Job Simulation/Work Sample Testing
Using Simulations/Work Samples to Select the Best Employees

Presented by
Jim Kuthy, Ph.D. & Heather Patchell, M.A.

Visit BCGi Online
• While you are waiting for the webinar to begin,
  – Don’t forget to check out our other training opportunities through the BCGi website.
  – Join our community by signing up (it’s free) and we will notify you of our upcoming free training events as well as other information of value to the HR community.

www.BCGinstitute.org

HRCI Credit
• BCG is an HRCI Preferred Provider
• CE Credits are available for attending this webinar
• Only those who remain with us for at least 80% of the webinar will be eligible to receive the HRCI training completion form for CE submission
BCGi is Sponsored by Biddle Consulting Group, Inc.

- Forty years of experience in Equal Employment Opportunity consulting
- Has represented hundreds of employers in litigation-related settings
- Has performed job analyses and validation studies for hundreds of employers
- Has created valid selection testing that have been used by thousands of employers

www.biddle.com

Job Simulation/Work Sample Testing

Using Simulations/Work Samples to Select the Best Employees

Presented by
Jim Kuthy, Ph.D. & Heather Patchell, M.A.

Contact Information

Jim Kuthy, Ph.D.
Principal Consultant
jkuthy@biddle.com
(800) 999-0438 x 239

Heather Patchell, M.A.
Consultant II
hpatchell@biddle.com
(800) 999-0438 x 155
The Presenters...

• Jim holds Masters and Doctorate Degrees in Industrial & Organizational Psychology
• Heather holds a Masters Degree in Psychology
• More than twenty-five combined years of experience in the employment selection field
• They have designed and/or validated work-sample tests for many employers, including conducting validation studies that have successfully passed review of federal agencies
• Jim taught Psychology and Business-related courses at the University of Akron and California State University, Sacramento and Heather is the Director of the BCCi Institute for Workforce Development and Managing Editor of the EEO Insight journal

Presentation Outline

• What is work sample/simulation testing?
• Benefits of work sample testing
• Reliability of work sample testing
• Administering work sample tests
• Scoring work sample tests and using results
• Thoughts about physically demanding work-sample testing
  – Americans with Disabilities Act

Work Sample Testing

• Work sample testing refers “hands-on” performance tests in which a job candidate is required to perform (or describe) activities/tasks that are performed on the job

Why should we consider using Work Sample Testing?

• It is widely accepted that one of the best predictors of future performance is past performance.
• During a work sample test the candidate performs a task similar to tasks performed on the job.
  – If properly designed, success on the test is likely to reflect future success on the job.
  – Work sample tests are generally better at predicting success in the short term than long-term overall successful performance.
  – However, since you cannot become a long-term employee unless you are a successful short-term employee first, this should not be too great of a concern.
• Properly-crafted work sample tests can potentially have less adverse impact than written tests.
  – Minimize written content if you are trying to reduce adverse impact against minority-group members.

What Can be Gained?

• Work sample tests add 24% increase in validity over cognitive ability tests alone.
• Similarly, they lead to a 24% increase in practical value (utility).
  – The relationships are directly proportional, but the smaller the selection ratio, the higher the utility.
  – They say this can be expressed as a 24% increase in the gain in dollar value of output or 24% increase in the percentage of increase in output produced by using cognitive ability testing alone.

What does this Really Mean?

- Utility analysis: The gain in terms of the improved job performance of new managers hired using a type of work-sample testing called “assessment center” over a four year period was estimated at $13.4 million (1986 dollars), or approximately $2,700 each year for each of the 1,100 people promoted in first-level management jobs.
- Must be weighed against the costs of setting up and conducting the testing.


Work Sample Tests can inform candidates about the job

- Work sample tests can provide a realistic preview of the job to candidates.
  - Realistic Job Preview (RJP)
- Research suggests that realistic job previews can reduce turnover rate of those who are hired.
  - Also, unqualified test takers may realize the job is not for them and remove themselves from the process.

Work Sample Test Examples

- **Mechanic**
  - Repair a problem on a car
  - Use a repair manual
- **Clerical**
  - Typing Test
  - Proofreading test
- **Airline Pilot**
  - Airplane simulator
  - Rudder control test
- **Truck Driver**
  - Sample of driving skill
  - Use a map
- **Computer Programmer**
  - Create a simple program
  - Debug a current program
- **Electrician**
  - Wire a simulated device
  - Troubleshoot a short circuit
- **Sewing Machine Operator**
  - Sew a simple garment
  - Create a buttonhole
- **Police Officer**
  - Writing test
  - Climbing, running

Adapted from “Selection Assessment Methods,” Pulakos 2005
Work Sample Test Example

Data entry or typing tests are examples of work-sample testing.

Work Sample Test Example

Welding Work Sample Test

Work Sample Test Example


Work Sample Test
Fidelity Levels

• **High fidelity**: Realistic tests using hands-on tasks or close simulations of important/critical tasks performed on the job
  – Can include well-constructed simulations (such as simulated computer programs)
  – Have high “face” validity (i.e., the look and feel of the job)
• **Low fidelity**: Less similar to actual job; less realistic
  – e.g., candidate indicates response verbally or in writing, rather than actually performing the task, such as a situational interview
  – Low fidelity testing can sometimes lack contextual clues

Work-sample tests can range from very high to very low fidelity

---

CAUTION

Just because a test “looks” like a job does not necessarily mean it will be a valid and useful test

---

Work Sample Test
Fidelity Level Examples

Lower Fidelity

High Fidelity
Simulations

- Imitations (mimicking) of work tasks are sometimes more appropriate than having a candidate perform actual job tasks
  - Simulations can be especially helpful when it is difficult, expensive, or unsafe to test in the actual work environment
  - High quality simulations are being used more frequently as technical advances are made in presentation options, such as videos and computers

Simulation Example

Aircraft simulation photo courtesy of www.simulator.com

Simulation Example

CritiCall

Use mouse to select correct agency

Who called? Click the "New Incident" button to go to the next question.
What to Measure during Work Sample Testing

• Measure critical or important observable work behaviors and/or work products
  – Identify importance during the job analysis
  – Measure knowledge, skills, or abilities (KSAs) needed the first actual workday on the job
    – Should not measure KSAs that will be learned on the job or that can be learned in a brief orientation

Keep it Simple

• Testing should be simple, if possible
  – Unless the job analysis shows that a minimally qualified employee on the first day of the job would need to be able to learn complex work-related tasks and/or follow complex instructions in a short period of time and then be able to immediately perform those tasks
• Simplicity eases potential confusion and aids with scoring
  – Simplicity generally results in greater reliability
Test Reliability

- Reliability is the degree to which test scores are consistent, dependable, or repeatable
- Standardized administration, instructions, and scoring help to improve reliability
- The appropriate type of reliability depends on the type of test and/or its purpose
  - Work sample tests are frequently supported using test/retest reliability
  - Internal consistency reliability may be appropriate for tests that have a number of different scoring criteria
  - Inter-rater reliability may be appropriate for determining the agreement of the evaluators
- Generally, work sample tests should have reliability values of >0.70 for each section of the test that applicants must pass

Administering & Scoring Work-Sample Tests

- Administrators/Raters should be trained on how to administer and score the test
  - Measuring how well the test administrators/raters have learned the process is a good idea
- Administrators should follow a word-for-word script to insure standardization across candidates
  - Pre-recorded instructions can help to insure this
  - Computer-based simulations minimize this need
- Demonstrate the tasks to be performed either in person or using a video

- Any pace, distance, weights, or other performance criteria during test events should be job related!
  - Generally, the task being tested should reflect acceptable performance on the job
    - E.g., Job requires heavy boxes moved from point A to point B. Some successful job holders can carry several boxes in one trip. Other successful performers carry fewer boxes and make more trips. Both approaches are acceptable on the job. Test should require candidates to carry the smallest amount that successful performers can justifiably carry on the job
**Time Limits**

- Unless there is evidence that speed of performance is job related, the time limit should be set so that all (or virtually all) qualified applicants can complete the test within the time limit that has been set.
  - The speed of testing should be similar to, or more lenient than, the speed at which the job is performed.
    - Keep in mind that in some instances the candidate will have little previous practice in the area being tested.
- If the test has multiple events and is scored using a time limit, you should schedule events so that timed events immediately precede or follow physically-demanding tasks ONLY if that occurs on the job.

**How Many Attempts should be Allowed?**

- It is recommended you allow for more than one attempt.
  - Unless it is obvious injury or harm would occur.
    - Fully document if candidate not allowed re-takes.
  - Maximum number of allowable attempts should be a realistic number to determine whether that job candidate would be able to successfully perform the task on the job.
    - Hint: Allowing practice on the testing task often eliminates the need for multiple attempts.

**Minimize “Test Learning” Time**

- Generally, work-sample testing process should require very little learning or practice for a candidate to succeed on the actual test event if that same person could succeed on the job without similar learning or practice.
  - The “learning curve” for some video games and computer devices is extremely high, and therefore may make them inappropriate for selection testing.
    - This is not the case when learning is one of the skills being measured during testing.
- The longer it takes test takers to learn to perform well on the test, the more likely the test may be measuring constructs that are not job related.
Practice for Testing

- Many employers find allowing practice in advance of testing helps with the perception of test fairness
  - Fully describe the events so that candidates can replicate them for practice
    - If event requires special tools or props, consider suggesting alternatives candidates can use during practice
  - For challenging testing that could benefit from practice, consider opening the test course in advance of testing to allow practice
    - Consider having those practicing sign a liability waiver

Test Scoring

- Scores should be based on clear and unambiguous observable performance
  - Observation checklists can be very helpful
- Ambiguous criteria, such as “appeared to be struggling when performing the task,” are not appropriate for scoring purposes
- The final decision should be based on whether the test taker successfully completed the task or not

Approaches to Scoring

Multiple Hurdle
- Each and every required event must be successfully completed

Compensatory
- Good performance on one required event can be used to compensate for unsuccessful performance on another event

Some employers find it best for testing to be stopped if candidate fails to complete a required event when multiple-hurdle approach is used
Test Use

• Using tests with Pass/Fail Cutoff Scores
  – Cutoffs are used to classify candidates into two groups – unqualified and qualified
  – Cutoff should be set “so as to be reasonable and consistent with normal expectations of acceptable proficiency in the workforce”

Using a test with Ranked or Banded candidates

• Ranked or Banded Duties or KSAs measured should be performance differentiating between “adequate” and “superior” job performance
  – Banding refers to assigning test takers into groups that have statistically-similar test scores
  – Ranking and banding typically requires a higher level of validity and reliability than when using pass/fail

Keep Good & Accurate Records

• Test scoring and related information should be recorded contemporaneously
• Carefully record applicants’ performance
• When appropriate, it is a good idea to indicate how the test taker failed an event, using non-subjective wording
  – For example, simply writing “failed event” is not as helpful as “failed to climb over wall even after five attempts”
  – Do not include subjective terms such as “struggled”
  – More detailed information will help with recollection at a later time
Caution!

• Do not make notes about non job-related factors associated with a candidate
  – For example, you should not make notes about the perfume/cologne or makeup someone is wearing, unless that is related to the job
  – Why? This is the type of potentially inappropriate information that can be used to bolster someone’s claim they were evaluated on non job-related factors

• We suggest not advising candidates about their pass/fail status onsite, even if it is obvious to you
  – That is best done later, once the scores have been reviewed

Hint

• Test takers frequently have an unrealistic recollection how well they performed during testing. For this reason, video and/or audio recording candidates' behavior during test events will often enhance an employer’s ability to successfully defend the elimination of unqualified job candidates
  – As with all test-related documents, recordings should be retained in the event of challenge by job candidates, which can sometimes occur years after testing has taken place
  – Recordings of job candidates should only be used for selection purposes

Work sample tests are best suited for...

• Job tasks that are not likely to change greatly and/or quickly over time
• Jobs where the candidate must possess an adequate skill level at the time of hire
• Jobs where the skill takes a very long time to learn/master and/or skills that not everyone can easily master
• Jobs where the ability demonstrated during testing is needed to perform important or critical job duties
Work Sample Testing can increase Satisfaction – Lessen Likelihood of Litigation

- Test takers generally report being more satisfied when testing is transparently similar to the job, such that those who would fail the test would realize they would not be able to successfully perform the job
  - It may sometimes be helpful to explain the relationship between the test and the job if it might not be immediately transparent to the test taker

Why should we care?

- EEOC v. Dial Corporation
  - Work sample test appeared very similar to the job, but was actually more difficult since it did not allow for the naturally-occurring breaks between actions that occurred on the job
  - Several women failed the test because they “struggled,” even though they completed the assigned testing tasks
  - Company ordered to pay $3.3 million to 52 rejected female job candidates
    - Upheld in 2006 by the United States Court of Appeals for the Eighth Circuit

Americans with Disabilities Act

- Reasonable accommodation is any modification or adjustment to a job or the work environment that will enable a qualified applicant or employee with a disability to participate in the application process or to perform essential job functions
  - Accommodations may be needed to ensure that tests or examinations measure the actual ability of an individual to perform job functions rather than reflect limitations caused by the disability
  - Tests should be given to people who have necessary, speaking, or manual impairments in a format that does not require the use of the impaired skill, unless it is a job-related skill that the test is designed to measure
- It may be difficult to quickly create reasonable accommodation for some types of simulations, especially those that are computer based
- It is a good idea to plan in advance how you would approach offering reasonable accommodation if necessary
**Americans with Disabilities Act**

The ADA limits “medically based” testing

- Measuring or recording the test takers’ physiological signs, such as heart or breathing rate, is considered a “medically-based test” under the ADA, even if that information is used for only for informational or safety/monitoring purposes
- Medically-based tests cannot be given to job candidates until after a bona fide offer of employment has been given to the job candidate

---

**Test Environment**

- Make it easy to follow each step
  - e.g., clearly mark path to follow, including starting and ending points
  - Any limits should be based on solid, job-related reasons
- Restrict spectators
- Minimize cheering, jeering, whistling, yelling, and/or interference or support from observers or other test takers
  - Administrators may provide similar level of support for all test takers

---

**Performance Techniques**

- Allow candidates to use their own, safe techniques for performing the actions required (such as climbing, carrying, dragging) unless there is a specific and/or safer technique required on the job
- Demonstrate safe test activity (e.g., lifting, dragging, etc.) methods
- If you offer suggestions, they should be the same for all candidates
**Record Keeping**

- Document performance during testing
  - *Do not wait until the end of the testing process*

- Suggestion: To help minimize potential later conflict, have test-completion form signed by administrator and test taker immediately once testing is completed
  - *Consider having test takers “check out” before leaving to insure all scores and activities have been properly recorded*

**Health and Safety**

- Emphasize safety *before* and *during* the test
- Make the test events and/or devices safe
- Immediately stop testing if candidate fails to follow a safety rule and/or safe working practice
  - *Explain how it should be done and/or demonstrate safe method*
  - *Allow testing to continue unless it is obvious that injury or damage would occur*
- Carefully document safety violations

**Documenting Safety Violations**

- If a candidate violates a safety rule or if during the physical ability task event they perform in a way that the administrator feels demonstrates that they might not possess the level of safety-related knowledge that a minimally qualified entry-level employee should possess, a complete and accurate description of that violation or unsafe work behavior should be presented to a panel of target-job experts (and/or supervisors and/or trainers of the target job) or safety-committee members for evaluation.
- Those experts shall determine whether the violation or unsafe work practice indicated that the candidate does not possess the level of safety-related knowledge or ability that a minimally qualified entry-level employee should possess.
Documenting Safety Violations

- The job candidate should be disqualified if it is determined by the panel that the candidate performed in a way that indicated they do not possess the level of safety-related knowledge that a minimally qualified entry-level employee for the target job should possess prior to any training or experience gained after hire.

Physically-Demanding Testing Suggestions

- Have test takers sign an informed waiver of liability
  - Check with your legal counsel about requiring a doctor’s clearance for this type of testing
- Have medical aid nearby if the event is physically stressful or potentially hazardous
- Provide guidance on safe techniques
- Monitor test performance without the use of medical devices, if test is given pre-offer
- Maintain safe testing environment
  - If appropriate, frequently check condition of testing equipment and/or props

Complimentary Testing Guide

Physical Ability and/or Work Sample Testing Development and Administration Hint:
Including a sample test description and blueprint, along with a validation checklist.
Final Thoughts...

• The closer the test mimics the content and context of the job, the more content valid it will be
  – Section 14C(4), Uniform Guidelines
• Minimize written-text content if you are attempting to lower adverse impact
• Use work-sample tests along with other selection devices to measure the “whole person”
• Take the time to do it right!

Questions?

Visit BCGi to learn about more free training opportunities!

www.BCGinstitute.org

or Contact the Presenters at:

jkuthy@biddle.com or hpatchell@biddle.com