilled into a “Nurse Effectiveness Scale” which demonstrated strong validity results.
A job analysis process was used to distill 19 job performance dimensions that met the following criteria:

- Observable aspects of job performance that were likely to relate to traits measured by the Situational Assessment;
- Represented qualifications that were needed on the first day of the job;
- Not geographically limiting;
- Not likely change over time; and
- Differentiated “best performers” in the nursing profession.

Emphasis on soft skills

Project Phase 1: Job Performance

<table>
<thead>
<tr>
<th>Nurse &amp; Conscientious</th>
<th>Problem Solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Communication</td>
<td>Patient Care (calm &amp; competent)</td>
</tr>
<tr>
<td>Administering Medications</td>
<td>Report Transitioning</td>
</tr>
<tr>
<td>Assertiveness</td>
<td>Continuous Observation</td>
</tr>
<tr>
<td>Change Adaptation</td>
<td>Developing Patient Relationships</td>
</tr>
<tr>
<td>Conflict Resolution</td>
<td>Patient Care Plan Management</td>
</tr>
<tr>
<td>Accountability</td>
<td>Patient Customer Service</td>
</tr>
<tr>
<td>Critical Thinking</td>
<td>New Technique Application</td>
</tr>
<tr>
<td>Multitasking</td>
<td>Delegating/Managing Patient Care</td>
</tr>
<tr>
<td>Following Clinician Instructions</td>
<td></td>
</tr>
</tbody>
</table>

Both Hospitals participated in the nurse rating process:

- Saint Francis Medical Center:
  - 470 nurses were rated by
  - 22 nurse supervisors
- Frederick Memorial Hospital:
  - 367 nurses rated by
  - 17 nurse supervisors

A total of 888 nurses were rated on 19 dimensions of job performance by 39 supervisors.
How to Build the Ideal Nurse Hiring Assessment

March 23, 2012

Project Phase 2: Situational Assessment Build

- “Situational Judgment Tests” (SJTs) pose hypothetical situations to applicants in a job-related context and then provides several plausible alternatives on the most effective and least effective way to handle the situation
- “Most effective” and “least effective” keys were determined using panel of 50 nurse supervisors
- Why SJTs?
  - SJTs measure valuable skills that are typically not measured in hiring processes, or are not measured very accurately or systematically.
  - Applicants perceive context-rich tests as more job relevant
  - SJTs are scored consistently and reliably, compared to when the same skills are measured using interviews
  - Because they are based on "consensus scoring," they are more objective than interviews

Project Phase 2: Situational Assessment Build

Step 1: Job analysis research was conducted at four medical/employment centers and 19 of 169 key KSAPCs were selected
Step 2: Developed SJT Scenarios (based on 19 key competencies)
Step 3: Refined SJT Scenarios & Converted to 50-Page Script
Step 4: Staged Sets & Filming
Step 5: Multi-point Keying Process was developed based on level of expert consensus
Step 6: Validation Study was conducted by correlating test scores to 19 job performance dimensions
  - Saint Francis Medical Center: 27 Nurse Supervisors rated 191 nurses
  - Frederick Memorial Hospital: 17 Nurse Supervisors rated 235 nurses

Criterion-related Validity: Demonstrated by empirical data showing that the selection procedure is predictive of, or significantly correlated with, important elements of work behavior
Interpreting Criterion-related Validity

U.S. Department of Labor Guidelines for Interpreting Validity Coefficients

<table>
<thead>
<tr>
<th>Coefficient Value</th>
<th>Interpretation</th>
<th>Percentage of Job Performance Explained (Coefficient Squared)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above .35</td>
<td>Very beneficial</td>
<td>12.3%+</td>
</tr>
<tr>
<td>.21 - .35</td>
<td>Likely to be useful</td>
<td>4.4% - 12.3%</td>
</tr>
<tr>
<td>.11 - .20</td>
<td>Depends on circumstances</td>
<td>1.2% - 4%</td>
</tr>
<tr>
<td>Below .11</td>
<td>Unlikely to be useful</td>
<td>&lt;1.2%</td>
</tr>
</tbody>
</table>


Project Phase 2: Situational Assessment Build

How well does the Situational Assessment Predict Job Performance?

Theoretical Expectancy Bands

<table>
<thead>
<tr>
<th>Band</th>
<th>Low Score</th>
<th>High Score</th>
<th>Cumulative % of Nurses Scoring in Band</th>
<th>Likelihood of Receiving “Above Average” Supervisor Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAIL</td>
<td>0</td>
<td>69</td>
<td>&lt;5%</td>
<td>&lt;42%</td>
</tr>
<tr>
<td>E</td>
<td>70</td>
<td>82</td>
<td>20%</td>
<td>43%</td>
</tr>
<tr>
<td>D</td>
<td>83</td>
<td>86</td>
<td>40%</td>
<td>54%</td>
</tr>
<tr>
<td>C</td>
<td>87</td>
<td>90</td>
<td>60%</td>
<td>65%</td>
</tr>
<tr>
<td>B</td>
<td>91</td>
<td>94</td>
<td>80%</td>
<td>75%</td>
</tr>
<tr>
<td>A</td>
<td>95</td>
<td>100</td>
<td>100%</td>
<td>85%</td>
</tr>
</tbody>
</table>

Note: The “base rate” for the expectancy table was set at 60% because this portion of the distribution received ratings >=5, indicating they were scored at the median (or higher) of job performance. The theoretical (corrected) validity used for this analysis was .30 (set roughly between the two corrected validities observed in the study).
Assuming 60% of the nurse applicants “show up qualified” (i.e., would receive average JP ratings) and the Situational Assessment is used to select highest scoring 20%...

- 67% of this “top 20% group” will exceed performance expectations on the job (opposed to only 60% by using no assessment)
- This is a 7% improvement in overall workforce qualifications (raised from 60% to 67%)
- In terms of people, say for instance your hospital hires 100 nurses per year, 7 additional nurses would be expected to exceed performance expectations (be successful on the job) than without using the assessment.

Situational Assessments Validity
Results Interpretation

The Clinical Assessments are designed to measure critical job knowledge domains that are necessary for clinical practice areas (e.g., Medical-Surgical, Emergency, etc.)

- They are developed and validated using a content validation methodology, which means:
  - Test content is mapped to a job analysis for the relevant practice area
  - Tests are balanced with respect to the questions measuring the specific aspects of each practice area
  - Cutoffs are established using a “Modified Angoff” process, where nurse experts set the cutoff score by evaluating the difficulty of each item on the test

- The Clinical Assessments are “Mastery Based” or “Criterion-Referenced” tests. This means:
  - Designed to distinguish between who “can” vs. who “cannot” perform the job
  - Test items focus on baseline competency skills and knowledge domains
  - Cutoff scores are set using nurse experts with respect to the “minimum competency level necessary to perform the job on the first day”

- Without using Clinical Assessments, hospitals stand the risk of hiring unqualified and/or under-qualified nurses who may not possess the key knowledge and skill competencies and may create liability.

- Notwithstanding this focus, the Clinical Assessments still demonstrated statistically significant correlations to job performance.
Project Phase 3: Clinical Assessment Build

**Average Correlations for both Hospitals Combined**

<table>
<thead>
<tr>
<th>Performance Dimension</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administering Medications</td>
<td>35%</td>
</tr>
<tr>
<td>Problem Solving</td>
<td>34%</td>
</tr>
<tr>
<td>Report Transitioning</td>
<td>33%</td>
</tr>
<tr>
<td>Critical Thinking</td>
<td>33%</td>
</tr>
<tr>
<td>Multitasking</td>
<td>32%</td>
</tr>
<tr>
<td>Assertiveness</td>
<td>32%</td>
</tr>
<tr>
<td>Following Clinician Instructions</td>
<td>29%</td>
</tr>
<tr>
<td>Continuous Observation</td>
<td>25%</td>
</tr>
<tr>
<td>Patient Care (Care &amp; Compassion)</td>
<td>22%</td>
</tr>
<tr>
<td>New Technique Application</td>
<td>22%</td>
</tr>
<tr>
<td>Patient Customer Service</td>
<td>21%</td>
</tr>
<tr>
<td>Accountability</td>
<td>19%</td>
</tr>
<tr>
<td>Delegating/Managing Patient Care</td>
<td>18%</td>
</tr>
<tr>
<td>Developing Patient Relationships</td>
<td>17%</td>
</tr>
<tr>
<td>Verbal Communication</td>
<td>16%</td>
</tr>
<tr>
<td>Change Adaptation</td>
<td>16%</td>
</tr>
<tr>
<td>Honesty &amp; Conscientious</td>
<td>14%</td>
</tr>
<tr>
<td>Conflict Resolution</td>
<td>13%</td>
</tr>
</tbody>
</table>

**Critical Thinking, Conscientious, Change Adaptation, Conflict Resolution**

**Soft Skills and Team Performance**

**HARD SKILLS & "SOLO" PERFORMANCE**

The test correlated to the JP dimensions just as one would expect, with the highest in the "hard skills" requiring "solo performance" and lowest with "soft skills" involving team performance.

Project Phase 4: Behavioral Assessment Build

- 344+ Nurses from Saint Francis Medical Center completed a 70-item personality/behavioral scale (scored using a 1-5 agreement scale) that measured:
  - Conscientiousness
  - Tough-Mindedness
  - Conventional
  - Extroversion
  - Stability
  - Teamwork
  - Good Impression

- 70-item scale was reduced to a 20-item “Nurse Effectiveness Scale” using a split-half validation methodology.
Project Phase 4: Behavioral Assessment Build

176 Nurses
“Calibration”
Sample
• Build the test
• Maximize correlations

168 Nurses
“Validation”
Sample
• Re-evaluate correlations

Behavioral Validity Coefficients

Conscientiousness Accountability Delegation/Mg Patients

Project Phase 5: Combining the Assessments

The tests were weighted using Structural Equation Modeling (SEM), allowing them to be combined in the most efficient way possible.

After correlating the three tests with avg. job performance, the remaining contribution of cognitive ability was non-significant.

This was because the Clinical and Situational Assessments were highly correlated with the cognitive ability test ($r = .16$ and $.29$).
Project Phase 5: Combining the Assessments

Percentage (%) of Job Performance Predicted/Explained by Assessment

<table>
<thead>
<tr>
<th>Percentage (%)</th>
<th>Job Performance Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td></td>
</tr>
</tbody>
</table>

Project Phase 5: Combining the Assessments

Top Six Ranked Performance Dimensions Predicted by Each Assessment

<table>
<thead>
<tr>
<th>Clinical</th>
<th>Situational</th>
<th>Behavioral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administering Medications</td>
<td>Patient Care (Calm &amp; Competent)</td>
<td>Accountability</td>
</tr>
<tr>
<td>Problem Solving</td>
<td>Developing Patient Relationships</td>
<td>Honest &amp; Conscientious</td>
</tr>
<tr>
<td>Report Tasking</td>
<td>Verbal Communication</td>
<td>Conflict Resolution</td>
</tr>
<tr>
<td>Critical Thinking</td>
<td>Patient Care Plan Management</td>
<td>Delegating/Managing Patient Care</td>
</tr>
<tr>
<td>Multitasking</td>
<td>New Technique Application</td>
<td>Change Adaptation</td>
</tr>
<tr>
<td>Assertiveness</td>
<td>Patient Customer Service</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conflict Resolution</td>
<td></td>
</tr>
</tbody>
</table>

Note (1): After controlling for the predictive variance of each of the other two tests.

- The final battery collectively measures each job performance dimension.
- The tests load on different aspects of job performance in a balanced way.
How Does This Translate to Actual Hiring Benefits?

**Practical Effectiveness of Using Prophecy Assessments**

<table>
<thead>
<tr>
<th>Base Rate of Qualified Nurses</th>
<th>Applicant-to-Hire Ratio</th>
<th>% Qualified Nurses Hired</th>
<th>Hiring Process Improvement Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>50%</td>
<td>89%</td>
<td>9%</td>
</tr>
<tr>
<td>80%</td>
<td>30%</td>
<td>92%</td>
<td>12%</td>
</tr>
<tr>
<td>80%</td>
<td>10%</td>
<td>95%</td>
<td>15%</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>63%</td>
<td>13%</td>
</tr>
<tr>
<td>50%</td>
<td>30%</td>
<td>69%</td>
<td>19%</td>
</tr>
<tr>
<td>50%</td>
<td>10%</td>
<td>77%</td>
<td>27%</td>
</tr>
</tbody>
</table>

Notes: (1) Assumes an average job performance of 50 and a SD of 10.
Running Through the Math…

Making the following assumptions…

– Hospital with 720 nurses
– Job Performance levels in the corporate nurse workforce is normally distributed, with half “above average” and half “below average”
– Turnover is 13.9% (per American Association of Colleges of Nursing/Bernard Hodes Group)
– No growth (to be conservative), although most research reports cite nurse vacancy rates exceeding 8% and the Bureau of Labor Statistics (BLS) projected that more than 581,500 new RN positions will be created through 2018, which is a 22% increase in the RN workforce

Let’s run through 3 scenarios:

– Using no test (or an invalid test) for 3 years
– Using Prophecy Assessments with a 50% cutscore for 3 years
– Using Prophecy Assessments with a 20% cutscore for 3 years

Changing the Workforce One Hire at a Time

Job Performance Levels of the Workforce
After 3 Years

Percentage of Above-Average Nurses in the Workforce
Percentage of Below-Average Nurses in the Workforce

Job Performance Levels of the Workforce After 3 Years of Using Prophecy (w/50% Selection Ratio)

Percentage of Above-Average Nurses in the Workforce
Percentage of Below-Average Nurses in the Workforce
Wrap-up

- The most important factor in building and using a valid assessment process is measuring the whole person (otherwise, valuable competencies are not considered in making critical employment decisions).
- The next most important factor is measuring the various competencies of that “whole person” using tests that have a proven, scientific connection to those various skills and competencies.
- The next most important factor is measuring those various competencies in reliable (consistent) and efficient (fast) ways.
- The final consideration is properly balancing (weighting) the various tests in a way that matches the job.
- Prophecy does all four.