



Legal Defensibility of the Prophecy Assessments

Prepared By:

**Biddle Consulting Group, Inc.
193 Blue Ravine Road, Suite 270
Folsom, CA 95630
(916) 294-4250
<http://www.biddle.com>**

September 10, 2012

Developed in Partnership by:
[Prophecy Healthcare, Inc.](#), [Biddle Consulting Group, Inc.](#), & [PeopleClues](#)

Contents

Background	2
The Partners	2
The Assessments	4
Relevant EEO Regulations	5
Validation Regulations and Standards	5
Adverse Impact Monitoring and Recordkeeping Obligations	6
Adverse Impact: The Trigger for EEO Liability.....	7
How Prophecy Assessments Addresses Validation Regulations.....	8
Clinical Assessments	8
Situational and Behavioral Assessments	10
Transportability Survey Tool.....	11
Appendix A: Background on Biddle Consulting Group, Inc.....	13
Appendix B: Validation Transportability Survey.....	14
Appendix C: FAQs Regarding the Defensibility and Compliance of the Prophecy Assessments	18
General Liability	18
Adverse Impact and Related Reporting	20
Validation and Defensibility	22

Background

This report was authored by [Biddle Consulting Group, Inc.](#) (BCG), one of the leading U.S. EEO Consulting firms (see Appendix A for more information on BCG) to provide an overview of the Equal Employment Opportunity (EEO) compliance issues surrounding the Prophecy Assessments.

The Prophecy Assessments were developed through a partnership of five agencies that participated in a three-year effort to develop the Prophecy Assessments, which were designed to constitute a comprehensive testing solution for hiring nurses in the healthcare industry.¹ These agencies included Amistaff Healthcare, Inc. (now “[Prophecy Healthcare](#)”), [Biddle Consulting Group, Inc.](#), [PeopleClues](#), [Saint Francis Medical Center](#), and [Frederick Memorial Hospital](#). The study was completed between 2008 and 2011 and involved over 900 professionals, including: 13 test development professionals and industrial-organizational psychology consultants, over 30 film professionals, 492 nurse staff from Saint Francis Medical Center (470 nurses and 22 nurse supervisors), and 384 nurse staff from Frederick Memorial Hospital (367 nurses and 17 nurse supervisors). A brief overview of the development partners is provided below.

The Partners

A brief biography of each study partner is provided below:

- **Amistaff Healthcare, Inc.** (now “[Prophecy Healthcare](#)”). Amistaff is a private organization that provides assessment services to the healthcare industry. One of the divisions of Amistaff is Nurse Testing which was founded in 2004 to create a standardized system of testing for nurses and healthcare professionals seeking

¹ See “Test Development & Validation Study Results” (dated June 6, 2012, by [Biddle Consulting Group, Inc.](#)).

employment at healthcare facilities and staffing agencies. Amistaff exclusively provided the Clinical (written) Assessment content included in this study, whereas the Situational Assessment content was provided through partnership with [Biddle Consulting Group, Inc.](#)

- [Biddle Consulting Group, Inc.](#) (BCG) specializes in Equal Employment Opportunity (EEO) consulting, litigation support, personnel testing software development, and Affirmative Action Plan (AAP) outsourcing and software. Since 1974, when known under the name Biddle & Associates, BCG has worked with thousands of employers in these areas, as well as providing litigation support as consultants or experts in over 200 state, federal, and circuit court of appeal EEO cases involving statistics and/or job-relatedness (test validity) analyses. BCG has developed and validated personnel tests in hundreds of situations that are used by thousands of employers. BCG's role in this partnership was to oversee the validation study and to jointly develop the Situational Assessments.
- [PeopleClues](#). PeopleClues is an international test publishing firm that provides modern and validated behavioral assessments through an online platform specifically designed for the commercial market. PeopleClues Assessments are used by thousands of companies to hire, train, and promote applicants and incumbents. PeopleClues' multi-faceted personality test item bank ("Clues"), which includes 70 personality-behaviorally based personality items, were included as part of this study.
- [Saint Francis Medical Center](#) (located in Cape Girardeau, Missouri) is a 258-bed facility serving more than 650,000 people throughout Missouri, Illinois, Kentucky, Tennessee, and Arkansas.

- [Frederick Memorial Hospital](#) (located in Frederick, Maryland) is a large medical facility that employs 2,700 medical professionals, including 375 doctors and several hundred nurses.

The Assessments

Prophecy includes three Assessments that are weighted and combined into a total score for applicants:

- **Clinical:** 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 includes between 30 and 70 multiple-choice items designed to measure the relative practice area in a balanced way. A total of 12 clinical practice area tests were included in the original validation study. These tests were originally developed and validated using a content validation strategy and validation tools (software and surveys) developed by BCG.
- **Situational:** The Situational Assessments were video-based, situational judgment tests designed to measure interpersonal competence in hospital-related situations. The final version of this test includes 21 video scenarios (30-120 seconds in length) that present complex interpersonal situations that occur in hospital environments (interacting with patients, physicians, and other co-workers) to the nurse applicant. Each video is 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. This test was originally developed through a joint partnership between Amistaff and BCG utilizing subject-matter experts from Saint Francis Medical Center.

- **Behavioral:** This 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. In addition, a 20-item “Nurse Effectiveness Scale” was specifically developed and validated through the original validation study. These scales were developed by PeopleClues.

Relevant EEO Regulations

Validation Regulations and Standards

Employment devices that are used in the screening and evaluation of job seekers are classified under the federal [*Uniform Guidelines on Employee Selection Procedures*](#) (1978)² as “tests.” When such tests are used in the hiring process by employers, they can possibly exhibit “adverse impact”—a legal term used to define situations where the success rates of two groups on a particular practice, procedure, or test are *substantially different*. If adverse impact is triggered, employers are then required to make a *demonstration* that their selection process is “job related for the position in question and consistent with business necessity.” This burden is typically met in compliance or litigation situations by addressing the fundamental validity requirements specified by the federal *Uniform Guidelines*, which have been adopted for such purposes by the U.S. Equal Employment Opportunity Commission (EEOC), the Office of Federal Contract Compliance Programs (OFCCP), and the Department of Justice (DOJ).

In addition to the *Uniform Guidelines*, personnel tests should also address the professional testing standards, including, but not limited to the *Principles for Validation and Use*

² Uniform Guidelines – Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, and Department of Justice (August 25, 1978), Adoption of Four Agencies of Uniform Guidelines on Employee Selection Procedures, 43 Federal Register, 38, 290-38,315

of Personnel Selection Procedures (published by the Society for Industrial and Organizational Psychology, 2003) and the *Standards for Educational and Psychological Testing* (published by the American Educational Research Association and American Psychological Association, 1999).

At a higher level, the Civil Rights Act of 1964, as updated by the 1991 Civil Rights Act (Title VII) requires that employers *make a demonstration* that the challenged test is “job related for the position in question and consistent with business necessity”³ whenever their practices, procedures, or tests exhibit adverse impact. The 1990 Americans with Disabilities Act (ADA) also provides regulations that enter the testing space, insofar as they relate to making reasonable accommodations during testing sessions and making sure that tests relate to the “essential functions” of position requirements.

Adverse Impact Monitoring and Recordkeeping Obligations

The federal *Uniform Guidelines* (above) require employers to compute annual adverse impact analyses on their selection practices for each group which constitutes either 2% or more of the total labor force in the relevant labor area, or 2% of more of the applicable workforce (Section 15A2).

In addition, the more recent (2006) OFCCP Internet Applicant Rule (41 CFR Part 60-1)⁴ provide additional data tracking requirements that pertain to the applicant tracking and onboarding processes of federal contractors.⁵ Specifically, these regulations require federal

³ 1991 Civil Rights Act, 2000e-2(k)[1][A][i].

⁴ Office of Federal Contract Compliance Programs (2005). 41 CFR Part 60-1, Obligation to solicit race and gender data for agency enforcement purposes; final rule, Federal Register, 70 (194), 58946-58963.

⁵ While the Uniform Guidelines are binding to all healthcare facilities, the 2006 OFCCP Internet Applicant Rule will only apply to some (the OFCCP has taken the position that it will determine their jurisdiction over healthcare facilities on a “case-by-case basis” due to recent legal developments—see OFCCP Directive, Number: 301 Date: Apr. 25, 2012).

contractors to track (including race and gender information) and maintain the results of searches that are conducted on “external applicant databases.” Federal contractors with fewer than 150 employees or a federal contract of at least \$150,000, the record retention period is one year (41 CFR 60-1.12). Contractors with at least 150 employees and a contract of \$150,000 are required to maintain the records for a period of two years.

The record keeping and adverse impact reporting requirements discussed above should be maintained by employers and may be used by federal regulatory agencies in the enforcement of Civil Rights and EEO regulations. In addition, employers that fail to maintain such data may be challenged by federal regulatory agencies under the “adverse inference” doctrine of the federal *Uniform Guidelines* (see Section 4D).

Adverse Impact: The Trigger for EEO Liability

Adverse impact occurs when two groups (e.g., men v. women) exhibit a substantially different success rate on a given practice, procedure, or test.⁶ Sometimes HR professionals make “quick evaluations” of adverse impact using the “80% test,” which evaluates whether the passing rate of the disadvantaged group (e.g., women) is less than 80% of the success rate of the comparator group (e.g., men). In most legal situations, however, a “substantially different” success rate is established using *statistical significance* tests. For example, consider the following table that displays the passing and failing numbers of men and women on an employment-related test:

Group	# Passing the Test	# Failing the Test
Men	110	90
Women	90	110

⁶ Adoption of Questions and Answers to Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures, 44 Federal Register 11,996-12,009.

In this table, the success rate of men is 55% and the success rate for women is 45%. When evaluated statistically, the likelihood of this difference occurring is only 5.2%, or odds of about 1-chance-in-20. To be “statistically significant,” in most legal settings, the statistical finding must be less than 5% (so this example would be *almost* statistically significant).⁷

How Prophecy Assessments Addresses Validation Regulations

The Prophecy Assessments were developed using both content and criterion-related validity techniques. Both techniques were applied for each part of the Prophecy Assessment (as described below) to address their respective requirements (Section 14C and 14B) of the federal *Uniform Guidelines*.

Clinical Assessments

The “Clinical Assessments” in the Prophecy battery represent a library of multiple-choice assessments that are design to measure the job knowledge required for various nurse practice areas (e.g., Medical-Surgical, Emergency Room, etc.). Because these assessments are designed to measure concrete and operationally-defined knowledge domains that are critical for the respective areas of nursing, they should be *content validated* by conducting local studies (especially when adverse impact is present).

Content validation procedures are completed by demonstrating a nexus between the content measured by the test and the content of the job (as shown through a job analysis). This

⁷ Test computed using Lancaster’s Mid-p Correction to the Two-Tail Fisher Exact Test. See Biddle, D. A. (2011). Adverse Impact and Test Validation: A Practitioner’s Handbook (3rd ed.). Scottsdale, AZ: Infinity Publishing. See also Biddle, D. A. & Morris, S. B. (2011). Using Lancaster’s mid-P correction to the Fisher’s Exact Test for adverse impact analyses. *Journal of Applied Psychology*, Jun 20, 2011. Adverse impact computations can be completed online using the tools located at: <http://www.disparateimpact.com>

also requires setting cutoff scores (or ranking/banding) in ways that are properly anchored to job success. While it is Prophecy Healthcare’s standard practice to conduct a content validation study for many of the Clinical Assessments on a national- or regional-scale, each healthcare facility should be assured that the specific Clinical Assessments are in fact relevant and job-related for the target nurse positions they are used (and this is a requirements whenever the tests exhibit adverse impact). Consulting assistance for this process is available through [Prophecy Healthcare](#) using the [Online Test Validation Analysis Program](#) (Online TVAP®) published by [Biddle Consulting Group, Inc.](#) (also licensed to [Prophecy Healthcare](#) and utilized by their consulting team for conducting validation studies).

While this content validation process will help insure local validity for the Clinical Assessments, the 2008-2011 study revealed that four of the Clinical Assessments showed various levels of correlations to overall job performance:⁸

- **General ICU Exam** (this test revealed statistically significant results when both hospitals in the study were combined, $p < .01$);
- **Dysrhythmia Exam** (this test revealed statistically significant results when both hospitals in the study were combined, $p < .01$);
- **Medical-Surgical** (this test revealed statistically significant results when both hospitals in the study were combined, $p < .05$); and
- **Neonatal Intensive Care Unit (NICU)** (this test revealed statistically significant results when both hospitals in the study were combined, $p < .01$).⁹

Based on the study results for these four Clinical Assessments, the validation evidence for these tests from the Prophecy study can be “transported” to new settings, provided that the

⁸ Sufficiently large sample sizes were not available in the original study to make conclusive criterion-related validity evaluations for each Clinical Assessment.

⁹ See the Validation Report for the original dates and version names of these assessments.

Assessments are assigned to positions where the tests are highly relevant.¹⁰ All other Clinical Assessments should be *content validated* using the [Online Test Validation Analysis Program](#) (Online TVAP®) program.

If the Clinical Assessment subjected to this content validation study survives the validation criteria applied in the [TVAP System](#), there is one additional requirement that needs to be addressed if Prophecy scores will be used to *rank* or *band* candidates who score above minimum-competency levels.¹¹ This has to do with insuring that the knowledges, skills, and abilities (KSAs) measured by the Assessment are *performance differentiating*. Performance differentiating KSAs are those that differentiate job performance when possessed by incumbents at above-minimum levels. For example, the KSA: “Ability to operate bed in facility” has not previously been rated by nurse job experts as “differentiating”; whereas the KSA: “Ability to verbally communicate clearly and effectively with others in English...” has typically been rated by nurse job experts as differentiating. To address the federal *Uniform Guidelines* requirements on ranking (29 CFR 1607.14C9), we recommend that at least 70% of the test items on a given Clinical Assessment are *linked* to differentiating KSAs (i.e., they measure KSAs which have been rated by job experts as ≥ 3.5 on the 5-point differentiating scale used in the Online TVAP System).

Situational and Behavioral Assessments

These two components were validated using criterion-related validity,¹² which allows other healthcare facilities (outside of the original study) to “transport” validation evidence to

¹⁰ This should be self-evident for most hospitals—e.g., use the Medical-Surgical test for practice areas by that same name, etc. The Dysrhythmia Exam, however, will likely be relevant across several practice areas.

¹¹ This requirement is not necessary if there is no adverse impact, and may be lessened if lower-scoring candidates “have little or no chance of being selected for employment” (29 CFR 1607.5H).

¹² The Situational Assessment was also validated using a content validity strategy.

their local setting. This type of validity evidence was amassed on these two assessments because only a single version of each Assessment is used across all clinical practice areas (e.g., Labor & Delivery, Medical-Surgical, etc.) and a sufficient sample size was available in the original study for conducting the statistical analyses necessary for a criterion-related validity study. In addition, the Behavioral Assessment measures “underlying” or “latent” traits that are not necessarily concretely defined, which requires using a validation method other than content validity.

Transportability Survey Tool

Healthcare facilities desiring to transport the validity evidence from the original Prophecy study for the Clinical Assessments (any/all of the four listed above), the Situational (for any nurse practice area), and the Behavioral (for any nurse practice area), can do so by having 10-12 job experts (2-3 nurse supervisors and 8-10 incumbents from various nurse practice areas) complete the attached “Job Comparability Survey” (Prophecy and BCG also have an online version of this survey available).

This Survey is based on the transportability requirements of the federal *Uniform Guidelines* (29 CFR 1607, Section 7B “Use of criterion-related validity evidence from other sources”) which states:

Criterion-related validity studies conducted by one test user, or described in test manuals and the professional literature, will be considered acceptable for use by another user when the following requirements are met: (1) Validity evidence. Evidence from the available studies meeting the standards of section 14B of this part clearly demonstrates that the selection procedure is valid; (2) Job similarity. The incumbents in the user's job and the incumbents in the job or group of jobs on which the validity study was conducted perform substantially the same major work behaviors, as shown by appropriate job analyses both on the job or group of jobs on which the validity study was performed and on the job for which the selection procedure is to be used; and (3) Fairness evidence. The

studies include a study of test fairness for each race, sex, and ethnic group which constitutes a significant factor in the borrowing user's relevant labor market for the job or jobs in question. If the studies under consideration satisfy paragraphs (1) and (2) of this section but do not contain an investigation of test fairness, and it is not technically feasible for the borrowing user to conduct an internal study of test fairness, the borrowing user may utilize the study until studies conducted elsewhere meeting the requirements of these guidelines show test unfairness, or until such time as it becomes technically feasible to conduct an internal study of test fairness and the results of that study can be acted upon. Users obtaining selection procedures from publishers should consider, as one factor in the decision to purchase a particular selection procedure, the availability of evidence concerning test fairness.

BCG recommends addressing this criteria by using the Transportability Survey attached prior to using Prophecy as a selection device (however, the same is *required* if the test exhibits adverse impact). This will help establish that the 19 job performance dimensions that were effectively predicted in the original study are relevant in the new local setting and target position(s).

Additional information pertaining to validation and use of the Prophecy Assessments is included in Appendix C.

Appendix A: Background on Biddle Consulting Group, Inc.

[Biddle Consulting Group, Inc.](#) (BCG) is a Human Resources consulting firm that specializes in the areas of [Equal Employment Opportunity \(EEO\) consulting](#), [litigation support](#), [personnel testing software development](#), and [Affirmative Action Plan](#) (AAP) technical support and software. Since 1974, BCG has worked with over 1,000 employers in these areas, as well as provided litigation support as consultants or experts in over 200 EEO state, federal, and circuit court of appeal cases involving statistics and/or job-relatedness (test validity) analyses. We have also developed or validated personnel tests in hundreds of situations that are used by thousands of employers.

BCG's employees have had professional articles published in several leading publications in the field of EEO compliance and test validation that deal with statistics, disparate impact, job-relatedness, and organizational behavior. BCG's books on [Adverse Impact](#), [Test Validation](#), and [Compensation Analyses](#) have become standard desk references in the field of EEO compliance and test validation, and have been used by federal enforcement agencies, law firms, and private industry.

BCG maintains a staff of 40-50 employees who specialize in these areas, including over 8-12 Masters and/or Doctorate level Industrial-Organizational Psychologists. Our leading partners are frequently on the national/regional speaking circuit in the areas of EEO compliance, affirmative action, and test development and validation. They have also provided statistical and test validation training for the executive, management, and compliance officer ranks in the U.S. Department of Labor (OFCCP).

Appendix B: Validation Transportability Survey

Job Comparability Survey

Instructions:

Thank you for agreeing to participate in our Job Comparability Survey. We need your input on this survey to determine whether a pre-employment Nursing test developed at another medical center can appropriately be used at this facility. Please read each of the 19 Major Work Behavior listed below and assign a 1-5 rating to indicate the similarity of each Major Work Behavior to a related Major Work Behavior for the Nurse position at your facility.

Name of Person Completing Survey: _____

Job Title of Person Completing Survey: _____

Phone Number: _____

Use This Rating Scale for Assigning Ratings Below

Rating	1	2	3	4	5
Description	<u>Not at All</u> Similar to my job	<u>Somewhat</u> Similar to my job	<u>Moderately</u> Similar to my job	<u>Mostly</u> Similar to my job	<u>Exactly</u> like my job

Major Work Behavior	Similarity Rating (1-5)
<p>Demonstrating honest and conscientious work practices (including being truthful, straightforward, and consistent). Including: reporting to work as scheduled and remaining on the job for the entire shift; completing all tasks as assigned in a timely fashion; respecting the property and well being of the hospital and others; appropriately maintaining the security and/or privacy of information; being self-motivating, taking action when necessary without external prompting; having pride in one’s own work and one’s own performance including the ability to consistently perform the best job possible.</p>	
<p>Verbally communicating clearly and effectively with others in English, both in person and by telephone/patient call system (or other non-visual devices), being attentive, understanding the points being made and appropriately responding, asking questions as appropriate, and not interrupting at inappropriate times. This includes the ability to effectively communicate and react at the appropriate level (e.g., elderly, confused, intoxicated, lacking education, age-specific, medical population-specific) while demonstrating tact and empathy. Includes the ability to introduce his/herself and his/her role and explain actions before they are taken. Ability to collaborate with other disciplines. Ability to explain information at an age- and condition-appropriate level.</p>	

Major Work Behavior	Similarity Rating (1-5)
Administering medications effectively and correctly to patients. Including but not limited to, medication attributes (e.g., classification, indications and actions, side effects, and adverse reactions, how to crush medications and when it is appropriate to do so, etc.), patient responses to medications (e.g., knowledge of procedure for assessing patients response to medications, handling allergic reactions, monitoring parameters when a patient is on anticoagulant therapy, assessing patient responses to titratable medications and titrating accordingly), and medication rules and regulations (e.g., knowledge of Medication Administration Record and documentation process, knowledge of standards in facility or protocols in facility for certain medications, etc.); knowledge of proper disposal of medications.	
Assertively taking actions when necessary without external prompting and standing by decisions appropriately, including recognizing and implementing patient care plans without constant supervision or guidance.	
Adapting to change without incurring overwhelming stress , including adapting to a different unit's setting when appropriately floated and working effectively in different units within the medical center.	
Effectively applying principles of conflict resolution , including coping and stress management techniques.	
Accepting accountability for one's actions , as well as the consequences of inaction.	
Utilizing critical thinking processes to perform patient head-toe assessment , including such actions as assessing patient status, analyzing data, altered mental status, identifying and intervening on abnormal assessment findings, formulating/prioritizing needed interventions and consulting/collaborating with other healthcare professionals. (Examples include but are not limited to, the assessment and management of edema/swelling, abnormal bleeding, circulation or lung/bowel sounds, gastric distention, emotional and behavior issues, etc.)	
Managing workflow by prioritizing responsibilities in a certain order or pattern, including safely performing more than one job-related task at a time (multi-tasking).	
Following clinician instructions and orders (face to face for bedside procedure).	
Quickly and accurately solving problems and/or make decisions using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems. Including applying general rules to specific problems to produce answers that make sense and to combine pieces of information to form conclusions (includes finding a relationship among seemingly unrelated events), and then to follow up or pursue additional information when information provided is not sufficient or does not make sense without prompting when appropriate.	
Providing patient care in a calm and competent manner under a variety of stressful conditions , including working with diverse patients/families/other staff who may be uncooperative, abusive, confrontational, and/or combative. Includes the ability to communicate and effectively address conflicts and problems. Includes effectively respecting and insuring patient's rights and responsibilities.	
Appropriately giving/receiving reports from other healthcare professionals regarding assigned patients; to obtain and convey all pertinent data on each assigned patient to clinician nursing staff or ancillary staff.	

Major Work Behavior	Similarity Rating (1-5)
Continuously observing and being aware of issues, actions, or factors in the environment that may affect himself/herself or others in a negative way. Includes being aware of his/her surroundings and recognizing unsafe/unhealthy activities or potential problems and acting/reacting appropriately or taking precautions. Includes the ability to react properly using crisis prevention intervention to verbally deescalate situations.	
Developing safe, therapeutic/trusting relationship with patients.	
Initiating, maintaining and revising plan of care for multiple and diverse groups of patients based upon physical head-toe assessment, patient history, and care needs. Accurately establish plan of care of new patients and assess needs for change and/or revision of care plan.	
Providing excellent customer service, especially by responding to customers' questions and needs in a timely and efficient manner and working cooperatively with other employees/departments. Includes responding to patients', relatives', co-workers', and others' questions and needs in a timely and efficient manner, maintaining patient privacy and confidentiality.	
Learning and applying new techniques or procedures with minimal stress, including, those relating to the care of multiple patients in the healthcare setting.	
Delegating and managing patient care activities to appropriate personnel (e.g., Patient Care Technicians, other assistant personnel), including communicating clearly and effectively in task delegation, ensuring tasks are completed, and leading/supervising personnel toward the achievement of goals.	
(A) Count the number of responses to each Major Work Behavior and enter the result in this box:	
(B) Sum the total number of rating points (e.g., 2, 3, 4, 3, 4) and enter the result in this box:	
(C) Divide B by A and enter the result in this box:	
(D) What percentage of your total job is represented by the 19 Major Work Behaviors above? (Write in between 0% and 100%). For example, if the 19 Major Work Behaviors above represent everything you do on your job, enter "100%" in the box. If they only represent one-half of you total Major Work Behaviors, enter 50%.	

Summary Instructions/Interpretation

After having a *representative sample* of Job Experts from the target position complete this survey, tally and double check the results. If the average value listed in Row C is three or higher **AND** if the average percentage in Row D is greater than 25%, the Major Work Behaviors listed above are substantially similar and represent the substantial portion of the target job, it may be concluded that the target job is sufficiently similar to the original validation position and the test validity may be transferred to the new position.

Appendix C: FAQs Regarding the Defensibility and Compliance of the Prophecy Assessments

General Liability

Can *individual test takers* challenge the Prophecy Assessments?

No. While Prophecy Healthcare will work cooperatively with test takers to review and discuss scores, individuals cannot bring challenges under current federal testing laws and regulations.

If my employer gets sued for using Prophecy Assessments, who is responsible for defending us?

The liability associated with using the Prophecy Assessments rests with employers.

Does Prophecy indemnify our employer against lawsuits?

Healthcare Facilities using Prophecy can be sued under two primary federal civil rights laws: *Disparate treatment* and *disparate impact*. Disparate treatment cases can emerge when a facility used the test in different ways for different groups. This type of lawsuit is rare, however, because facilities typically use Prophecy tests in a uniform fashion.

Disparate impact cases can be brought by either private plaintiff groups (e.g., a class of people who share a certain protected characteristic—e.g., race or gender) or state/federal government enforcement agencies, such as state departments of fair employment or the federal Equal Employment Opportunity Commission (EEOC), the Department of Justice (DOJ), or (in the case of a federal contractor), the Office of Federal Contract Compliance

Programs (OFCCP). When such a case is brought by one of these plaintiff groups, they bear the burden of proving that the use of the Prophecy Assessments exhibited a statistically significant “adverse impact” on the at-issue group. Such disparity occurs whenever the facilities actual use of the Prophecy scores to hire applicants resulted in a significant disparity using a two-by-two table analysis (see www.disparateimpact.com). Note that such analysis is different than the test exhibiting different averages between various groups, or the tests generally exhibiting differences across multiple hiring situations or employers. If the plaintiff group can successfully prove the “adverse impact” burden, the burden then shifts to the facility to prove that the Prophecy test(s) are “job related for the position in question and consistent with business necessity” (as per the 1991 Civil Rights Act, Section 703[k][1][A]). This generally means providing evidence that the test(s) address the key requirements of the federal [Uniform Guidelines](#). While the Biddle Consulting Group, Inc. (BCG) has prepared a general report demonstrating the validity of the Prophecy Assessment components in two hospital settings, the facility will need to either transport this validity (using a method that addresses Section 7B of the *Uniform Guidelines*) or conduct a local validation study that addresses section 14C and/or 14B of the *Uniform Guidelines*. Because the validity of tests is always contextual and situationally-specific, Prophecy does not guarantee validity. If the situation emerges where validity services are needed, they can be provided by the Prophecy team on a time-and-charges basis.

Adverse Impact and Related Reporting

Do the Prophecy Assessments exhibit adverse impact?

The subgroup performance levels of the Prophecy Assessments vary for each of the Assessments. Because adverse impact is generated based upon an employer's specific *use* of test scores, the level of adverse impact exhibited on the Prophecy Assessments will depend on the number of applicants screened and the specific cutoff, banding, or ranking decisions used.

Does Prophecy Healthcare make reporting tools available for employers so they can complete adverse impact analyses on their hiring practices?

Yes. Employers that use the Prophecy Assessments can run their hiring numbers through BCG's exclusive website for computing adverse impact: www.disparateimpact.com. This tool will help employers comply with the record keeping and adverse impact requirements of the federal *Uniform Guidelines* (41 CFR Part 60-3). When adverse impact is exhibited by their use of the Prophecy Assessments for a given situation, the validation steps described herein will need to be completed.

The Office of Federal Contract Compliance Programs (OFCCP) has recently taken the position that it will determine their jurisdiction over healthcare facilities on a “case-by-case basis” due to recent legal developments.¹³ What recordkeeping and validation requirements do

¹³ OFCCP Directive, Number: 301 Date: Apr. 25, 2012.

healthcare facilities have based on whether they are considered “federal contractors or subcontractors”?

The “adverse impact” and “validation” requirements discussed herein apply to all healthcare facilities that have more than 15 employees (i.e., they are federal requirements for all U.S. employers). However, the *recordkeeping obligations* of healthcare facilities that are considered federal contractors are more stringent. While all employers are required under the *Uniform Guidelines* to track race, gender, and score-level data for applicants that are selected or referred into an employer’s hiring process, federal contractors also need to comply with the 2006 *OFCCP Internet Applicant Rule*. Therefore, if Prophecy Assessment scores are used in any way in the hiring process, even if they are only used to “refer” applicants (based on certain scores) into a healthcare facility’s existing hiring program, they are still required to track and maintain race, gender, and score-level data of job seekers accessed using Prophecy. This is because Prophecy is considered as an “employment-related test” under both the *Uniform Guidelines* and the 2006 Internet Applicant Rule:

- “OFCCP does not deem employment tests to be basic qualifications under the final rule and contractors must continue to collect and maintain records related to the impact of employment tests that are used as employee selection procedures, without regard to whether the tests were administered to Internet Applicants.”¹⁴
- “Employment tests used as employee selection procedures, including on-line tests, are not considered basic qualifications under the final rule. Contractors are required to retain records about the gender, race and ethnicity of employment test

¹⁴ Federal Register [Rules and Regulations], Vol. 70, No. 194, October 7, 2005, p. 58947.

takers who take an employment test used to screen them for employment, regardless of whether test takers are Internet Applicants under section 60–1.3. For example, if 100 job seekers take an employment test, but the contractor only considers test results for the 50 who meet the basic qualifications for the job, demographic information must be solicited only for the 50 job seekers screened by test results because the test was used as a selection procedure only for those individuals. By contrast, if the contractor used the test results from 100 test takers to narrow the pool to 50 job seekers whose basic qualifications are considered, the test is used as a selection procedure and demographic information from all test takers must be solicited.”¹⁵

Based on these requirements, federal government contractors that use Prophecy need to retain the race, gender, and score-level data of those queried, referred, or evaluated; they are also advised to annually compute adverse impact analyses.

Validation and Defensibility

Does Prophecy comply with federal testing regulations (the *Uniform Guidelines on Employee Selection Procedures, 41 CFR, 60-3*)?

Because validity pertains to what a given test score *means* relative to a certain job, test publishers cannot make generalized, one-size-fits-all validity assertions for tests. For example, a 5-minute typing speed test might be “valid” for a receptionist position when used with 45 words-per-minute cutoff and “valid” when used with an 80 words-per-minute cutoff for a professional word processing position, but not vice versa.

¹⁵ Ibid, p. 58954.

The federal *Uniform Guidelines* allow employers to conduct local validation studies on their tests using three possible strategies: content validation, criterion-related validation, and construct-related validation. Prophecy provides helpful tools for employers to utilize the first two strategies (content and criterion-related validity).

How can Prophecy Assessments be validated using content validity?

For the Clinical Assessments, employers can coordinate with [Prophecy](#) to use the [Online Test Validation Analysis Program](#) (Online TVAP®) published by [Biddle Consulting Group, Inc.](#) (also licensed to [Prophecy Healthcare](#) and utilized by their consulting team for conducting validation studies). Using this tool will also help the employer set localized, custom cutoff scores for the Clinical Assessments. While the Situational Assessment has been validated using both content- and criterion-related validity, we recommend that employers insure local validity by using the Transportability Survey made available by [Prophecy Healthcare](#). The same is required to transport validity for the Behavioral Assessments. This automated survey tool was exclusively developed to transport the validity evidence gathered in the original Prophecy validation studies.

Can the Prophecy Assessments be specifically validated at my employer?

Employers desiring a stronger level of validation evidence can coordinate with Prophecy to conduct local validation studies. For the Clinical Assessments, this involves convening a representative sample of job experts, conducting a job analysis workshop, and completing surveys regarding the connection between the Prophecy Assessments and the target job requirements. For the Situational and Behavioral, this requires

administering the tests to incumbents, administering the Job Performance Survey to Nurse Supervisors (to gather their job performance ratings on the nursing staff), and conducting a statistical analysis followed by a report. This type of study requires at least 150 subjects¹⁶ and a fee (alternatively, the study can be completed by the employer under Prophecy’s direction).¹⁷

Can the Prophecy Assessments be validated using a “Validity Generalization” (VG) strategy?

No. In employment situations where statistically significant adverse impact is exhibited, the 1991 Civil Rights Act requires that employers *make a demonstration* that the challenged test is “job related *for the position in question* and consistent with business necessity.”¹⁸ This job-specific doctrine of *situational specificity* requires conducting a local validity study. However, employers are encouraged to conduct a “Validation Transportability” study (see above).

Does Prophecy comply with professional testing standards, such as the Joint Standards and the SIOP Principles?

Yes. The *Standards for Educational and Psychological Testing* (1999) (endorsed by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education) and the *Principles for the Validation and Use of Personnel Selection Procedures* (2003) (endorsed by the

¹⁶ Criterion-related validation studies can be conducted with fewer incumbents (or applicants if a predictive model is being used); however, employers are advised that statistical power is diminished with smaller samples (e.g., power is 90% to detect a .25 validity coefficient when the sample size is 134).

¹⁷ This study can also be completed using applicants, rather than incumbents.

¹⁸ 1991 Civil Rights Act, 2000e-2(k)(1)(A)(i).

Society for Industrial and Organizational Psychology, Inc.) collectively represent the set of professional standards in the testing community. These standards were addressed when developing the Prophecy Assessments.