

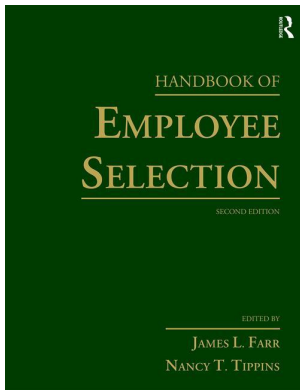
This article was downloaded by: 10.2.97.136

On: 21 Sep 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



## **Handbook of Employee Selection**

James L. Farr, Nancy T. Tippins, Walter C. Borman, David Chan, Michael D. Coovert, Rick Jacobs, P. Richard Jeanneret, Jerard F. Kehoe, Filip Lievens, S. Morton McPhail, Kevin R. Murphy, Robert E. Ployhart, Elaine D. Pulakos, Douglas H. Reynolds, Ann Marie Ryan, Neal Schmitt, Benjamin Schneider

## **Blue-Collar Selection in Private Sector Organizations**

Publication details

<https://test.routledgehandbooks.com/doi/10.4324/9781315690193-34>

Robert P. Michel, Shannon Bonner

**Published online on: 22 Mar 2017**

**How to cite :-** Robert P. Michel, Shannon Bonner. 22 Mar 2017, *Blue-Collar Selection in Private Sector Organizations from: Handbook of Employee Selection* Routledge

Accessed on: 21 Sep 2023

<https://test.routledgehandbooks.com/doi/10.4324/9781315690193-34>

**PLEASE SCROLL DOWN FOR DOCUMENT**

Full terms and conditions of use: <https://test.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## BLUE-COLLAR SELECTION IN PRIVATE SECTOR ORGANIZATIONS

---

ROBERT P. MICHEL AND SHANNON BONNER

This chapter focuses on selection for blue-collar jobs. There are several aspects of blue-collar jobs and blue-collar work environments that make selection efforts different from those for white-collar or other types of jobs. Without a clear understanding of these contextual considerations, the selection practitioner's efforts at developing selection systems for blue-collar jobs can be undermined. Our goals in this chapter are to highlight the issues that make blue-collar selection unique and provide guidance to selection practitioners who deal with blue-collar selection, as well as suggest possible areas for future research that would help inform blue-collar selection efforts.

### DEFINITION/BOUNDARIES

The term "blue-collar" has been used since the early 20th century to describe working-class jobs and contrast them to "white-collar" jobs that are professional or managerial in nature and typically occur in an office environment. Similar terms have been added to the lexicon over the years to describe other classes of workers (e.g., pink, green, and gold collar), but blue and white collar remain the primary distinctions when differentiating classes of workers. Because of the type of work involved in blue-collar jobs, several stereotypes have developed over time, some positive (e.g., work ethic, pride, and loyalty) and some negative (e.g., unskilled and unrefined), but we use the term neutrally, simply to describe a specific class of jobs.

What constitutes a blue-collar job varies depending on the source. For our purposes, we define blue-collar jobs as those that involve some type of manual or physical labor, often involving the use of tools or specialized equipment, and typically occurring in "non-office" environments that are sometimes hazardous. Clearly falling within this definition are skilled trades or craft jobs in manufacturing and construction industries, such as mechanics, machinists, electricians, welders, plumbers, and carpenters. However, blue-collar jobs can range from completely unskilled (e.g., manual laborer) to very specialized, highly skilled jobs (e.g., nuclear reactor operator).

As with the broad range of skill required by blue-collar jobs, the complexity and autonomy varies widely as well. Jobs such as helper, assistant, or laborer may entail following direct and concise instructions. These jobs have relatively low information processing demands and involve minimal planning or decision making. At the other end of the spectrum, jobs such as locomotive engineer, nuclear reactor operator, and demolition expert are highly complex and have extensive information processing demands, as incumbents must plan carefully, evaluate a myriad of possible outcomes, and coordinate the work of others to ensure safe and effective outcomes.

Characteristic of many skilled trade occupations is the use of an apprenticeship system, in which employees are hired as apprentices (or even pre-apprentices) and progress to journeymen and then, in some cases, masters. The purpose of the apprenticeship is for the employee to learn the trade through on-the-job training (OJT), typically complemented by classroom training. Given the hands-on nature of most blue-collar jobs, OJT is a perfect fit. Apprenticeships typically last three to six years, at which point a successful apprentice would become a journeyman. While the use of the terms “apprentice” and “journeyman” by employers in the U.S. is not regulated, official recognition of apprentice or journeyman status is regulated. Many states and jurisdictions have formal requirements to obtain a journeyman license, and the Department of Labor has a formal Registered Apprenticeship program, through which the worker receives a nationally recognized certification.

### CONTEXTUAL CONSIDERATIONS

A handful of contextual factors are important to consider when designing selection programs for blue-collar jobs. Each factor is not necessarily unique to blue-collar jobs, but the issues are more salient than they are for white-collar jobs. When one or more of these factors is a consideration for a particular blue-collar job, it can impact both the tools that are included in the selection process and the amount of resources dedicated to selection system design.

#### Work Environment

Just as the level of skill can vary across jobs classified as blue collar, so, too, can the work environment. Some blue-collar jobs, such as iron worker or roofer, may entail performing tasks outdoors in varying weather conditions, and others, such as firefighter and police officer, may experience extreme temperatures and loud noises. At the other end of the spectrum are jobs such as high-technology manufacturing operator, where workers operate computers and robots in clean, climate-controlled environments. Despite the variation in blue-collar work environments, they all differ from the typical white-collar environment, where much of the workday is spent either sitting at a desk in front of a computer or in meetings.

Another common characteristic of many blue-collar work environments is potentially hazardous working conditions, such as extensive physical demands (e.g., firefighter and police officer), use of heavy equipment and machinery (e.g., construction worker), working with dangerous materials (e.g., demolition expert), and working at heights (e.g., line worker). For this reason, safety is often a central component of the work culture for these jobs. Research shows meaningful individual differences in the extent of safety behaviors in which people typically engage (O’Connell & Delgado, 2011), so safety is often an important focus of the selection process for blue-collar jobs. Depending on the nature of the job, this might entail using assessments that evaluate requisite physical abilities (Campion, 1983), specialized knowledge (Hoffman, Jacobs, & Landy, 1995), or safety awareness (Vredenburg, 2002).

A final component of some blue-collar work environments worth noting is shift work. In many blue-collar industries, 24-hour operations are either necessary (e.g., utilities and emergency response occupations) or desirable for efficiency (e.g., manufacturing). This means that some employees must work at times other than traditional working hours, including swing and night shifts. Some organizations use rotating shifts, in which an employee works on each of the various shifts over the course of several weeks. While this helps to distribute the burden of non-traditional work hours across all employees, it can also be disruptive to one’s personal life, and some research suggests that rotating shifts may even be damaging to one’s health (De Bacquer et al., 2009).

#### Tenure

Another important consideration when designing selection programs for blue-collar jobs is the average tenure of employees. In many blue-collar industries, employees stay with their employer

longer than in other industries. For example, according to the Bureau of Labor Statistics (2015), over the past decade manufacturing employees have consistently had the highest median tenure for any private sector industry, at about six years. When you drill down to specific industry sectors, the average tenure is even higher, with paper and printing (9.7 years) and utilities (9.2 years) leading the way. Moreover, because these are average tenures across entire industries or sectors, they include both blue- and non-blue-collar jobs. Undoubtedly, if these tenure figures were isolated by specific blue-collar job groups, the median tenure for many groups would be even higher.

Logically, the longer an employee stays with an organization, the impact of his or her performance on overall workforce productivity becomes more pronounced over time. This is accounted for in utility formulas by the inclusion of a multiplier for the average tenure of those hired (Schmidt & Hunter, 1983b). The economic value to an organization of a specific selection procedure depends on multiple factors, including the dollar value of job performance variability. Determining the variability in job performance for a specific job can be time consuming and complex, but Schmidt and Hunter (1983a) estimated that the lower-bound standard deviation of employee contributions in dollars is 40%. Using a selection tool with a predictive validity of .30 for a job with a starting salary of \$64,000 and a selection ratio of 40%, this translates into a utility benefit of about \$7,500 per hire in the first year. Actual utility gains across an employee's tenure are complicated and involve multiple factors, including the availability and cost of replacement employees (Cascio & Boudreau, 2011), but for example's sake, if we assume the impact is linear, the cumulative effect on an organization's bottom line for a job with an average tenure of 10 years is \$75,000 per hire. When this is multiplied by the number of hires per year, the effect of average tenure on selection utility becomes quite pronounced, so it makes sense to consider investing more in the selection process for high-tenure jobs to ensure you reap the benefits.

## Labor Unions

Many blue-collar workers are represented by labor unions, which exist to protect the interests of their members, particularly around wages, benefits, job security, and working conditions. The best way for unions to achieve their goals is by having influence or control over organizational policies and procedures. Agreements regarding workers' rights are negotiated between union leadership and company management and are codified in labor contracts or collective bargaining agreements. Concessions made by either side during the negotiation process become part of the labor contract and generally cannot be renegotiated until the labor contract expires.

Because unions exist to protect the interests of their members, who are already company employees, they are generally more concerned with internal selection practices (e.g., promotions and transfers) than the hiring of external applicants for entry-level jobs. However, in some instances the union may also care about external selection practices since those hired are their future members. The greater the impact the quality of hires has on the union's well-being, the greater their interest will be. For example, in one organization with which the authors are familiar, company management reserves the right to outsource certain functions if specific performance targets are not met. In this instance, the quality of hires has a direct bearing on the union's well-being, and they take a keen interest in the external selection process.

As far as internal selection, from the union's vantage point the fairest and most objective way to handle selection decisions is through seniority (Bownas, 2000). When employee seniority is the sole factor for determining promotions and transfers, the union never has to favor one member over another and union loyalty is rewarded. To the extent that additional selection procedures are used, the union will want them to be as objective and job-relevant as possible, preferably based on current performance (e.g., training success, apprenticeship completion, or completion of verifiable goals). If some form of testing is used, the preference will be for tests that have strong fidelity with the job, such as work samples.

In contrast to the goals of union leaders, line managers want fast and accurate selection of competent performers to ensure the productivity of their work unit. Because of the inherently competing priorities of labor unions and management, there is a tradition of conflict and distrust between the two groups. The relationship can vary greatly from organization to

organization, or even from union to union within a specific organization, and can range from downright acrimonious to positive and productive. For the selection professional working with unionized jobs, understanding the perspectives of both labor and management is critical. As Bownas (2000) astutely pointed out, the two sides are really no different in what motivates them. They are both looking out for their self-interests, which is only natural. The selection professional's aim, then, should be to remain an objective third party, whose goal is to develop the most effective selection process possible that helps, at least partly, address the concerns of each side.

Effectively working with both parties also requires an understanding of the history of the union–management relationship and what, if anything, has been incorporated into the labor agreement that would impact the ability to implement and manage a selection process. Organizations should take care during the negotiation process to retain the flexibility to administer and score tests as needed. In some instances, the union may advocate including specific tests or specifying cutoff scores in the collective bargaining agreement. This can be problematic in a number of ways. If the job changes, new selection tests may be warranted yet impossible to implement due to the agreement. Likewise, in situations where cutoff scores are specified in the labor agreement, organizations lose the flexibility to adjust to changing labor markets or organizational needs. In the worst-case scenario, stipulations in the labor agreement are no longer legally defensible. Since the union is not responsible for defending the selection procedure against a legal charge, they have little motivation for considering the future legal ramifications of decisions made during the negotiation process. Consequently, it is prudent for the organization to either include a selection professional in these decisions or ensure that those responsible for making the decisions are fully informed.

Another important consideration when working with unions is whether to include union leadership and their members in the selection tool development and validation process. To increase buy-in and promote transparency, it is certainly desirable to do so, but union members' willingness to participate will depend largely on the relationship between the union and company management, as represented employees are very unlikely to participate in any type of data collection without the approval of union leadership. Without their participation, the success of the process depends on either the availability of non-represented employees in the same or similar positions or the ability to conduct a predictive validation study with applicants, which will take much longer and increase costs. Direct union involvement also increases buy-in, helps ensure the ultimate quality of the selection process, and helps avoid questions and issues post-implementation.

Throughout this process, candor and clear communication are critical. If the union suspects that the selection professional is hiding something, cooperation will become almost impossible. This does not mean sharing every aspect of the selection process; rather, it means being clear about what you will (and will not) share and why. For example, many companies consider selection test scoring formulas and cutoff scores to be secure information, and only individuals with a true need to know and who have been trained in how to interpret the information are allowed access to them. In instances like this, the important thing is being clear about why you cannot share the information and explaining that many other stakeholders (e.g., applicants, employees, hiring managers, and recruiters), not just the union, are not given access to the information either. In addition, you have to tailor your messaging to the audience. While company management will typically be interested in how the selection process provides value to the organization, union leadership will want to understand how the process impacts their members.

If prior union–management relations have been constructive, both sides are included in the development of the selection program, and communication is clear throughout the process, it should result in a rigorous selection process and a smooth implementation. However, if one or more of those three components is missing, union cooperation is unlikely, and if implementation succeeds at all, grievances are likely to follow.

### **Applicant Population Issues**

During the early phases of designing a selection system, test developers should critically evaluate the probable applicant pool. For example, an organization with a large, well-qualified applicant pool can afford to be more selective and still meet hiring needs, so it could choose to develop

a rigorous selection process with multiple hurdles. In contrast, if there is a dearth of qualified applicants, the organization may need to limit its selection criteria to only the most critical factors to ensure enough applicants can pass all steps of the selection process.

Another applicant pool factor that can impact selection system design is the level of job-relevant experience or skill applicants are likely to possess. Aptitude tests might be most appropriate if the pool is filled with inexperienced workers, but if an organization anticipates mainly experienced, journeyman-level applicants, job knowledge and work sample tests could be used.

The level of test-taking experience and skill may also be important for many blue-collar applicant pools. The average age of some blue-collar applicant pools may skew higher, meaning that many of the applicants have not been in traditional educational or learning environments for quite some time. The goal of the selection process is to measure job-relevant knowledge, skills, and abilities (KSAs), rather than contaminating factors such as test savviness, so for these applicant populations, it is particularly critical to clearly outline what to expect in the hiring process. If aptitude or knowledge tests are used, you should also provide as many resources as possible (e.g., descriptive test brochures or practice tests) to allow applicants to familiarize themselves with the test format and ensure that only job-relevant KSAs are assessed once they test.

Selection system developers can also benefit from considering likely future changes to their applicant pools. For example, an increase in education standards or advances in technology can lead to a shift in the knowledge and skills that applicants possess. Likewise, a change in the generational composition of the applicant pool can lead to different applicant values and expectations. Advance consideration of these issues can allow developers to create a selection system with greater long-term utility.

### ***English Language Proficiency***

In certain regions of the country, a large portion of the applicant population for some blue-collar jobs may lack English language fluency. This raises the question of which language should be used for any selection tests. Regulatory bodies, such as the Equal Employment Opportunity Commission (EEOC), have filed lawsuits against companies that have required English proficiency without justifying it as needed for safe and effective performance of the job (Equal Employment Opportunity Commission, 2014). At a minimum, an organization should gather job analysis data to determine the language skills required on the job. If English proficiency cannot be established as a job requirement, it may mean that selection tests should be offered in multiple languages, or that nonverbal tests, such as Raven's Progressive Matrices, should be considered.

Assuming English proficiency can be established, it is essential that care be taken to match the reading level required on the job with that required by the test. This can be accomplished by collecting and evaluating training manuals and other written job materials. Patterning test content after these materials will ensure that the reading level required by the test is not harder than that required by the job itself. This approach will also help with stakeholder buy-in and ensure legal compliance, should the selection procedure ever be challenged.

### ***Partnerships Among Industry, Education, and Government***

In the global economy, many employers are under increasing pressure to reduce costs and enhance efficiency. This can translate into a reluctance to hire unskilled workers and invest in years of training or apprenticeship. Instead, many employers seek workers who are already skilled in a particular craft. Because the pool of skilled applicants is often smaller than the hiring needs of large organizations, one potential solution is to create a pipeline of skilled applicants through partnerships with educational institutions or government training programs. In such partnerships, the employer helps identify the skills that are needed for successful performance, and the educational institution or government program designs curriculum to support the development

of those skills. Students graduate from these programs with a base level of job skill that allows employers to spend less time and money engaging in on-the-job training.

From a selection standpoint, the obvious benefit of such partnerships is a pipeline of better-qualified applicants. Applicants coming from such partnerships may also reflect greater diversity than already skilled applicants. Of course, this only pays dividends to the organization if it hires and retains enough of the graduates to offset its investment in the partnership. Understanding this requires a supply-chain approach to staffing (Cascio & Boudreau, 2011), in which all steps of the staffing process, from an understanding of the labor pool to employee retention practices, are evaluated. Potentially more important, but less tangible and more difficult to quantify, are benefits such as company brand and goodwill in the community.

At the same time, these types of partnerships can present unique challenges to the organization. The programs providing the training, whether educational institutions or government bodies, have a vested interest in their graduates' success, so one potential difficulty is a partner who starts teaching to the employer's selection requirements rather than the full range of KSAs needed for successful job performance. This can be problematic since most selection processes target a subset of critical KSAs rather than the entire job domain. If the partner teaches specifically to the selection requirements, individuals hired from the program may have unexpected gaps in their abilities.

Organizations considering entering into partnerships with schools or government programs should be mindful of potential benefits and drawbacks. When successful, organizations can fill positions with a diverse group of well-qualified applicants, often drawn from local communities. This can have a positive impact on the organization's reputation as well as on hiring and training costs. Alternately, an organization's reputation can be harmed if it engages in these partnerships but is unable or unwilling to hire expected numbers of program graduates. The decision not to hire program graduates may be the result of changes in hiring needs or budgets, or may occur because the training program, once executed, lacked sufficient rigor to produce qualified applicants. Whatever the reason, failure to hire program graduates can lead to dissatisfaction and backlash from both the graduates and the education/government partner, ultimately harming community relations.

### SELECTION TOOL CONSIDERATIONS

In this section, we review common selection tools frequently used for blue-collar jobs. Our discussion focuses specifically on considerations for using these tools with blue-collar jobs. For a more complete treatment of many of the topics, we refer the reader to the relevant chapter(s) in Part III of this book (Categories of Individual Difference Constructs for Employee Selection).

As noted earlier, the complexity of blue-collar jobs varies greatly. Given the importance of aligning a selection system with the complexity and demands of the required work, particular care must be taken when identifying selection tools appropriate for predicting successful performance in blue-collar jobs. Measuring skills or abilities that are too complex for the position can result in inappropriate restrictions in applicant flow, at best, and legal risk, at worst. Alternately, measuring KSAs that are too simple for the position can result in new hires who cannot perform the work effectively. It is also important to identify the KSAs that employees must possess upon hire. Those KSAs that employees are trained on shortly after hire add no value to the selection process.

### Interviews

Interviews are a ubiquitous selection tool, for blue-collar jobs or otherwise. Research on interviews has consistently shown that adding structure and standardization is critical to maximizing their validity (Campion, Palmer, & Campion, 1997; Huffcutt & Arthur, 1994). In blue-collar environments, structured interviews that focus on the technical knowledge and skills needed on the job are particularly appealing. They can be less labor intensive to develop and administer than

knowledge or work sample tests, yet still allow employers to assess the knowledge and skill of external job applicants in a manner similar to job skills checklists or observational assessments, which are often used for internal blue-collar selection. With a structured, technical interview, job analysis data are used to identify critical tasks and the necessary inputs to and outcomes of those tasks. Interview questions are then developed to describe the task scenario and provide the input. Interviewees are scored on the extent to which they provide the correct outcome. Unlike many behaviorally based situational interviews, such as those described by Latham, Saari, Pursell, and Campion (1980), technically focused structured interviews can have objectively scored right and wrong answers. When such objectively scored interviews are used in combination with multiple trained interviewers, inter-rater reliability is enhanced, defensibility is improved, and candidate perceptions of fairness may increase.

Organizations seeking to use a technical interview must carefully evaluate the extent to which subject matter experts are available to administer the interview. With technical interviews, the interview panel must be fully knowledgeable of the technical subject, in order to effectively administer and score the interview. In addition, due to the open-ended nature of spoken responses to technical interview questions, the test development team must dedicate a substantial amount of time clearly defining scoring anchors and training interview panelists during the validation process.

For internal selection in union environments, the use of interviews can be particularly challenging. Depending on union–management relations, union leadership may perceive interviews as a method to allow management to circumvent the seniority system. Unions that are unwilling to accept behaviorally based interviews may be open to the use of technical interviews as a reasonable replacement for work samples or knowledge tests. If feasible, including knowledgeable union members on the interview panel will further help alleviate any suspicion on the part of the union. This also provides additional subject matter expertise and can strengthen union–management relations.

## Cognitive Ability Tests

It is a firmly established finding that cognitive ability is the best single predictor of training success and job performance for most jobs (Ghiselli, 1973; Hunter & Hunter, 1984; Schmidt & Hunter, 1998). Schmidt and Hunter estimated the average predictive validity of general mental ability (GMA) tests to be .51 for overall job performance and .56 for training success. The validity of GMA is largely driven by its impact on the acquisition of job knowledge, and this is evident in increasingly higher validities for more complex jobs (Schmidt & Hunter, 2004). As noted earlier, blue-collar jobs vary widely in the complexity of the work, from unskilled to very complex. Based on the research evidence, we would expect cognitive ability to be a valid predictor for all blue-collar jobs, but particularly so for those that are more complex (e.g., locomotive engineer, nuclear reactor operator, and demolition expert) or require significant training in order to become proficient (e.g., most skilled trades).

One meta-analysis focused specifically on blue-collar construction and skilled trades jobs in the electric utility industry (Jones & Gottschalk, 1988). Mean corrected validities for specific cognitive abilities ranged from .30 (memory) to .53 (mechanical ability). Consistent with previous research, the correlations were even higher for training criteria, ranging from .55 (memory) to .77 (quantitative ability). Other meta-analytic studies that included blue-collar jobs have also found consistently significant validities (e.g., Ghiselli, 1966; Schmidt & Hunter, 1978; Schmidt, Hunter, Pearlman, & Shane, 1979).

The importance of the complexity or learning curve of many blue-collar jobs is compounded by the need for safe performance. An inability to properly evaluate the consequences of different actions or to learn the information needed to perform the work safely can have disastrous consequences. For example, five workers were killed at an Illinois chemical plant in 2004 when a worker failed to follow instructions and opened the wrong reactor. Similarly, a 2008 explosion that killed 14 and injured hundreds at the Imperial Sugar refinery in Georgia occurred, in part, because cleaning workers did not understand the consequences of sugar dust build-up. As these



examples illustrate, for many blue-collar jobs it is critical that those who are hired have the aptitude to learn the work and perform it effectively and safely. The most effective and efficient way to do this, at least for entry-level selection, is using relevant measures of cognitive ability.

Although the utility of cognitive ability tests for blue-collar selection is firmly established, their use presents several challenges. First, given the adverse impact associated with cognitive ability tests, users should anticipate legal challenges and grievances. This mindset should guide the job analysis, test development, and validation processes to ensure they are planned properly, executed carefully, and documented thoroughly. Strategies for reducing adverse impact, such as those outlined by Ployhart and Holtz (2008), should be considered. Because a reduction in adverse impact will typically come at the expense of some predictive validity, the challenge is one of comparing different implementation strategies and determining what level of tradeoff is acceptable. In the context of evaluating how to weight multiple specific abilities in a predictor battery, Wee, Newman, and Joseph (2014) demonstrated one particularly innovative strategy for doing this based on Pareto optimization.

The face validity of cognitive ability tests with blue-collar populations is another challenge that must be considered. Applicant reactions may be negative if the link between the test and the job is not clear. This can both dissuade good applicants and increase the chances for challenges. Consequently, even if an efficient GMA test is a valid predictor, it may make sense to use a longer test battery that incorporates the specific abilities needed for successful job performance and has items that are contextualized to the work.

### Knowledge and Experience Tests

Many blue-collar jobs require specialized knowledge. Jobs such as electrician, chemical technician, or nuclear power plant operator cannot be performed safely and effectively in the absence of specialized knowledge. Placing a worker who lacks the requisite knowledge in these types of roles could have terrible consequences on safety, both for the worker and for the general public. Selection systems that measure job-relevant knowledge, through written knowledge tests, experience checklists, or job simulations, can help identify applicants who have sufficient knowledge to perform such critical work safely and effectively (Burke, Sarpy, Tesluk, & Smith-Crowe, 2002).

Knowledge tests, thus, often make sense in lieu of cognitive ability tests for the external selection of experienced blue-collar workers. Internally, they are often used for promotion in skilled trades or similar blue-collar jobs. The knowledge one needs to learn in order to become proficient in the trade is dictated by the trade itself and is typically very defined (e.g., electrical knowledge for electricians or mechanical knowledge for mechanics), so it's important to ensure that employees possess the requisite knowledge before moving into higher-level roles. In general, while knowledge tests provide an easy method to identify already skilled workers, they can be labor-intensive to develop and maintain. Defining the relevant content domain and writing items can require many hours of subject matter expert time, and, particularly in industries where technology drives rapid change, the content may need to be updated regularly. Thus, test developers are advised to consider the tradeoffs between the efficacy of knowledge testing as a way to identify skilled workers and the cost of development and upkeep.

### Work Samples

In many cases, especially in union environments, there can be heavy resistance to knowledge tests for skilled positions. When the work entails hands-on performance, job applicants often perceive written knowledge tests as too far removed from the work or too esoteric to be appropriate. In these situations, work samples or other job performance tests may be deemed more palatable by both test takers and the union (Steiner & Gilliland, 1996).

Because work samples are designed to reflect critical job content, they must be based on a thorough job analysis that includes stakeholder input and carefully weighs which tasks to include. One

question test developers must answer is whether to include rarely used but very critical tasks in the work sample. For example, line workers who perform their jobs dozens or even hundreds of feet up in the air on utility poles or towers may never be called upon to rescue an injured colleague, but the ability to perform that task can mean the difference between life and death. In this case, the criticality of the task may be sufficient to merit evaluating job applicants' ability to perform it.

For work samples that use raters to evaluate applicants' performance, consideration must be given to the specific behaviors to be scored, as well as to the qualification, training, and calibration of the raters. In order to effectively score any type of hands-on assessment, the raters must typically be experts in the subject, and time must be spent to ensure they are calibrated and will assign the same ratings after observing the same behaviors. Test developers can reduce rater variability by incorporating standardized equipment into the work sample and creating measurable evaluation criteria. For example, a machinist might be given specifications, raw materials, and tools and be asked to manufacture a part to the specifications. The assessment could be scored based upon the extent to which the part matches the specifications, as determined using calibrated measurement tools. Such an assessment should be highly reliable and provide a standardized test-taker experience.

For internal applicants, another job performance test that can be effective is a job skills checklist that requires applicants to demonstrate proficiency on a pre-selected set of skills on their current job and obtain "sign-off." Checklists are especially valuable in promotional situations, when the higher-level job requires skills that can be learned and demonstrated in the current, lower-level job. In these situations, the employee is observed performing specific tasks over the course of days, weeks, or even months. A knowledgeable assessor, such as a trainer or supervisor, makes a note when the individual has successfully completed a task on the checklist. When a certain number of tasks have been "checked off," the individual is deemed to be qualified. Such assessments have the benefit of evaluating proficiency in performing a wide range of tasks. One potential concern with job skills checklists is whether those being evaluated will have ample opportunity to perform all of the tasks on the checklist. Some tasks, although sufficiently critical to merit inclusion on a checklist, may only occur on certain shifts or at certain times of the year. This can be inherently unfair to employees who are on different shifts or trying to obtain promotions at other times of the year. Because employees who are dissatisfied with the fairness of the promotional process are more likely to engage in behaviors that are detrimental to organizational goals (McCarthy, Hrabluik, & Jelley, 2009), the test developer must carefully consider the availability of opportunities to perform tasks when creating job skills checklists for promotional purposes.

## Physical Ability Tests

Physical ability testing can be instrumental in identifying qualified job candidates for blue-collar jobs with specific physical requirements. However, test developers face unique challenges when creating physical ability tests, and their development should be approached carefully and knowledgeably. First and foremost, test developers must clearly understand the actual physical demands of the job in order to develop measures that evaluate the ability to meet those demands. They then must decide whether to develop complex physical work samples that reflect relevant portions of the job or more straightforward measures of specific physical capacities needed for successful performance. Physical capacity tests (e.g., maximum amount of weight that can be lifted or speed in running a specified distance) are often relatively simple to administer and have been shown to predict future job performance (Henderson, Berry, & Matic, 2007), but they may lack fidelity with the tasks performed on the job, which can make them subject to greater legal scrutiny (e.g., *Berkman v. City of New York*, 1982). Physical capacity tests require the test developer to engage in more comprehensive research to justify how the capacity measured relates to the tasks performed on the job. This may be necessary in situations where physically demanding job tasks are only learned after extensive on-the-job training and cannot practically be included in the selection process. Physical work samples (e.g., carry a 150-pound victim down a flight of

stairs or remove and re-set a cross arm on a utility pole), on the other hand, may be more complex to administer but are more easily linked to job task performance.

With any type of physical testing, care must be taken to protect test takers from injury and avoid inappropriate collection of medical data. If there is concern over the physical safety or health of test takers, test developers must determine if pre-testing will be used before allowing individuals to take the test. They should also identify and provide clear guidance on any situations for which the test administrator should intervene, including when a test taker engages in behaviors that are likely to cause injury.

A final challenge with physical testing is adverse impact. Physical tests, especially those focused on muscular strength and cardiovascular endurance, frequently produce adverse impact against women (Hough, Oswald, & Ployhart, 2001). Addressing adverse impact can be challenging, especially when working with incumbent populations that contain relatively few group members for whom impact is most likely. In order to include under represented group members in validation studies, test developers may have to solicit participation from outside the sponsoring organization. Test developers can further enhance the chances of selecting under represented group members by offering training programs that give potential job applicants an opportunity to build the physical skills needed for the test and the job. This not only has the potential to improve applicant diversity, but it also demonstrates the organization's commitment to creating a level playing field, which can be important should a legal challenge arise. Further discussion of physical ability tests is outside the scope of this chapter, but interested readers are encouraged to read Chapter 12 of this volume for a more in-depth treatment of the subject.

## Personality Tests

The use of personality tests for high-stakes, high-volume selection is a common practice. Several influential meta-analyses in the 1990s (Barrick & Mount, 1991; Ones, Viswesvaran, & Schmidt, 1993; Tett, Jackson, & Rothstein, 1991) brought about a resurgence in their use as a pre-employment selection tool after the field had largely abandoned them for several decades. Only the Barrick and Mount research evaluated validities separately for different occupational groups. For the skilled/semi-skilled group, Conscientiousness was the only Big Five dimension with a lower-bound credibility value that was not negative. However, the skilled/semi-skilled group included both jobs we would consider blue collar (e.g., production worker, assembler, and truck driver) and those we would not consider blue collar (e.g., clerical, flight attendant, and nurse's aide), making it difficult to draw firm conclusions.

Although personality tests are now commonly used for pre-employment selection, their practical usefulness is hotly debated (cf. Morgeson et al., 2007; Ones, Dilchert, Viswesvaran, & Judge, 2007; Tett & Christiansen, 2007). Much of the debate centers around the potential for intentional response distortion on the part of applicants. Attempts at reducing or eliminating response distortion through alternative item designs (e.g., forced-choice scales) or measurement approaches (e.g., those based on item response theory) have shown promise but continue to show mixed results. (We refer the interested reader to Ziegler, MacCann, & Roberts (2012) for a thorough discussion of the topic.) One practical approach to dealing with response distortion in high-volume selection contexts is to use personality tests in a select-out rather than a select-in manner. Mueller-Hanson, Heggstad, and Thornton (2003) found that response distortion in a simulated selection setting was much more prevalent at the top end of the distribution, and criterion-related validity was significantly higher at the lower end of the distribution. Given this finding, rather than selecting the highest scorers by rank ordering or using a high cutoff score, an approach that may yield greater utility is to use the test as an early screen in the process to eliminate the lowest scorers and then use other methods to identify the best candidates.

In addition to considering a select-out approach, we would give blue-collar selection practitioners two other pieces of advice when evaluating personality tests for pre-employment selection. First, pay close attention to predictor-criterion alignment when validating personality tests. Because of the breadth of blue-collar jobs, the importance of contextual performance factors varies greatly.

## Robert P. Michel and Shannon Bonner

For example, some blue-collar work is carried out in a team environment or involves extensive customer contact, whereas other blue-collar jobs are less dependent on interpersonal interactions than on solitary performance. Some blue-collar work involves tasks that require attention to detail and careful focus (e.g., electrician), and, as discussed earlier, safety is an important consideration for many blue-collar jobs (e.g., public safety jobs, utility jobs, and manufacturing jobs). Consequently, the use of personality tests for blue-collar selection requires carefully aligning predictor constructs with facets of job performance that are important to a particular job or job family.

Second, do not rely on concurrent validation designs as the only source of validity evidence. While research on the magnitude of the loss of validity caused by intentional response distortion is mixed, most researchers agree at this point that applicants can and do distort their responses. The Army's experience transitioning from research to operational settings is a stark example of the false security that validation with incumbent samples can provide (White, Young, Hunter, & Rumsey, 2008). If a pure predictive validation study is not possible, we advise that you at least follow up a concurrent study with predictive evidence by tracking those hired once the test is implemented and collecting performance data at some later point (e.g., after six months or one year on the job).

### Biodata

The only non-cognitive predictor other than personality for which we could find validity evidence specific to blue-collar jobs is biodata. Jones and Gottschalk (1988) found little support for biodata as a predictor of training or job proficiency criteria in their meta-analysis, but in his literature review, Pannone (1994) found mostly successful examples of biodata predicting important criteria for blue-collar jobs, including tenure, job performance, training performance, and test performance (in later stages of the selection process). Another study demonstrating impressive results for biodata was conducted by Jacobs et al. (1996). In a large-scale investigation of bus operator performance, two different biodata scales together significantly predicted both subjective (i.e., supervisor ratings of safety, attendance, and customer service) and objective (i.e., absences) criteria.

One important consideration when using biodata, or other non-cognitive measures, with blue-collar populations is face validity. Pannone (1994) posited that biodata face validity is a more critical concern with blue-collar than white-collar applicants, making prior training and work experience key components of any biodata measure used for blue-collar selection. He also highlighted a potential dilemma this poses, since face-valid biodata forms tend to be more easily faked. This can be particularly challenging for internal selection, as unions will almost always oppose measures that bear little resemblance to actual work behavior (Bownas, 2000).

### Realistic Job Previews

Given the unique working conditions for many blue-collar jobs, it is particularly important that applicants develop a clear understanding of the work environment during the selection process. In fact, selection practitioners should consciously approach selection system design for blue-collar jobs as if two decisions are being made—the organization deciding who they want to hire and applicants deciding if the job is the right fit for them. The best way to ensure that applicants fully appreciate the work environment is through the use of realistic job previews (RJPs). An RJP can take many forms, and the specific demands of a particular blue-collar job may make certain approaches more effective than others. In some instances, a simple written RJP that clearly outlines the working environment may be sufficient. In other instances, an RJP video is the best way to demonstrate the working environment. If practical, an interactive simulation in which the applicant answers questions based on information provided in the RJP may be even more effective. Finally, if the job lends itself to it and the consequences of a poor hire are particularly costly to the organization, more resource-intensive RJP approaches, such as ride-alongs or job shadowing, may be warranted.

RJPs can also be incorporated into other selection tools. Interviews are a particularly easy way to add RJPs to the selection process, by having the interviewer describe the work environment or specific job demands and asking applicants for their reactions or having them describe similar environments in which they have worked. Another possibility is the use of high-fidelity computer simulations that measure job-relevant KSAs while simultaneously giving the applicant a preview of the actual work.

Raising applicants' understanding of the actual work environment through RJPs can accomplish two goals. First, it can discourage individuals who prefer not to work in the job's true environment from continuing with the application process. Second, it can clarify potentially misguided perceptions and attract applicants who might otherwise avoid working conditions they perceive to be undesirable (Premack & Wanous, 1985). Both of these RJP goals are important for blue-collar jobs. Clearly, the unique working conditions for many blue-collar jobs are not for everyone. For applicants for whom this is true, it is in their and the organization's best interests to realize this during the selection process, before they have each dedicated time and resource on what ultimately turns out to be a poor fit. On the other hand, stereotypical perceptions of blue-collar work environments may lead some applicants to believe that working conditions are worse than they actually are or to overlook potential benefits of the work environment that they had not considered. Providing applicants with as much information as possible about the work environment can accomplish both goals.

### SELECTION PROCESS CONSIDERATIONS

Chapter 16 of this volume provides a general overview of issues to consider when administering assessment tools. In this section, we address several test administration considerations that are particularly salient in blue-collar selection contexts.

#### Standardization

Standardization of the testing process should be a primary concern for any selection system, and this is particularly so for blue-collar selection. First, for large organizations with blue-collar jobs, having a standardized selection process allows the organization to have confidence in the skill level of employees who move from one location to another. Second, many blue-collar jobs are unionized, and differences in details as small as the specific model of calculator or spell checker used by test takers (*Delta Twp. v. F/F Assn. of Michigan*, 1998), the noise level in the room during testing, or the location of test administration (*Palm Beach County Sheriff's Office v. PBC Police PBA*, 2005) can lead to grievances that can undermine the entire testing program.

Test developers can take several steps to facilitate standardization. First, test developers should exercise care in clearly specifying all aspects of test administration protocol. This includes everything from room configuration (e.g., spacing between workstations) to testing aids (e.g., model of calculator, number of pieces of scratch paper) to scripts for test administrators to read aloud to candidates when scheduling and administering tests. Test developers should also create a standardized process to communicate test results. Once the entire test administration process has been thoroughly documented, test developers can conduct training for test administrators. In situations where accuracy is most critical, such as union environments or large-scale hiring, organizations should consider developing a test administrator certification process and conducting spot checks to confirm that the entire protocol is being followed as designed.

#### Test Security

Given the high-stakes nature of most blue-collar selection and the development cost associated with many selection tools, test security breaches can be particularly costly. Test developers can take

many precautions to reduce the risk of security breaches. First, when developing a test, caution should be taken in selecting participating job experts. Job experts, especially those involved in test creation, should be vetted to ensure they have no history of dishonesty and no conflicts of interest related to the test. This may mean excluding trainers or educators, who are often evaluated on the extent to which their students can pass tests. In addition, all participating job experts should be made aware of the importance of test security and sign confidentiality agreements that articulate what information about the test development process can be shared and with whom.

Additional steps that can be taken during the test development phase to maximize test security include creation of multiple or adaptive versions of the test (Guo, Tay, & Drasgow, 2009) or item pools from which different content is drawn for each test administration (Zhang, Chang, & Yi, 2012). Because creating multiple or adaptive forms entails extensive additional work, such an approach is recommended for high-volume situations and situations where sufficient data can be collected during the test validation process to effectively create multiple versions. For high-volume, ongoing hiring, using adaptive testing or multiple, equated test forms is recommended. For positions that are filled in classes, such as police officers, there may be a bit more flexibility in variance between test forms if new forms are developed for each class.

Test administration protocols can also have an impact on test security. Although unproctored testing is increasingly popular because of the efficiency benefits, it increases item exposure and creates an opportunity for test takers to utilize assistance on the test. Using proctored testing, and training proctors to actively monitor for cheating, greatly reduces both of these concerns. Testing policies, such as limiting the number of attempts on a test or requiring a waiting period before a retest is allowed, can also help improve test security.

Another, sometimes overlooked, threat to test security is unwittingly allowing test content, scoring keys, or scoring formulas to enter the public domain in response to a request or subpoena from a regulatory agency, union, or plaintiff. Without the proper protections in place, these materials are potentially accessible by third parties, such as competitors or applicants. In the case of an agency demand, once the government possesses the materials, they can be accessed through a simple Freedom of Information Act (1967) request. Agencies, unions, and plaintiffs do not have the same sensitivity or motivation to protect secure test material, so it becomes the responsibility of the test developer and test user to preserve security, a professional and ethical responsibility outlined clearly in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). If the request for test materials is unreasonable or irrelevant to the issues being investigated, test users are within their rights to refuse disclosure. However, if the request is legitimate, a common remedy is to execute a protective order, limiting disclosure of the secure materials to qualified testing experts under secure conditions.

Preventing test security breaches should always be the primary goal, but identifying them once they have occurred is also important. Test developers should monitor test performance over time to look for changes in pass rates that could indicate test security breaches (Guo & Drasgow, 2010). Monitoring should account for differences within locations, administrators, or versions of the test, using software to help analyze patterns of test answers in search of evidence of possible cheating.

## Technology

Technology can aid in various aspects of test administration, and organizations involved in blue-collar selection must weigh issues of complexity, fidelity, and cost when deciding which technologies make sense to incorporate in the selection process. Part VIII of this book deals with the role of technology in employee selection, and we encourage blue-collar selection practitioners to read the relevant chapters to develop a more thorough understanding of the issues involved. In this section, we discuss the most commonly encountered issues in our experience.

Although research findings are mixed, there is support that technology can be used to administer tests in ways that garner positive test-taker reactions (Bauer, Truxillo, Mack, & Costa, 2011).

Many types of tests, particularly those that use multiple-choice items (e.g., aptitude, knowledge, and non-cognitive tests), easily lend themselves to computer-based test (CBT) administration. CBT delivery can also be used for more complex assessment types that incorporate videos or interactive simulations, and many CBT delivery systems offer the ability to create a robust test-taker experience. At even higher levels of sophistication, adaptive tests can reduce test length by more efficiently and precisely determining a test taker's skill level by selecting test items based upon real-time test performance (Tonidandel, Quinones, & Adams, 2002).

CBT systems can also reduce test administrator burden. Some systems display test instructions on screen or announce them via pre-recorded audio, eliminating the need for a human test administrator. CBT systems can also manage the timing of the test and score tests without human interaction. These features can improve the accuracy and reliability of test scores by eliminating or reducing human error during administration or scoring.

Although computerized testing offers many benefits, its use with some blue-collar jobs can present challenges. In terms of efficiency, CBT is most effective when used in an unproctored setting, eliminating the need for test administrators and, in some cases (e.g., unproctored Internet testing), space and computer resources. However, many types of tests used for blue-collar selection (e.g., aptitude and knowledge tests) require strict test security and a need to minimize opportunities for cheating. Technology (e.g., computer adaptive testing and remote monitoring) and follow-up verification testing provide ways to address some of the concerns, but these approaches are not foolproof. They also require extensive resources to develop, making them cost prohibitive for many organizations. In addition, even if these approaches are successful at eliminating cheating at the group level (e.g., the impact on validity is minimal), in many blue-collar selection environments the testing process is under constant scrutiny, and a single incident of confirmed cheating or a security breach can create significant problems for the organization.

Another issue with the use of technology in blue-collar selection environments is the availability of computer resources in some locations. In many organizations that employ a large number of blue-collar workers, much of the hiring takes place at dispersed or remote operational sites. In other instances, affordable technology may not be capable of withstanding the necessary environmental conditions. Mobile devices can help overcome some of these challenges (e.g., using tablets or cell phones to complete job skills checklists in the field), but they may not be suitable for many types of assessments. Consequently, test developers should pre-test technological solutions to validate that they work as intended and reduce the burden for end users. Otherwise, paper-and-pencil alternatives should be considered.

A final challenge with CBT in blue-collar selection environments is the fidelity of the technology to the job. While computer use is increasingly common in the workplace, many blue-collar workers have little daily interaction with computers. In these situations, requiring job applicants to use a computer for selection testing, when a computer will not be used on the job, can pose a problem. This is especially true if the use of technology is more likely to screen out older workers, who may have had less exposure to technology during their education. Test developers should consider the match of the technology to the job, and to the applicant pool, when evaluating test delivery technology.

### SELECTION POLICY CONSIDERATIONS

It is good practice for any organization that uses employment tests to have a defined selection testing policy, but for several reasons this can be particularly true for organizations employing blue-collar workers. The progression paths, higher-than-average tenure, and presence of labor unions associated with many blue-collar jobs can lead to a myriad of situations for which the "right" answer is less than obvious without clear guidelines. A good selection policy outlines the roles and responsibilities associated with the testing function and defines key terms. Most selection policies incorporate issues such as retest intervals and how long test results are good for. Examples of additional issues that could be included are discussed below, and a sample selection testing policy is shown in Figure 34.1.

**A. Background and Purpose**

Pre-employment testing makes a significant contribution to fair, valid, and objective employee selection and placement. This Testing Policy is adopted to ensure the proper use of tests and protect the rights of individual test takers. Compliance with the letter and spirit of this policy will enable the company to adhere to legal requirements, professional standards and guidelines, and standard business practice.

All persons involved in the testing function, and especially those who have direct contact with applicants, tests, scoring keys, and test scores must uphold a high level of integrity and honesty. It is the policy of the company to treat employees and applicants equally, with dignity, fairness, and respect. Persons working in testing must be fair and impartial, and it is important to remember that even the perception of unfairness, dishonesty, or discrimination by applicants or others can have devastating effects on the testing program. If anyone encounters a situation where these perceptions might exist, he/she should immediately contact the Testing Program Manager.

**B. Key Definitions**

*Accommodation*

A modification or special arrangement in the testing process for persons with disabilities that allows the person to participate in the testing process.

*Legal Requirements*

Various laws, guidelines, and executive orders enforced by agencies that govern one or more aspects of the testing process or enforced by another federal or state agency. These agencies include the Equal Employment Opportunity Commission (EEOC), the U.S. Department of Labor, Office of Federal Contract Compliance (OFCCP), and the State and Local Fair Employment Practice Agency.

*Professional Standards*

Various guidelines published by professional associations such as the American Psychological Association (APA), the Society for Industrial and Organizational Psychology (SIOP), and the Society for Human Resource Management (SHRM).

*Reliability*

The degree of consistency in test scores.

*Retest Period*

The amount of time that must pass before an applicant can take a test again. This period varies for different types of tests depending on issues of reliability, validity, practice effects, variation in applicant test performance, and variation in administration procedures.

*Selection Procedure*

Any assessment device used as the basis for employment decisions. These include not only tests, but also minimum qualifications, interviews, reference checks, probation periods, seniority, etc.

*Test Administrator*

The person who administers and scores tests. This person is normally a non-union employee.

*Test Result*

An interpretation of an individual's test scores in terms of expected work performance or other business-related outcome (e.g., pass/fail, accept/reject, and probability of success).

*Validity*

The job-relatedness of a test or selection procedure.

**C. General Requirements**

1. Application and Scope

This Testing Policy applies to all employment tests used by the company. All employees involved in the administration and use of employment tests must be alert to possible abuse, misuse, or other problems that arise involving the use of the tests.

2. Problems

Any problems, including complaints or grievances, should be brought immediately to the attention of the Testing Program Manager.

3. Enforcement

The Testing Program Manager is responsible for the enforcement of the Testing Policy. The testing function will be monitored through: 1) audits conducted by the Testing Program Manager, 2) statistical monitoring of test data, and 3) other procedures as appropriate.

**FIGURE 34.1 Sample Selection Testing Policy**



#### 4. Violation

To maintain the integrity of the testing program, appropriate corrective action will be taken in response to any policy violations. The Testing Program Manager (and direct supervisor of the violator as appropriate) will review each alleged case of violation and determine appropriate action. Depending on the intent and severity of the violation, corrective action may include suspension from testing activities, re-training, de-certification, and/or discipline, up to and including discharge by the direct supervisor.

#### **D. Certification of Test Administrators**

All Test Administrators must be certified. The certification process includes attendance in Test Administrator training conducted by the Testing Program Manager or his/her designee. Certification also requires passing the Test Administrator certification test.

#### **E. Test Administration**

Tests must be administered only by certified Test Administrators. Administration instructions are to be followed verbatim to ensure that all applicants are treated the same. The Test Administrator should remain in the testing room at all times. This practice ensures the security and integrity of the testing session and test materials.

Questions, other than routine ones, raised by applicants should be referred to the Testing Program Manager.

#### **F. Security and Storage**

All support staff that handle tests share the responsibility for maintaining test security. Handling of tests and materials must be accomplished in a manner that safeguards their security.

- All test booklets, answer sheets, and scoring keys must remain locked in the test storage room when not in use. When tests are being administered, it is essential that test materials remain in the possession of the person administering the tests. As soon as practical, the tests should be returned to their normal storage location.
- The test storage room and testing rooms that contain test materials or equipment must be kept secure at all times. No one outside of certified Test Administrators should have access to test storage areas. Any other requests for access should be directed to the Testing Program Manager. Keys to these areas must remain in the possession of certified Test Administrators, and should not be duplicated.
- When appropriate, test materials must be securely destroyed (i.e., shredded) in the presence of a certified Test Administrator.
- Any loss or compromise of tests or test materials must be reported immediately to the Testing Program Manager.
- Tests should not be lent or shown to anyone other than certified Test Administrators (except applicants during bona fide testing sessions). Requests for test materials by anyone else should be directed to the Testing Program Manager.

#### **G. Reporting Test Results and Scores**

Users of test results must be trained in their interpretation. They must understand the test results with which they work, and they must protect the privacy of the individual test taker. Test results and scores should be reported on a *need-to-know* basis, and should be reported only by and to the appropriate person(s). No person is entitled to test results by virtue of rank, position, or title.

#### **H. Retest Periods and Results Expiration**

Each company selection procedure has a specified *retest period*. Retest periods are based upon issues of reliability, validity, practice effects, applicant developmental activities, and administrative issues. Skill and knowledge tests have a 3-month retest period, and all other tests have a 6-month retest period.

Skill and knowledge test results remain in effect for 1 year, unless the skill or knowledge being tested (and thus required for successful job performance) changes. The results for all other tests are good indefinitely unless there is a significant change to a test, in which case it will be considered a new test.

#### **I. Grandparenting and Exemptions from Testing**

Employees currently in a job progression when a selection test is implemented do not have to meet the testing requirement. Except for skill and knowledge tests, employees who leave a job after a test is implemented and later repost for the same job are exempt from meeting the testing requirement, as long as they had held the job for 6 months or more AND had a performance rating of satisfactory or higher when they left the job. In the case of skill and knowledge tests, they are exempt from meeting the testing requirement if they return to the same job within 1 year of leaving it OR unless the skill or knowledge being tested has changed, as defined in Section H.

#### **J. Test Accommodations under the Americans with Disabilities Act (ADA)**

Prior to scheduling testing, and again prior to administering any tests, all applicants must be provided with an opportunity to request an accommodation in the testing process under the ADA. Any applicant who requests an accommodation must be referred immediately to the Testing Program Manager. The Testing Program Manager will review the request with the company's accommodation review team, and take the necessary steps to ensure that appropriate accommodations are made.

## Grandparenting Rules

While all external applicants should be required to take a selection test once it is implemented, the same is not necessarily true of internal candidates. Unlike external applicants, the employing organization already has firsthand knowledge of its current employees' job performance, and if an employee is already performing substantially similar work to that for which a selection test is designed, it likely does not make sense for him/her to have to meet the selection requirement. To deal with this, organizations typically employ "grandparenting" rules. That is, the day an employment test is implemented, all internal employees currently in the line of progression are exempted from ever having to pass the test. This can also apply to employees who are not currently in the line progression but who have previously performed the work successfully.

One gray area around grandparenting rules concerns jobs that are not in a particular line of progression but entail substantially similar work. Assuming the KSAs required for successful performance are the same in two progressions, it would seem to make sense that grandparenting rules would extend to candidates moving between the lines of progression. The challenge is defining what constitutes "substantially similar," which is not always straightforward.

## Journeyman-level Hiring/Experience Exemptions

Because of employee development and morale benefits, as well as the lower recruiting and hiring costs, many organizations employing blue-collar employees in jobs with defined progression structures prefer to fill higher-level openings internally. However, this isn't always possible, and sometimes the organization must seek experienced applicants to fill higher-level openings. In these situations, the organization must decide whether the same selection process for entry-level employees will apply to experienced applicants. The employer has three options in this case: (1) require experienced applicants to complete the same selection process as entry-level applicants, (2) modify the selection process by substituting components that better apply to experienced applicants, or (3) waive specific components of the selection process.

The first option ensures the most consistency, but it may not make sense in all instances. For example, it is likely desirable to at least modify the interview to include questions that focus on the applicant's relevant experience. It also may make sense to use a knowledge test in lieu of, or in addition to, any tests that entry-level applicants must take. Waiving selection requirements for experienced applicants may seem intuitively appealing to some in the organization, but this is rarely the most appropriate route. Those who argue for this approach typically lack an appreciation of the potential impact of different selection, performance, and promotion standards of other employers for whom applicants have worked. One instance where waiving selection requirements might make sense is when states or localities have very standard, verifiable criteria for obtaining a journeyman card.

## Selecting for Higher-level Jobs

For some blue-collar jobs, entry-level positions are sometimes unskilled or semi-skilled positions that become "feeder" jobs for skilled craft positions within the organization. In these instances, it is in the organization's best interest to ensure that those hired into the feeder jobs have the ability to successfully learn and perform the work of the skilled craft positions they will ultimately assume. The best way to do this is to hire entry-level workers using selection procedures validated for the higher-level job. However, if a selection procedure has adverse impact, it is only acceptable to do this if the majority of employees in the feeder job progress to the higher-level job in a reasonable amount of time (*Uniform Guidelines on Employee Selection Procedures*, 1978), and it is the organization's responsibility to evaluate this.

Establishing that the promotion rate supports selecting for the higher-level job is easy when there is a specified training or probationary period after which all employees in the feeder job

move into the higher-level job. However, if progression isn't so automatic, it requires an evaluation of historical promotional data. This can be fairly straightforward if employee movement has been regularly and carefully logged in an applicant tracking system, but it can become quite challenging if recordkeeping has been shoddy or inconsistent. It is also important to be cognizant of "seasoned" versus "unseasoned" data. For example, to determine the specific percentage of employees who have progressed into a higher-level job within a five-year timeframe requires using employees who were hired at least five years ago. This is because it remains unknown whether those hired since then who are still in the feeder job will progress to the higher-level job within five years.

Once it is established that the progression rate supports the use of a selection procedure for a higher-level job, the organization must monitor the progression rate over time. It is possible that the progression rate supports the use of the procedure at one point in time but may not 5, 10, or 20 years later. This can be due to internal organizational changes (e.g., progression structures) or external economic changes (e.g., downturn in the economy that impacts the retirement rate in the higher-level jobs creating fewer opportunities for the entry-level job).

### Contingent Workers

Many organizations employing blue-collar workers use contingent workers (e.g., contractors and temps) to address short-term or periodic needs. When hiring contingent workers, one question is whether to put them through the same selection process as regular full-time employees occupying the same position. If the same process is used, it can take longer for the employer to fill needed slots, removing some of the flexibility associated with using contingent workers. However, if contingent workers are not required to complete the same selection process, it can create several land mines for the employer.

First, if full-time employees are working alongside contingent workers who did not have to complete the same selection process, it becomes easier to question the business necessity of the process. This may not be an issue if contingent workers are used to complete a short-term assignment, since they likely will not be performing the full scope of the job. However, the risk increases the longer the contingent workers are retained, so if the employer knows it will be a longer-term assignment, it makes sense to either put contingent workers through the full selection process or intentionally narrow the scope of their work up front to avoid any challenges to the business necessity of the selection process down the line.

Another political issue that can arise is when a supervisor really likes a contractor or temp and wants to bring them on full time. If the contingent worker did not have to complete the selection process when he/she was initially hired and is now unable to meet the selection criteria to become a full-time employee (e.g., fails a pre-employment test), the selection practitioner should expect pushback. The supervisor may ask for an exception to the hiring criteria or may openly question the value of the selection tool. In either instance, it can quickly undermine the entire process. In addition, the Internal Revenue Service (IRS) has recently settled several major co-employment cases with large employers (e.g., Microsoft in 2007). The IRS's rules on co-employment are complex, so organizations are well served to seek legal counsel when considering how, or if, to apply internal selection procedures to contingent workers.

### CONCLUSION

The basic principles of selection for blue-collar jobs are similar to those for other types of jobs, but contextual, process, and policy considerations unique to blue-collar environments must be attended to when designing selection programs for blue-collar jobs. The work environments for many blue-collar jobs are much different than those for white-collar jobs, and blue-collar employees tend to stay with their employers much longer than other workers do. In addition, many blue-collar jobs are unionized, which necessitates a different approach and philosophy to

selection system design. The selection practitioner must also be aware of applicant population characteristics for some blue-collar jobs that differ from white-collar applicant populations in important ways (e.g., education level, test savviness, and language proficiency).

All of these issues can impact the types of selection tools that are useful or appropriate for blue-collar selection. For entry-level selection, cognitive ability tests have particular utility for blue-collar jobs because learning and knowledge acquisition is often a key component of these jobs, particularly skilled trades that use a journeyman structure and complex blue-collar jobs with far-reaching safety implications. For experienced hires, knowledge tests and work samples are typically more appropriate, and for internal selection in union environments, objective selection procedures that look like the job will be most preferred.

As far as selection policy, the defined promotional path for many blue-collar jobs has important implications. It can necessitate focusing on jobs beyond the entry level when designing and validating selection processes and require implementing grandparenting policies to define which internal candidates must complete the selection process and which are exempt. The use of contingent workers in many blue-collar industries also has important implications for selection policies, as does the hiring of experienced employees in journeyman progressions.

Employee selection research specific to blue-collar jobs is scant and often inconsistent. Given the prevalence of blue-collar jobs in the workforce and their importance to the economy, more research is warranted. There are three areas where more research would be particularly beneficial to the blue-collar selection practitioner. The first is more systematic evidence regarding the efficacy of non-cognitive predictors for blue-collar work. Ability, knowledge, and skills tests are often used for blue-collar selection, and there is significant opportunity to complement these with non-cognitive tests if more consistent research can be brought to bear to help guide these efforts for practitioners. Second, with the high average tenure in many blue-collar industries, research on the longitudinal effectiveness of different predictors would be instructive. That is, as blue-collar employees gain experience and become acculturated to the work environment, do some predictors show more (or less) effectiveness as time passes. Third, given the prevalent use of cognitive ability and physical ability tests for blue-collar selection, continued research on adverse impact reduction would be useful. Recent efforts at developing and expanding theories of adverse impact (e.g., Cottrell, Newman, & Roisman, 2015; Outtz & Newman, 2010) should help guide these efforts.

## REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance. *Personnel Psychology, 44*, 1–26.
- Bauer, T. N., Truxillo, D. M., Mack, K., & Costa, A. B. (2011). Applicant reactions to technology-based selection. In N. T. Tippins & S. Adler (Eds.), *Technology-enhanced assessment of talent* (pp. 190–223). San Francisco: Jossey-Bass.
- Berkman v. City of New York, 536 F. Supp. 177 (E. D. N. Y. 1982).
- Bownas, D. (2000). Selection programs in a union environment: A commentary. In J. Kehoe (Ed.), *Managing selection in changing organizations* (pp. 197–209). San Francisco: Jossey-Bass.
- Bureau of Labor Statistics. (2015). *Union members summary [Economic news release]*. Retrieved from www.bls.gov/news.release/union2.nr0.htm
- Burke, M. J., Sarpy, S. A., Tesluk, P. E., & Smith-Crowe, K. (2002). General safety performance: A test of a grounded theoretical model. *Personnel Psychology, 55*, 429–457.
- Campion, M. A. (1983). Personnel selection for physically demanding jobs: Review and recommendations. *Personnel Psychology, 36*, 527–550.
- Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology, 50*, 655–702.
- Cascio, W. F., & Boudreau, J. W. (2011). Utility of selection systems: Supply-chain analysis applied to staffing decisions. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol 2: Selecting and developing members for the organization* (pp. 421–444). Washington, DC: American Psychological Association.

- Cottrell, J. M., Newman, D. A., & Roisman, G. I. (2015). Explaining the black-white gap in cognitive test scores: Toward a theory of adverse impact. *Journal of Applied Psychology, 100*, 1713–1736.
- De Bacquer, D., Van Risseghem, M., Clays, E., Kittel, F., De Backer, G., & Braeckman, L. (2009). Rotating shift work and the metabolic syndrome: A prospective study. *International Journal of Epidemiology, 38*, 848–854.
- Delta Twp. v. F/F Assn. of Mich., M.E.R.C. # A97 J-0057, 111 Lab. Arb. Rep. (BNA) 936 (1998) (Sugerman, Arb.).
- Equal Employment Opportunity Commission. (2014). *EEOC sues Wisconsin Plastics for discrimination against Hmong and Hispanic employees* [Press release]. Retrieved from <http://www.eeoc.gov/eeoc/newsroom/release/6-9-14.cfm>.
- Freedom of Information Act, 5 U.S.C. § 552 (1967).
- Ghiselli, E. E. (1966). *The validity of occupational aptitude tests*. New York, NY: John Wiley & Sons.
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26*, 461–477.
- Guo, J., & Drasgow, F. (2010). Identifying cheating on unproctored internet tests: The Z-test and the likelihood ratio test. *International Journal of Selection and Assessment, 18*, 351–364.
- Guo, J., Tay, L., & Drasgow, F. (2009). Conspiracies and test compromise: An evaluation of the resistance of test systems to small-scale cheating. *International Journal of Testing, 9*, 283–309.
- Henderson, N. D., Berry, M. W., & Matic, T. (2007). Field measures of strength and fitness predict firefighter performance on physically demanding tasks. *Personnel Psychology, 60*, 431–473.
- Hoffman, D. A., Jacobs, R., & Landy, F. (1995). High reliability process industries: Individual, micro, and macro organizational influences on safety performance. *Journal of Safety Research, 26*, 131–149.
- Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment, 9*, 152–194.
- Huffcutt, A. I., & Arthur, W. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology, 79*, 184–190.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72–98.
- Jacobs, R. R., Conte, J. M., Day, D. V., Silva, J. M., & Harris, R. (1996). Selecting bus drivers: Multiple predictors, multiple perspectives on validity, and multiple estimates of utility. *Human Performance, 9*, 199–217.
- Jones, D. P., & Gottschalk, R. J. (1988). *Validation of selection procedures for electric utility construction and skilled trades occupations: Literature review and meta-analysis of related validation studies*. Washington, DC: Edison Electric Institute.
- Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. *Journal of Applied Psychology, 65*, 422–427.
- McCarthy, J., Hrabluik, C., & Jelley, B. (2009). Progression through the ranks: Assessing employee reactions to high-stakes employment testing. *Personnel Psychology, 62*, 793–832.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*, 683–729.
- Mueller-Hanson, R., Heggstad, E. D., & Thornton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology, 88*, 348–355.
- O'Connell, M. S., & Delgado, K. (2011). Safer hiring. *Industrial Management, 53*, 24–30.
- Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology, 60*, 995–1027.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance [Monograph]. *Journal of Applied Psychology, 78*, 679–703.
- Oultz, J. L., & Newman, D. A. (2010). A theory of adverse impact. In J. L. Oultz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 53–94). New York, NY: Routledge.
- Palm Beach County Sheriff's Office v. PBC Police PBA, AAA Case No. 32–390–100713–04, 121 Lab. Arb. Rep. (BNA) 1624 (2005) (Smith, Arb.).
- Pannone, R. (1994). Blue collar selection. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Bio-data handbook: Theory, research, and use of biographical information and selection and performance prediction* (pp. 261–273). Palo Alto, CA: CPP Books.
- Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology, 61*, 153–172.
- Premack, S. L., & Wanous, J. P. (1985). A meta-analysis of realistic job preview experiments. *Journal of Applied Psychology, 70*, 706–719.

- Schmidt, F. L., & Hunter, J. E. (1978). Moderator research and the law of small numbers. *Personnel Psychology*, *31*, 215–231.
- Schmidt, F. L., & Hunter, J. E. (1983a). Individual differences in productivity: An empirical test of estimates derived from studies of selection procedure utility. *Journal of Applied Psychology*, *68*, 407–414.
- Schmidt, F. L., & Hunter, J. E. (1983b). Quantifying the effects of psychological interventions on employee job performance and work-force productivity. *American Psychologist*, *38*, 473–478.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274.
- Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Applied Psychology*, *86*, 162–173.
- Schmidt, F. L., Hunter, J. E., Pearlman, K., & Shane, G. S. (1979). Further tests of the Schmidt-Hunter Bayesian validity generalization model. *Personnel Psychology*, *32*, 257–281.
- Steiner, D. D., & Gilliland, S. W., (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology*, *81*, 134–141.
- Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology*, *60*, 967–993.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*, 703–742.
- Tonidandel, S., Quiñones, M. A., & Adams, A. A. (2002). Computer-adaptive testing: The impact of test characteristics on perceived performance and test takers' reactions. *Journal of Applied Psychology*, *87*, 320.
- Uniform Guidelines on Employee Selection Procedures*. 29 C.F.R. § 1607 *et seq.* (1978).
- Vredenburg, A. G. (2002). Organizational safety: Which management practices are most effective in reducing employee injury rates? *Journal of Safety Research*, *33*, 259–276.
- Wee, S., Newman, D. A., & Joseph, D. L. (2014). More than *g*: Selection quality and adverse impact implications of considering second-stratum cognitive abilities. *Journal of Applied Psychology*, *99*, 547–563.
- White, L. A., Young, M. C., Hunter, A. E., & Rumsey, M. G. (2008). Lessons learned in transitioning personality measures from research to operational settings. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *1*, 291–295.
- Zhang, J., Chang, H., & Yi, Q. (2012). Comparing single-pool and multiple-pool designs regarding test security in computerized testing. *Behavior Research Methods*, *44*, 742–752.
- Ziegler, M., MacCann, C., & Roberts, R. D. (Eds.) (2012). *New perspective on faking in personality assessment*. New York, NY: Oxford University Press.