Assessing Federal Job-Seekers in a Delegated Examining Environment

A Report to the President and the Congress of the United States by the U.S. Merit Systems Protection Board

U.S. MERIT SYSTEMS PROTECTION BOARD

Office of the Chairman
1615 M Street, NW

Washington, DC 20419-0002

Chairman

December 2001

The President
President of the Senate
Speaker of the House of Representatives

Dear Sirs:

In accordance with the requirements of 5 U.S.C. 1204(a)(3), I am pleased to submit this Merit Systems Protection Board report "Assessing Federal Job-Seekers in a Delegated Examining Environment."

Today nearly 700 agency organizations with examining authority delegated by the Office of Personnel Management assess applicants seeking their first Federal jobs. These delegated examining units make key decisions such as what assessment tools to use and how to use them. Through these decisions they significantly influence the merit basis of Federal hiring. This report examines how Federal agencies assess job applicants in a largely decentralized environment. It also explores the issue of balancing the need for rigorous assessment with the demand for timely hiring, especially when agencies compete with private sector employers in tight labor markets. The report focuses on three widely used pre-appointment assessment tools and the post-appointment probationary period.

Our report raises questions about how to protect Federal merit hiring in today's decentralized environment. Despite the training and oversight that the Office of Personnel Management provides, staffs of delegated examining units vary widely in capability. In part this is because agencies hold varying views of the importance of pre-appointment assessment. Agencies' assessment capabilities also vary because of differences in the funding that agencies can afford to put into developing and using assessment tools. Negative outcomes of poor employee selections—a likely result of using poor assessment tools—include undermining merit hiring and increasing the cost of Government (e.g. lower productivity, higher turnover, and need for additional training). Our report includes recommendations aimed at preventing these outcomes.

We hope that you will find this report and its recommendations useful as you consider steps to improve the selection of new employees in the Federal civil service.

Respectfully,

Beth S. Slavet

U.S. Merit Systems Protection Board

Beth S. Slavet, Chairman

Barbara J. Sapin, Vice Chairman

Susanne T. Marshall, Member

Office of Policy and Evaluation

Director
John M. Palguta

Project Managers
Harry C. Redd III
Carol A. Hayashida

Project Analysts
Karen K. Gard
Jamie J. Carlyle, Ph.D.

Table of contents

Executive summary . . . vii
Background . . . viii
Findings . . . viii
Conclusions . . . ix
Recommendations . . . x

Introduction and background . . . 1
Introduction . . . 1
Background . . . 1

Study methodology . . . 5

Good selection is good business . . . 7
The argument . . . 7
A business case example . . . 8

Managing in today's environment . . . 11
Changes affecting the environment . . . 11
The managers' balancing act . . . 11
Particular concerns of Government . . . 13

The candidate assessment process . . . 15
Assessing before appointment . . . 15
Assessment after appointment . . . 18

The source for information about applicants . . . 21

Review of the assessment tools . . . 23
Reviewing applications . . . 23
Written or performance tests . . . 24
Private sector use . . . 24
Federal agencies' use . . . 26
Training and experience ratings . . . 27
The best approach . . . 28
Less valid approaches . . . 29
Still other approaches . . . 30
Multiple approaches raise concerns . . . 31
The implications of automated systems . . . 31
Interviews . . . 32
Probationary assessment . . . 35

Conclusions . . . 37

Recommendations . . . 39

Appendix 1: A discussion of selection tool validity . . . 41

Appendix 2: An inventory of written and performance tests used by or available to Federal agencies . . . 43

Executive summary

One of the most important jobs of any Federal manager is deciding whom to hire. Federal managers are as interested as their private sector counterparts in hiring the best employees possible, and have an additional incentive to do so since they work in an environment that by law must be based on merit. While much has been written about the need for effective recruitment efforts in building a quality workforce, how one determines which candidate to hire from among those recruited may be of equal or greater importance. This study examines the approaches most often used to assess job candidates under the Federal merit-based hiring system. It finds that these approaches are not all equally effective. The report concludes with recommendations for improving the ability of Federal agencies to identify those job applicants best matched to the requirements of the jobs being filled and the needs of the Federal service.

The manner in which the Federal Government fulfills the statutory requirement to select from among job candidates "solely on the basis of relative ability, knowledge, and skills," has changed dramatically over the last 20 years. Perhaps the biggest change has been the movement from a largely centralized applicant assessment process operated by the U.S. Office of Personnel Management (OPM) to a largely decentralized approach managed by nearly 700 examining units within the various Federal departments and agencies across the country. The authority for those units to examine job applicants is delegated to them by OPM. These units make many decisions, including which assessment tools or methods to use in evaluating applicants. They do this within a framework of standards and regulations developed by OPM, which itself trains examining unit staffs and periodically reviews their operations.

There are a number of advantages associated with the decentralization and delegation of examining authority to individual Federal agencies, including the ability to tailor the approaches used to the needs of each agency and the potential to adapt more quickly to changes in the available labor pool. As this report details, however, there are also some potential disadvantages that must be taken into account if they are to be avoided or minimized. Among the latter are the lack of specialized expertise in many agencies to develop and maintain valid, effective applicant assessment methodologies and, (in some cases) agencies' lack of sufficient funds to purchase such tools or services elsewhere.

Since there is also a shortsighted view in many agencies that candidate assessment is a cost rather than an investment, the end result is that the quality of the employee selection process continues to vary widely among agencies. Some agencies have

and will undoubtedly commit themselves to the development and maintenance of effective, merit-based employee selection tools and processes. Others, however, will just as assuredly have the desire but not the capacity to act. In the latter instance, not only will the quality of the Federal workforce suffer as a result, but the risk of non-merit factors entering into the selection process will increase.

Background

Two 1996 amendments to the Civil Service Reform Act of 1978 authorized OPM's director to delegate to agencies examining authority for all occupations except administrative law judge (ALJ), and to charge agencies fees for assessing job applicants and performing other staffing services. As a result, except for ALJs, OPM only assesses and refers candidates today on a fee for service basis. Most job applicants are assessed by the 684 delegated examining units (DEUs) to which OPM has formally delegated examining authority (larger departments and agencies typically have multiple examining units). These DEUs select and develop assessment tools, perform the assessments, and refer candidates to selecting officials, with OPM periodically reviewing their activity.

Many factors affect agencies' capabilities to develop and use effective assessment instruments. Among the most important are agency culture, organizational structure, budget, staff expertise, court orders, and the number and kind of jobs an agency fills each year. Agencies that regularly engage in substantial hiring for key occupations are more likely to develop high-quality assessment tools than agencies that hire infrequently and in small numbers. Agencies conducting substantial hiring may also be more inclined—and perhaps more able—to pay for the expertise and infrastructure necessary to develop and apply the better assessment tools. While OPM has professional staff skilled in developing assessment instruments, they work on a cost-reimbursement basis and some agencies report that they cannot afford OPM's assistance.

Findings

- **The process used to assess and rank applicants for Federal jobs has changed significantly over the last 20 years. This has brought benefits but it has also raised concerns about the quality of some hiring decisions in today's environment.**

  The move to a predominantly decentralized applicant assessment system over the last two decades was motivated by a need for a hiring process that was faster, more flexible, and yet still effective in identifying the best candidates for the jobs to be filled. Delegating examining authority to nearly 700 DEUs has contributed to definite improvements in the speed and flexibility of the Federal hiring process. However, the quality of the assessment process varies according to the skills, resources, and commitment of the various examining units and their parent organizations. The quality of the assessment process being used also varies in response to some environmental factors. In recent years, for example, demand has exceeded supply for workers in a number of occupations and labor markets, and employers have struggled just to find enough minimally qualified job applicants to fill their vacancies. This has created a mindset among some managers and organizations that there is no need to assess applicants beyond establishing their basic qualifications for the job. They reason that with so few applicants, little or nothing would be gained by further assessing the qualified applicants.

- **Current Federal agency applicant assessment practices—with some notable exceptions—appear to be most often driven by a desire for a process that is fast and "inexpensive." Although federal agencies also want applicant assessment tools that will help them select the relatively best person for the job, in actual practice the weight given to that criterion too often appears to come in a distant third to speed and cost savings.**

There are three basic approaches to pre-appointment applicant assessment:

1. Written and performance tests;
2. A review of each applicant's training and experience (done manually or via an automated system); and
3. Interviews and reference checks.

Federal managers and HR staffs in the delegated examining units typically use some combination of these three methods to help decide whom to hire. The value of each method, or combination of methods, in predicting the relative success of each applicant on the job varies widely based on the quality of the assessment tools being used and the skills of those using the tools. Written or performance tests tend to have the highest relative predictive validity but are the least often used—in large part because they are also more time and resource intensive.

- **There is also a post-appointment assessment process for most new Federal hires—the probationary period—during which the process of removing employees is relatively straightforward. While removal during probation is an effective last resort in dealing with poor selections, it is also one of the most expensive for the organization and—emotionally and financially—for the employee.**

The importance of the post-appointment probationary period increases as the quality of the pre-appointment assessment process or of the applicant pool (or of both) decreases. In the Federal Government the probationary period is the final step of the examining or assessment process, allowing observation of on-the-job performance. Each probationary separation, however, contributes to increased agency costs. Time and resources spent on the initial recruitment and selection process will have to be repeated and a low return will be realized on the salary and time spent on each probationary employee who has to be removed. Finally, a high financial and emotional cost is often paid by the employee who fails to satisfactorily complete the probationary period. While some probationary separations are probably inevitable, relying on the probationary period to catch the increased number of poor selections likely to be caused by poor or minimal pre-appointment assessment, almost certainly will increase the costs of Government.

Clearly, the probationary period is a necessary assessment tool, and one that is quite valuable in helping an organization cut its losses when a selected applicant proves to be a poor match for the job to be done. The probationary period, however, should never be seen as the primary assessment device, or as a surrogate for other assessment processes. The goal—and one in keeping with the requirements of the statutory merit system principles—should always be to use the best pre-appointment assessment process practicable.

Conclusions

Good applicant assessment is both an art and a science. If it is done well, it is a highly valuable asset for any organization. It is also a necessary element for the existence of a non-partisan, merit-based, Federal workforce. In retrospect, the weak link in the otherwise sound decision to delegate examining authority to Federal agencies has been the difference between expectations and reality with respect to each department's and agency's development and maintenance of high-quality applicant assessment processes. Some have and use high-quality processes; others do not. Those lacking the resources to meet the expectations either use less valid tools or must pay another organization to do the work for them. OPM is a logical organization to which agencies should be able to turn for help in developing valid applicant assessment tools and systems, but it is not funded to provide that assistance on other than a reimbursable basis. Thus, the "have not" agencies face serious difficulties in improving their assessment methods.

The assessment methods covered by this report, commonly used by Federal agencies, are similar to those used by other employers. Each of these methods—whether an interview, a written test, or an evaluation of an individual's training and experience—can vary significantly in its ability to predict future job performance. Whether a method is good or not so good depends on the expertise of the people crafting and subsequently using it. Agencies that treat applicant assessment as a business investment, and that make available the resources to fund it, are far more likely to see a good return on their investment than are agencies that treat assessment as a burden to be tolerated. In addition to being consistent with merit system values, good assessment is good business.

Recommendations

The Director of the Office of Personnel Management should—

1. Measure the gap between the requirement that all agencies use valid assessment tools and the reality that some agencies cannot afford to develop and apply them. Incorporate into OPM's strategic plan a strategy for eliminating that gap.
2. Seek to secure the support and resources necessary to eliminate the gap. For example, propose legislation, request additional funding, publish regulations, provide models, authorize demonstration projects.
3. Devise a method for addressing the costs of developing assessment tools so that all delegated examining units have access to the most valid and practical tools irrespective of their agencies' internal expertise or financial capability to buy such devices. Consider proposing a return to appropriated funding for OPM's centralized development and validation of good candidate assessment tools that agencies could acquire and use at little or no cost.
4. Emphasize and expand OPM's current effort to identify factors (either competencies or knowledge, skills, and abilities) that are important to measure for different occupations at different grade levels and to develop assessment tools appropriate to measure those factors. Disseminate the information and tools to agencies once they are available.
5. Focus greater attention, during oversight reviews of delegated examining units, on how well those units are observing OPM's regulatory requirement to validate the assessment tools they use.
6. Review automated staffing systems used or proposed for use by agencies to ensure that the assessment processes that they incorporate meet the requirements established by law and OPM's regulations governing employment practices.
7. Strongly focus attention on the importance of the probationary period as the final step in assessing new employees.

Agency heads should—

1. Cooperate with and encourage OPM in its efforts to develop and make available valid candidate assessment tools, and ensure that the candidate assessment methods being used or considered by their delegated examining units meet all legal and regulatory requirements.
2. Take a strategic view of the role of candidate assessment, treating candidate assessment methods as a business investment by:
   - Budgeting for the development of the best practicable assessment tools;
   - Providing adequate and adequately trained staff for their delegated examining units;
   - Training managers, supervisors, and human resources staff members in the use of those tools;
   - Holding managers, supervisors, and their supporting HR staffs accountable for properly using the best practicable instruments when assessing candidates for employment; and
   - Holding managers and supervisors accountable for effectively using the post-appoint-

ment probationary period to assess new employees, and for separating nonperformers during the probationary period instead of allowing them to gain employment protections that rightfully are earned only upon successful completion of the probationary period.

3. In the interest of preserving merit as an essential component of the hiring system, focus managers' and supervisors' attention on the need to strike the proper balance between selecting the best job candidates and hiring quickly in a tight labor market.

Introduction and background

Introduction

This is a report about a study that examined assessment methods[1] Federal agencies widely use when hiring individuals into their first Federal jobs. The Merit Systems Protection Board conducted this study in partial fulfillment of its statutory responsibility to study and report on the civil service and other executive branch merit systems.

Today, organizations everywhere face stiff competition for skilled workers. Like their private sector counterparts, Federal managers are competing for a share of the labor pool, often after undergoing years of staff reductions and workplace changes that have affected the skills their organizations need. Frequently the labor pool is shrinking. In such circumstances Federal managers clearly understand the benefits of making good hiring decisions, as well as the costs of making bad ones. Less clear, however, is whether they fully appreciate the importance of the assessment tools they use when deciding among qualified job applicants.

In the last 20 or so years, the choice of assessment devices and who uses them has changed. Federal agencies today have substantial control over the assessment tools that are used to differentiate among the applicants. During this time span, MSPB periodically has reported on significant staffing actions or events. The Board undertook this current study to:

- Summarize changes that have taken place and how those changes have affected assessment methods and their use;
- Identify the assessment devices Federal agencies are now most often using;
- Explore how well those assessment devices contribute to meeting managers' needs; and
- Explore how well assessment tools contribute to meeting the Federal civil service goal of hiring based on merit.

This report uses information from past Board studies as well as new information on the Federal civil service and on private sector hiring practices and procedures.

Background

The changes covered by this report began with the Civil Service Reform Act of 1978. Among its many provisions, that act defined statutory merit system principles and identified prohibited personnel practices to guide the actions of Federal managers and management officials. Two related goals of that act were particularly significant to issues covered by this report: one to decentralize personnel administration in the civil service and the other to increase delegations of authority to agencies to manage their human resources. These goals supported the view that Federal employees and the public they serve would benefit from managers and supervisors, who

[1] We use the terms "methods," "tools," "devices," and "approaches" interchangeably in this report as we discuss and analyze assessment.

are closer to the work being performed, making the decisions concerning the hiring, developing, motivating, rewarding, promoting, disciplining, and firing of employees.

The 1978 reform act abolished the U.S. Civil Service Commission and made the newly created Office of Personnel Management responsible for leading Federal human resource management and for centrally administering some human resource programs. The OPM director was specifically authorized to redelegate to agencies any of the personnel management authorities which the President delegated to the director, with two significant exceptions:[2]

1. The director could not redelegate authority to examine for administrative law judge positions, and
2. The director was prohibited from redelegating authority for competitive examinations for most positions common to Federal agencies except in exceptional cases where economy and efficiency made such delegation necessary and would not weaken merit system principles.

The first restriction remains in effect today; the second was changed in 1996 to permit the OPM director to redelegate to agency heads any of his or her personnel management functions, including competitive examining.[3] Concurrently, another 1996 change in law ushered in the fee-for-service approach that OPM now follows when it conducts examining or other staffing services for agencies.[4]

The practice of agencies examining job applicants was not new, although the past practice was much more controlled. Under the Civil Service Commission, agencies had conducted examining under certain circumstances, but the assessment methods and procedures they used were developed and directed by the Commission. This structure fostered consistency in the candidate assessment methods and processes used to fill a job regardless of who did the examining or where it was conducted. This approach continued until delegated examining by OPM introduced a new and significant dynamic in the 1980s. In an October 1989 report the Board noted:

> * * * in delegated examining, OPM turns over full responsibility to an agency, including the development of examining instruments and the development of internal procedures for handling applications.[5]

With the 1996 changes in law permitting delegation of competitive examining authority, this agency autonomy had to be somewhat reduced. Thus, OPM establishes standards and prescribes the procedures that delegated examining units must follow, trains their staffs, and periodically reviews their actions for procedural correctness and their operations for conformity with OPM's standards and regulations and with the merit system principles. The operations review also aims to ensure that the assessment tools used by the units are appropriate for the jobs being filled and were developed following a careful job analysis. OPM does not, however, prescribe the specific assessment methods that delegated examining units use.

Before the 1996 changes in law, approximately 200 offices across the country exercised delegated examining authority. These offices were limited to examining for jobs that were unique to the serviced agency, or that were above GS-9 and were hired in very small numbers by each of many organizations.

---

[2] Civil Service Reform Act of 1978, Pub. L. No. 95-454, § 1104(a), 92. Stat. 1111, 1120 (1978).

[3] 5 U.S.C §1104(b)(1).

[4] For additional information about the events leading to the 1996 changes, see U.S. Merit Systems Protection Board, "The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service," Government Printing Office, Washington, DC, August 1999, p. 2, and the footnote on that page.

[5] U.S. Merit Systems Protection Board, "Delegation and Decentralization: Personnel Management Simplification Efforts in the Federal Government." Government Printing Office, Washington, DC, October 1989, p. 8.

Following the 1996 changes in law, widespread delegation of competitive examining to agencies occurred, and by May 2001 there were 684 delegated examining units in operation. Most of the growth occurred in the first 2 years after the law changed. These units are in virtually every Federal civil service agency. Together with OPM's employment service centers, delegated examining units conduct all examining for hiring new entrants (except administrative law judges) into the civil service.[6] Collectively, delegated examining units process thousands of employment actions annually; individually they may process a mere handful or hundreds each year. These units are responsible for developing their own assessment devices.

In many cases—usually involving jobs where substantial hiring is being conducted—assessments are conducted through well constructed tests. In these instances the extent to which inferences can be made about future job performance based on performance on the assessment tool (known as the tool's validity) is relatively high. However, assessments for jobs common to many agencies or for jobs filled relatively infrequently or in small numbers, are more likely to be conducted through ratings of training and experience. As we show later, assessments based on training and experience vary widely in their ability to predict future job performance.

Delegated examining units are subject to OPM regulations[7] that govern employment practices, a term that includes assessment methods used when making employment decisions. In a merit-based environment the validity of an assessment device is particularly important. If the assessment device doesn't add substantially to a manager's knowledge of each individual's likelihood of success on the job, it is a useless burden on the process. But if managers believe an assessment method provides useful information and it really does not, then actual harm can be done to our system of selecting on merit, even as we believe we are acting meritoriously.

Another basis for concern is whether different categories of individuals (e.g., men, women, or different races or ethnic groups) perform differently on different assessment tools. Such disparate impact is inconsistent with our merit principle values, and every possible action should be taken to prevent its occurrence.

Since 1978, Federal organizations have also been subject to the requirements of the Uniform Guidelines for Employee Selection Procedures.[8] Adopted jointly by the Equal Employment Opportunity Commission, the Civil Service Commission (later OPM), the Department of Labor, and the Department of Justice, the Uniform Guidelines present a single set of principles guiding Federal practices for using tests and other selection procedures. Together, the Uniform Guidelines and OPM's regulations establish a minimum threshold for Federal employment practices, including use of assessment tools.

[6] They may also conduct examining for competitive (usually promotion) actions involving persons who already are Federal employees when the action is initiated, but those actions are outside the scope of this report.

[7] 5 C.F.R. Part 300, Subpart A—Employment Practices, §§ 300.101–300.103.

[8] 43 FR 38290, August 25, 1978.

Study methodology

To gain information for this report we reviewed the results of our own literature search and several conducted for us by the Corporate Leadership Council[9]; earlier MSPB reports on employee selection and turnover and OPM's oversight activity; data from the Board's Merit Principles Survey 2000; and OPM and General Accounting Office reports that addressed Federal staffing or selection issues. We also conducted interviews or had extensive conversations with OPM officials in Employment Service and the Office of Merit Systems Oversight and Effectiveness; current and former officials from five Federal departments or agencies in the Washington, DC area; and human resources representatives from three large private sector employers in the Washington Metropolitan area.

We also analyzed data contained in OPM's Central Personnel Data File.

Finally, with the cooperation of four Federal Executive Boards (FEBs) in Baltimore, Philadelphia, New York, and Chicago, we conducted a series of focus group meetings with 44 line managers and field office human resources officials in those cities. We administered a survey to the FEB participants and have included information from that survey and their group meetings in this study.[10]

[9] The Corporate Leadership Council is a private business organization that, on a paid subscription basis, provides best practices research and executive education to human resources executives. It is a component of a larger organization known as the Corporate Executive Board.

[10] These meetings involved 44 FEB participants from 29 separate organizational components of 17 departments and independent agencies. Thirty-five participants were line executives or managers; nine were field HR officials. In NewYork an additional 6 FEB members joined the focus group participants in completing a survey for this study, giving us survey responses from 50 FEB participants.

Good selection is good business

The argument

Making good employee selections is a good business practice, leading to high productivity, development of a cohesive work group, reduced employee turnover, and other positive outcomes. Poor selections undermine the conditions leading to these outcomes and ultimately drive up costs while lowering productivity.

The absence of a profit and loss statement in most Federal organizations is an obstacle in quantifying the effects of good or poor employee selections. Nonetheless, Federal agencies are including quantitative measures as part of their workforce planning. One example that offers promise is found in the Internal Revenue Service's overall strategic planning effort, which includes using a balanced measures approach for evaluating its human resources programs. It includes aspects such as candidate pool sufficiency, recruit acceptance rate, retention of new hires, and performance rating after 1 year of employment, and analyzes these measures by employment source. This effort should help IRS improve its hiring practices and give that agency a leg up in meeting its future staffing needs. It is an example of the kind of planning and analysis that Federal agencies need to successfully compete in the marketplace.

There is substantial agreement that poor selection decisions can be costly to an organization and that poor assessment approaches can contribute to poor selection decisions. A mismatch between a job's requirements and what an employee brings to that job can add to both an organization's budget outlays and its human capital costs through:

- Increased training costs to bring the employee up to speed (if it is even possible to do so);
- Excessive managerial or supervisory time to coach or counsel the employee;
- Additional workload on other employees to do or redo the employee's work;
- Reduced organizational productivity and effectiveness; and
- Ultimately, in some cases, loss of the employee and the need to refill the vacant position.

We can get a sense of the cost of hiring the wrong person from table 1, which presents one private sector organization's estimates based on the costs of "wasted salary, benefits, severance pay, headhunter fees, training costs, and hiring time. "[11]

Table 1. Estimated costs of hiring the wrong person

For an entry-level employee	$5,000 to $7,000
For a $20,000/year FTE	$40,000
For a $100,000/year FTE	$300,000

Private sector "average cost per hire" data for 3 recent years[12]—shown in figure 1—further

[11] Corporate Leadership Council, Literature Review, Washington, DC, "Employee Selection Tests," March 1998 (Catalog No. 070-198-213), p. 2.

[12] Ibid., p. 2.

Figure 1. One estimate of the average cost per hire in the private sector in 1991, 1993, and 1995

Source: Corporate Leadership Council, Literature Review, Washington, DC, "Employee Selection Tests, March 1998 (Catalog No. 070-198-213), p.2.

stresses the importance of making a good selection with each hiring opportunity.

Also instructive are the efforts of the Saratoga Institute, which has "developed and tested a standard formula for calculating turnover [cost]. It includes the cost of termination, replacement, vacancy, and learning curve productivity loss. These four variables generally cost a company the equivalent of at least six months of a nonexempt person's pay and benefits and a minimum of one year's worth for a professional or manager."[13] Turnover, of course, is not just a consequence of poor selection. However, when all other factors are equal, turnover is likely to be lower if the employee selection process makes the best possible match between employees and jobs. Good assessment tools help accomplish this.

Determining the costs of poor employee selections and turnover clearly is not a science. In the end, the figures obtained depend on the assumptions used. Regardless of how an organization's hiring and turnover costs compare to any of the figures we have cited, it is evident that such costs are substantial. While organizations obviously cannot avoid all of the costs associated with hiring new employees, high-quality selections can help maximize the benefit obtained from these costs. And, again, good assessment methods are part of that equation.

Finally, in today's increasingly automated workplaces, it's dangerous to discount the importance of making good employee selections. While automation can reduce the number of employees needed, in a workforce with fewer employees the importance of each employee's contribution increases. A small number of poor selections can have a disproportionately large adverse effect on the quality of an organization's performance. The importance of staff costs in an automated environment is captured in the following quote from a report prepared by the Federal Chief Information Officer Council:

> A recent Gartner Group study reports: 'Analysis of a typical large distributed computing project over a five-year cycle indicated technology (e.g., client and server hardware, software, upgrades and maintenance) represented only 15 percent of the total, while labor represented 70 percent of total distributed costs.'[14]

A business case example

The following example demonstrates the business case for developing and using valid selection tools. Border patrol agent is a bread and butter occupation for the INS. This is a law enforcement occupation and all successful applicants must complete a residential training program conducted at the Border Patrol Academy. For many years INS faced a painful problem: an intolerably high (11.3) percent of all newly hired border patrol agents failed one critical component of the training program—Spanish language fluency. By the time each failure was confirmed, the agency had invested considerable time and money in the individual's training. Collectively, each failure's expense to the taxpayer (training costs and salary), the INS staffing needs left unmet because one of every nine new hires failed training solely because of the language component, and the failed candidate's disrupted plans,

[13] Dr. Jac Fitz-ens, "The ROI of Human Capital, " American Management Association, New York, 2000, p. 34.

[14] "ROI and the Value Puzzle," Capital Planning and IT Investment Committee, Federal CIO Council, January 1999, Washington, DC, p. 17.

were judged to be unacceptably high costs. Something had to be done. Something was.

INS devised and implemented an artificial language test that measured job seekers' ability to learn a foreign language. The effects of this new test were remarkable. The new hires' failure rate for Spanish language training was reduced by 76 percent, from the original 11.3 percent to 2.7 percent. During just its first 5 years of use, the artificial language test produced $6.5 million savings in training costs alone, according to Border Patrol Academy staff. In addition, INS benefited from the increased number of agents who completed the training period and entered the job. Knowing more about job candidates obviously helped INS make better selections. This helps explain why that agency has committed significant resources to develop and administer two of the kinds of assessment tools discussed in this report—written tests and structured interviews—for occupations with large-scale hiring, such as border patrol agent. For the INS, the development and use of valid tests is a business investment, not just a cost.

We recognize that relatively few Federal agencies have jobs with residential training programs for new hires, and even fewer have key occupations that require foreign language fluency. However, such requirements and settings simply make it easier to identify poor job-employee matches quickly. By investing in better methods for assessing job candidates, there probably is not an agency that would not realize long-term savings from better individual and organizational performance, lower training and learning curve costs, reduced managerial or supervisory coaching and counseling time, and less turnover of new hires. Valid assessment tools used by trained managers and HR specialists are a key means to achieve that goal.

Managing in today's environment

Changes affecting the environment

Federal managers participating in focus group meetings for this study told us about environmental changes that have affected their role in hiring and raised their awareness of the costs associated with the hiring process. The changes that they emphasized were:

- The increased personal time and attention they as managers must give to recruiting and hiring in today's environment;[15]
- OPM's charging agencies for providing staffing services; and
- Agencies' decisions to operate their own delegated examining units (with the administrative and staff costs such units require).

A change less often mentioned in our groups is agencies' increased use of automated systems to accomplish HR processes. Many agencies have turned to automation to reduce the costs and processing times for a wide range of staffing tasks, such as announcing vacancies, receiving applications, determining basic qualifications of applicants, assessing the qualified candidates, and even preparing the lists of candidates who may be considered for hiring. Automation can speed up these processes substantially and may reduce total processing costs, but automation does not necessarily lead to the best—or even good—assessments of candidates. Agencies' current emphasis on applying automated staffing systems has the potential to improve, maintain, or lower the quality of Federal hiring choices, and should be approached with great caution. The Board has recently initiated a study of automated staffing systems to help determine how they affect merit in Federal hiring.

The managers' balancing act

While most managers agree that it makes good business sense to hire the best person for the job, they often must balance that goal against the imperative to "get the job done now." Thus, they usually give speed of hiring significant weight in the hiring equation. Managers must also consider the marketplace when hiring new employees. The most meritorious hiring process in the world, replete with the best predictors of future job performance, can't help them if the job applicant pool is empty or meager despite their best efforts to make it robust.

Even in an environment that requires due diligence to ensure that hiring is based on relative merit, speed of hiring and job market considerations remain important, a point emphasized by several managers in our focus groups. During a discussion of the desirability of using written or performance tests to assess job candidates, one executive emphasized that "When there are almost no candidates

[15] To compete with their private sector counterparts, Federal managers may find it increasingly necessary to spend proportionately more of their time on recruiting, and may need to accept that recruiting is one of their major job requirements.

for the job, it's overkill to test the few you can find. We have to hire quickly or risk losing even the marginal candidates." This view mirrors one expressed by a private sector business executive quoted in a recent trade magazine article: "To attract more applicants and literally fill seats, some companies are streamlining their hiring practices by getting rid of pre-employment * * * testing."[16] Since private sector businesses hire *and fire* at will,[17] the immediate and potential long-term consequences of such an approach almost certainly would be less lasting for them than for government organizations.

We understand the frustration managers face when confronted with too few or marginally qualified job applicants. However, before they resort to selecting from among shallow or meager applicant pools, they should do everything possible to improve both the number and the quality of applicants in those pools. Planning and cooperation among line managers, supervisors, and their supporting human resources staffs may have a significant payoff here. We recognize, though, that there are situations where such efforts will not make a difference, and that managers may then have to decide whether to hire from a limited applicant pool or do without the needed new employees.

As an agency charged with safeguarding Federal merit system principles, MSPB is reluctant to endorse a hiring approach that is not based on effective pre-employment assessment methods, even to solve problems related to limited applicant pools. Instead, we prefer to support a solution to the conundrum of quality of assessment versus speed of hiring that we have suggested in earlier Board reports,[18] i.e., take advantage of advances in technology to use valid assessment tools more innovatively and more quickly. Particularly for entry-level professional and administrative occupations, we think existing but currently unused written tests can be used in ways that would satisfy both merit and speed considerations. OPM has recently begun delegating administration of two assessment devices with the same name—ACWA[19]—to agency delegated examining units that request it. One of these devices is a written test,[20] and this delegation may open the way to using written tests at different points in the assessment process.

The success of any approach aimed at speedy application of the most valid assessment tools (which often but not always will be written tests) will depend, however, on several factors:

- Showing that the more capable tool is sufficiently better than the currently used alternatives to be worth using in their place;
- Showing that the tool can be used quickly enough to satisfy managers' timeliness concerns; and
- Making the most valid tool available to agencies at a cost they can afford.

---

[16] Bob Calandra, quoting Matt Halpern, in *Test Case*, "Human Resource Executive," December 2000, p. 80.

[17] We recognize that the private sector's freedom to hire and fire at will is not absolute. It is constrained by Federal Equal Employment Opportunity laws, state employment laws, and sometimes by negotiated employment agreements. Nonetheless, fewer due process requirements apply to employment in the private than in the Federal sector.

[18] U.S. Merit Systems Protection Board, "Entering Professional Positions in the Federal Government," Government Printing Office, Washington, DC, March 1994, pp. 29-31, and U.S. Merit Systems Protection Board, "Restoring Merit to Federal Hiring: Why Two Special Hiring Programs Should be Ended," Government Printing Office, Washington, DC, January 2000, pp. 21-25.

[19] ACWA stands for Administrative Careers With America. One version is a currently unused written test with six scoring keys that assesses entry-level individuals for 96 different occupations grouped into 6 groups with shared attributes. A seventh group of 16 occupations requiring specific education is covered under ACWA through assessment of the individuals' education, training, and experience. In June 1990 the ACWA written test replaced an earlier written test known as the PACE, or Professional and Administrative Careers Examination. The second version of ACWA is a self-rating schedule used for the same occupations as the written test. It replaced the ACWA written test after November 1994 and since then has been used by OPM service centers conducting examining for covered occupations.

[20] The written version of the ACWA has not been used since November 1994, when OPM stopped administering it. However, it has remained in OPM's inventory of tests and thus is available for agencies' use under OPM's recent delegation decision.

We recognize, however, that realities such as very low unemployment rates and stiff competition for skilled employees in selected occupations still may argue for a hiring approach based on minimal candidate assessment. In our view, such hiring could be done without compromising the integrity of a merit-based hiring system, but only if managers effectively use the final step of the examining process—the post-appointment probationary period all new employees must undergo—to remove those who prove unable to perform their jobs in a satisfactory manner during that probationary period.

Even judicious reliance on post- instead of pre-appointment assessment to validate hiring decisions is very likely to prove a risky—and costly—option. Such a hiring approach is likely to be easier to use if the jobs involved have limited training requirements. When the jobs being filled have lengthy training requirements, managers may decide to do without new employees instead of selecting marginal ones. This is yet another balancing act that managers must perform. Further, managers reluctant to acknowledge performance deficiencies in employees they themselves selected may not consider reliance on the post-appointment probationary period as the key assessment device to be a viable option.

Still, labor market conditions may lead Federal managers to decide to accept marginal employees. While we prefer even in these circumstances that the best assessment tools be applied, managers and their supporting HR professionals may be persuaded to conduct minimal assessments. In such cases the challenge for managers would be to find ways to develop these minimally qualified employees into successful ones or to separate them if they don't work out. Federal employees' lack of appeal rights under most conditions during the probationary period should be a strong incentive for managers to use this final assessment step effectively. Allowing marginal employees to successfully complete the probationary period (and thus gain appeal rights and other employment protections) before trying to separate them vastly increases the dollar and emotional cost of relying on post-appointment processes to deal with poor selections.

Particular concerns of Government

When Federal managers and their private sector counterparts compete for the best available job applicants, they often use the same kinds of methods to assess job applicants. These methods include interviews, evaluation of training and work experience, and sometimes written or performance tests. Nonetheless, many Federal managers still find themselves at a hiring disadvantage, sometimes because the Government's cumbersome white-collar pay system gives them less flexibility than their private sector counterparts to set or adjust pay rates in response to labor market changes.

Equally often, however, the disadvantage lies in pieces of the hiring process that are particular to the public sector. Both Federal and private sector managers want to hire the best employees but Federal managers operate within a system established and governed by laws and regulations. It is rooted in the principles that:

- All qualified citizens should have opportunity to compete for Federal jobs,
- Selection and advancement will be based solely on merit, and
- The Federal workforce should be representative of the public it serves.

Most of the time, Federal managers hiring new employees from outside the Government must follow a process that includes two requirements established by laws that typically do not apply to private businesses. These requirements are "the rule of three" and veterans preference.[21]

[21] The rule of three applies to hiring for positions in the competitive civil service, but does not apply to positions in the excepted service, which include those for which it is impracticable to conduct examining (e.g., attorney or chaplain) and positions of a confidential or policy-making nature. Hence the "Most of the time" in the first sentence of this paragraph.

The rule of three requires managers to select from among the top three available candidates on the list the examining office gives to them. In a December 1995 report the Board noted that this law has unintended and undesirable consequences,[22] and presented alternatives to its continued use.[23] The hiring choices of managers are also affected by the public policy decision to grant hiring preference to certain veterans of military service and to family members of certain other veterans. Persons eligible for this preference ("preference eligibles") have additional points added to their earned passing scores, and then are listed ahead of persons without preference points whose earned scores are equal to the augmented scores of the preference eligibles. Under certain circumstances preference eligibles are placed ahead of all eligible candidates. The Board's December 1995 report noted that:

> Interaction between the rule of three and the current approach to veterans preference too often produces results that are not in the best interests of managers or job candidates, including candidates with veterans preference.[24]

As the Board has noted in earlier reports,[25] these procedural requirements are not the only way, or even the best way, to ensure that candidates are referred on the basis of merit while ensuring that individuals eligible for veterans preference receive that preference. An alternative approach, which groups eligible candidates into broad categories instead of giving them discrete numerical scores, has been shown to improve the hiring of individuals with veterans preference, and we have supported broadening the application of that approach. We continue to believe that at least the option to score eligible candidates through broad categories instead of numerical scores should be permitted through a change in law. Nonetheless, the requirements of the rule of three and veterans preference continue to represent realities that often limit managers' selection choices.

---

[22] U.S. Merit Systems Protection Board, "The Rule of Three in Federal Hiring: Boon or Bane?," Government Printing Office, Washington, DC, December 1995, transmittal letter (p. i.). These unintended and undesirable consequences include: using random numbers to determine the rank order of candidates with tied scores (essentially using chance to decide who may be referred and in what order); and restricting managers' choices to the top three candidates instead of ensuring them that they can exercise choice in their selection in lieu of having to choose the single top-scored candidate. (see pp. ix-x.)

[23] Ibid., pp. xi-xii. Options suggested by the Board include allowing candidates to be referred in broad categories without specific numerical scores, or allowing managers to select from among an "adequate number of candidates" and leaving the definition of that term to the agency.

[24] Ibid., p. xi.

[25] U.S. Merit System Protection Board, "Entering Professional Positions in the Federal Government," Government Printing Office, Washington, DC, March 1994, pp. 12-13, 59, 62; and "The Rule of Three in Federal Hiring: Boon or Bane?, " December 1995, pp. 5-8, 31-36.

The candidate assessment process

Candidate assessment is normally carried out in the four steps summarized below in table 2, with only the key first step being outside the scope of this report. As the table shows, step 1—done by a delegated examining unit or an OPM service center—involves measuring applicants against the appropriate qualifications standard established or approved by OPM to determine whether they meet the basic qualifications for the job. While the standards for some fast-changing occupations are criticized as being out of date, and others are criticized as being so broadly defined as to be nearly meaningless, uniform qualifications standards allows all examiners to base their initial determination on common criteria. Certain other reviews also usually occur at step 1 but sometimes are conducted later in the process. These include reviews of medical or physical examinations and reviews of background or personal history information that may bear on an individual's eligibility, suitability, or fitness for Federal employment.

Table 2. The four steps in matching federal job applicants to vacancies

Step	Performed by	Assessment tools used
1. Determine whether each applicant is eligible for the job.	OPM Employment Service Center or Delegated Examining Unit.	OPM Qualification Standards.
2. Score candidates and determine their referral order.	OPM Employment Service Center or Delegated Examining Unit.	Vary by assessing office, job, and skill level (grade) being filled.
3. Select from among the referred candidates.	Manager with authority to make selections.	Typically, interviews and reference checks.
4. Retain or separate during probation.	Supervisor or higher manager.	Job performance.

Assessing before appointment

Applicants still under consideration after step 1 have been found basically qualified for the vacant job and are now candidates. Under step 2 they are further assessed to determine their relative qualifications to perform the work of the job. Using assessment tools considered appropriate for the job to be filled, the examiner scores the candidates to determine the order in which they will be referred for selection. It is not unusual for subject matter experts from the hiring organization to take part in or even actually conduct this step, particularly in hiring for jobs in higher grades. Score order ("ranking") is critical to candidates' opportunities to be hired because of its importance to the operation of the rule of three and veterans preference, discussed earlier.

Decentralization and delegation of examining authority have affected step 1 only with respect to who performs the step. The effect on step 2 has been far more profound, allowing agencies to decide what assessment devices to use. The merit system implications of this change potentially are significant.

As we noted earlier, as of May 2001 there were 684 Federal delegated examining units, more than three times the number that existed before the 1996 law authorized widespread delegation of examining.

Each delegated examining unit (and OPM service center) is responsible for threshold assessment decisions: how to assess candidates and what tools to use for that purpose. While delegated examining units do their best, they are not all equal in terms of their expertise or quality of staff. This fact, plus the current delegated examining structure, creates an environment where the variation in choices and application of assessment methods is greater than in the past. Candidate assessment previously was conducted by a much smaller number of offices operated under direct control of, or guidance from, a central personnel authority, the Civil Service Commission (later the Office of Personnel Management).

OPM's oversight of delegated examining units, outlined earlier in this report, is a safeguard against the use of egregiously poor assessment tools or the misuse of good ones. As one OPM official noted, their review also helps ensure consistency within each delegated examining unit. The review provides a disinterested party's review of the relationship between each job being filled and the tool(s) used to assess the job applicants.

Delegated examining has increased the number of offices deciding what assessment tools to use, which in turn has increased the likelihood of greater variations in the quality of those tools. In relatively rare cases, agencies are staffed with people trained in assessment techniques and have developed and validated assessment instruments (often tests) that they use when hiring for their mainstream occupations. However, most agencies assess job candidates through review of their training and experience, an assessment approach whose usefulness in predicting future job performance can vary widely, as we show later in this report. Because agencies vary widely in their expertise in developing assessment tools, it is unlikely that all 684 delegated examining units are equally good or bad at developing assessment tools based on training and experience.

The level of guidance and oversight OPM provides to delegated examining units represents that organization's effort to achieve balance between costs and outcomes (and between the ideal and the practical) while protecting merit-based hiring. Finding the right balance point when there are nearly 700 fulcrums is difficult, to say the least.

We think OPM's guidance and level of oversight with respect to delegated examining are close to adequate for their purpose. A weakness persists with respect to verifying the validity of each assessment method used—something that requires expertise not generally found in, and time not usually available to, the oversight staff. We also think that OPM, with its staff of experts trained in developing valid assessment devices located elsewhere in the agency, could do more to ensure that all Federal agencies—regardless of the size of their workforces or budgets—have access to the best possible assessment instruments. Then the delegated examining units could choose the most practicable one(s) to use. Because OPM's Personnel Resources and Development Center (PRDC, the component with the expert assessment skills) operates on a cost-reimbursement basis, many agencies and delegated examining units are unable to benefit from the expertise available there.

We should mention here that a significant change will be occurring in the assessment tools used at step 2. Traditionally, candidate assessment has been focused on the extent to which individuals possessed the knowledge, skills, abilities, and other factors ("KSAOs" but more often called only "KSAs") required to do the job. Representing criteria determined through job analysis to be important to the job and to distinguish among candidates, KSAs are defined as follows by OPM:[26]

[26] U.S. Office of Personnel Management, "Delegated Examining Operations Handbook," Washington, DC, Inst. 5, October 1999, Glossary.

- Knowledge is a body of information applied directly to the performance of a function;
- Skill is an observable competence to perform a learned psychomotor act; and
- Ability is a competence to perform an observable behavior or a behavior that results in an observable product.

While not defined in OPM's current delegated examining handbook, "other" has traditionally included less easily measured traits such as promptness and honesty.

KSAs are the underpinning of OPM's qualifications standards. This contributes to the fact that those standards typically assess qualifications using measurements expressed in terms of months or years of experience at a particular work level. (They also usually allow education to be substituted for experience.)

Since the beginning of the 1990s, attention increasingly has been focused on measuring "competencies." One source defines competency as "an underlying characteristic of a person which results in effective and/or superior performance in a job."[27] That source also reports that in U.S. corporations the impetus for selecting individuals on the basis of competencies has been to focus on personal characteristics that predict superior performance.[28] In the experience of the authors reporting this information about competencies, they tend to be measured in "frameworks [that] often include a mixture of behaviors, values, tasks, aspirations and personality characteristics."[29]

OPM defines competency as an "observable, measurable pattern of skills, knowledge, abilities, behaviors, and other characteristics that an individual needs to perform work roles or occupational functions successfully."[30] OPM also offers a definition of "competency-based job profile," an item that will be important in Federal assessments in the years ahead and that already is important in agencies piloting the use of competencies for OPM. That definition is:

> A statement of the general and technical competencies required for optimal performance in an occupation or job family. Competencies identified as critical for a job provide a basis for developing applicant assessments and related products. When fully implemented, the profile approach will replace the OPM Qualifications Standards operating manual that currently describes minimum qualifications for Federal jobs.[31]

This definition contemplates significant changes in the process for determining the fit between individuals and jobs once OPM finishes work on developing competency-based job profiles for all occupations or job families. Qualifications standards will disappear, to be replaced by profiles expressed and measured in terms very different from those currently used. However, in this report we do not distinguish between assessments based on competencies and ones based on KSAs. We expect that the sequence of assessment steps won't change although the factors assessed and how they are measured may be different.

Actual selection—step 3 of table 2—is the pre-appointment assessment step least changed by decentralized and delegated examining authority. Federal managers have been responsible for deciding which eligible candidate(s) to hire for as long as there has been a civil service, and the assessment tools they personally use to make that final decision

---

[27] Robert Wood and Tim Payne, Competency Based Recruitment and Selection: A Practical Guide," John Wiley and Sons, New York, 1998, p. 24, (quoting in turn from G. O. Klemp, Jr.'s 1980 report to the National Institute of Education titled "The assessment of occupational competence.")

[28] Ibid., p. 26.

[29] Ibid., p. 27.

[30] U.S. Office of Personnel Management, op. cit., Glossary.

[31] Ibid., Glossary.

do not appear to have changed very much over the years. The most common way managers gain additional information about job candidates is through interviews, which may be either structured or unstructured. (We will discuss these two kinds of interviews later in this report, and additional information will be provided in a forthcoming Board report dealing specifically with interviewing.) Other ways managers gain information about job candidates include review of candidates' employment applications and supplemental materials and interviews with current and former supervisors.

Assessment after appointment

The assessment of newly hired employees does not end with their appointment. Most employers, Federal and private sector alike, require a post-appointment trial period during which new employees have very limited employment protections. In the Federal Government, this trial period is called the probationary period. Except in rare instances where agencies have been given permission to change its length (usually to make it longer), the probationary period for Federal employees lasts 1 year. During this time, new employees are subject to firing with very limited appeal rights. At the end of the period they gain a greater range of due process rights, making adverse actions that might be appropriate in cases of poor performance—including termination, the last resort—much more difficult to accomplish.

The post-appointment probationary period—identified in table 2 as step 4 of the assessment process—has not been affected by decentralization and delegation. It was and remains the responsibility of supervisors or managers.

Based largely on anecdotal evidence, Federal managers have long been viewed as reluctant to remove unsuccessful employees during the probationary period, even though they will never have an opportunity to do so with fewer due process requirements. Factually, this perceived reluctance may not be correct. For an earlier Board study on poor performers[32], we used data from OPM's Central Personnel Data File,[33] or CPDF, to calculate the percentage of permanent, career-conditional accessions in FY1993 who were subsequently identified in the CPDF as being terminated during their probationary period. We calculated the FY 1993 probationary termination (i.e., fired or otherwise involuntary separated) rate to be 4.0 percent. We repeated the calculations for this study, basing our calculations on similar accessions for both FY 1998 and FY 1999. Our calculations for each of these years yielded a higher probationary termination rate of 6.0 percent.

We have no basis for judging whether a probationary termination rate of 6.0 percent means Federal managers adequately use the probationary period although, in truth, the figure was higher than we expected. Indeed, for FY 1999, that figure represented terminations of 2,515 individuals. If good selection methods were used to differentiate among a pool of well-qualified candidates, then a 6-percent separation rate probably means the probationary period was being used effectively. Similarly, if employees "shaped up" after being informed of poor performance, then this separation rate may be appropriate. However, if poor selection methods initially were used, if the quality of the applicant pool was low, or if small numbers of applicants led to hiring virtually every basically qualified candidate, then a higher separation rate might have been expected. What is clear is that not all Federal managers are reluctant as a last step to use the probationary period to remove unsuccessful performers, which is good news.

---

[32] The FY 1993 figure was calculated for but ultimately not reported in the MSPB report "Federal Supervisors and Poor Performers," which was published in July 1999.

[33] This is a computerized data base maintained by OPM that contains information on over a million Federal civilian non-postal employees in the executive branch.

The significance of properly using the probationary period to remove unsuccessful performers is borne out by figures in the Board's annual report for fiscal year 2000.[34] During that time, 296 probationers appealed their termination. Of those appeals, 266 were dismissed for lack of jurisdiction. Of the remainder, 28 were settled and only 2 were adjudicated by the Board. Such small numbers subject to third-party intervention should further encourage managers to use the probationary period to deal with poor performers.

With this background into who performs which steps to match job applicants with Federal jobs, we now turn to a discussion of how job applicants make their interest known and how agencies typically distinguish among qualified applicants.

[34] U.S. Merit Systems Protection Board, "Annual Report, Fiscal Year 2000," Washington DC. The figures in this paragraph were taken from the table on page 26.

The source for information about applicants

Job applicants usually use a resume or other application form to show interest in a job, typically in response to a vacancy announcement which should provide comprehensive information about the job. These applications usually provide information about the job seekers' previous work experience, education, training, and special skills or abilities, licenses or certificates, and interests, all in an effort to show how they are qualified for jobs they are seeking.

For many decades applicants were required to complete a standard government form (SF-171) to apply for a Federal job. Abolished in January 1995, the SF-171 has been replaced by an optional form (OF-612). This optional form is useful to Federal job applicants because it contains prompts for certain required information (e.g., country of citizenship or a veterans' preference claim) that may not routinely be expected on an application for a job in the private sector. Sometimes agencies require applicants to submit resumes or complete special forms to accommodate automated procedures.

As we discuss shortly, regardless of its form the application is usually an information source during at least two stages of the assessment process. Therefore, its accuracy and clear relevance to the job being filled is critical both to the job seeker and the hiring organization.

While the resume or other initial application form is essential to the application and assessment processes, few managers are going to make a hiring decision based solely on information contained in this source document. In this respect, Federal organizations appear to be no different from private sector organizations. Employers want to know more than what the application tells them. That is why assessment tools, some of which are discussed next in this report, exist.

Review of the assessment tools

In this section we discuss the tools Federal agencies most often use when assessing job applicants after receiving their resumes or other job applications. Table 3 identifies the four most commonly used assessment tools. It also identifies the step or steps in the assessment process when each of these tools is likely to be used.

Table 3. The four pre-employment assessment tools that Federal agencies most often use

Tool	Used by	Purpose
1. Review of resume or other application form.	A. Examining office.	1. Match qualifications to job requirements. 2. Score the qualified applicants and determine their relative order for referral.
	B. Managers.	Assess referred candidates for selection.
2. Written or performance test.	Examining office.	1. Match qualifications to job requirements. 2. Score the qualified applicants and determine their relative order for referral.
3. Training, experience, and/or biodata rating.	Examining office.	1. Match qualifications to job requirements. 2. Score the qualified applicants and determine their relative order for referral.
4. Interview.	A. Examining office.	Score the qualified applicants and determine their relative order for referral.
	B. Managers.	Assess referred candidates for selection.

Reviewing applications

Regardless of the format of the application submitted to show interest in a job vacancy, the information it contains always should be sufficient for a reviewer in the receiving organization to carry out the first assessment step—determining whether the individual is basically qualified for the job being sought.

For entry-level applicants with little or no previous work experience, this may be the only purpose the resume or other application form can serve. For experienced individuals applying for more advanced jobs, the initial application may serve both to determine their eligibility for the job and to make distinctions among them for referral purposes. As we noted earlier, agencies using automated procedures to fill jobs may require applicants to complete specified application forms or supplemental forms. These forms help ensure that applicants provide specific information organized in a specific manner. This helps ensure that all applicants provide information needed to assess them, and also simplifies review of the applicants' information. Often such forms are designed to permit machine reading and automated scoring of the applications. As we noted earlier in this report, automation certainly can increase processing speed and may reduce costs, but it does not necessarily lead to good assessments. Thus, automated assessment systems should be carefully examined before they are applied.

The single greatest difficulty with relying on resumes or other job applications to assess job candidates is identified in the accompanying cartoon.

© 2000 Thaves. Reprinted with permission. Newspaper dist. by NEA, Inc.

The value of a resume or other initial application form should lie in the clear presentation of its content. However, its substance may be blurred—and qualifications overstated—by applicants eager to present themselves and their qualifications in the most positive light. Conversely, if applicants don't have a clear picture of a job's requirements, they may understate their qualifications or place emphasis on the wrong factors.

Written or performance tests

Well-constructed general mental ability (also known as cognitive ability) tests are among the best assessment devices in terms of predicting future job performance. According to testing experts, "Work sample measures are slightly more valid but are much more costly and can be used only with applicants who already know the job or have been trained for the occupation or job."[35] Research has shown that general mental ability tests predict as much as 26 percent of the variability among different employees' performance on the job.[36] Twenty-six percent may not sound very impressive, but in the business of predicting future job performance using instruments other than actual performance on the job, it is in fact quite high. As we shall see later, some of the more widely used assessment tools are far less predictive.

Here's what an article written by two experts has to say about general mental ability tests:

> In the pantheon of * * * personnel measures, general mental ability (also called general cognitive ability and general intelligence) occupies a special place for several reasons. First, of all procedures that can be used for all jobs, whether entry level or advanced, it has the highest validity and lowest application cost. * * * Second, research evidence for the validity of general mental ability measures for predicting job performance is stronger than that for any other method. * * * Third, general mental ability has been shown to be the best available predictor of job related learning. * * * Fourth, the theoretical foundation for general mental ability is stronger than for any other personnel measure.[37]

Private sector use

The example set by many of America's top businesses demonstrates that there is an important place for written and performance tests when hir-

[35] Frank L. Schmidt and John E. Hunter, "The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings," *Psychological Bulletin,* The American Psychological Association, Inc., Vol. 124, No. 2, September 1998, p. 264.

[36] Ibid. This 26-percent figure is based on a validity of .51, as reported in the article's table 1, p. 265. See appendix 1 of this report for an explanation of how this and subsequent figures used in this report to express predictive variability in performance on the job were determined.

[37] Ibid., p. 264.

ing staff. The American Management Association (AMA) annually conducts and reports on a survey of workplace testing. The survey is sent to AMA member and client organizations, which are largely among the top 5 percent of U.S. businesses in terms of annual sales and total employees, and who collectively employ one-fourth of the U.S. workforce.[38] Table 4, which summarizes information from the reports released in 2000 and 2001, covers three kinds of tests administered over a 4-year period only to job applicants as opposed to current employees seeking internal placement actions. The first test, Cognitive ability—spatial, verbal, math, is most similar to the entry-level written tests used by or available to Federal agencies. The third test, physical simulations of job tasks, may be similar to some of the tests Federal agencies use for skills testing.

Table 4. Percent of private sector employers reporting that they administered any of three kinds of tests to job applicants during 1998-2001

Kind of test administered	1998 use	1999 use	2000 use	2001 use
Cognitive ability—spatial, verbal, math	32.4	20.3	20.7	17.7
Managerial assessments	22.9	11.8	14.3	11.9
Physical simulation of job tasks	18.9	10.7	12.1	8.9

Source: American Management Association, "2000 AMA Survey on Workplace Testing: Basic Skills, Job Skills, and Psychological Measurement, Summary of Key Findings," New York, 2000, p. 4 (for 1998, 1999 and 2000 data); American Management Association, "2001 AMA Survey on Workplace Testing: Basic Skills, Job Skills, and Psychological Measurement, Summary of Key Findings," New York, 2001, p. 2 (for 2001 data).

In both its 2000 and 2001 survey reports, AMA commented on the declining use of all forms of psychological measurement, something that is very evident for the three kinds of tests covered by table 5. The 2000 report offered the following explanation:

> [T]hese findings would support the following scenario: faced with the need to expand their workforces in a growing economy, skill shortages were so severe that companies made less use of any form of measurement in their need to fill open positions.[39]

Similarly, the 2001 report said:

> The share [of responding firms employing one or more forms of psychological measurement] has dropped precipitously over the past three years, from 48% in 1998 to 33% in 1999 and 2000, to 29% in 2001. Skill shortages may explain this; when mission-critical positions go unfilled in a tight labor market, companies may well take shortcuts in the application and evaluation process.[40]

Unlike Federal agencies, the private sector businesses can hire and *fire* at will. Two aspects of that environment make it relatively easy for their managers to risk lowering productivity by using less stringent assessment tools:

- The private sector environment has far fewer (and may have no) procedural requirements governing the termination of employees who don't work out, and
- Making significant employment decisions, including withholding pay increases and terminating employees, are within the normal scope of private sector managers' duties.

Nonetheless, it is important to recognize that AMA concludes that reduced use of more stringent selection methods has been in response to specific labor market conditions, and not for philosophical rea-

[38] American Management Association, "2001 AMA Survey on Workplace Testing: Basic Skills, Job Skills, and Psychological Measurement, Summary of Key Findings," New York, NY, 2001, unnumbered page titled "About this survey".

[39] American Management Association, "2000 AMA Survey on Workplace Testing: Basic Skills, Job Skills, and Psychological Measurement, Summary of Key Findings," New York, NY, 2000, p. 4.

[40] American Management Association, "2001 AMA Survey on Workplace Testing: Basic Skills, Job Skills, and Psychological Measurement, Summary of Key Findings," New York, NY, 2001, p. 2.

sons. Although private sector businesses may not stress the importance of merit-based hiring to the extent that Federal agencies do, they are driven by a need to make a profit, and having the best possible employees improves the odds of doing that. The AMA reports highlight the business imperative of having flexibility when faced with fluctuating job markets. Being able to vary applicant assessment strategies in the face of labor pool changes increases an organization's competitiveness, a point articulated by participants in our focus groups.

Federal agencies' use

Although use of tests has declined in Federal hiring since OPM stopped administering the written test known as the ACWA[41] (Administrative Careers With America) in November 1994, written tests have not totally disappeared. In fact, the recent decision by OPM to grant delegated examining units authority to administer both versions of the ACWA upon their request could lead to a resumption in use of its written test version.

Today, instead of relying on a single broad-spectrum test covering many occupations, a number of Federal agencies have developed tests specifically applicable to one or more of their key or mainstream occupations. In appendix 2 we list the tests that OPM identified for us as currently available for use by Federal agencies. It is an extensive list containing a mix of agency-specific and Governmentwide tests. For professional and administrative occupations, most of the tests are for entry-level grades. For clerical and similar white-collar support jobs, many of the tests measure skills. The blue-collar apprentice and helper tests measure learning ability. It's clear from appendix 2 that many Federal jobs are subject to some form of test. We note that some of the tests listed in the appendix are in the form of structured interviews, indicating that the definition for the term "test" is fluid.

Why, then, are tests not used more often in Federal hiring? As we noted in our January 2000 report:[42]

- Many managers associate written tests with the centralized hiring conducted in the past, and many have memories of unsatisfactory experiences with centralized examining during periods of significant hiring.
- Agencies are predisposed against written tests because of their possible adverse impact on one or more classes of applicants.
- For entry-level professional and administrative jobs in more than 100 occupations, managers have access to court-approved noncompetitive hiring processes that give them substantial control over hiring without requiring any form of assessment beyond determining that each applicant meets the qualifications requirements for the job being filled.[43]

In the Board's Merit Principles Survey 2000 we asked 311 supervisors who had filled entry-level professional and administrative vacancies within the past 2 years about 10 possible information sources. Specifically, we wanted to know: "In your opinion, to what extent is each of these sources of information a good predictor of professional/administrative job performance?" Respondents showed a clear disdain for written tests, giving "Written test scores" the second-lowest vote of confidence after "Reputation of educational institution attended" (see table 5). In light of the AMA survey information cited earlier, the disdain of Federal supervisors and managers toward written tests seems unwarranted.

---

[41] For more information on the ACWA, see footnotes 19 and 20.

[42] U.S. Merit Systems Protection Board, "Restoring Merit to Federal Hiring: Why Two Special Hiring Programs Should Be Ended," Washington, DC, Government Printing Office, January 2000, p. 22.

[43] These processes have been in effect since 1981, and most Federal agencies are very conversant with their provisions.

Table 5. Federal supervisors' opinions of 10 information sources as good predictors of professional/administrative job performance

Information source	Percent saying "to a great extent"	Percent saying "to a moderate extent"	Percent saying "to a minimal extent" or "to no extent"
a. Written test scores	6	30	64
b. Prior work experience	57	39	4
c. Level of education achieved	28	53	19
d. Reputation of educational institution attended	4	20	77
e. Major field of study	28	34	38
f. College grade point average	13	34	53
g. Interviews	61	32	7
h. Quality of the application or resume	34	45	21
i. Reference checks	36	42	22
j. Personal recommendations	22	37	41

Source: MSPB "Merit Principles Survey 2000," question 77B.

In our January 2000 report we recommended using written tests later in the assessment process, to provide additional information about candidates, so they are not a hurdle that screens candidates in or out.[44] We believe managers would view tests more favorably under this approach and again encourage OPM and agencies to explore this option.

While a well-designed written test can help reduce unsuccessful hiring, developing one is an expensive process requiring time and expertise. Without some minimum number of jobs to be filled, developing a written test may also be impractical. Faced with tight budgets, reduced HR staffs, and in many cases calls to further shrink their workforce numbers, it is not surprising that most agencies are either unwilling or unable to support the development of such instruments. This unwillingness or inability contributes to agencies' using the other assessment methods discussed in this report, and may also contribute to significant unevenness in their ability to develop those other methods. The likely net effect is a great variation in the quality of the assessment tools Federal agencies are using to identify the best job candidates. The variation has been confirmed in OPM oversight reports.

Training and experience ratings

Many, if not most, Federal jobs are filled after assessing candidates on the basis of their training and experience, the two factors (or eduation as their surrogate) that OPM's qualifications standards are built around. Typically, this assessment process involves developing a plan whereby points are assigned for certain kinds of work experiences or training that demonstrably are related to aspects of the work of the vacant position. Pragmatically, it is difficult to fault this process, one that has long been used by Federal agencies both when hiring new employees as well as when promoting internal candidates. Many individuals initially selected through this process, and many others subsequently promoted through it, are enjoying or have enjoyed successful careers.

One reason assessing training and experience is widely practiced is that Federal managers and supervisors put a premium on applicants' job-

[44] Ibid., pp. 23–25.

related experience when hiring. When we surveyed 2,600 Federal supervisors,[45] 82 percent said that job-related experience was very important (the highest rating possible on the survey) when they were hiring people from outside the Government. The next highest rated factor, documented qualifications, was cited as very important by 65 percent of the respondents. Information about both, of course, should appear on the application form and should be measurable through appropriate assessment methods, but one thing assessment experts advise is that candidate assessments through review of training and experience "tend to be more appropriate for jobs requiring prior preparation, and less appropriate for entry-level jobs."[46] In addition, there are several ways to carry out assessments of individuals' training and experience, and they vary widely in their ability to predict future job performance.

The best approach

The most valid approach to assessing training and experience involves the greatest amount of work on the part of both those being assessed and those doing the assessment. Called "the behavioral consistency method," this involves the use of supplemental application forms addressing important job requirements. Applicants' responses to these supplemental questionnaires are rated against anchored rating scales developed by job experts using actual examples of achievements for each requirement being assessed.[47] This approach has been found to predict as much as 20 percent of the variability in employees' performance on the job,[48] which is relatively high. Developing an assessment tool that allows this level of inference about future job performance is both expensive and time consuming, as the following quote explains:

> In terms of time and cost, the behavioral consistency procedure is nearly as time consuming and costly to construct as locally constructed job knowledge tests. Considerable work is required to construct procedure and the scoring systems; applying the scoring procedure to applicant responses is also more time consuming than scoring most job knowledge tests and other tests with clear right and wrong answers. However, especially for higher level jobs, the behavioral consistency method [for rating training and experience] may well be worth the cost and effort.[49]

If we could be sure that all 684 delegated examining units consistently used this method when developing and applying scales for rating applicants' training and experience, we could be sure that their applicant assessments are good predictors of future job performance. However, as the following excerpts from five different OPM oversight reports show, delegated examining units are not equal with respect to their competence or outcomes.

> * * * the definitions of the ranking levels were vague and did not provide adequate distinctions among the levels. Also, employees expressed concern that ranking and selective factors in vacancy announcements did not reflect current work conditions. * * * In addition, we could not reconstruct rating criteria because case files did not document job analysis that would support a crediting plan. Documenting the job analysis would help HR staff develop more accurate crediting plans.
>
> In some staffing cases at both the * * * and * * * offices, we found the quality of crediting

[45] See "Filling Jobs—What do Supervisors Look For?," *Issues of Merit,* Office of Policy and Evaluation, U.S. Merit Systems Protection Board, May 1998, p. 5.

[46] "Recent Innovations in Public Sector Assessment," Assessment Council of the International Personnel Management Association, Personnel Assessment Monographs, Vol. 2, No. 2, April 1990, p. 34.

[47] Ibid., p. 31.

[48] Ibid., p. 32; Schmidt and Hunter, op. cit., Table 1, p. 265 (reflecting a validity of .45).

[49] Schmidt and Hunter, op. cit., p. 268.

> plans used to determine the best qualified candidates needs improvement. * * * [W]e found that most * * * offices do not develop the necessary crediting plans for multiple grade levels at which jobs are announced.
>
> * * * we identified the need for significant improvements in: * * * the quality of rating plans; the clarity of application procedures, and documentation of examining procedures.
>
> * * * we found some crediting plans that inadequately described knowledges, skills, and abilities required to do the work, thus producing inaccurate lists of eligibles. We also found other crediting plans that were too restrictive, improperly limiting competition. Both of these types of errors could reduce the ability of managers to select from among the best qualified, and erode merit system principles.
>
> The DEU studies the quality of new hires on an ongoing basis, obtaining feedback from supervisors to correct any weaknesses in * * * appointees. * * * Managers, program officials, and personnelists were involved in developing a new standardized [recruiting and examining] procedure. This included analysis of job requirements and preparation of qualification presecreening and selection interview questions. [The DEU plans to track hires for five years to assess whether this procedure identifies selection criteria that result in high performers.]

Based on sheer numbers, this is not surprising. Despite efforts by OPM and agencies to ensure the best possible assessment outcomes, it would be very unusual for nearly 700 discrete delegated examining units to develop and use equally high-quality training and experience assessment tools.

In both the handbook for delegated examining units and the training it provides to those units, OPM stresses the importance of using rating scales that are developed from good job analysis that gets to the essence of the jobs to be filled. It also checks for this during periodic oversight reviews. OPM also stresses in the delegated examining handbook that assessment approaches must be valid; be practical with respect to the job being filled, the labor market, and the capability of the agency staff; and conform to laws and regulations affecting the staffing process, including the Uniform Guidelines.[50]

Further, agencies almost always rely on people with expert knowledge of the job being filled to develop the assessment criteria and scoring keys. In addition, they often rely on these subject matter experts to conduct the assessments. These actions help ensure that each training and experience assessment tool is valid.

Nonetheless, agencies still vary widely in terms of factors important to developing and applying assessment tools. These factors include size, discretionary funding, staff capability and expertise, and organizational culture. Some agencies—particularly those that see the effort as an investment with a high and continuing return to them—are somehow better able or more willing to "go the extra mile" to develop and apply high-quality assessment instruments.

Although many units are probably using the best training and experience assessment tools they are capable of developing, if the tools being used do not meet the more strenuous test of the behavioral consistency method, what else might they use and how useful are they?

Less valid approaches

One less valid approach to assessing training and experience is called the point method. One research source[51] has reported that this approach is used by local, state, and federal government organizations (but we don't believe that Federal agencies use it to

[50] U.S. Office of Personnel Management, "Delegated Examining Operations Handbook," §2.3, p. II-2.

[51] Assessment Council of the International Personnel Management Association, op. cit., pp. 30, 34.

distinguish among applicants already deemed qualified). The researchers explain that this method gives candidates points based on number of years of experience and training and the relevance of each to the job being sought.[52] Although the point method may offer a tempting choice to agencies that view assessment tools only as costly impediments to speedy hiring, as the researchers describe it, it is more like the step 1 process used to measure applicants against OPM's qualifications standards.

A second less valid approach to assessing training and experience, called the modified point method, is likely to be used by Federal delegated examining units in place of the behavioral consistency method. The modified point method starts with a job analysis that identifies important job requirements and specific tasks applicants might have performed in the past to show that they meet those requirements. Scoring procedures are then developed to credit the information obtained.

While better than the point method (which does not require a job analysis and uses a very mechanical scoring approach) this modified approach to rating training and experience is far less rigorous than the behavioral consistency method. Consequently, it is correspondingly less useful to its users. Studies indicate that this modified point method approach accounts for only about 1 percent of the variability in employees' performance on the job.[53] If applied to individuals seeking entry-level jobs, where work experience is often minimal or nil, its value as a predictor of future job performance would be even less. Because this modified point method begins with a job analysis, it has a surface resemblance to the behavioral consistency method (essentially it has some of the latter method's form without its substance). This gives it the appearance of being a better predictor of future performance than it actually is.

We do not know that any delegated examining units have developed their training and experience assessment instruments using this approach but, after reviewing OPM oversight reports containing findings such as those quoted earlier, we are concerned that some may have done so. Our concern rises from: this method's surface similarity to the more rigorous behavioral consistency method; the fact that this method is developed following techniques required by OPM's delegated examining handbook; and problems with delegated examining units' assessment methods that staff in OPM's Office of Merit Systems Oversight acknowledged in a meeting with Board staff. The wide variation in agencies' organizational cultures and resource commitments to candidate assessment adds to the plausibility that some delegated examining units have employed this approach, probably without even realizing they are not receiving much benefit from their candidate assessment efforts.

Still other approaches

Another procedural variation in assessing training and experience is the self-rating method. In this approach applicants self-rate their skill levels on a series of factors pertinent to the job being sought. This is usually done on a form designed for subsequent optical scanning or by direct keyboard input. Although potentially subject to applicant overstatement (and hence misrepresentation), well-designed self-assessment tools include provisions to verify applicant claims. When properly constructed and applied to occupations with specific task requirements, self-rating schemes have been found to account for as much as 4.5 percent of the variability of employees' performance on the job.[54] This assessment approach is attractive to employers because of its potential cost savings, but its predictive value clearly does not make it one of the best assessment methods.

---

[52] Ibid., p. 30.

[53] Ibid., p. 32; Schmidt and Hunter, op. cit., p. 265 (reflecting a validity of .11).

[54] Assessment Council of the International Personnel Management Association, op. cit., p. 38 (reflecting a validity of .21).

An example of a self-rating schedule is the version of the ACWA that OPM administers on a reimbursable basis to fill entry-level professional and administrative jobs in approximately 118 occupations covered by a court-approved consent decree. Because the ACWA self-rating schedule focuses on entry-level jobs, where applicants typically don't have specialized work experience (and perhaps no work experience at all), it uses life experiences to assess job applicants. The rating factors include many biographical data elements. As we noted earlier, OPM has begun entering into agreements with delegated examining units that will allow them to administer this and the written test versions of the ACWA.

Assessment can also be performed on the basis of biographical data, or biodata.[55] Psychologists actually consider this a distinct form of assessment, but for purposes of this report we include it as a variation of training and experience assessment. Well-constructed biodata assessments have been found to account for as much as 12 percent of the variability in employees' performance on the job,[56] making them better than self-rating schedules of training and experience alone. As we noted above, the ACWA self-rating schedule contains many biodata elements, so it could be said to be either a self-rating or a biodata tool. The written test version of the ACWA has a biodata component. However, as we noted in footnote 19, that written test has not been used since November 1994, when it was replaced by the self-rating schedule with the same name.

Multiple approaches raise concerns

Clearly, there are many ways to use training and experience to assess job applicants and candidates. To us, the substantial difference in the ability of the most and least valid of these various approaches to predict future job performance—and the high probability that not all of the 684 delegated examining units are using the most valid of these approaches—is cause for concern. Our concern is that managers in some unknown number of Federal organizations are not benefiting from high-quality assessment instruments when selecting new employees. We do not express this concern to indict either agency delegated examining units or OPM's stewardship over them. Instead, we express it as a prelude to the following observations:

- Staff within OPM's Personnel Resources and Development Center have considerably more expertise in developing valid assessment tools than is found in most Federal agencies;
- As a result of a budget-driven policy decision, OPM's Personnel Resources and Development Center is funded solely through reimbursement for its services;
- Agencies vary widely in their ability to develop and apply good training and experience (or any other) assessment instruments. Agencies with little in-house expertise in this field, and little or no discretionary money to pay OPM or anyone else for the needed expertise, are at a distinct disadvantage; and
- In such an environment, the resulting assessment tools are likely to vary considerably in quality. When the quality of assessment tools varies, consistent application of those tools by a delegated examining unit is not a virtue if it simply means that the unit is applying unworthy assessments across-the-board.

The implications of automated systems

Assessments conducted through training and experience are well adapted to computer-based applications. Agencies increasingly are buying commercial software packages designed to carry out many steps

[55] Collected through structured multiple-choice questionnaires, biodata are factual biographical information, "such as demographic information, education, work experiences, interests and social activities, habits, hobbies, family history, attitudes, values, achievements, and personal characteristics." (Assessment Council of the International Personnel Management Association, op. cit., p. 42.)

[56] Schmidt and Hunter, op. cit., Table 1, p. 265 (reflecting a validity of .35).

of the hiring process, including candidate assessment. Some of these packages "read" narrative applications through optical scanning and score them (often by searching for keywords or phrases); others score forms specially created to assess candidates for particular occupations at specific grade levels. In some cases agencies find they must adapt the commercial packages to meet their needs; in other cases vendors have been selling their products to Federal agencies long enough to have built into their products many civil service procedural requirements. Some agencies have developed their own equivalent "packages" and some of these agencies compete with commercial vendors for the opportunity to process applications and assess candidates for other agencies. The validity of all of the commercial and agency-developed computer-based systems has not been established. In addition, whether systems that assess candidates based on searches for key words or phrases provide Federal job applicants and candidates—and indeed the public—the merit-based assessments required of Federal agencies is an open question.

Because automated systems make hiring speedier and more convenient, Federal managers generally applaud them. With carefully crafted training and experience assessment tools, managers gain these added benefits of automation without losing the ability to make meaningful distinctions among job candidates. When less well-developed tools are used, managers may gain speed and convenience but risk selecting from a poorly drawn candidate pool and selecting without knowing as much as they can about the likely job success of the selected individual.

As the President's human resources advisor and agent, OPM should have a stronger presence with respect to the proliferation of automated assessment tools, since assessment approaches are critical to the Government's charge that Federal hiring will be based solely on relative merit. As a minimum, OPM should require evidence of, or conduct its own reviews to assure, proper validation of the assessment approach used by each automated tool, whether developed by a government organization or a commercial vendor.

Interviews

The interview is an almost universal assessment tool. In fact, the interview process is so widely used that MSPB will examine it separately in more detail in a soon-to-be released MSPB report. Our intent here is to put the interview into perspective with the three other pre-employment assessment tools covered by this report.

Virtually all managers and supervisors interview job candidates before selecting new employees. These interviews may be either structured or unstructured. However, as we noted in table 3, the interview may be used at more than one step of the assessment process. For certain kinds of jobs—notably those in law enforcement—a structured interview conducted by a panel may routinely be included in the rating process that ultimately ranks the eligible candidates (decides their relative standing for employment consideration).

In a structured interview the questions usually are determined through a careful job analysis, all candidates are asked the same questions in the same order, and there is a fixed procedure for scoring the responses.[57] According to testing experts, "The average validity of a structured interview is .51,"[58] equal to the validity of written tests. This means that a structured interview accounts for as much as 26 percent of the variability in performance among employees on the job, making it one of the very best assessment tools.

The structured interview process is neither cheap nor simple. For example, OPM has developed a 1-day orientation course on the basics of structured

[57] Ibid., paraphrased from p. 267.

[58] Ibid., p. 267.

interviewing, which it has piloted with around 400 agency personnel for two specific occupations (Information Technology and Accounting). These agency participants were expected in turn to teach the techniques to other agency personnel, so that over time the numbers of course participants would be vastly increased. In addition, agencies such as Customs Service and the Immigration and Naturalization Service consider "how to" training as a prerequisite for managers who will conduct structured interviews, and provide 2- or 3-day training courses to those managers.

In addition to the time required for basic instruction, each structured interview requires a job analysis, a determination of the competencies (or knowledge, skills, and abilities) to be measured, and development of both questions to measure them and a rating scale against which to judge the answers. A conscientious manager could easily spend a week developing and preparing to use his or her first structured interview—a large block of time for a person already likely to be pressed for time. Even experienced interviewers may require several days to develop a structured interview in situations where jobs are not filled often, or where the competencies required of the job have changed since it was last filled. Further, when developing structured interviews most managers require assistance from HR specialists and/or subject matter experts, adding to the staff and time costs.

Unstructured interviews could not be more different from structured ones. For example, unstructured interviews have neither a format nor a prescribed set of questions; instead the interviewer usually asks each individual different questions that may or may not be job related. Further, "responses to individual questions are usually not scored, and only an overall evaluation (or rating) is given to each applicant, based on summary impressions and judgments."[59] Unstructured interviews are less expensive to develop and use than structured interviews, but even if very well constructed (which includes maintaining a clear job-relatedness focus) account for only about 14 percent of the variability in employees' performance on the job.[60] While not bad, this shows that they clearly are inferior to structured interviews.

In the Board's survey of Federal supervisors,[61] 54 percent of the respondents rated performance in interviews very important when they considered outside applicants. Only job-related experience and documented qualifications were considered more important among the 16 factors supervisors were asked to judge for outside hiring.

The interview at the selection step can serve multiple purposes. We identified four of those purposes in a short survey we administered to participants in our focus group meetings, and asked the participants to rate the importance of each. Table 6 shows the results.

Table 6. Interview purposes and how important managers said each was

	Percent of survey respondents saying:		
Purpose is to determine:	**Very important**	**Somewhat important**	**Not at all important**
Candidate's level of interest in the job	86	10	2
How well the candidate will fit into the work group	84	12	2
Candidate's qualifications for the job	66	26	6
How well I think the candidate and I will work together	62	32	4

Our focus group participants also said that the selection interview offers an opportunity to directly assess candidates' oral communications skills,

[59] Ibid.

[60] Ibid. (Based on a validity of .38).

[61] "Filling Jobs—What do Supervisors Look For?," *Issues of Merit,* (op. cit.), p. 5.

Source: Reprinted with special permission of North American Syndicate

including listening skills. Several managers also selected an "other" response choice and wrote in that this interview also presents an important opportunity for them to explain the job and job setting to their candidates. Thus the selection interview offers managers an opportunity both to learn more about job candidates and to give the candidates information. This is a positive use of interviews, so long as the focus remains squarely on exchanging job-related information.

Interviewing is a complex communications process, requiring attentive listening as well as good speaking skills. What is asked is critical; so is having a scheme for assessing the information elicited. Since effective interviewing is not necessarily intuitive, with training most people can improve their interviewing technique.

In our small focus group sample, 54 percent of the participants said they had received training in conducting interviews while 46 percent said they had not. This is not a particularly heartening distribution. However, training in structured interviewing is an integral part of a current OPM initiative aimed at improving the assessment process. This effort is being piloted in two occupational fields: accounting and information technology. The accounting pilot involves 7 agencies, while the information technology pilot involves 11.[62] For both pilots OPM Employment Service officials report that OPM staff have provided assessment training to representatives in all participating agencies, including training in the development and use of structured interviews. More than 200 information technology managers and supervisors and HR staff members participated in the training for the IT project; a somewhat smaller number of accounting managers and supervisors and HR staff members participated in the training for the accounting project.

The cartoon above makes the serious point that there is more to any interview than just oral communications. Interviewers and interviewees alike must be attentive not just to the substance of what is said but also to the style of the speech and the body language and other visual signals that accompany it. Good interviewers will avoid sending signals that guide the responses of the persons being interviewed. Further, interviewers must be careful not to make snap judgments, but instead must let the process lead them to discover information about the candidate, and must take pains to make sure the information being discovered is directly

[62] The agencies in the accounting occupation pilot are the Departments of Energy, Interior, Labor, the Treasury, and Veterans Affairs, and the General Services Administration and Office of Personnel Management. The agencies in the information technology pilot are the Departments of the Navy, State, Agriculture, Justice, and Veterans Affairs; the Bureau of Labor Statistics within the Department of Labor; the Internal Revenue Service and Bureau of the Mint in the Department of the Treasury; the Office of Personnel Management, the Pension Benefit and Guaranty Corporation, National Aeronautics and Space Administration, and General Services Administration.

related to the job to be filled. They also must be careful not to judge the interviewees against their own image; that is to say, interviewers must avoid judging more favorably some candidates simply because they are like themselves in appearance, mannerisms, and dress, or because their presentation style is like their own. Structured interviews conducted by trained interviewers are an excellent way to reduce interviewer bias that might otherwise affect the interview results. Both training and interview structure are important to removing or reducing such bias, making OPM's recent training on structured interviewing all the more important.

A final note: information gained through interviews should be used to confirm information gained from other sources, and interviewers should also use other sources to verify information gained in the interview. Interviewing is usually the last pre-employment assessment step conducted by managers or supervisors, who should then tie together the information collected through all of the assessment tools before making a hiring decision.

Probationary assessment

The probationary period is a manager's safety valve against a poor selection decision. This post-appointment assessment approach is based on observation of actual performance on the job, which is the most valid of all predictors of future work performance. This assessment is conducted by each new employee's supervisor or manager.

Every new Federal employee is subject to a probationary period, and this period is the final step in the examining process. We observed earlier that the Federal probationary period is 1 year long, except for rare instances involving organizations that have received authority to establish probationary periods of different (usually longer) periods of time. However, a uniform period for all employees is not the only way to construct a probationary period. For example, in the Public Service of Canada (the equivalent of our Civil Service) the length of the probationary period varies based on the nature of the job, ranging from a period of a few to as many as 36 months.

It could also be argued that managers should have the flexibility to establish the length of probation for a new employee within a fixed time range (so that different people in similar jobs might have different length probationary periods) so long as they establish that length at the time of hiring. Under such an approach the length might be influenced by, for example, the training and experience the new employee brings to the job. Finally, it could be argued that managers should have the flexibility to end an employee's probationary period early after determining that the person's performance has met or exceeded expectations, just as managers can end it early by separating the employee after determining that his or her performance is unsatisfactory.[63] Refinements of this nature would acknowledge that Federal agencies and work settings are not all the same.

Whatever its length, proper use of the probationary period is critical. The most important practical consideration for managers is one we noted earlier: until employees successfully complete their probationary periods, they generally have no right to appeal decisions to terminate their employment. Once past the probationary period, Federal civil servants have job protection rights that often make their firing for poor performance burdensome—contributing to the myth that civil servants cannot be fired. In fact, as noted in a recent MSPB publication:

> Nearly 8,400 federal employees were fired (or quit in lieu of being fired) in fiscal year 2000 for reasons related to poor performance, suit-

[63] While supervisors and managers should attempt to improve the performance of poorly performing probationary employees, in the end the probationary period must conclude with one of two affirmative decisions: judge the performance to be satisfactory and allow the employee to continue working, or judge the performance to be unsatisfactory and terminate the employee. That is why nearly all organizations have a probationary period with fewer employee rights and safeguards.

ability, conduct, or conduct-related deficiencies such as abandonment of position or falsifying a job application. Most actions such as these are not appealed to MSPB, and when they are, federal agencies are highly successful in defending their decisions.[64]

Reinforcing the critical importance of proper use of the probationary period are Board turnover studies showing that many individuals who stay with the Government beyond their first 3 years of employment are likely to stay for a career. This means that a mismatched individual who gets through the probationary period because a manager did not properly use that post-appointment assessment period could be a problem employee somewhere in Government for a long time. There are many anecdotes concerning managers and supervisors who kept poorly performing new employees beyond their probationary period, and subsequently dealt with these poor performers by inflating their capabilities and passing them off to other Federal organizations, sometimes with resultant promotions. Consequently, managers, supervisors, and coworkers end up dealing with marginal or unacceptable performance seemingly forever. This is an unacceptable but all too real consequence each time a manager decides not to use the probationary period effectively to deal with poor performers.

[64] "OPE Focus on the Facts," *Issues of Merit,* Office of Policy and Evaluation, U.S. Merit Systems Protection Board, November 2000, p. 1

Conclusions

If agencies were uniformly using adequate tools to assess job candidates, we would not argue for better tools. However, all too often agencies hire new employees using assessment methods that appear not to be very good. The Board is also concerned that the variation in agencies' attitudes toward, and resource commitments to, candidate assessment too often result in poor assessment methods being used. This puts merit-based hiring at risk. While risk cannot be entirely eliminated, steps should be taken to reduce it whenever possible. OPM is the agency best situated (with respect to staff resources and assigned mission) to provide the leadership agencies need to keep the risk of poor employee assessment to an acceptable minimum.

The current decentralized approach to hiring raises questions about the balance between merit and expediency in Federal hiring. OPM has delegated examining authority to 684 delegated examining units. OPM establishes the standards for these units' operations, trains their staffs, and periodically conducts oversight reviews of their operations. However, agencies decide what assessment tools they will use and are responsible for developing them (or having them developed), and they vary in their expertise to develop (or pay for) valid ones. Potentially, the Federal Government applies 684 "rubber yardsticks" of differing quality when assessing job candidates for occupations common across Government agencies. Identifying and eliminating or improving the least valid of these assessment instruments would strengthen the Federal civil service, but doing so has budget implications for OPM and agencies.

Meanwhile, OPM continues to develop assessment tools for agencies' use only on a reimbursable basis. We think it is important that OPM expand its efforts both to develop good assessment tools and to make them available to all Federal agencies. And we think the current policy of always requiring benefiting agencies to reimburse OPM for the development of assessment instruments is detrimental to the Government's interest in hiring based on merit. In the interest of merit staffing, OPM should have the capability and resources to make the best practicable candidate assessment tools available to agencies regardless of the agencies' ability to pay for those tools.

With nearly 700 delegated examining units, differences in their capability to assess candidates is a certainty. Based on information contained in OPM oversight reports, discussions with oversight staff, and our own observations about agencies' differing views of the importance of candidate assessment, we believe that the difference between the best and worst assessments being conducted is substantial. And we believe that delegated examining units whose assessment tools are at the lower end of the range represent a serious threat to hiring based on merit, even though they may believe they are operating in a meritorious manner. Stronger efforts to identify weak assessment tools during reviews of DEUs, and steps to make valid assessment tools available to agencies needing them, are ways OPM

can strengthen merit while helping DEUs better serve the agency managers they support.

Agencies increasingly are turning to automated systems to improve the speed and lower the cost of hiring. It is not clear that agency and commercial providers of these systems have demonstrated either that their systems are valid selection tools or that they meet the special requirements associated with Federal hiring. The proliferation of automated staffing systems only raises the level of concern over these issues.

Good applicant assessment tools are an investment, not just costly and time-consuming impediments to the speedy hiring of new employees. Federal managers should demand high-quality assessment tools as adamantly as they demand a speedier hiring process. Delegated examining units themselves are a business investment: the cost of that investment is best recouped by providing the units with the best instruments practicable. Agencies pursuing this goal and thwarted by resource limitations would then be in a stronger position to seek assistance from OPM without cost or at a negotiated reduced cost, and OPM's argument for providing such assistance would be strengthened.

The fact that OPM determined that it should include training in developing and using structured interviews as part of its current competencies initiatives for the accounting and information technology professions reinforces our concern that Federal managers too often rely on unstructured interviews when they assess job candidates. Structured interviews are far more valid, but as we have seen, training managers and HR staffs in their construction and conduct requires a significant time and dollar commitment.

Many Federal agencies appear to disdain or fear written tests. That well-constructed written tests produce better selections—and thus save money by helping reduce turnover costs—is well documented. Although written tests are not appropriate in every case, opportunities to improve candidate assessment by using them are being lost.

Using a combination of assessment tools can improve the likelihood of selecting the best available person, but not if the tools are of poor quality, measure the same things, or are used in lieu of better ones that are available.

Federal agencies need the ability to recognize labor market fluctuations in a timely manner and the flexibility to adjust their candidate assessment strategies to match those fluctuations. When labor pools are severely limited, assessing candidates beyond determining that they are basically qualified may not be practicable. In contrast, when dozens or hundreds of applicants may apply for only a few jobs, more stringent assessment is called for. Managers and their supporting human resources staffs need to develop a cooperative working relationship so they can monitor labor market conditions and determine the best assessment strategies to use when hiring. Full achievement of changes contemplated by this conclusion probably will require training for most managers, supervisors, and HR specialists.

Good use of the post-appointment assessment tool—probation—can help reduce the negative consequences of a poor selection. If agencies continue to use less predictive assessment tools, or if they shift to less rigorous assessment strategies to accommodate shallow applicant pools, then increased attention should be focused on the probationary period. However, over-reliance on the probationary period to remedy the consequences of using poor assessment methods inevitably will lead to increased turnover costs, including the costs required for repeat hiring and training, and the costs of lost productivity.

Recommendations

The Director of the Office of Personnel Management should—

1. Measure the gap between the requirement that all agencies use valid assessment tools and the reality that some agencies cannot afford to develop and apply them. Incorporate into OPM's strategic plan a strategy for eliminating that gap.
2. Seek to secure the support and resources necessary to eliminate the gap. For example, propose legislation, request additional funding, publish regulations, provide models, authorize demonstration projects.
3. Devise a method for addressing the costs of developing assessment tools so that all delegated examining units have access to the most valid and practical tools irrespective of their agencies' internal expertise or financial capability to buy such devices. Consider proposing a return to appropriated funding for OPM's centralized development and validation of good candidate assessment tools that agencies could acquire and use at little or no cost.
4. Emphasize and expand OPM's current effort to identify factors (either competencies or knowledge, skills, and abilities) that are important to measure for different occupations at different grade levels and to develop assessment tools appropriate to measure those factors. Disseminate the information and tools to agencies once they are available.
5. Focus greater attention, during oversight reviews of delegated examining units, on how well those units are observing OPM's regulatory requirement to validate the assessment tools they use.
6. Review automated staffing systems used or proposed for use by agencies to ensure that the assessment processes that they incorporate meet the requirements established by law and OPM's regulations governing employment practices.
7. Strongly focus attention on the importance of the probationary period as the final step in assessing new employees.

Agency heads should—

1. Cooperate with and encourage OPM in its efforts to develop and make available valid candidate assessment tools, and ensure that the candidate assessment methods being used or considered by their delegated examining units meet all legal and regulatory requirements.
2. Take a strategic view of the role of candidate assessment, treating candidate assessment methods as a business investment by:
   - Budgeting for the development of the best practicable assessment tools;
   - Providing adequate and adequately trained staff for their delegated examining units;
   - Training managers, supervisors, and human resources staff members in the use of those tools;

- Holding managers, supervisors, and their supporting HR staffs accountable for properly using the best practicable instruments when assessing candidates for employment; and
- Holding managers and supervisors accountable for effectively using the post-appointment probationary period to assess new employees, and for separating nonperformers during the probationary period instead of allowing them to gain employment protections that rightfully are earned only upon successful completion of the probationary period.

3. In the interest of preserving merit as an essential component of the hiring system, focus managers' and supervisors' attention on the need to strike the proper balance between selecting the best job candidates and hiring quickly in a tight labor market.

Appendix 1: A discussion of selection tool validity

The "validity" of a selection tool describes the relationship between that tool and some job-related factors. Examples include the relationship between scores on the selection tool and measures of job performance, or the relationship between the content of the instrument and the content of the job being filled. Whether the selection tool is an actual test or a scheme for evaluating training and experience, it may be possible to calculate several measures of validity for it based on the purpose of the selection tool. A selection tool may have a strong relationship to job performance measures when used for one kind of job, but very little relationship when used for another kind of job. These measures of relationship focus on both what the tool measures and how well the tool measures.[65]

Typically, discussions about validity in the context of selection instruments used in the Federal sector are concerned with evaluating the usefulness of a method for predicting how well a candidate will perform on the job. For example, if a tool is valid for selecting realty specialists, we should be able to use the tool to predict which candidates are most likely to succeed in realty specialist jobs. Not only should this instrument provide information about the candidates' performance on factors relevant to the work of realty specialists, but it should also make distinctions among the candidates who were evaluated using it (that is, some candidates should score better on this device than others). To be useful a selection tool needs to predict differences in how the candidates will perform in the job. But it is also essential that the selection tool evaluate factors relevant to the job being filled. It is not very useful, for example, to use a selection instrument that evaluates factors relevant to the work of a public affairs specialist when the job being filled involves realty specialist work.

We express predictive validity (that is, the relationship between the score a candidate receives on the selection tool and some measure of job performance) as a numerical value between 0 and plus or minus 1.0 (anything between 0 and –1.0 would indicate an inverse relationship between the device and job performance). The closer the value is to 1.0, the stronger the relationship between the score received on the selection tool and favorable performance in the job. But, as we noted in a recent Board report, no single selection instrument has achieved a validity measure of 1.0, and "relatively few achieve a value as high as .40 or .50. Among the best commonly used tools in the Federal sector are work sample tests (.54), general mental ability tests (.51), [and] structured interviews (.51), * * *."[66]

These expressions of predictive validity are used to estimate how much of the variability of an employee's performance on the job can be predicted by performance on the assessment tool. This estimate is made by squaring the validity measure to arrive at the percentage of variance in job performance that can be accounted for by the test. Using the work sample value of .54 from above, applying this formula shows that performance on the work sample test would predict 29 percent of the vari-

[65] W. F. Cascio, "Applied Psychology in Human Resource Management," fifth ed., Prentice Hall, Upper Saddle River, NJ, 1998, p.99.

[66] U.S. Merit Systems Protection Board, Washington, DC, "Restoring Merit to Federal Hiring," January 2000, p.16.

ability of how well people will do on the job (.54 x .54 = .2916, or 29 percent). In the section of this report addressing training and experience ratings, we report that one approach to assessing candidates through that method has a validity of .11. Squaring the figure leads us in that narrative to the discovery that only about 1 percent of job performance variability can be predicted through the rating (.11 x .11 = .0121, or 1 percent). Throughout the report's text we express the validity of the various assessment tools in terms of the variability in performance they can predict.

Federal managers have several reasons to be concerned with the validity of their selection methods. For one thing, the first statutory merit system principle says that "selection and advancement should be determined solely on the basis of relative ability, knowledge, and skills," which is what selection tools should identify. In addition, there are legal requirements concerning the validity of selection instruments that must be met. But most importantly of all, from the standpoint of effective resource management it just makes good sense to use the best tools we have, in the most appropriate way, to select the highest quality job candidates. When making selections, managers who rely on inadequate information—or information that is not appropriate for the job being filled—face an increased risk of making bad selection decisions. As the Board has stressed in a number of earlier reports, the consequences of a bad hiring decision can persist for many years after the individual is hired, and "ultimately affect the performance of the entire work unit."[67] From a business sense, it is better to spend time and other resources during the selection process to find good employees than later to spend time and resources trying to improve the performance of a poorly selected employee or to get rid of an ineffective performer.

[67] U.S. Merit Systems Protection Board, "Federal Supervisors and Strategic Human Resources Management," Washington, DC, June 1998.

Appendix 2: An inventory of written and performance tests used by or available to Federal agencies

Identification:	Applies to applicants for:	Grades used at:
Blue-collar occupations		
Test 100, Apprentice Trades and Helper-Trainee	Applicants for apprentice and helper trainees in crafts and trades	
Test 101, Apprentice Printing Trades	Printing trades apprentice applicants	
Test 108, Apprentice Non-Printing Trades	Non-printing trades apprentice applicants	
White-collar occupations		
Administrative Careers With America (ACWA) written test	Approximately 112 professional and administrative occupations at entry-level (GS-5 and –7) grades	Remains in OPM's inventory but is not used
Administrative Careers With America (ACWA) self-rating schedule.	Approximately 112 professional and administrative occupations at entry-level (GS-5 and –7) grades	Used in place of the written test
Test 11, Food Inspector, Reading Comprehension	Food Inspection, GS-1863	GS grades 5 and 7
Test 23, Museum Technician, Spatial Perception	Museum Technician, GS-1016	GS grades 2 and 3
Test 24, Scientific Aide/Technician, Computation and Test 35, Scientific Aide/Technician, Reading Comprehension *Continuation of* Test 24, Scientific Aide/Technician, Computation and Test 35, Scientific Aide/Technician, Reading Comprehension	Safety Technician, GS-019; Psychology Aid and Technician, GS-181; Biological Science Technician, GS-404; Plant Protection Technician, GS-421; Range Technician, GS-455; Soil Conservation Technician, GS-458; Irrigation System Operation, GS-459; Forestry Technician, GS-462; Nursing Assistant, GS-621; Rehabilitation Therapy Assistant, GS-636; Health Aid, GS-640; Nuclear Medicine Technician, GS-642, Medical Technician, GS-645; Pathology Technician, GS-646; Diagnostic Radiologic Technologist, GS-647, Therapeutic Radiologic Technologist, GS-648; Medical Instrument Technician, GS-649; Respiratory Therapist, GS-651, Pharmacy Technician, GS-661; Orthotist and Prothestist, GS-667; Dental Assistant, GS-681; Dental Lab Aid, GS-683; Environmental Health Technician, GS-698; Animal Health Technician, GS-704; Engineering Technician, GS-802; Construction Control, GS-809; Surveying Technician, GS-817; Engineering Drafting, GS-818; Electronics Technician, GS-856; Industrial Engineering Technician, GS-895; Office Drafting, GS-1021; Physical Science Technician, GS-1311; Hydrologic Technician, GS-1316; Meteorological Technician, GS-1341; Cartographic Technician, GS-1371; Geodetic Technician, GS-1374; Mathematics Technician, GS-1521, Cryptanalysis, GS-1541; Agricultural Commodity Aid, GS-1981	GS grades 2 and 3 for all listed occupations except: Orthotist and Prothestist; which is at GS-3 only

Identification:	Applies to applicants for:	Grades used at:
Test 41, Name and Number Checking	Not specified	Not specified
INS-132, Border Patrol: Spanish language proficiency	Border Patrol Agent, GS-1896	Not specified
INS-133, Border Patrol: Translation in an artificial language	Border Patrol Agent, GS-1896	Not specified
Test 134, Border Patrol, Judgment, Law Enforcement Problems	Compliance Inspection and Support, GS 1802	GS grades 2, 3, 4 Not used for Detention Enforcement Officers in the GS-1802 series
Test 135, Following Oral Directions	Entry-level food inspector applicants	
Test 146, Arithmetic Computation	Entry-level food inspector applicants	
Test 148A, Object Perception	Entry-level food inspector applicants	
Test 149A and B	Entry-level food inspector applicants	
Test 157, Abstract reasoning—Symbol classification and letter series	Not specified	Not specified
Test 175A and B, Museum Technician; Messenger, Clerical (Name and Number Checking and Coding)	Museum Technician, GS-1016	At GS grades 2 and 3
Test 219, Typing performance	Not specified	Not specified
Test 220, Automated version of typing performance test	Not specified	Not specified
Test 433 (Unnamed) and Test 354 (Unnamed)	Shorthand Reporter, GS-312 and Closed Microphone Reporting, GS-319	GS grades 6, 7, 8, 9
Test 480, Shorthand.	Reporting Stenographer, GS-312	GS grades 5 and 6
Test 510, Job specific learning ability, miniature training and experience	Not specified	Not specified
Test 512, Job knowledge	Not specified	Not specified
Test 527, Critical thinking, arithmetic reasoning, and writing skills (Customs Service) and Test 529, Job-specific structured interview and Test 615, Applicant Assessment (Customs) Social Skills	Customs Inspection Canine Enforcement Officer, GS-1890	GS grades 5 and 7
Test 528, Job-specific structured interview	Not specified	Not specified
Test 535, Job-specific learning ability, miniature training and experience	Computer Specialist, GS-334	GS grades 5 and 7
Test 546, Logical reasoning, mathematical reasoning, and judgment and Test 547, Job-specific structured panel interview competencies and Test 620, Job-specific Suitability Assessment	United States Marshal, GS-082 or GS-1811	GS grades 5 and 7
Test 548 (Customs), Logical reasoning, mathematical reasoning, and judgment and Test 616, Applicant Assessment (Customs)	Criminal Investigator (Customs), GS-1811	GS grades 5 and 7
Test 548, Logical reasoning, mathematical reasoning, and judgment	Treasury Enforcement Agent, GS-1811 and Special Agent (Wildlife), GS-1812	GS grades 5 and 7 GS grade 7
Test 549, Name not specified	Treasury Enforcement Agent, GS-1811	GS grades 5 and 7

Identification:	Applies to applicants for:	Grades used at:
Test 590, Job-specific structured panel interviewing skills.	Not specified	Not specified
Test 610, ATF Job-specific Suitability for Law Enforcement Work	Alcohol, Tobacco and Firearms Inspection, GS-1854	GS grades 5 and 7 (Used in addition to ACWA)
INS-611, Job-specific applicant Suitability Assessment	Border Patrol Agents, GS-1896	GS grades 5 and 7
INS-612, Job-specific applicant Suitability Assessment and INS-736, Verbal (Vocabulary and Logical Reasoning and Reasoning with Probabilities)	Civil Aviation Security Specialist, GS-1801; Center Adjudications Officer, GS-1801; District Adjudications Officer, GS-1801; Criminal Investigator (INS), GS-1811; Immigration Inspection, GS-1816	GS grades 5 and 7 in INS
INS-613, Job-specific Applicant Suitability Assessment and INS-734, Verbal (Vocabulary and Logical Reasoning)	Detention Enforcement Officer, GS-1802	GS grades, 2, 3, 4, and 5
Test 640 (Unnamed)	Social Insurance Claims Examiner, GS-993	Only for retirement and unemployment claims examiners at GS-4
Test 650, Perceptual Speed (name and number checking); Verbal (vocabulary and reading comprehension); Reasoning (abstract reasoning and number series)	Environmental Protection Assistant, GS-650 Contract Representative, GS-962 General Claims Examiner (1-grade interval only), GS-990	GS grades 2, 3, 4 GS grades 3 and 4 GS grade 4
Test 715, Perceptual Speed (name and number checking); Verbal (reading comprehension); Mathematical Reasoning	Police, GS-083; Park Police, GS-083 Police (Secret Service), GS-083; Security Guard, GS-085	GS grade 2 GS grade 5 GS grades 4 and 5 GS grade 2
Test 716, Achievement and Experience—Bio-data and social skills	Police, GS-083; Police (Secret Service), GS-083; Security Guard, GS-085	GS grade 2 GS grades 4 and 5 GS grade 2
Test 730, Verbal (Vocabulary and Reading)	Customs Patrol Officer, GS-1884	GS grades 5 and 7
Test 731, Border Patrol, Vocabulary and Logical Reasoning	Border Patrol Agent, GS-1896	GS grades 5 and 7
INS-737, Verbal (Vocabulary, Logical Reasoning, Mathematical Reasoning and Reasoning with Probabilities)	Not specified	Not specified
Test 740, Job-specific Verbal (Reading Comprehension and Grammar, Arithmetic Computation and Mathematical Reasoning)	Railroad Retirement Claims Examining, GS-993	GS grades 5 and 7
INS-742, Job Specific Structured Interview Skills	Not specified	Not specified
INS-743, Job Specific Structured Interview Skills Measured in Telephone Interviews	Not specified	Not specified
INS-744, Job Specific Interview Skills	Not specified	Not specified
INS-749, Job Specific Structured Interview Skills	Not specified	Not specified

Identification:	Applies to applicants for:	Grades used at:
Test 800, Perceptual Speed (name and number comparison, coding, and simple arithmetic computation) and Test 801A, Verbal (Vocabulary, Reading, Grammar, Spelling and Verbal Analogies) Test 800, Perceptual Speed (name and number comparison, coding, and simple arithmetic computation) and Test 801A, Verbal (Vocabulary, Reading, Grammar, Spelling and Verbal Analogies)	Fingerprint Identification, GS-072; Security Clerical and Assistance, GS-085; Psychology Aid and Technician, GS-181; Social Service Aid and Assistant, GS-186; Recreation Aid and Assistant, GS-189, Personnel Clerical and Assistance, GS-2033; Military Personnel Clerical and Assistance, GS-204; Messenger, GS-302; Misc. Clerical and Assistant, GS-303; Information Receptionist, GS-304; Mail and File, GS-305; Correspondence Clerk, GS-309; Clerk-Stenographer, GS-312; Secretary, GS-318; Clerk-Typist, GS-322; Office Automation Clerical and Assistance, GS-326; Computer Operation, GS-332; Computer Clerk and Assistant, GS-335; Management and Program Control, GS-344; Equipment Operator, GS-350; Printing Clerical, GS-351; Data Transcriber, GS-356; Coding, GS-357; Electric Accounting Machine Operation, GS-359; Telephone Operating, GS-382; General Telecommunications, GS-392; Communications Clerical, GS-394; Financial Clerical and Assistance, GS-503; Accounting Technician, GS-525; Cash Processing, GS-530; Voucher Examining, GS-540; Civilian Pay, GS-544; Military Pay, GS-545; Budget Clerical and Assistance, GS-561; Tax Examining, GS-592, Insurance Accounts, GS-593; Medical Records Technician, GS-675; Medical Clerk, GS-679; Legal Instruments Examining, GS-963; Claims Clerical, GS-998; General Arts and Information, GS-1001; Language Clerical, GS-1046; Editorial Assistance, GS-1087; Procurement Clerical and Technician, GS-1106; Property Disposal Clerical and Technician, GS-1107; Production Control, GS-1152; Library Technician, GS-1411; Archives Technician, GS-1421; Statistical Assistant, GS-1531; Education and Training Technician, GS-1702; Supply Clerical and Technician, GS-2005; Sales Store Clerical, GS-2091; Freight Rate, GS-2131; Transportation Loss and Damage Claims Examining, GS-2135; Dispatching, GS-2151	1. All are at GS grades 2, 3, 4 except: a. Psychology Aid and Technician, GS-181; Social Service Aid and Assistant, GS-186; and Recreation Aid and Assistant, GS-189 are at GS grades 2 and 3 b. Clerk-Stenographer, GS-312, is at GS grades 3, 4, 5 c. Secretary, GS-318, is at GS grades 3 and 4 2. The following occupations also require a performance test or self certification: a. Clerk-Stenographer, GS-312 b. Clerk-Typist, GS-222 c. Office Automation Clerical and Assistance, GS-326 d. Data Transcriber, GS-356 3. Sales Store Clerical, GS-2091, does not require Test 801A
FAA Written Tests	Airway Transportation System Specialist, GS-2101	GS grades 5 and 7
FAA Tests	Air Traffic Control GS 2152	GS grades 5 and 7

Other Selected MSPB Publications from the Office of Policy and Evaluation*

- The Federal Merit Promotion Program: Process vs. Outcome
- The U.S. Office of Personnel Management in Retrospect: Achievements and Challenges After Two Decades
- Growing Leaders: The Presidential Management Intern Program
- Assessing Federal Job Seekers in a Delegated Examining Environment
- Competing for Federal Jobs: Job Search Experiences of New Hires
- Restoring Merit to Federal Hiring: Why Two Special Hiring Programs Should Be Ended
- The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service
- Federal Supervisors and Poor Performers
- Civil Service Evaluation: the Evolving Role of the U.S. Office of Personnel Management
- Federal Supervisors and Strategic Human Resources Management
- The Changing Federal Workplace: Employee Perspectives
- Adherence to the Merit Principles in the Workplace: Federal Employees' Views
- Achieving a Representative Workforce: Addressing the Barriers to Hispanic Participation
- Fair and Equitable Treatment: A Progress Report on Minority Employment in the Federal Government
- The Rule of Three in Federal Hiring: Boon or Bane?
- Sexual Harassment in the Federal Workplace: Trends, Progress, Continuing Challenges
- Leadership for Change: Human Resource Development in the Federal Government
- Whistleblowing in the Federal Government: An Update
- A Question of Equity: Women and the Glass Ceiling in the Federal Government

Selected Current Projects

The use of interviews
Governmentwide Merit Principles Survey 2000
Competence in overseeing federal contracts
Federal government recruiting
Automated tools for candidate assessment

For a copy of any Office of Policy and Evaluation publication, contact:

U.S. Merit Systems Protection Board
Office of Policy and Evaluation
1615 M Street, NW
Washington, DC 20419

E-mail: studies@mspb.gov
Phone: (202) 653-6772 extension 1350
or 1-800-209-8920
V/TDD: (202) 653-8896 (TTY users, please use the Federal Relay Service at 800-877-8339 to contact this office.) Fax: (202) 653-7211

* Most of these reports and the Office of Policy and Evaluation newsletter, *Issues of Merit*, can be downloaded from the STUDIES page of the MSPB web site: mspb.gov