



PARTNERSHIP FOR PUBLIC SERVICE

Asking the Wrong Questions

A Look at How the Federal Government Assesses and Selects Its Workforce

A REPORT IN THE FEDERAL HIRING SERIES
OCTOBER 2004

The Partnership for Public Service is a nonpartisan, nonprofit organization dedicated to meeting the needs of government by transforming the federal workplace into an environment that nurtures talent and inspiring a new generation to serve. In its activities and publications, the Partnership focuses its energy on achieving leverage through entrepreneurial partnerships and collaboration. With an emphasis on action, accomplishments and measurable results, it acts as a catalyst for change.

The Partnership gratefully acknowledges the contributions of Pan, the Performance Assessment Network, which sponsored this report on assessment practices in the federal government.

EXECUTIVE SUMMARY

Consensus in Washington, D.C. is more elusive than ever, but one thing that everyone can agree on is that the federal hiring process is broken and must be fixed. At the same time, doubts remain as to what exactly that means.

When most people hear talk of fixing the hiring process, they think you are referring to the time it takes to navigate the system, which commonly runs up to six months and can last more than a year. Clearly these delays must be addressed. But the key point to understand about the flaws in the federal hiring process is that they are about much more than just speed. Government needs to hire fast, but more important, it needs to hire well.

Nothing has driven this point home more forcefully than the recently-released 9/11 Commission report, which showed that, in some instances, having the right person in the right job can literally be a matter of life and death. The report details multiple incidents in which one federal employee simply doing his job well was able to thwart terrorist activities. As the report says, “The quality of the people is more important than the quality of the wiring diagrams.”

Putting the right people in place requires a continuum of effective practices. Our government needs an effective process for targeting potential workers. It needs one for making the sale once a candidate is interested. And it needs to create work conditions that will enable it to retain workers.

But one of the most important things that the federal government must do to ensure that it has the right people in the right place is accurately assess the skills and competencies of potential workers. This report analyzes the federal government’s assessment practices and shows that its

policies for assessing talent are not only outdated; they are unacceptable and in need of dramatic reform. Key findings include:

- ★ According to independent analyses, the primary method used by the federal government to assess new hires – assigning points based on applicants’ self-reported training and experience – is the least effective available predictor of job performance.
- ★ Other assessment tools have been found to be up to five times more accurate as predictors of performance than the method most widely used by the federal government.
- ★ Only 39 percent of federal employees agree that their work unit is hiring people with the right skills.
- ★ Persistent myths, in particular a misreading of a 23-year-old consent decree, are an obstacle to effective talent assessments.
- ★ Some agencies, such as the Internal Revenue Service and the Transportation Security Administration, have introduced effective assessment practices.

Two centuries ago, Thomas Jefferson said, “No duty the Executive had to perform was so trying as to put the right man in the right place.” Jefferson’s sentiments still ring true. Getting the right people into government will not be easy. It cannot be done with the stroke of a pen. But government’s ability to do so will determine the quality of services it delivers to the American people. Effective assessment tools are essential to any successful plan to guarantee a high-performing federal workforce. This report will examine what needs to be done to improve federal assessment practices so that when critical decisions are being made, it will be the right person making the call.

INTRODUCTION

Faced with an impending wave of retirements and the need to fill thousands of newly-created jobs to handle new responsibilities, such as those related to homeland security and the administration of expanded Medicare benefits, the federal government has a once-in-a-generation opportunity to hire a new cohort of federal employees whose skills and competencies are well-matched to the challenges ahead. Unfortunately, the methods used by the government to evaluate whether a potential applicant has “the right stuff” are outmoded and ineffective.

Without a concerted effort to foster the development and use of improved procedures for selecting who gets hired, the federal government will likely hire the wrong people for the wrong jobs, wasting its resources and losing productivity while it inadvertently overlooks some of its best candidates.

As a nonpartisan, nonprofit organization dedicated to revitalizing the federal civil service, the Partnership for Public Service has often focused on the government’s need for talent.¹ This report – part of a series focused on federal hiring – examines the urgent need to overhaul the way the government evaluates external job applicants and selects new employees.

To complete this study, the Partnership conducted a series of interviews with federal agency leaders and private sector and academic experts, reviewed current assessment practices for external hiring and examined existing literature on the effectiveness of various assessment techniques in both the public and private sector. This report discusses the significance of the challenge, the principal barriers that have discouraged broader development of the next genera-

tion of federal assessment tools and the kinds of assessment techniques being used in the private sector and in select federal agencies. It also lays out the Partnership’s recommendations for improving the federal government’s ability to hire the right person for the right job.

I. THE CASE FOR CHANGE

Improved Assessment Techniques Are Urgently Needed

On average, over the past three years, the federal government has hired more than 100,000 new employees each year into permanent, full-time positions, primarily to replace employees who retired or resigned.² Given current federal employee turnover projections, this need is likely to continue for many years to come.

The best evidence available suggests that many government agencies are not effectively assessing applicants for these positions. Only 39 percent of federal employees agree that their work unit is able to hire people with the right skills.³ And when hiring competitively, most government agencies try to identify the “best” candidates with “rating and ranking” schedules that simply assign points based on the applicants’ *self-reported* past training and experience.⁴ This predominant method has been found to be one of the *least effective* ways to predict job performance.

Here is an example of how most new competitive federal hires are actually assessed. Suppose an agency wants to hire an accountant. The agency would review all applications and look at the candidates’ education and work experience to see if they meet the Office of Personnel Management’s (OPM) minimum qualification standards. For those candidates who meet the minimum requirements, the agency would look once again at each applicant’s educational back-

ground in accounting and award points based on the amount of relevant schooling. Then the reviewer would look at the level of experience working as an accountant and assign points for that too. There is virtually no evaluation of the quality of the work experience or the education. Also, there is little room for separation in the scoring system, and the candidates tend to get bunched up. And again, the information provided is self-reported, making it prone to embellishment.

Separately, about 40 percent of new federal hires are assessed with written tests, which have proven to be rather effective predictors. This 40 percent consists primarily of GS-2, 3 and 4 entry-level clerical and technical hires.⁵

The challenge of accurate and timely assessment is especially acute for the government because applicants for federal jobs frequently outstrip the number of positions available. In 2002, for instance, 1.7 million people applied for 62,000 jobs as screeners for the Transportation Security Administration, 47,000 people applied for 900 FBI special agent positions, 20,000 people applied to an interagency information technology job fair, and 1,500 people applied for 25 intern positions at the Environmental Protection Agency.⁶ Although government still faces recruiting challenges with many positions, these numbers clearly illustrate the need for government to be able to assess candidates efficiently.

Poor employee assessment and selection practices mean poor-quality services for the American people from their government. More specifically, questionable assessments result in bad employee/job matches that have a real cost in terms of lower productivity and increased turnover.⁷ For example, if just 10 percent of the new *professional* and *administrative* hires each year leave prematurely because of poor assess-

ment practices, there is an unnecessary added cost to the government of at least \$150 million annually.⁸

Negative consequences of poor assessment tools include:

- ★ **Turnover.** Employees who are not well-matched to their jobs are more likely to leave or be terminated. Human resource experts estimate turnover costs to range from 50 to 200 percent of the exiting employee's salary.⁹ A study by pharmaceutical company Merck, for instance, found that turnover costs amounted to 1.5 to 2.5 times the exiting employee's salary.¹⁰
- ★ **Lost productivity.** A study of white-collar jobs in the federal government found a positive relationship between the use of robust selection procedures and increased job performance, which tracked a 17 percent increase in productivity.¹¹
- ★ **Absenteeism.** Poorly chosen employees tend to be less engaged in their work and more likely to have higher rates of absenteeism. Each day a federal employee takes leave costs the government an average of \$190 in direct pay and benefits costs alone.¹² That does not even factor in the costs of lost productivity.

II. CURRENT STATE OF FEDERAL EMPLOYEE ASSESSMENT

A. Government Typically Uses Worst Predictors of Job Performance To Select Employees

Many federal agencies now rely on what independent analysts have found to be the least effective method to select among applicants. As illustrated earlier in this report, to hire competitively, most agencies use the "training and expe-

Chart A: *Validity of Common Selection Procedures in the Prediction of Overall Job Performance*¹³

Selection Procedure	Validity Score ¹⁴
WORK SAMPLE TEST	0.54
STRUCTURED EMPLOYMENT INTERVIEW	0.51
GENERAL MENTAL ABILITY TEST	0.51
TRAINING AND EXPERIENCE BEHAVIORAL CONSISTENCY METHOD (<i>requires applicants to describe their major achievements in several job-related areas and evaluates these statements against a scale</i>)	0.45
INTEGRITY TEST (<i>measures conscientiousness, agreeability, emotional stability</i>)	0.41
UNSTRUCTURED EMPLOYMENT INTERVIEW	0.38¹⁵
ASSESSMENT CENTER (<i>various observed activities, role playing, etc.</i>)	0.37
BIODATA MEASURES (<i>questions about past life experiences, such as early life experiences, hobbies, school experiences; measures are first empirically tested for correlation with job performance</i>)	0.35
SITUATIONAL JUDGMENT TEST	0.34¹⁶
REFERENCE CHECKS	0.26
TRAINING AND EXPERIENCE POINT METHOD	0.11

Source: *Psychological Bulletin*, Vol. 124; *Journal of Applied Psychology*, Vol. 86

rience point method” to arrive at “a numerical score for each applicant by rating the education and experience described in the information submitted by the applicant against the evaluation criteria for the position.”¹⁷ Over half of all competitive new hires selected are referred to the selecting officials based on a rating of their training and experience.

This “training and experience point method” is one of the least effective predictors of job performance compared to other methods. For decades, experts have compared the results of different selection procedures to actual job performance, allowing them to determine how well a selection procedure predicts future job performance. A survey of those experts’ findings by Frank Schmidt and John Hunter for the American Psychological Association found that

the training and experience point method fares poorly. Schmidt and Hunter explain that this approach has “low validity,” but its use is widespread in the government:

The point method of evaluating previous training and experience is used mostly in government hiring – at all levels, federal, state, and local. A major reason for its widespread use is that point method procedures are relatively inexpensive to construct and use.... All point method procedures are credentialistic; typically an applicant receives a fixed number of points for (a) each year or month of experience on the same or similar job, (b) each year of relevant schooling (or each course taken), and (c) each relevant training program completed, and so on. There is usually no

*attempt to evaluate past achievements, accomplishments, or job performance; in effect, the procedure assumes that achievement and performance are determined solely by the exposures that are measured.*¹⁸

As the Merit Systems Protection Board recently found, there is “a shortsighted view in many agencies that candidate assessment is a cost rather than an investment,” with “the end result . . . that the quality of the employee selection process continues to vary widely among agencies.”¹⁹ “All too often,” the Board concluded, “agencies hire new employees using assessment methods that appear not to be very good.”²⁰

To determine how well a particular assessment approach predicts job performance, experts have compared the actual job performance of individuals who have been selected for their positions to their scores on an assessment method – such as work sample reviews, structured interviews or the training and experience point method widely used inside the government. As you can see from Chart A, Schmidt and Hunter’s “meta-analysis” of previous studies, the training and experience point method so widely used in government hiring is *the least* effective method for predicting employee performance.

Using the “validity scores” in Chart A as a measure, it appears that a concerted effort by the government to invest in the development of new assessment tools could result in new assessment practices that are up to five times more effective.

B. The ACWA Assessment Tool Is Flawed

In addition to the general weakness of federal assessment programs, there is one assessment tool that deserves special critical scrutiny of its value and the reasons for its continued use.

Hiring problems are particularly severe – and the recommendations of this report correspondingly urgent – with respect to a broad swath of federal jobs that are generally filled by a poorly designed assessment test that the government developed in response to a court order issued in 1981 – the *Luevano* consent decree.²¹ The assessment tool (known as ACWA) that government developed as a temporary response to the court order is not an effective assessment tool and its use continues to distort many federal hiring decisions to this day.

By any measure, the ACWA self-rating exam has not proven to be a successful assessment tool. The General Accounting Office (GAO) recently surveyed federal Human Resource directors and other officials and found the widespread view that ACWA is “cumbersome, bureaucratic and labor intensive” and that it is “not an effective screen to identify quality candidates.” The National Academy of Public Administration similarly found that ACWA “has not met agency needs for providing timely certificates and quality candidates.”²²

Several other concerns about ACWA stand out:

- ★ ***Lack of “face validity.”*** ACWA’s questions do not appear job-related to applicants, giving the test low “face validity.” Managers report that “applicants were ‘turned off’ to federal employment by the lack of relevance of many of the exam questions to the specific jobs for which they were applying.”²³ Chart B contains some examples of questions asked of applicants.
- ★ ***“Faking.”*** ACWA’s exclusive reliance on self-reporting allows applicants who misrepresent their past experiences to rise to the top.

Chart B: Examples of Questions on the ACWA Test

ACWA asks Park Ranger applicants for the National Park Service:

★ *The high school grade I most often received was:*

- A. A
- B. B
- C. C
- D. D or lower
- E. I do not remember

★ *On my present or most recent job, my supervisor rated me as:*

- A. outstanding
- B. above average
- C. average
- D. below average
- E. not employed or received no rating

★ *Have you successfully taught a writing course or worked as a professional journalist, editor or writer?*

Yes No

★ *Have you successfully determined space layouts or assignments for equipment, furniture, displays, etc.?*

Yes No

★ *Have you taught a nonacademic skill to others for pay (for example, horseback riding, playing a musical instrument, martial arts, first aid, etc.)?*

Yes No

ACWA asks Management Analyst applicants for the Navy, in addition to questions on school performance:

★ *Have you successfully done work where your primary responsibility was to help others work out their problems (for example, worked as a therapist)?*

Yes No

★ *Have you written a play, script or novel that was sold, published or produced?*

Yes No

★ *Have you successfully completed an apprenticeship in trade or craft?*

Yes No

★ *Have you successfully done work where you had to coordinate vacation schedules, lunch breaks, etc., with other workers?*

Yes No

★ *Have you successfully learned a hobby or leisure activity requiring extensive study or use of complex directions (for example, constructing stereo or computer systems, building a car, making stained glass objects, etc.)?*

Yes No

★ ***Bias in favor of experience over talent.***

ACWA's focus on past experience creates a bias against recent graduates or others who may have superior capabilities but less experience than other applicants.

Dissatisfaction with ACWA has also produced a host of secondary problems, with many agencies adopting counterproductive workarounds to avoid its use. Indeed, from 1991 to 1998, agencies used ACWA for just one percent of entry-level jobs.²⁴ Instead of hiring competitively into *Luevano* positions using the ACWA test, agencies used alternatives with some undesirable side effects. These include:

★ ***Misuse of Outstanding Scholar hiring.*** The U.S. Merit Systems Protection Board has found that agencies rely improperly and excessively on the Outstanding Scholar hiring authority that was authorized by the *Luevano* decree as an exception to competitive hiring requirements. Over time, agencies have increasingly used this authority as a fast and easy way to fill positions, rather than the limited purpose for which it was created, which was to counteract, on an interim basis, any adverse impact arising from the use of the ACWA test. The expanding use of this non-competitive authority is inconsistent with the statutory merit system principle of fair and open competition.²⁵

From 1993 to 1997, agencies relied more heavily on Outstanding Scholar hiring than on competitive hiring.²⁶ Contrary to the decree's purpose in establishing the program, however, agencies "have not concentrated [its use] on hiring African-Americans and Hispanics."²⁷ In 1997, African Americans accounted for no greater a proportion of Outstanding Scholar hiring than of competitive hiring (11 percent of both), and

Hispanics were more likely to be hired through competitive means (11 percent) than through Outstanding Scholar hiring (7 percent).²⁸ At the same time, there are some agencies such as the Social Security Administration which have done an exceptional job of using the Outstanding Scholar Program to improve diversity.²⁹

★ ***Filling jobs at a higher salary level than needed.***

Since the consent decree only controls hiring at the GS-5 and 7 entry levels, some agencies choose to fill positions covered by the *Luevano* decree at the GS-9 level to avoid having to use the ACWA examination. Many human resource officials "do not ask for the test-based certificates and fill ... their professional and administrative jobs at the GS-9 level."³⁰ To the extent that these positions could be filled effectively at the GS-5 or 7 levels, hiring at the GS-9 level is wasteful.

★ ***Avoiding external hires.*** The ACWA requirement also creates incentives for managers to use internal promotions to fill vacancies because agencies do not need to use ACWA when promoting internal candidates.³¹ From 1991 to 1998, federal agencies filled only one out of five entry-level professional and administrative positions through open competition.³² During that same period, they filled the largest number of these positions – 43 percent – through internal promotions.³³ Over-reliance on internal promotions reduces the odds that the talent pool for these positions will contain a sufficient number of individuals with the potential to be high performers. Research has shown that, as a group, those individuals promoted internally into entry-level professional and administrative jobs do not advance as far as employees hired by other methods.³⁴

Taken together, these workarounds – misuse of the Outstanding Scholar authority, hiring at higher grade levels than needed and excessive reliance on internal promotions – mean that the ACWA test is distorting the way the federal government hires and promotes even when the test is not being broadly administered. Since the

III. ASSESSMENT INNOVATIONS ACROSS GOVERNMENT

The Partnership’s interviews with agency representatives reveal that many agencies continue to use ACWA because they feel that the *Luevano* decree prevents them from implementing alternatives. That is a myth. Agencies have

THE ORIGIN OF THE ACWA ASSESSMENT TEST: THE *LUEVANO* CASE

Luevano was a lawsuit brought by African American and Hispanic job applicants challenging as discriminatory a written assessment test then being used by the government. The federal government settled the suit in 1981, agreeing to a consent decree that required the elimination of this test and the development of new non-discriminatory selection procedures for filling more than 100 entry-level professional and administrative positions when they are filled at the GS-5 or 7 levels, ranging from park rangers to IRS revenue officers. Hiring into positions covered by the *Luevano* suit, therefore, represents a considerable portion of entry-level hiring. Since at least 1993, the number of people hired into *Luevano* positions has steadily increased, rising from just over 4,000 in 1993 to more than 12,000 in 2001.³⁵

In addition to abolishing the old test and committing the government to establishing new assessment procedures, the *Luevano* decree also allowed the government to fill *Luevano* positions through the Outstanding Scholar program. Under this program, federal managers could hire an applicant without opening the job to competition so long as the new hire had a G.P.A. of 3.5 or higher. The government agreed to use this authority temporarily (for two to five years) after it implemented its new assessment procedures to counteract any adverse impact on African-Americans or Hispanics that might result.

To implement the *Luevano* decree, OPM developed the tool that is now the principal test for competitive hiring into positions still covered by the decree: the Administrative Careers with America self-rating schedule (ACWA), which is a 156-part questionnaire asking applicants to report, in multiple-choice format, about selected past experiences and accomplishments.³⁶

ACWA test is the only approved assessment tool developed by the Office of Personnel Management for the covered positions, agencies unwilling to use ACWA engage in counterproductive workarounds to fill those vacancies.

full authority to develop their own alternatives to ACWA, so long as the agency can “validate” any selection procedure that has an adverse impact on the basis of race, national origin or gender. Validation studies determine whether a particular assessment or selection procedure is job-related. Validation is a standard part of

developing applicant assessment and selection procedures. A specialist who develops a selection procedure validates it as a matter of course, and private companies validate to protect themselves against potential liability for unlawful discrimination.

Indeed, several agencies have already developed selection procedures as alternatives to ACWA with no hindrance from the *Luevano* decree. As a human resource official at the Bureau of Citizenship and Immigration Services explained, “we would have done the same thing with or without *Luevano*.”³⁷ The following examples illustrate three different approaches to the development of valid assessment tools that can be used to replace ACWA.

★ ***Bureau of Citizenship and Immigration Services.*** The Bureau of Citizenship and Immigration Services (BCIS) developed and validated new test procedures for several positions covered by the *Luevano* consent decree. Under the new procedures, applicants take a logical reasoning test and fill out a biodata questionnaire. For some positions, applicants also take a writing test or appear before a panel for a structured interview. The BCIS validated these procedures by giving a test battery to current employees and selecting the tests that best correlated with actual job performance and minimized adverse impact.³⁸

★ ***Bureau of Labor Statistics.*** The Bureau of Labor Statistics has also developed a new selection procedure to hire economists covered by the *Luevano* consent decree. Under this procedure, applicants take a written test in two parts, assessing logical and quantitative reasoning. To develop the test, assessment specialists reviewed job materials that economists use at the Bureau of Labor Statistics and worked with a panel of Bureau experts to

identify and target the type of reasoning required for the job.³⁹

★ ***Customs and Border Protection.*** In 1988, Customs implemented a new selection procedure for the position of customs inspector, which is covered by the *Luevano* decree. Now, applicants begin by taking three tests: logical reasoning, quantitative reasoning and an “integrity” test that assesses counterproductive behaviors. Those who pass proceed to a video-taped situational judgment “interview” in which applicants watch a work-related situation unfold on video and then act out their response. After implementing the new selection procedure, Customs surveyed managers and found an “overwhelming” number believed the new hiring program had increased the quality of hires.⁴⁰

Customs validated this test battery by giving the tests to current employees and correlating the results with their performance in training. Customs took careful steps to minimize adverse impact by fine-tuning the tests. To do so, Customs identified test questions that produced significant adverse impact and modified those questions. This procedure substantially reduced adverse impact while still identifying superior candidates.⁴¹

OPM Qualification Standards Misused

In addition to using poor tests, many agencies compound the problem by conflating their assessment process with a set of minimum standards promulgated by the Office of Personnel Management. These standards – known as the OPM Minimum Qualification Standards or “min quals” in the vernacular of federal HR – were never intended to be used as assessment instruments. Instead, they define minimum educational or experiential qualifications for

each federal job. Their basic purpose is to rationalize compensation levels so that persons with similar experience performing similar tasks are compensated in similar fashion. There is considerable debate about whether the standards even accomplish that purpose very well, but there is no dispute that they do not constitute testing tools.⁴²

The minimum standards, for example, have never been tested for validity in predicting job performance. OPM's own Operating Manual makes it clear that the minimum standards "are not designed to rank candidates, identify the best qualified applicants for particular positions, or otherwise substitute for a careful analysis of applicant's knowledge, skills, and abilities."

Notwithstanding their lack of validity, many agencies have integrated the minimum qualification standards into their assessment process by awarding a "passing score" of 70 points, out of 100, to candidates simply for satisfying the minimum qualifications. This widespread phenomenon represents an exceedingly poor assessment practice. Candidates who pass muster under the minimum qualifications should always be sent "back to the starting line" before an agency's assessment tools are utilized.⁴³ To achieve the maximum utility out of an assessment tool, it should score applicants' talents and abilities on a point scale which is "1-100" both in name and in use. To automatically award 70 points solely because an applicant, for example, has a college degree in the relevant field, means that every other gradation of skill, talent and ability must take place in the context of the remaining 30-point scale. This kind of compression significantly erodes the power of any assessment tool to make meaningful distinctions in likely candidate performance.

To take the example of a more familiar testing context, it's a little bit like awarding students

1,120 out of a possible 1,600 points on the Scholastic Aptitude Test (SAT) just for spelling their own names correctly. The resulting selection approach places emphasis on minimal qualifications instead of optimal or high performance qualifications. In this way, the government is not using superior performance as a template for selection and, instead, it is hiring for mediocrity.

For any new assessment tool to be as effective as it can be, agencies must reengineer their hiring processes and implement a scoring strategy which allows for significant variation so that candidates will be able to distinguish themselves.

IV. EFFECTIVE ASSESSMENT AND SELECTION PRACTICES

Improving selection procedures requires three key steps: (1) identifying the skills and competencies needed to do the job right; (2) designing a system to select based on those skills; and (3) validating the system by demonstrating that it measures characteristics reasonably related to job performance.

Identify Skills Needed To Do the Job Right. Federal managers need to work seriously with human resource experts to pinpoint the skills and competencies that distinguish high performers in the job. Managers should focus not simply on experience or expertise in specific areas, but on characteristics such as communication skills, drive, leadership abilities, analytical thinking and the ability to develop other skills.⁴⁴ As a review of practices at leading private sector companies found, "[f]or many of the best companies, the skills most desired as they search for new employees are not those of technical expertise.... [D]esirable candidates also need to have skills in communication, leadership and interpersonal relations."⁴⁵

Target Candidates that Possess Key Skills. Once the necessary skills and competencies have been identified, there are many ways to design a system that accurately selects for those competencies. Effective “tests” not only include pencil and paper exams, but also encompass innovative approaches like job role-plays; interactive, computer-based questions about past behavior; or video/computer simulations.

Validation. Agencies are required by law to “validate” any selection procedure that has an adverse impact on the basis of race, national origin, or gender. Validation studies determine whether a particular assessment or selection procedure is job-related.

Studies have shown that a wide range of selection procedures – from structured interviews to work samples to written tests – can be powerful predictors of successful job performance. These tools are readily available and have been successfully used in the both the private and public sectors. High-performing companies in the private sector generally use some form of testing or structured interviews – or a combination – when selecting employees. A review of hiring practices at 25 high-performing companies (as well as our own review of select federal agencies) found several different types of effective employee selection procedures, which were frequently employed in combination to obtain the best result.

Samples of Effective Hiring Practices

1. Testing

Credit card company Capital One, for example, places great emphasis on “the power of testing, measurement, logic, and analysis.”⁴⁶ Call center applicants complete a multistage process beginning with 17 questions on an automated telephone voice-response system. This system is

credited with reducing the cost-per-hire by 45 percent and increasing the pass rate of new employees by 30 percent.⁴⁷

Leading Internet innovator Cisco greets Internet applicants with a web “profiler” that gives applicants profiles of individuals who would be successful in particular jobs and, if the applicant is still interested, asks a series of questions about their capabilities and behavior. Cisco “hired one person out of every four who submitted profiles, as compared to the 20,000 resumes we were putting into our resume tracking system each month.”⁴⁸

2. Structured Interviews

In a structured interview, a skilled interviewer asks all applicants the same or similar questions, carefully prepared to identify job competencies, and scores the results for comparison. A carefully conducted structured interview by a trained interviewer is substantially more effective than an unstructured interview as a predictor of future job success.⁴⁹

Employers typically develop questions for a structured interview by interviewing current employees to identify “critical incidents” that have emerged on the job in which high performing employees distinguished themselves. The employer then asks applicants how they responded to similar situations in the past, or how they would respond to such situations in the future.⁵⁰

Top performing companies such as Starbucks, Amgen, Intel, Harley-Davidson and Merck use structured interviews as their primary assessment method for selecting employees.⁵¹ The “key element” of these interviews is “a list of well-prepared questions ... designed to reveal a candidate’s competencies.... Usually, each applicant is asked the same set of questions, and replies

are scored or otherwise evaluated so that a record of the interview exists and comparisons among candidates can be made.”⁵²

Another study benchmarking hiring practices at six major companies – Ernst & Young, Hewlett-Packard, IBM, KPMG, Lockheed Martin, and Universal Systems and Technology, Inc – similarly found heavy reliance on structured interviews as an important part of the selection process. “Because interviews are such an important and visible component of the hiring process,” this report found, “many companies insist that their representatives ... complete specific training in behavioral interview techniques.”⁵³

3. Work Samples

Work samples, which ask applicants to engage in a simulation that captures key elements of the job, are strong predictors of job performance and highly regarded by human resource experts. Specific examples include:

★ *IRS Call Center Representatives.* The Internal Revenue Service (IRS) recently overhauled hiring for many of its call centers. Now, those seeking jobs as call center representatives field “live” calls from “taxpayers” for an hour. Turnover rates at centers using this approach dropped by thirty percent.⁵⁴

Applicants for these centers now begin by reviewing background information and then answer calls from assessors posing as taxpayers. These assessors are actually employees of a human resource consulting firm calling from Long Island, New York. One “taxpayer” might ask, for instance, why she has not yet received her refund. A strong candidate will respond helpfully, elicit the additional information needed and tell the taxpayer when to

expect her refund.

The assessors rate the applicants on their ability to organize information, solve problems, listen well, speak clearly and provide effective customer service. Within fifteen minutes of the simulation, the IRS has the applicant’s score and is prepared to make a decision.⁵⁵

★ *IRS Revenue Agents.* Revenue agent applicants similarly demonstrate their skills in practice. Applicants are asked to scrutinize a sample financial statement and then call the “business filer” (a trained assessor) to discuss the problems they identified. The assessors rate the applicants on accounting and interpersonal skills – exactly what hirers need to know.⁵⁶

4. Situational Judgment Tests

“Situational judgment” tests are similar to work samples. These tests present applicants with job related scenarios – in writing, in an interview, or on a video – and ask applicants how they would respond. They can ask applicants for narrative or role-played responses or use a multiple-choice format. A police officer applicant might be asked how she or he would respond to a domestic abuse situation; a teacher applicant might be asked how he or she would respond to an angry parent.

5. Biodata Questionnaires

Based on research showing that some past experiences can predict future job performance, biodata questionnaires seek information about past behavior and achievement – such as leadership, problem-solving and writing – relevant to skills required by the job.⁵⁷

6. Cognitive Ability Tests

Cognitive ability tests are paper and pencil or

individualized assessment measures of an individual's general mental ability or intelligence. The SAT is an example of a cognitive ability test.

According to testing experts, well-constructed cognitive ability tests are among the most effective assessment tools in terms of predicting future job performance.

7. Assessment Centers

Assessment centers bring applicants into a central location for multiple assessment selection procedures targeting a full range of competencies. An assessment center approach has been used successfully in the Transportation Security Administration, which hired almost 60,000 security screeners in 10 months. To select these

employees, TSA used a multiple hurdle approach that assessed a full range of skills required by the job.⁵⁸

In many cases, employers gain the best evaluation of a candidate by combining several different approaches in a test battery or single, blended test or questionnaire. As the Department of Labor advises, “[Employers should] not rely too much on any one test to make decisions. Use the whole-person approach to assessment.”⁵⁹ Screening procedures that are relatively simple but which may not be as predictive of job success as desired should be treated as only one stage in the selection process based on a “successive hurdles” approach to winnowing down the applicant pool to identify those best matched to the job.

CONCLUSION

The federal hiring process is a chain of events that starts with identifying workforce needs and continues through ensuring that newly hired employees are fully oriented to and engaged in the work of their agencies. In the concluding section to this report, the Partnership makes several recommendations for actions that are urgently needed to strengthen what, for far too many federal organizations, is clearly the weak link in that chain.

IV. RECOMMENDATIONS

The Administration should:

1. **Raise awareness** of the strategic importance of effective employee assessment and selection practices by specifically including this as a “critical success factor” in the periodic evaluation of each agency’s progress on human capital under the President’s Management Agenda.

A “performance indicator” for each agency on this factor would include not only a measurement of how quickly it hires but also a measurement of the quality of its hires tracked over time. A separate “performance indicator” would specifically evaluate the soundness and validity of each agency’s assessment and selection practices.

2. Work with agency leaders to *create incentives* for agency managers and agency Chief Human Capital Officers (CHCOs) to work collaboratively and to make the commitment of time and resources required to develop and implement effective employee assessment and selection techniques.

At a minimum, it should be made clear that effective employee recruitment, assessment and selection is a basic management responsibility – with the active assistance of the CHCO and human resources staff – and that managers will be held accountable in a tangible way for demonstrating successful outcomes based on sound practices.

Federal departments and agencies should:

3. Ensure that the need to develop effective applicant and selection practices is *included in agency workforce plans*.

Not only should workforce planning efforts identify the talent needs of each agency in the future and any current talent gaps, but those workforce plans should also specifically identify the most promising sources for that talent and the assessment and selection process or processes that will be used to identify the best candidates for the jobs to be filled.

4. *Expand their use of student internship programs* and the federal career intern appointment authority where feasible.

One of the best assessment tools is the actual observation of the individual on the job for an extended period of time. Student intern programs, such as the Student Career Experience Program (SCEP), provide such an assessment opportunity and allow agencies to offer the most promising candidates permanent appointments without further competition. They also provide an excellent opportunity to address diversity needs. The Partnership issued a report in July 2002 highlighting the benefits of government internships both to students and federal agencies.⁶⁰

Specifically, the federal career intern program allows agencies to use streamlined procedures to hire promising candidates to two-year “excepted appointments” and – at the option of the agency – to convert successful candidates to permanent appointments. Unless earlier removal action has been taken, unsuccessful candidates will simply have their appointments expire at the end of two years.

The Partnership’s report recommended that federal agencies include student employment programs in their strategic workforce planning efforts as a talent pool for future hiring. The report found that federal agencies only converted 17 percent of SCEP participants into permanent federal employees in 2001. That number has since grown to 19 percent.⁶¹

5. ***Hold managers accountable*** for making use of the one-year probationary period for new permanent hires as an extension of the assessment process.

No hiring system will ever be perfect, and there will inevitably be bad hires. The federal probationary period offers an opportunity for an additional screening to make sure that the right hiring decision has been made. Managers need to place greater emphasis on talent assessment during this period to help correct poor hiring decisions.

OPM Should:

6. Work with federal agencies to ***identify the government's greatest hiring needs*** over the next three to five years and actively promote the development of future assessment and selection procedures for relevant job “families” or occupations.

Such a targeted approach could easily yield improved assessment and selection procedures for 60 to 70 percent of the government’s future hiring needs. OPM should be sure to include in a meaningful way interested stakeholders, such as veterans groups, federal employee representatives, “good government” groups and the civil rights community.

7. ***Sponsor pilot projects*** that offer funding for the development or “field testing” of new or improved assessment and selection procedures.

Part of the difficulty facing some federal agencies that would like to improve their assessment and selection procedures is that the development and validation of a “state of the art” assessment tool can be quite costly. In addition, while agencies were delegated greater HR authority and responsibility in the early 1990s, they were also asked to significantly reduce the size of their HR staffs as part of a goal to reduce the proportion of “overhead” positions in each agency. Consequently, many federal agencies do not have the expertise or the resources needed to buy the expertise to develop better assessment tools.

Finally, there is a persuasive argument to be made that development of better assessment tools – particularly for common occupations – is most cost effectively managed on a centralized basis. Use of those tools, of course, can still be managed on a decentralized basis. In this context, OPM could also promote development of initial “screens” that multiple agencies could use as a “first cut” to be complemented by the use of agency-specific assessment approaches once the applicant pool has been winnowed down to a more manageable size.

8. ***Establish a resource center or other sharing mechanism*** that expands the reach of effective assessment and selection procedures or approaches developed by individual agencies or by private experts on behalf of those agencies. OPM should coordinate the development of those tools with third-party vendors for potential application throughout government. At a minimum, OPM should serve as a clearinghouse for information on effective assessment options and strategies that also meet all federal merit system requirements, including provisions for the application of veterans' preference. That information could be useful to both agencies and third-party providers trying to adapt their products to the federal environment.
9. ***Move beyond the Luevano consent decree.*** The problems with the decree and its continued negative impact on federal assessment and selection procedures have been well documented by the Merit Systems Protection Board, the General Accounting Office, and the National Academy of Public Administration. After 23 years under a "temporary" decree, it is time to develop a plan of action with realistic time frames for ending the need for the decree.

Congress Should:

10. ***Fund the development of valid applicant assessment tools*** for delegated use by a variety of federal agencies where this makes the most sense.

A primary reason agencies use the least effective assessment tools is that they are also the least expensive. Congress should make sure that funding will be available for departments and agencies that want to use more effective and expensive assessment tools. In the long term, these investments would pay for themselves many times over.

11. ***Unleash OPM*** by easing restrictions on their ability to assist federal departments and agencies in developing more effective assessment tools. OPM has made fixing the hiring process a priority, which has been vital to the prospects for serious reform, and Congress should make sure that they have the flexibility they need to specifically address the assessment issue.
12. Use its oversight authority to ***monitor progress*** in the executive branch's efforts to make needed improvements to its applicant assessment and selection practices.

Endnotes

- 1 For example, “Mid-Career Hiring in the Federal Government: A Strategy for Change,” February 22, 2002; “Tapping America’s Potential: Expanding Student Employment and Internship Opportunities in the Federal Government,” July 10, 2002; “Homeland Insecurity: Building the Expertise to Defend America from Bioterrorism, 2003.”
- 2 In FY 2001 there were 94,161 permanent, full-time, new federal hires; in FY 2002, there were 135,978 hires including those for the then newly formed Transportation Security Administration; and in FY 2003 there were 88,293 permanent, full-time, new hires according to the Office of Personnel Management’s Central Personnel Data File (CPDF).
- 3 U.S. Office of Personnel Management, 2002 Federal Human Capital Survey.
- 4 U.S. Merit Systems Protection Board, *The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service*, (August 1999): 5.
- 5 *Ibid.*
- 6 “Recruitment,” *Government Executive Magazine* (May 2003).
- 7 Marc Drizin, *Workforce Engagement and Its Impact on Organizational Performance: The 2004 National Benchmark Study*. Technical White Paper, Performance Assessment Network, Inc.
- 8 This is based on the conservative estimate of 50 percent of annual salary as the replacement cost of a professional or administrative employee. Since there was an average of over 42,000 new hires made annually into full-time, permanent professional and administrative positions over the last three years and the average salary for professional and administrative employees in the federal government was \$72,810 as of December 2003, ten percent turnover at half salary ($4,200 \times \$36,405 = \152.9 million).
- 9 Wayne F. Cascio, *Costing Human Resources: The Financial Impact of Behavior in Organizations 23-25* (4th ed. 2000); U.S. Merit Systems Protection Board, *Assessing Federal Job Seekers*, 8.
- 10 *Ibid.*, 23-25.
- 11 *Ibid.*, 263 (citing F.L. Schmidt et al., *The Economic Impact of Job Selection Methods on Size, Productivity and Payroll Costs of the Federal Workforce: An Empirically Based Demonstration*, 39 *Personnel Psychology* (1986): 1-30).
- 12 OPM, *Work Years and Personnel Costs: Fiscal Year 2001*, 65.
- 13 Table results from Frank L. Schmidt & John E. Hunter, “The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings,” 124 *Psych. Bull.* (1998): 262.
- 14 When the validity number is squared, the resulting number represents the proportion of variability in job performance that the measure explains. A measure with a validity of .20 explains 4 percent of variability in job performance ($.20 \times .20 = .04$).
- 15 Some experts regard this validity coefficient for unstructured interviews as potentially misleading and Schmidt and Hunter acknowledge that the validity would be “undoubtedly lower for carelessly conducted unstructured interviews.”
- 16 Michael A. McDaniel et al., “Use of Situational Judgment Tests to Predict Job Performance: A Clarification of the Literature,” 86 *J. Applied Psych.* (2001): 730.

Applicant Assessment

- 17 U.S. Merit Systems Protection Board, *The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service*, (August 1999): 5.
- 18 See Frank L. Schmidt & John E. Hunter, “The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings,” 124 *Psych. Bull.* 262 (1998).
- 19 U.S. Merit Systems Protection Board, *Assessing Federal Job-Seekers in a Delegated Examining Environment* (2001): vii.
- 20 *Ibid.*, 37.
- 21 A consent decree is a negotiated settlement of a civil case that is then entered as an order of the court so that the court may assert jurisdiction over any alleged breach of the terms of the settlement.
- 22 National Academy of Public Administration (NAPA), *Entry-Level Hiring and Development for the 21st Century: Professional and Administrative Positions* (1999): 13.
- 23 General Accounting Office (GAO), *Human Capital: Opportunities to Improve Executive Agencies’ Hiring Processes* (2003): 19-20.
- 24 NAPA, *Entry-Level Hiring* xiii.
- 25 U.S. Merit Systems Protection Board, *Restoring Merit to Federal Hiring: Why Two Special Hiring Programs Should be Ended* (2000): 15-20.
- 26 *Ibid.*, 7.
- 27 *Ibid.*, 11.
- 28 *Ibid.*, 10.
- 29 Partnership for Public Service, *Case Study – SSA: Recruiting for Careers, Not Jobs* <www.ourpublicservice.org/solutions>.
- 30 U.S. Merit Systems Protection Board, *The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service* (1999).
- 31 5 Code of Federal Regulations, part 335.
- 32 NAPA, *Entry Level Hiring*, 9.
- 33 *Ibid.*
- 34 *Ibid.*, xv.
- 35

Year	1993	1994	1995	1996	1997	1998	1999	2000	2001
HIRED INTO LUEVANO JOBS	4,084	4,534	5,140	5,863	5,966	7,155	7,882	8,809	12,329

U.S. Merit Systems Protection Board, *Facts About Outstanding Scholar Hiring During 2000 and 2001*.

- 36 OPM also developed another instrument, the Administrative Careers with America test, a general mental ability test customized to different job categories, but agencies use that test rarely, if at all. Interview with Ernie Paskey, Director, Competency Assessment Branch, Center for Talent Services, OPM (2003).
- 37 Interview with David Pollack, Research and Development Branch, Office of Human Resources and Development, Department of Homeland Security (2003).
- 38 Interview with David Pollack and Mary Anne Nester, Research and Development Branch, Office of Human Resources and Development, Department of Homeland Security (2003).
- 39 Interview with Magda Colberg, Logos, Inc. (2003); Magda Colberg & M.C. Varon Cobos, *Documentation Report on the Development and Validation of the Economist Test at the Bureau of Labor Statistics* (Jan. 2001) (submitted to the BLS in fulfillment of Contract No. J-9-J-0-0031) (available from author).
- 40 Interview with Patricia Harris Thomas, Director, Personnel Research and Assessment Division, Customs and Border Protection, Department of Homeland Security (2003).
- 41 *Ibid.*
- 42 Interview with Donna Gregory, Former Director, Personnel Research and Development Center, Office of Personnel Management (August 2004).
- 43 *Ibid.*
- 44 Robert Wood & Tim Payne, *Competency-Based Recruitment and Selection* (1998): 27.
- 45 Leon Martel, *High Performers: How the Best Companies Find and Keep Them* (2002): 95.
- 46 *Ibid.*, 111.
- 47 *Ibid.*
- 48 *Ibid.*, 107.
- 49 See Chart A
- 50 Lyle M. Spencer & Singe M. Spencer, *Competence at Work: Models for Superior Performance*, 114-125 .
- 51 Martel, *High Performers*, 112-116.
- 52 *Ibid.*, 113.
- 53 NAPA, *Entry-Level Hiring*, 89.
- 54 Interview with Richard Lilienthal, Personnel Psychologist, Internal Revenue Service (October 2, 2003); see also Brian Friel, "Interview with a Computer," *Government Executive Magazine* (June 15, 2003).
- 55 Interviews with Marilyn Gowing, Vice President for Public Sector Consulting and Services, Aon Consulting (2003).
- 56 Interview with Richard Lilienthal (Oct. 3, 2003).
- 57 Robert M. Guion, *Assessment, Measurement, and Prediction for Personnel Decisions* (1998): 598-606.

Applicant Assessment

- 58 A more detailed description of the historic recruitment effort undertaken to establish the Transportation Security Administration is available in the “Solutions Center” portion of the Partnership for Public Service’s website (www.ourpublicservice.org) under “Recruitment.”
- 59 U.S. Department of Labor, Employment and Testing Administration, *Testing and Assessment: An Employer’s Guide to Good Practices* (1999).
- 60 Partnership for Public Service, *Tapping America’s Potential: Expanding Student Employment and Internship Opportunities in the Federal Government*, (July 2002).
- 61 Office of Personnel Management, Central Personnel Data File, FY2003.



PARTNERSHIP FOR PUBLIC SERVICE

1725 Eye Street NW, Suite 900
Washington DC 20006

202 775 9111 phone
202 775 8883 fax
www.outpublicservice.org

ISBN 0-9762425-0-8



\$9.50