8

STAFFING AND TRAINING

In this chapter, we describe two of the purposes that are dearest to the hearts of industrial and organizational psychologists, namely, staffing and training. Some have referred to these functional activities under the broader label of talent management. Staffing involves attracting people to the organization and placing them in jobs where they will be of benefit to the organization. In large companies, staffing also concerns plans for the future. It is important to have a pool of people that we are developing for jobs that have increasing responsibility. In staffing, the job is considered essentially fixed or given, and people are chosen so as to match the job.

In training, people are given experiences that enhance their knowledge, abilities, and skills. Thus, rather than being picked for their match with the job, people are changed to better fit the job. More than any other purpose in human resources, training has a well-developed systems approach that covers the entire process from design to evaluation.

STAFFING

A company without people is an empty shell. People make the place (Schneider, 1987), and finding the *right* people to fill the company is the point of staffing. There are three main staffing functions:

1. Recruitment
2. Selection
3. Placement (some aspects of which are referred to as "on-boarding")

Recruitment usually involves getting people from outside the organization to apply for a job inside the organization. However, in large companies there is often a system in which jobs are first advertised inside the company and then advertised outside the company if no suitable internal candidates are found. Thus, recruiting can be both from inside and outside the company. Recruiting is done to attract job applicants. Selection involves choosing people to hire from among the pool of job applicants. The organization tries to discover information about the applicants that will tell the company who will be best suited for the job. The applicants try to discover what it will mean to work at the organization. If all goes well, the organization and the applicant choose each other, and the applicant becomes an employee. Selection is about choosing the right person for the job. Placement, on the other hand, is about choosing the right job or position for the person, as well as taking the necessary actions to smooth the transition of the hired individual from an outsider to a productive employee.

If an organization hires someone believed to be useful to the organization but considers that person for more than one job at a time, then placement is an issue. Outside the military, placement across jobs is not usually a concern because people are hired for specific jobs rather than for a range of jobs. Therefore, we will have little to say about placement of this type in this chapter. We also do not discuss on-boarding since this involves a variety of actions beyond job analysis, such as acquainting the new hire with his or her benefits and meeting peers and company officers. We describe recruitment and selection in more detail in the rest of this chapter. We describe the functions that recruitment fulfills and show how job analysis provides the necessary information to write job specifications that communicate job requirements to job applicants.

Selection has an extensive history in industrial and organizational psychology. Books by Gatewood, Feild, and Barrick (2016), Guion (2011), Heneman, Judge, and Kammeyer-Mueller (2019), Ployhart, Schneider, and Schmitt (2006), and Schmitt and Borman (1993) offer more detail than we can provide here. Selection is closely associated with testing, so we outline procedures used for choosing and constructing tests and methods of knowing whether the tests are working for their intended purpose. We distinguish between tests that are intended to measure specific constructs (e.g., color vision) and those that are intended to provide work samples (e.g., a typing test). Finally, we show how several different methods of job analysis can be used profitably to support personnel selection activities.

Staffing is about matching people to jobs. From a staffing standpoint, jobs are typically considered fixed or set. People are to be found that fit the jobs rather than the jobs being tailored to fit people. The trick in staffing is to know what is required by the job in the way of human knowledge, skills, abilities, and other characteristics (KSAOs). Once we know what is required by the job, we can find the people with the characteristics needed to meet those requirements. Finding people that fit a job is not the only way to enhance performance; training changes people, leaders motivate people, and job redesign alters jobs (Chapter 7).

Because quite a few industrial and organizational psychologists (and many more people in businesses) have been looking at the fit between people and jobs for many years, you might expect that the techniques for creating a solid match have been worked out. You might expect that there is a reference like a dictionary where one can go to look up the job and see the associated human requirements (job requirements). Although several such systems have been devised (e.g., Ability Requirements Scales, job element method), none are yet considered definitive. Part of the problem is that although many human capacities are well understood (for example, perception of length and color), their specific applications are too varied to catalog. For example, both a cabinet installer and a graphic artist make use of perception of length and color, but the installer uses the information to fit pieces of wood together, whereas the artist uses the information to produce eye-catching compositions.

Some important human attributes are not well understood or measured. Take resistance to monotony or boredom, for example. Such a trait is very important in inspection work where the vast majority of products pass inspection, but exceptions have nasty consequences (consider poultry and tires, to name only two). In theory, all jobs can be described in terms of elements that require some human capacity. In practice, jobs are dynamic abstractions. Think of the jobs of drone operator and social media manager, which did not exist a few years ago. If there were a dictionary linking jobs to job requirements, it would either be incomplete or constantly under revision.

Whenever we move away from the performance of the task itself and instead consider what the task requires in terms of KSAOs, we have to make an inferential leap from the task to the KSAOs (Dierdorff & Morgeson, 2009). The more abstract the KSAO and the less the KSAO resembles the task itself, the greater the inferential leap and thus the greater the possibility of landing off base. It requires a fairly small leap to infer that accountants must have knowledge of addition and subtraction of numbers. On the other hand, it is a considerably larger leap to infer that an accountant job requires a certain number of years of experience, a high level of verbal ability, or a high level of resistance to stress. What we need is a classification system that simultaneously treats the essential attributes of both jobs and people. Despite some notable attempts, the required taxonomy is not yet available, so we have to proceed as best we can on a case-by-case basis.

Harvey (1991) presented an extended discussion of the problem. He described two main inferences that are typically made in selection: (1) from the job to the KSAOs and (2) from the KSAOs to the tests or other methods used to assess differences in applicants on the KSAOs. A further complication is how to synthesize and use these differences in applicants to make actual hiring decisions. On top of that, there is the sometimes mystifying process by which the chosen ones elect to take the job. Because staffing involves such additional complexities, doing even the best job analysis does not guarantee a successful selection outcome, though it does enhance the probability of success.

Recruitment

When employers have jobs to fill, they encourage people to apply for them—that is, they recruit applicants. Recruiting methods are as varied as companies;

some rely on word of mouth, and some are required by law to place written advertisements in specific publications. Accepted recruitment methods include classified ads in newspapers or trade journals, walk-ins, job fairs, college campus recruiting events, online job boards, and whatever clever activities recruiters can devise. For example, a recruiter for a software company puts her business cards in C++ (a programming language) books in bookstores (Harris & Brannick, 1999). Another puts her recruiting ads on drink coasters at a restaurant frequented by people who are likely to be good applicants (Harris & Brannick, 1999). Sometimes current employees are encouraged to recommend applicants they think would make good additions to the company. Nowadays there are many web-based recruiting and job application sites, including Indeed.com, CareerBuilder.com, Glassdoor.com, Google for Jobs, LinkedIn.com, and Monster.com, to name but a few. Recruitment serves at least the following functions:

1. To let people know that there is a job opening, including people of diverse backgrounds as part of equal employment opportunity (EEO) efforts.
2. To entice people (especially those suited to the job) to apply (i.e., to market the company as a good place to work).
3. To inform potential applicants about the requirements of the job and the nature of the company or agency.

The first function is easy to fulfill with most any approach, as long as the message is clear. Take a look at the job ads on any of the web-based recruiting sites mentioned above. A recent one reads as follows:

> Hiring Truck Drivers with CDL - Class A license for open dedicated positions with competitive pay, flexible home time, and great benefits.

This job ad clearly lets us know that the employer is looking to hire someone to drive trucks (18-wheel, semi-tractor trailers). The parts about competitive pay, flexible home time, and great benefits are intended to entice the applicant. The job requirement is a Class A commercial driver's license (CDL). The job title clearly suggests driving skills and experience are expected, although that is not explicitly stated (but perhaps embedded in the licensure requirement). Formal communications such as advertisements or job listings should ideally both entice applications and screen applicants. Recruiting has a marketing function (selling the organization to the potential applicant) that has received some attention from industrial and organizational psychologists but really deserves more. Part of the lack of attention may be because the enticement function has more to do with the company (its culture, compensation package, locations, etc.) than with the job itself. Instead, job analysis tends to be most closely related to the discovery and communication of job requirements in recruitment. On the other hand, especially when unemployment is low, recruiting is very important. Psychological research ought to be able to help employers find ways to reach or create pools of qualified applicants.

Job Specification (Job Spec)

A job specification is a written description of job requirements. Job specifications offer a way of communicating what the organization is looking for. The job specification is typically written from the job description (Wernimont, 1988). It may include exactly the same types of KSAOs that are used for testing (discussed in the selection section of this chapter), such as verbal ability or resistance to stress. It may also include degrees and work experience required or other essentials, such as the insurance required for our truck driver.

Statements contained in such job specifications should tell applicants and those hiring about the requirements that the job imposes in terms of KSAOs. Thus, the obvious descriptors include licensing, job context, personal job demands, worker activities, and especially human attribute requirements. However, personnel selection is heavily regulated by law in the United States (and in some other countries as well), and the kind of job analysis required to support personnel requirements is influenced by legal considerations (e.g., Gatewood, Feild, & Barrick, 2016; Gutman, 2000; Thompson & Thompson, 1982). Job analysis is employed wherever there is interest in merit-based selection (i.e., suitability for a job based on KSAOs rather than such attributes as race, sex, or ethnicity, which are not typically related to job success).

Key Considerations for Recruitment

Job analysis for recruitment has to communicate the essential nature of the job and the most important job requirements (worker attributes). Want ads nearly always give the job title but usually give little more in the way of a job description. Want ads typically seek people with experience in the job for which they are recruiting; qualified applicants may be presumed to have a good idea about the nature of the job. On the other hand, college recruiters often give more information about the jobs they are trying to fill. Both the recruiters and the potential job applicants will want more information to assess the fit between the person and the job.

A decent job description will provide a descriptive job title and a description of the purpose and major duties of the job. However, the job description is not likely to provide a good list of worker requirements unless the job analysis was designed to do so from the start. In practice, job specifications may be written solely by looking at task descriptions without taking a look at the job itself (Wernimont, 1988). Such a practice is risky. Expect trouble if the requirements in the specification cannot be linked directly to the job analysis. For example, it is very difficult to defend requirements of degrees and length of experience (e.g., requires high school or GED and 2 years of experience). Levine, Maye, Ulm, and Gordon (1997) discussed such requirements at length. If you remember that an example of how to set these specifications (also called *minimum qualifications*) was presented way back in Chapter 1, reward yourself with a trip to your favorite coffee shop. The results of a worker-oriented job analysis (see Chapter 3) can be used to create job specifications by summarizing the most important worker requirements. Such a "direct estimation" method for determining the needed KSAOs has become increasingly popular (Morgeson & Campion, 2000) and is the basis of most of the KSAO measurement in O*NET. In any event, good job specifications allow applicants to self-screen. If the potential applicant sees a match, the applicant will throw his or her hat into the ring. If not, the applicant searches elsewhere.

Selection

Selection is about choosing the best applicants. Stated more formally, *selection* is the process of choosing from among the applicants those whom the company wishes to hire. The logic of selection is to examine the job to find the tasks and KSAOs, examine the applicants to find their standings on either the tasks or the KSAOs, and then to offer the job to those with the best (or at least some minimally qualified) predicted job performance. (It gets easier if the applicant is close kin to the company president, but that's another story.) Although the process is very simple in theory, it gets rather complicated in practice. Part of the practical difficulty comes from laws about hiring practices, and part comes from what we know about psychological tests and their use.

Test Validation

"So what are psychological tests?" you wonder. They come in a wide variety. The most common are the paper-and-pencil variety (although most are presented these days via computer, so perhaps point-and-click would be a better description). Of the point-and-click variety, we can distinguish two major types: those intended to assess maximum performance and those intended to tap typical performance. In the maximum-performance category, we have all sorts of ability tests, such as the infamous SAT and/or professionally developed tests of analogies, vocabulary, arithmetic reasoning, and so forth. Tests of specific bodies of knowledge (e.g., real estate law for realtors, local geography for cab drivers) would also fall under this category. The item types typically present a problem and ask you, the test taker, to choose the best answer from among those available. An item might be something like "Keats is to Shelley as Byron is to [Blank]," and the correct response would be "Some other dearly departed British poet." For the typical performance category, we have tests of personality, interest, and so forth. The typical item presents a statement, and the test taker indicates its applicability in some way. An item might be "I am the life of the party," and you would be asked to indicate how accurately that statement describes you. Another popular type of test is the work sample. A welder might be asked to weld something that can then be tested (perhaps by an X-ray) to see if the weld is done properly. All sorts of simulations would apply here as well, such as the assessment center, which includes a series of business simulations often used to select upper-level managers. Even though the interview is often not really a professionally developed psychological test (although it could be if one were to design a structured interview), it is a test according to the law because it is a means of gathering and evaluating information to make a selection decision. Therefore, the rules that apply to psychological tests also apply to employment interviews, both structured and unstructured.

The law allows employers to use psychological tests to help make employment decisions unless the tests discriminate against people in protected classes such as race and sex. (We presented the legal basis for job analysis in some detail in Chapter 6. Bet you're just itching to go back there and have another look!) The way testing actually works, though, is a bit more complicated. Initially, employers can use any test they want. But if the test shows "adverse impact," meaning that it tends to disproportionately exclude members of a protected group from employment opportunity, then employers are in trouble (they can lose a costly lawsuit) unless they can show that the test is job related and consistent with business necessity. Many tests are likely to show adverse impact against one group or another. For example, vocabulary tests tend to favor whites over blacks. Physical tests such as grip strength tend to favor males over females. Employers may choose not to use such tests, or they may attempt to justify their use in terms of job relatedness and business necessity.

Psychologists tend to talk about validity rather than job relatedness; validity is actually a broader concept, but it will be the key to demonstrating job relatedness. *Test validation* is what psychologists call the process of discovering the meaning of test scores. In the case of employment testing, test validation is about investigating the job relatedness of the test for the job(s) of interest. For example, we might investigate whether a vocabulary test is associated with individual differences in the job performance of bank clerks or whether a grip strength test predicts the job performance of auto mechanics. Psychologists have devised several different methods of test validation, but they are all aimed at understanding the meaning of test scores. Next, we describe the most commonly used validation strategies for employment testing. Try to think of test validation as an argument supported by theory and data. Some arguments are stronger than others, and some arguments fit certain situations better than others.

Content-Oriented Strategies

Content-oriented test validation (wow, what a compound noun) aims to link the content of the test to the content of the job. This is easiest to do when there is an obvious link between the job tasks and the test.

The argument in support of the test provided by content-oriented methods is strongest when selection is to be based on work samples, simulations, or tests of task knowledge required by the job. In such a case, KSAOs can be defined in terms of the tasks of interest and do not require much of an inferential leap from the task to the test. If the job requires welding, for example, one can devise a welding test that samples the important tasks on the job. When evidence for validity flows from judgments about test content, the job analysis should generally focus on tasks, and the major interest should be work activities. Other descriptors may also be of interest, including products and services; machines, tools, equipment, and job aids; and work performance indicators and standards. Job analysts should be used, and, from a legal standpoint, a large sample of incumbents and supervisors is ideal. Several methods of data collection are possible, including observations, interviews, technical conferences, questionnaires, and doing the work. The units of analysis should include the tasks and rating scales used to assess the tasks. According to the Uniform Guidelines on Employee Selection Procedures (EEOC, 1978; see Chapter 6), the tasks should be rated for importance or criticality, and most or all of the critical tasks should be included. The basic idea is to show that the personnel requirements match or represent the important tasks on the job.

Recall, if you will, our receiver job (Table 7.3). The list of 10 essential duties and responsibilities resulted from a process in which the job analysts first observed and interviewed incumbents; they also did part of the work (all but driving the forklift). The analysts then wrote a list of tasks and shared them with distribution center managers and human resource professionals at the company, who revised and approved the list.

To recapitulate the gist of the job, the receiver's function is to receive merchandise from the manufacturer and store the merchandise in the distribution center. There are five main steps to the job (if you skip ahead to Table 8.5, you will see each of these steps broken down into finer detail):

1. Unpack the truck. The receiver has to unload the parts from the truck using a forklift.
2. Unpack the skid. The receiver needs to verify the parts against the packing slip to ensure that the parts sitting on the loading dock are exactly what was intended and what will be paid for.
3. Receive the parts into the computer system. The receiver then enters code numbers into the computer so that the warehouse inventory is updated and the computer system can place each part into a bin, also for inventory.
4. Do paperwork and labeling. The receiver prints labels using the computer and printer and then affixes the labels to the parts.
5. Put away parts. Finally, the receiver will move the parts from the loading dock to bins specified by the computer system in the distribution center. (The parts will be subsequently pulled from the bins and shipped to a retail outlet, but that is another job.)

Suppose that we have decided to select receivers using two work sample tests. The first test will involve verifying parts against a packing list. The applicant will be given a packing slip that has the names, identification numbers, and quantities of a series of parts that are laid out on a receiving dock. The applicant has to count the number of parts of each kind and compare them to the packing list, noting any discrepancies. Some of the parts will be bundled with shrink wrap, requiring the applicant to multiply the number of bundles times the number of parts in each bundle to find the correct quantity. Some parts will be bundled so that the applicant must subtract parts from the total in order to get the correct count. There will also be some parts that are on the dock but not on the packing list, and vice versa. We will then score the test by examining the number and nature of mistakes, if any, made by the applicant.

The second test will be a forklift operation test. This test will require the applicant to pick up a full pallet using a forklift and move the pallet to a labeled bin. Once at the bin, the applicant must use the forklift to store the pallet on the top shelf. It just so happens that high storage with a forklift is hard to do, and anyone who does this kind of work often will be able to judge the skill of a forklift operator by watching this operation. We will then score this test by using a rated judgment of the applicant's competence to operate a forklift.

In Table 8.1, we have reproduced the receiver's essential duties and responsibilities from the job description in Table 7.3. We have indicated with an "X" the job content that appears on either of the two tests. You can see that the tests represent some, but not all, of the job's content. Because of their obvious relevance to the job, work sample tests are well received by managers and applicants. Such tests are relatively unlikely to be legally challenged. From the standpoint of the Guidelines, however, for a test to be justified on the basis of test content, it must sample all or nearly all of the job's content. This makes little sense to us because the tests possess obvious relevance to the job. Part of the job (knowledge specific to the operation of the computer program) would be trained rather than selected for, so there would be little point in representing that in a selection test. However, scientific practice and legal practice do not always coincide. Note also in this case that no quantitative ratings of criticality or importance were collected. Rather, the set of essential functions used by the company for the job description were considered to be the important duties and responsibilities.

Let us consider a second example of an attempt at developing a test to be defended by its content as related to the job's content. An insurance company was interested in developing a test of the decision-making of claims supervisors to be used for development (i.e., feedback to employees about their skills). In consultation with the client, the organizational consultant decided to develop a situational judgment test (SJT). An SJT presents work scenarios in which problems arise and then lists several possible courses of action. The test taker must decide which of the actions are best and worst. Items looked like this:

> You receive a call from a grumpy customer. Her claim was not resolved, and the claims representative who handled her claim is no longer with the company. What should you do?
>
> a. Find the old claims representative to find out what happened, then process the claim yourself.

TABLE 8.1 ■ Example Content Validity Linkage Between Essential Duties and Tests for the Receiver Job

Essential Duties	Packing List	Forklift
Compares identifying information from freight bills and packing lists by counting, adding, subtracting, and multiplying shipped items.	X	
Unpacks and examines products for damage and shortages.		
Contacts shipper to notify of damages and shortages.		
Records product quantities and identifying information into computer.		
Verifies computer records against physical counts and investigates and reports any discrepancies.		
Identifies location for storage of received items.		X
Prints barcode labels and put-away documents using computer and printers.		
Operates forklifts and other machines to move and lift shipments from receiving area to storage area.		X
Records quantities and location of items stored manually or with scan gun and enters information into computer.		
Works closely with packing department to ensure customer satisfaction through proper handling of product orders.		

b. Apologize to the customer. Give the claim to your best claims representative.
c. Give the claim to your newest claims representative as a learning opportunity.
d. Review the file to see what additional information is needed. Get a claims representative to collect any needed information, then resolve the claim yourself.

If you were taking this test, you would pick the best and worst choices. (This is a paraphrased version of an actual item, and sorry, we do not know the keyed responses. We guess that response *c* is the worst choice.)

Experts provide the answer keys (that is, the actual best and worst choices for the situation) that are used to score the test taker's responses. The experts are chosen because they are recognized as the best or most knowledgeable about the subject matter. In this case, top company executives with experience in the claims supervisor job were used to generate the answer keys.

For the job analysis part of the project, two analysts began by reviewing job descriptions. Then they interviewed 12 incumbents and four supervisors of the job. The interviewers asked the incumbents and supervisors about the job in question, focusing on the kinds of decisions made on a daily, monthly, quarterly, and annual basis. They recorded lots of examples. The examples were analyzed for themes. The result was essentially a taxonomy of decision-making for claims supervisors. The types of decisions were labeled (1) *identifying talent*, (2) *developing talent*, (3) *achieving business results*, (4) *creating a positive work environment*, and (5) *handling claims*. The job analysts wrote at least 10 scenarios to represent each of the five decision types. The analysts sorted each other's scenarios into the five categories to ensure that the scenarios matched the decision type. The surviving scenarios were edited and revised by supervisors of the job and representatives from the human resources department. The resulting SJT had five items per decision type. The SJT could then be linked to the content of the job through the job analysis. The idea is that the scenarios in the test represent (sample) scenarios encountered on the job.

In both these cases (i.e., the receiver job and the claims supervisor job), the tests represent a part of the content of the jobs in question. In neither case does the test represent all of the job's content. The two types of tests also vary in the fidelity with which they represent the focal job. The receiver work sample tests can be made to be quite faithful to the job by sampling actual parts and packing lists, using the actual distribution center, forklift make and model (they do come in various shapes and sizes), and so forth. The claims supervisor SJTs abstract problem situations rather than present problems as they actually occur in the workplace.

In some validation efforts where test content is at issue, experts are used to judge the degree to which the test content represents the job. Such techniques are not often used when work samples form the basis of the test in question but are more likely when a domain of knowledge is being tested. For example, someone who teaches history in high school must know a great deal about U.S. history. One might argue to test potential history teachers using a test of history (actually teachers are certified in other ways, but this example helps make clear the intent of the validation effort). Experts might be used to judge the degree to which the content of a test of U.S. history represents the domain of interest (i.e., the body of knowledge in U.S. history that might be taught). A test that was judged to be representative of the domain of U.S. history might be supported with content-oriented evidence.

In our view, the argument for validity based on test content is strongest when the test content is sampled directly from the job. For example, the receiver job clearly requires some skill in arithmetic. To verify the packing list, the receiver needs to count, multiply, and subtract. If we test such skills by having the applicant verify a packing list, the content validity argument carries maximum force. However, if we develop a very good test of arithmetic and give that to the applicant, then content validity is less persuasive because the analyst has made an inference about qualities that are necessary to do the job, and the test is justified based on that inference rather than on the performance of the task itself. It sometimes happens that KSAOs that are "obviously" important are not required by the job. For example, we once analyzed the job of photo processor and discovered that, contrary to what we believed going into the job analysis, color-blind people could do the job successfully.

Although the argument in favor of a test based on its content is strongest when it is applied to tasks that are sampled from the job, the argument is never as compelling as it might be because the inferences we make about people are based on test scores. Organizations decide whom to hire based on test scores. The content argument, however, only applies to test stimuli (i.e., the contents or items of the test) and not to the scores based on the responses to those items. Our inference is that the scores will be meaningful or valid because the test content reflects important components of the job. Therefore, support for test validity is stronger when inferences based on content-oriented strategies are supplemented with criterion-oriented strategies. This is particularly true when tests are chosen based on KSAOs inferred from the job rather than tasks sampled from the job. Criterion-oriented strategies involve examining empirical relationships between test scores and important outcomes such as job performance.

Criterion-Oriented Strategies

In many cases, the personnel requirements are not simply a sample of the job tasks. For example, we might be interested in the conscientiousness of a mechanic. There is nothing about the steps involved in adjusting brakes that indicates conscientiousness per se, but careless work in adjusting brakes can have terrible consequences. Because of the consequences of poor performance, there is a natural connection between the work information and the worker information, that is, from the tasks to the KSAOs. In such cases, it is still necessary to collect information about the tasks. However, task information alone is no longer sufficient to support the personnel requirements. In such cases, criterion-related validity evidence obtained from a validation study might be used.

A criterion (singular; plural is *criteria*) is any outcome measure of interest. In personnel selection, this is most commonly a measure of job performance (or a lack of performance, as when someone quits). In the criterion-oriented test validation strategy, the researcher shows how test scores are related to scores on a criterion measure. For our medical assistant job from Chapter 7 (see Table 7.2), for example, we might (confidentially) review patients' records over some period of time to determine the accuracy with which the assistant recorded information. We might test each medical assistant with a measure of conscientiousness (attention to detail, punctuality, personal discipline), which will provide our test scores. Then we can compute a correlation (typically the correlation coefficient; see Chapter 9) to show the association between the test and criterion scores. We would expect to see that more conscientious medical assistants would have fewer errors. Or we might get confidential evaluations from patients regarding their experiences with the medical assistants. Such patient reactions could provide a different sort of criterion and would likely correlate with scores from tests of a different attribute, such as extraversion.

Construct-Oriented Strategies

We have not yet covered the industrial psychologist's favorite method of validation, construct validation. It's our favorite because we can say in good conscience that novices should not try this method at home. In construct validation, we are trying to make several complex judgments. One is whether a construct, a variable that exists in the psychologist's imagination, fits nicely into a good theory. For example, the variable anxiety may in our theory be linked with anger or depression as well as some kinds of job performance and job satisfaction. We can then collect data to test hypotheses drawn from the theory. For example, does a measure of anxiety predict angry outbursts at work?

Construct validity deals with the meaning of test scores by examining a pattern of relationships between our focal measure and other measures that are theoretically relevant. For example, let's say we have developed a measure of anxiety for surgeons, which includes such items as "I get nervous when I see blood" and "I get distracted when other people are watching me work." We could give our new measure of anxiety to a group of surgeons and also give them another older, established test of anxiety. We would expect there to be a substantial correlation between the two anxiety measures. We could also give the surgeons a job satisfaction survey that may index things like the social support they receive. We would expect a substantial but negative correlation between anxiety and job satisfaction because high anxiety is associated with low job satisfaction. Because construct-oriented evidence of validity is concerned with a pattern of relationships among variables, it subsumes criterion-related evidence. That is, you can think of criterion-related evidence as a special case of a construct-oriented strategy. We expect our measure of anxiety to be related to a measure of job performance. For example, surgeons who have high anxiety may have unsteady hands. We could examine the relationship between anxiety and a measure of surgical performance (straight cuts, clean sutures, time to completion, etc.) taken from a simulation at a medical school. (Surgery being what it is, we prefer to avoid tests using live patients.) Sharp-eyed readers will make the connection that the types of evidence we are describing here fall under the heading of "evidence of validity based on relationships with measures of other variables" from the *Principles*. This is one of the accepted sources of validity evidence described in the *Principles* and the *Standards*, although they do not call it construct validity for some reason. We emphasize once again that there are not different kinds of validity but rather different types of evidence that bring meaning to scores on tests and other methods to assess the attributes of people.

Research yielding evidence bearing on psychological constructs is the most interesting type of research from a scientific standpoint. It offers the most information helpful in developing psychological theories. It is, therefore, the industrial and organizational psychologist's favorite approach. Construct-oriented studies are rarely conducted as such in organizations because of practical constraints. The organization wants to improve its workforce and to improve its bottom line. Organizations are typically not interested in learning more about the meaning of test scores other than whether they successfully predict job performance.

Applicant Job Experience and Validation

Applicants differ in their job experience. Our truck driver ad is intended for licensed truck drivers. In other cases, most or all of the applicants may lack experience in the job. Beginning managerial and professional jobs (manager trainee, medical residents, etc.) do not have experience in the job when they are hired. Vocational training programs for jobs such as the truck driver are also aimed at less experienced applicants. When the applicant pool contains lots of experienced people, it makes sense to evaluate them on the specific knowledge and skills required by the job. Under such circumstances, content-oriented validation strategies are reasonable. When the applicant pool is inexperienced, applicants will not have the specific knowledge and skills needed for the job. In such cases, it makes sense to evaluate the applicants on more general abilities and other characteristics needed for the job. For the truck driver, we might be concerned with multilimb coordination and interest in realistic occupations that require operating vehicles. General knowledge of transportation might also be a good construct to measure—not so much because of the specific job skills needed but because it indicates interest in the general domain of trucking.

Predictive and Concurrent Designs

There are two types of criterion-oriented strategies: predictive and concurrent. They differ primarily in the way in which data are collected, which is referred to as the "design" of the validation study. In the predictive design, job applicants are given the test. The applicants who survive the selection process are hired, and at some later point, those hired are evaluated on their job performance. Levine and Weitz (1971) offered advice based on their research that suggests how long to wait before measuring job performance. For simple jobs, they said to wait only a short time. For more complex jobs, they said to wait a good bit longer. At any rate, after the wait time has passed, we measure job performance. Then the correlation between test scores and job performance is computed. In one form of the predictive design (let's call it option A), the test under study is not used as part of the selection process. Instead, the test response sheets are locked up somewhere and later scored after the criterion measure is obtained. If this design is used, then we can see what the impact of testing is likely to be. Note that some people who would presumably have failed the test will be hired, and if the test is job related, those people should not do well on the job. Note also that if the tests are locked away, they cannot have an influence on the incumbents' supervisor or the incumbents themselves. The other (more commonly used) predictive design (let's call it option B) is to use the test for selection purposes so that some applicants are rejected for poor scores on the test. Job performance measures are gathered at a later time and compared to the test scores. Such a design provides poorer information about the value of the test for selection because people who are predicted to do poorly are rejected and therefore not available for the job performance measure. You might be thinking, "But that's the whole point of the test! You want to avoid hiring people who will perform poorly!" There is a certain force to that argument, which is why option A is not used routinely. We merely note that there is a difference between a test working and *showing* that a test is working. It will be easier to show that the test is working if you use option A. Also, in the predictive design in which the test is used, supervisors may know the test scores and therefore may treat the incumbents differently. A common source of job performance measures is supervisory ratings, and such ratings could be contaminated by knowledge of incumbent test scores.

In the concurrent design, the researcher tests incumbents and gets job performance measures from the same people at the same time (hence, the term *concurrent*). The concurrent design is nice because it allows the researcher to collect and analyze the test validation data in relatively short order. However, because current employees rather than applicants are tested, the results of the study may not apply as well to future applicants as do the results of a predictive study. Concurrent designs are more commonly used than predictive designs for practical reasons. On the other hand, unless forced by legal challenges, actual or potential, employers are not likely to carry out validation studies. This does not go over well with industrial and organizational psychologists, who love to do such research.

Synthetic Validity

All criterion-oriented strategies involve collecting data on a number of people for both tests and a criterion. The data are usually summarized with a correlation coefficient. It turns out that the correlation coefficient is not very stable or accurate in a statistical sense when it is based on small numbers of people. Although statistics books are reluctant to give minimum numbers, in our experience validation studies based on fewer than 100 people are suspect. It is good to have 250 or more people for a criterion-related validation study. The typical validation study includes people from one job. As you might guess, it is unusual for a company to have more than 100 people doing the same job. Of course, such jobs do exist, including bank tellers, clerical assistants, cashiers, and so forth. However, for many jobs and many companies the criterion-oriented strategy for a single job is not feasible. The idea in synthetic validity is that the performance of different jobs or tasks requires some of the same human attributes. For example, counting buttons and counting cash both require counting. Training bank tellers and serving customers both require some interpersonal skills. Graphic design and watch repair both require good vision. Watch repair and surgery both require precise, controlled movements of the hands and fingers, and so forth. So the idea in synthetic validity is to use different jobs or tasks to indicate individual differences in the same human attributes. Adding people across jobs can create large enough numbers of incumbents to make a criterion-oriented study feasible.

A very different version of synthetic validity comes about when a test is validated in a setting other than that of the target job. If the job on which the test is validated can be shown to be very similar to our target job, the Uniform Guidelines on Employee Selection Procedures (EEOC, 1978) allow this. A more abstract approach not anticipated by the Guidelines is called *validity generalization*. In validity generalization, many studies of the validity of a test type—say, a cognitive ability test—are pooled across different jobs and settings. If the test type works in predicting performance across the board, we then make a judgment that the test type should work for our job also. Obviously, this is a controversial approach. More conservative scientists will not believe the results unless the jobs and settings in which the test has been validated are very close to the one in which we are interested.

Still another approach that might be classified as synthetic validity is the J-coefficient described as part of the *job element method* in Chapter 3. If you remembered this and are of age, have a glass of fine champagne on us—or rather, on our publisher.

Signs and Samples

Consider our receiver job once again. Suppose we are considering two tests: a paper-and-pencil test of conscientiousness (one of the Big Five personality dimensions) and a forklift driving work sample test. Consider the conscientiousness test first. Let's first assume that our distribution center handles various computers and associated computer accessories like monitors and printers. Clearly, we want distribution center employees who move and stock expensive computers to be careful rather than sloppy in their work. Mistakes are invariably irritating and may be costly as well. Although the test is labeled "conscientiousness," the activities involved when taking the test do not appear anything like receiving and putting away computers. Thus, the test is a *sign*, or index, of *conscientiousness* that summarizes a lot of different behaviors, such as showing up to work on time and counting parts accurately. Taking the test, however, requires reading some material and marking an answer sheet with a pencil. On the other hand, our forklift driving test looks just like what receivers do on the job. The test corresponds well to a specific task that is part of the job. In a sense, the test is a *sample* of the work required for the receiver job.

Many people feel that there is a qualitative difference between the two types of tests (signs vs. samples). In particular, people are often more comfortable with samples as tests for selection than they are with signs. Although not a formal strategy for validation, we sometimes use the term *face validity* to refer to the degree of apparent similarity between the test and the job. In practice, it is more common to justify samples with content-oriented strategies and to justify signs with criterion-oriented strategies. For example, for the receiver job, we might use a content-oriented strategy to demonstrate the job relatedness of the forklift driving test, but a criterion-oriented strategy to demonstrate the job relatedness of the conscientiousness test.

Despite the obvious difference in the two types of tests, the interpretation of both is essentially the same: Differences in test scores are interpreted as predictions of job performance. Thus, in our example, scores that indicate highly conscientious behavior and scores that indicate skill in driving a forklift would be interpreted as indicating (predicting) good job performance for receivers. Whether different test validation procedures are really necessary for the two types of tests is somewhat controversial. For example, compare the views of Harvey (1991) and Wernimont (1988).

Comparison of Designs

As we mentioned earlier, there are two main inferences in the job analysis process used for selection: (1) from the job to the KSAOs and (2) from the KSAOs to the tests or other methods used in selection (Harvey, 1991). The criterion-oriented validation design offers a means to confirm the accuracy of the aforementioned inferences using data. By comparing test scores with job performance measures, we can essentially see whether our inferences were correct. If there is an association between test scores and job performance, this bolsters our argument that we figured out the job requirements and developed tests that measured them.

The content-oriented method is both better and worse than the criterion-oriented method, depending on your point of view. The content method attempts to match the test directly to the job. Thus, especially with work samples, it can actually bypass the determination of KSAOs. For example, if we devise a forklift-driving work sample, the KSAOs required to do the task do not have to be named. They are whatever is required to do the task on the job and on the test. This is helpful because it essentially eliminates one of the inferences in the selection process. On the other hand, validation is about understanding the meaning of test scores. Content validation strategies don't deal directly with test scores. Instead, they deal with test stimulus materials. Thus, we could devise two different tests for determining a receiver's knowledge of driving a forklift. In one test, we could use a multiple-choice format rather like a driver's license exam and ask about forklift safety (What is the maximum load? Maximum speed?). Or we could create a simulation in which receivers would be asked to move some pallets quickly and would then be graded on how well they did or the number of mistakes they made. A content-oriented strategy would proceed in essentially the same way for either test. Thus, the content-oriented approach can avoid the job-to-KSAOs inference, but it necessitates an inference (or assumption) about the relation between the test scores and job performance that is not verified against data. Even with a good simulation, it is impossible to duplicate the job in all its detail using a test, so we still have to assume that whatever KSAOs are needed for the job are manifested in the test we are using.

Research on Judging KSAOs for Validation

The central problem in selection is the matching of people to jobs, and the key to doing so is to infer the KSAOs that are necessary for successful job performance. How good are people at inferring KSAOs? The answer to this question is complex, and it is based on rather skimpy data. Overall, we can say that people who are well informed about the job and who use a systematic procedure when making judgments can infer KSAOs with some degree of accuracy.

Reliability. The reliability (particularly interjudge agreement) of KSAO judgments is important because the reliability of such judgments sets a limit on the validity or accuracy of the judgments. Put another way, if two people make independent judgments of the KSAOs required by a job and they agree, fine. If the two people make judgments of the KSAOs required by a job and they disagree, we have a problem. Although data on the issue are scarce, existing data suggest that individual judges are not very reliable in their judgments of required KSAOs. For example, Hughes and Prien (1989) had job experts examine 343 task statements regarding the job and then make ratings of KSAOs. They found that the average agreement of importance for the KSAOs was rather low at .31, and they concluded that interjudge reliability was a serious problem. Cain and Green (1983) studied trait ratings provided in the *Dictionary of Occupational Titles* (U.S. Department of Labor, 1977). They also found rather low reliability of ratings. The reliability of ratings due to differences in judges can be largely controlled by using panels of judges rather than individual judges (Van Iddekinge, Putka, Raymark, & Eidson, 2005). If a panel of about 10 expert judges can be assembled, their mean judgment of the KSAO importance will tend to be much more reliable than that of the individual judges. For example, if the correlation between two judges is .30, then we would need 10 judges to achieve a reliability of .80 for the mean importance rating (see Nunnally & Bernstein, 1994, for the appropriate computations). Of related interest, Langdale and Weitz (1973) found that personnel interviewers agreed with each other more about the suitability of applicants when the interviewers were armed with a job description than when they were given only a job title.

Validity of KSAO judgments. A study by Trattner, Fine, and Kubis (1955) examined how well job analysts could predict the test scores of job incumbents in 10 different jobs. There were two panels of eight job analysts. Each panel reviewed information about each of the 10 jobs and estimated aptitude test scores. One panel examined job descriptions; the other also got to observe the jobs in question. The job analysts' ratings were averaged within panels and then correlated with the mean test scores of incumbents in each of the 10 jobs. The panels were able to predict mean test scores fairly well for mental and perceptual tests. They were less able to predict physical test scores. Observations of people doing the job appeared to help the analysts make more accurate judgments than simply reading a job description. One must regard these results as preliminary surrogates compared to whether analysts can estimate the actual relationship of the test scores to job performance.

Parry (1968) asked job analysts to estimate validity coefficients rather than mean test scores. He found only moderate validity for the judgments. Later work by Schmidt, Hunter, Croll, and McKenzie (1983) indicated that panels of people can estimate the validity of tests to a reasonable degree. They noted that the prediction of actual test-criterion correlations is difficult because such correlations vary quite a bit due to the small sample sizes typically found in industrial settings.

The U.S. Air Force (USAF) has conducted several studies (cited in McCormick, 1979) that are relevant to job analysis for selection. USAF researchers found that psychologists and instructors could estimate the relevance of various tests for predicting success in technical training. They also found that psychologists could judge aptitude requirements for tasks and that these judgments could be pooled to estimate the ability requirements for jobs. Note that this is the same idea behind the job element approach.

As far as we know, there are no studies that examine the ability of people to judge a person's competencies. Note that the judgment required for competencies is qualitatively different from the judgment required for a typical job. Competencies are usually considered in light of business goals. A company thought to be innovative might value creativity, for example. Therefore, competencies are thought of not typically as job requirements but more as attributes on which to judge the fit of the individual to the organization.

To examine the validity of such judgments, we would need research that spans multiple organizations. This research would need to demonstrate that creative people were more successful, productive, or satisfied in Company A than they were in Company B. On the other hand, as we have seen, competencies are often shorthand for managerial functions such as planning, financial management, and representation to people outside the organization. Such competencies would appear to be required for success in most managerial jobs.

Position Analysis Questionnaire studies. McCormick (1979) reports a study of the PAQ used for predicting General Aptitude Test Battery (a test battery, mostly cognitive, then used by the federal government) scores and validity coefficients (test-criterion correlations) for a sample of 163 different jobs. The database used dimension scores for each job as predictors and either mean test scores or validity coefficients as the criterion. Results indicated that the job incumbents' mean test scores across jobs were quite predictable for the cognitive tests (intelligence, verbal, numerical, spatial, form perception, and clerical perception). The correlations between the PAQ and the mean cognitive test scores ranged from .61 to .83. The correlation between the PAQ and the mean motor coordination test score was also impressive (.73). Correlations between the PAQ and test scores for finger and manual dexterity, however, were lower (.41 and .30, respectively). The correlations between the PAQ and *validity coefficients* were much smaller, ranging from .04 to .39. Validity coefficients for the cognitive tests and the motor coordination test were much less predictive from the PAQ than were the mean test scores. Correlations between the PAQ and validity coefficients for the manual and finger dexterity test were .15 and .39, respectively.

A similar study was reported by Gutenberg, Arvey, Osburn, and Jeanneret (1983), who looked at a database of test validation studies. They used dimension scores from the PAQ to predict the size of test–criterion correlations. The logic of the study was that if the PAQ results indicate that a job requires a good deal of cognitive ability, then the correlation between a criterion and a cognitive ability test for that job should be large. If the PAQ indicates that a job does not require much cognitive ability, then the correlation between a criterion and a cognitive ability test should be small. Results of the study indicated that the PAQ did predict the size of the correlations for cognitive tests, but less so for perceptual and physical tests. As we mentioned earlier, validity coefficients are hard to predict because most studies are based on small samples of data, so the PAQ findings are impressive.

Summary of judging KSAOs. Panels of job analysts armed with good job information can provide reliable and valid estimates of job requirements. Job analysts are better at judging mean test scores of incumbents than they are at judging the validity of tests (but there may be several technical reasons for this apart from the judgment task itself). Although there is little in the way of empirical data on the issue, we suspect that reliability and validity of judgments about KSAOs will increase to the extent that the KSAOs are relatively concrete and closely tied to specific tasks (Dierdorff & Morgeson, 2009). For example, consider the job of surgeon and the KSAOs "hand and finger dexterity" and "verbal reasoning." Although verbal reasoning tests may predict performance in surgery quite well, this is not obvious from considering what happens in a typical surgery as seen on your favorite medical television drama.

Key Considerations for Selection

For selection, you need information about worker attributes, that is, the KSAOs needed to successfully complete the job. These days, a variation on this theme is "competencies," which are sometimes seen as more business-oriented and broader versions of KSAOs. Whatever descriptor type is chosen, worker information is used to pick tests, create interview questions, and generally communicate what the job requires for success. To complete a validation study, however, also requires work attributes, mainly task information. Information about tasks is needed for at least one of four purposes. First, if we are going to conduct a content-oriented validation study, we have to know the task content of the job so that we can relate the tasks on the job to the tasks on the test. Second, if we are going to conduct a criterion-oriented validation study, we have to know the task content of the job to select one or more criteria. Third, if we intend to compose a work sample test, tasks can provide the necessary content. Finally, we need to collect task information to satisfy federal legal requirements for conducting validation studies (the Guidelines). There are exceptions. If the criterion is an objective measure such as percent scrap or turnover, task information may not be needed.

One approach to job analysis that meets the requirements for selection is the *combination job analysis method* (C-JAM; see Chapter 4 on hybrid methods). C-JAM begins with a comprehensive list of tasks and follows up with a list of KSAOs that are needed to complete the tasks successfully. C-JAM uses panels of expert judges to develop both those tasks and KSAOs. The tasks can be used for criterion development or to form the basis of judgments of content validity. The KSAOs can form the basis for the choice of tests to be used in selection.

Another method that has been used for selection is the *critical incident technique* (see Chapter 2 on work-oriented methods). The critical incident technique requires that job experts recall specific incidents of either very good or very poor performance on the job. Each incident begins with a statement of context (what led up to the incident), followed by what the worker did, and finally what the result of the worker's action was. When a large number of incidents have been compiled, they may be sorted into dimensions that represent tasks, KSAOs, or both. The task dimensions can inform the creation of rating forms that are used to collect judgments such as managers' or peers' judgments of job performance; these job performance judgments could be used as criteria in test validation studies. The KSAO dimensions can be used to select or develop selection tests.

The *job element method* (see Chapter 3 on worker-oriented methods) is similar to the critical incident technique in that it takes specific bits of the work and combines them to form dimensions. Because the terminology used in job elements is that of the worker or job expert, resulting dimensions could be used to both pick or devise tests and to create rating forms.

The main difference between critical incidents and job elements is that the critical incident technique begins with examples of good and poor performance, and the job element method begins directly with judgments of what is required to be successful on the job.

Generic Traits

Structured methods such as the Position Analysis Questionnaire (PAQ), the Threshold Traits Analysis System (TTAS), and the Ability Requirements Scales (ARS; see Chapter 3 on worker-oriented methods) have also been used for selection, particularly to justify the choice of tests. (Although newer, O*NET could also be used.) Such methods require the job analyst to review the job and then to complete a standardized questionnaire that describes the job mostly in terms of worker requirements. Because the PAQ, TTAS, ARS, and O*NET provide a great deal of information on the worker side and because the information provided is directly comparable across jobs, such methods have promise for furthering the scientific basis of selection. Unlike C-JAM, critical incidents, or job elements, the PAQ, TTAS, ARS, and segments of O*NET (such as basic skills and cross-functional knowledges) are not tailored to a specific job and therefore apply the same attributes to each job. Because they are structured to be standardized across jobs, the information that they provide is more comparable across jobs.

TABLE 8.2 ■ Example KSAOs List for the Receiver Job

Knowledge	Skills
Product identification Arithmetic (count, add, subtract, multiply) to find quantity Vendor and part identification **Knowledge of paperwork** Meaning of items on packing list (e.g., part numbers, payment numbers) Meaning of items on freight bill Meaning of items on put-away document **Knowledge of machines** Knowledge of operation of computer system • Names and commands for different screens • Screen places (e.g., part numbers) • Error messages Knowledge of operation of printers Knowledge of operation (rules) of forklifts Knowledge of operation of scanner **Knowledge of Washington, DC, layout** **Knowledge of safety regulations and procedures**	**Operation of forklift** **Reading and writing words and numbers**

Abilities	Other Characteristics
Perceptual speed and accuracy (visual)	**Conscientiousness** Attendance
Eyesight (acuity)	Punctuality
Physical	
Strength	**Attention to detail**
• Lifting	**Work motivation (pace and endurance)**
• Dragging	
• Pulling	**Assertiveness (signing freight bill)**
Agility	**Integrity**
• Stooping and bending	Theft avoidance
• Climbing	Drug and alcohol avoidance
• Reaching	Interpersonal hostility and violence avoidance
• Twisting	
Depth perception (vehicle operation)	**Vigilance (resistance to monotony)**
Spatial orientation and visualization (packing and storage)	**Team orientation (agreeableness and customer orientation)**
Multilimb coordination (vehicle operation)	
Hearing (warning signals)	

Note: KSAOs = knowledge, skills, abilities, and other characteristics.

On the other hand, they do not provide much in the way of descriptions of the work itself. A list of KSAOs derived from a study of generic traits required for the receiver job is shown in Table 8.2. Note that there is a large number of KSAOs listed in the table. Not all of the KSAOs identified during a job analysis will be used for selection. A subset of the KSAOs and associated tests is shown in Table 8.3. This table highlights an interesting fact: Several different kinds of tests may be used to measure a single KSAO.

TABLE 8.3 ■ Example KSAs and Associated Tests for the Receiver Job

	Application/ Interview	Paper & Pencil	Simulation	Other
KNOWLEDGE OF: Arithmetic (count, add, subtract, multiply) to find quantity	1. Interview	1. Customize 2. Saville & Holdsworth (Published test)	1. Number of items in a carton 2. Number of items missing from a carton	
SKILL IN: Reading and writing words and numbers	1. Application 2. Interview		1. Read packing slip 2. Read PC screen	
ABILITY IN: Eyesight (acuity)		1. Number checking		1. Eye exam
ABILITY IN: Perceptual speed and accuracy (visual)		1. Number checking 2. Guilford-Zimmerman Aptitude Survey (Published test)	1. Find and select parts from bins matching part numbers to written orders	

Note: KSAs = knowledge, skills, and abilities.

Research Comparing Efficacy of Job Analysis Methods for Selection

One might expect that there are many studies designed to see which job analysis methods work best for selection. Unfortunately, this is not the case, perhaps because such studies are potentially expensive and difficult to carry out. Fortunately, Levine and colleagues conducted a couple of studies that took on this challenge. We hope that you, dear reader, may be inspired to take on the challenge as well.

Levine, Bennett, and Ash (1979) surveyed experts to compare the PAQ, the critical incident technique, task analysis, and the job element method for personnel selection. They found that the experts most often used task analysis and the job element method. However, the experts were not highly favorable in their evaluations of any of the methods. Further research on the four methods by Levine, Ash, and Bennett (1980) showed that the critical incident technique resulted in slightly better test plans than the other methods, but the PAQ was the least expensive to use. The PAQ was not as popular for the development of test plans, probably because it does not provide task information. Levine et al. (1980) further noted that none of the methods provide information about minimum qualifications.

Summary Thoughts on Staffing

In this section, we have discussed each of the several strategies for developing (hopefully) valid tests. Each relies in some fashion on information yielded by job analysis. It would be a mistake, however, to view these different strategies as either/or. Selection is a high-stakes process. It significantly impacts an organization's success and the career success and well-being of its applicants. As such, judgments about the effectiveness and psychometric quality of instruments used for selection and decisions based on the instruments are critical. The more evidence we can marshal in support of this information, the better. So when practical, using more than one or all of the strategies we have described (i.e., content-oriented, criterion-oriented, and construct-oriented) would make for sounder judgments. When a multipronged strategy is employed for assessing measurement quality, particularly validity, a combination of job analysis methods will be most appropriate.

TRAINING

Like staffing, training typically considers the job to be a fixed entity to which people are fit. In staffing, good fit between jobs and people is achieved by selecting the person who best matches the requirements of the job. In training, people are developed to fit the job—that is, people are trained so they come to possess the knowledge, skills, or other requirements of the job, such as certification. Because they are more enduring, abilities and personality characteristics are rarely trained.

Consider our receiver job. Receivers must quickly and accurately compare names and numbers on packing slips with names and numbers on delivered merchandise. Therefore, we might select them for perceptual speed and accuracy (an ability). Receivers have to compare the quantity of items on a packing slip to the quantity received, so they need to be able to count, add, subtract, and multiply quickly and accurately (often, parts come in boxes that are only partially full, so it is much more efficient to multiply and subtract than to count). People can be either selected or trained for arithmetic (a knowledge). Most likely, in this example, they will be selected because it is reasonable to assume that applicants will have the needed background in arithmetic. On the other hand, many receivers will be trained and certified in operating a forklift (knowledge of its operation and skill in actually driving it). Most applicants will likely know how to drive a car, but not a forklift.

The kind of training we are talking about in this book is training designed for jobs at work. Rather than being aimed at workers in general, such training is intended to facilitate job performance on tasks that constitute a job held, or aspired to, by incumbents. For example, a general class in human relations skills offered to all employees in a bank is not likely to be the sort of training we are describing. On the other hand, human relations skills in the context of instruction about how to conduct performance appraisal meetings for managers in the same bank would be considered training. The distinction is one of tailoring training to specific tasks that belong to a job currently held, or shortly to be held, by trainees. We applaud corporate efforts to further employee development through general education and skill development such as human relations skills classes, reimbursement for college credit, and so forth. Such instruction, however, is not typically based on a careful analysis of the job, nor are the instructional objectives defined in behaviors of interest to the job. Although general education is likely to benefit employees and ultimately prove beneficial to employers, it is not designed with the goal of immediate transfer to the job.

Some form of training is necessary for virtually every job. If nothing else, people have to know where their workstations and tools are. Like recruiting, training comes in endless varieties from formal to informal. Formal training may take place away from the workplace over a long period of time. For example, mechanical flight trainers are used to train pilots in how to respond to emergencies such as engine fires. Or training may be rather informal and take place on the job. For example, a supervisor may ask a subordinate to instruct a new employee in how to use the telephone system.

Training is expensive. Employees participating in off-site training are away from the job but still get paid. The training instructors need to be paid, and any space and equipment used in the training has to be supported. For example, the cost of using a mechanical flight simulator for training may be more than $1,000 per hour. If the training is done on the job, then someone else usually has to take time away from his or her usual job to provide the training. Such time away creates a loss in efficiency—either some work fails to get done, or it gets done more slowly.

On the other hand, the cost of *failing* to provide training generally outweighs the cost of training. Many jobs will be done poorly or simply will not be done without some training. A computer programmer who is denied access to an account will not be effective. An untrained sales representative can lose business in many ways. People who handle hazardous materials can injure themselves or their coworkers. People who operate powerful machinery can create a real menace. Therefore, training is necessary to promote successful job performance and to avert disaster. From the employee's perspective, training can reduce stress by promoting feelings of confidence and competence.

We begin this section with an outline of what we call the "training cycle." The training cycle is a big-picture approach to solving an organizational problem through training. Training is unusual in its comprehensive problem-solving approach. Such an approach is a very good way to think about all applications of job analysis (e.g., staffing, compensation, and job design) and not just training.

After we present the training cycle, we describe the kinds of information that a job analysis needs to provide in order to answer questions about training. As you will see, training is similar to selection in the kind of job information required, but training usually requires more specific information about the tasks (work attributes) than does selection.

The Training Cycle

Training cycles of various kinds have been developed by different authors (e.g., Gagné & Briggs, 1979; Goldstein & Ford, 2002). They are all alike in that they provide a means–ends analysis for meeting organizational goals. The cycle that we describe was developed in Goldstein and Ford (2002, which builds upon the work of Goldstein, 1993) and is shown in Table 8.4. The table contains four columns labeled *Needs Assessment*, *Training and Development*, *Evaluation*, and *Training Goal*. Each is briefly described.

Needs Assessment

Needs assessment is called different things by different people responsible for training. Needs assessment is sometimes called "front-end analysis" because it must be done *before* the training is developed. Needs assessment is crucial to

TABLE 8.4 ■ Training Cycle

Needs Assessment	Training and Development	Evaluation	Training Goal
1. Needs Assessment a. Organizational Support b. Organizational Analysis c. Task and KSA Analysis d. Person Analysis 2. Instructional Objectives	1. Selection and Design of Instructional Programs 2. Training	1. Development of Criteria 2. Use of Evaluation Models a. Individual Difference b. Experimental c. Content	1. Training Validity 2. Transfer Validity 3. Intra-Organizational Validity 4. Inter-Organizational Validity

Source: Republished with permission of Cengage Learning, from Goldstein, I. L. (1993). *Training in organizations: Needs assessment, development, and evaluation* (3rd ed., p. 21). Pacific Grove, CA: Brooks/Cole Publishing Company. Copyright; permission conveyed through Copyright Clearance Center, Inc.

Note: KSA = knowledge, skill, and ability.

developing a good training program. Of the four phases of the training cycle, needs assessment is the phase that is most closely tied to job analysis. The basic idea is that before we begin training, we need to have a good idea of what the training is supposed to accomplish. If the ultimate goals of training are not specified prior to training, there is a good chance that they will not be met. For example, Goldstein and Ford (2002) pointed out that firefighters may face dramatically different conditions and must do different things when fighting fires in high-rise buildings than when fighting fires on the docks of a river. Failure to specify goals regarding both types of fire combat will produce inadequate training programs.

Needs assessment involves three main entities: the organization, the tasks and KSAs, and the people to be trained. *Organizational analysis* considers the job to be part of a larger organization. The job exists to fulfill some function. For example, our medical assistant job prepares treatment rooms for patient examinations. If our customers (i.e., patients) are complaining about disorderly and dirty rooms, it could be because the medical assistant does not know how to organize and clean the rooms properly. We could have disorderly and dirty rooms for lots of other reasons, too, but the point is that the medical assistant job provides an output that is used by another job (the doctor uses the room). The output provides some objectives that can be used to inform training.

Another way of saying this is to consider what management wants the training to achieve and to consider whether training is the best option, as compared, for example, to putting up signs or simply asking the incumbent to do some task differently. The organizational analysis should tell us how to know whether our training is effective. For example, organizational analysis of grocery clerks might point to customer satisfaction with the speed of service. Organizational analysis informs the training objectives; in other words, organizational analysis answers the question of *why* training takes place.

Task and KSA analysis provides training content. All job-oriented training must consider the job itself, that is, the tasks contained in the job or common to several jobs. For example, there are aspects to driving a forklift that are determined by the machine and its use. The forklift driver needs to be able to steer the forklift to place the tines of the fork into a pallet, to operate controls that lift the pallet, and so forth. For simple jobs, the tasks that are found in the job description may be sufficient for understanding what needs to be trained. For most tasks, however, additional knowledge, skills, or abilities must be trained. For example, the forklift operator needs to know rules for safe operation of the forklift, such as speed limits, maximum weight for lifted loads, maximum heights and angles for lifting, and so on. Many years ago, we experienced what could happen if these rules are not followed. At a soda bottling plant, forklift drivers stacked pallets, each five soda cases high, up to three pallets high. An inexperienced driver failed to use the right angle in placing the top pallet at the end of a row. As a result, an entire row of soft drinks came tumbling down like dominoes. It took two full shifts of three workers to clean up the mess. Task and KSA analysis provides the substance of the training; that is, task and KSA analysis (in part) answers the question of *what* is to be trained.

Person analysis concerns the KSAs that trainees bring with them. A native French speaker will have little use for a course in conversational French. A student enrolled in calculus without first mastering algebra will be unprepared to handle the new material. It is a waste of time to train people on things they already know. The receiver job does not contain an arithmetic training module. Incumbents are expected to know how to add, subtract, multiply, and divide before they are hired. On the other hand, training will not be useful if the training incorrectly assumes that trainees possess skills that they do not. If most of the incumbents are native Spanish speakers with little exposure to English, the elaborate English instruction manuals may be largely wasted. Person analysis requires that information be gathered about the potential trainees before the training is developed. Person analysis helps to make sure that training is efficient by providing no more or less than what trainees need to achieve the training goals. Person analysis helps to answer (in part) the question of *what* is to be trained. It may also help us learn who is deficient in what skills, so the question of *who* should take our course is also addressed.

Together, organizational analysis, task and KSA analysis, and person analysis help to determine the *instructional objectives* for the training. Instructional objectives are statements that describe the behavior of properly trained people. Well-written instructional objectives contain enough context so that both the training environment and the appropriate response are clear. Furthermore, good instructional objectives also contain statements about standards, that is, statements that specify a minimum acceptable level of performance. For example, in a course on the operation of a forklift, the immediate goal of the training is for the trainee to pass a certification exam and receive a certificate (license) to operate a forklift. The instructional objectives of the course are the things that a person has to do to pass the test. For example, the operator might have to use a forklift to pick up a pallet and move it from one end of a warehouse to the other within 2 minutes without collisions or loss of objects from the pallet.

You are wondering about the term *organizational support* that appears in Table 8.4 but has not been described (you *are* sharp this late in the chapter!). Goldstein and Ford (2002) noted that gaining acceptance and support of top management and other organizational members as well as their willingness to commit time and resources to the training is critical to the whole needs assessment enterprise. However, it is less central to our mission here of focusing on the role played by job analysis.

Training and Development

Training and development activities are the heart of the training itself. The needs assessment provides instructional objectives. Instructional objectives say what properly trained people are supposed to do. Training and development provide the setting and activities that get the trainees to learn what it is that they are supposed to do, and if the training program works well, the learning will be maintained on the job. This phase involves both the selection and design of the instructional programs and the training itself. Lots of questions have to be answered in the design of the instruction.

- Where will the instruction take place?
- What media will be used?
- Who will provide the training?
- What instruction is already available? How suitable is it? How expensive is it?
- How much time and money are available for training?

Decisions about the training get made and the training gets developed one way or another. As we mentioned earlier, training programs vary enormously. Instruction on a single topic can be done as a lecture or a film, in small-group discussions, through interaction with a machine such as a computer or simulator, by reading a book or other printed material, through other media (e.g., closed-circuit TV), by observation of a worker completing the task or practice on the job (preferably accompanied by well-timed feedback), or through some combination of these.

Evaluation

It is easy to say that we favor effective and efficient training. It is harder to know that we have it. How can we show the effectiveness of training? This is the question of evaluation. Part of the answer to such a question is provided by the instructional objectives. The objectives should provide statements about the following:

- The context or stimulus situation.
- The behavioral requirement (what the person does).
- The minimally acceptable response.

Training should involve testing so that the trainee can demonstrate what has been learned. If assertiveness is being trained, the trainee should demonstrate assertive behaviors in a social situation. If the training is about how to operate a forklift, the trainee should use a forklift to complete a task.

Evaluation models. The evaluation models listed in this phase of the training cycle include *individual difference*, *experimental*, and *content*. Each of these models provides a data collection strategy that generates information about the quality of training. The individual difference model involves collecting data about a trainee's relative standing in training and a trainee's relative standing on job performance and relating the training data to the job performance data. For example, we might compare police academy grades with subsequent patrol officer performance measures. If the two measures are correlated, training evaluations can be used for selection. Unfortunately, such correlations do not necessarily mean that much was learned in training.

The experimental model involves using trained and control groups and/or pretest–posttest designs to show the effects of training. We would expect that knowledge tests given at the beginning and end of training would show gains. For example, we might give bank tellers a test of the required bits of information needed to cash a check both before and after training. If the training is effective, we expect to see better performance on the test after training than before training. Performance on a test for a group that has just received a new kind of training may be compared with the performance of a group that received the old (status quo) form of training (or no training). By so doing, we can see how the new training compares with the prior method. Note that in both the individual difference and experimental evaluation models, the comparison is between people (correlations show standings relative to the mean; ANOVA [analysis of variance] shows mean differences between groups). This is different from a comparison to a standard of performance on the task itself as we required for the instructional objectives.

For example, suppose we are instructing people to serve in tennis. Our method involves giving people practice and feedback in serving the ball. Our test of proficiency is the number of serves out of 20 tries that are hit from behind the baseline into the service box at a speed of at least 50 miles per hour. At the end of training, each trainee gets tested via 20 tries that are timed by a speed gun and judged by a line person as either in or out. We then count the number of balls that land "in" at more than 50 miles per hour. The individual difference model would involve comparing our service scores with later indices of tennis proficiency, such as tournament standings, points won, aces, and so forth. The experimental model might use the same test both before and after training to see whether the serve improved over the course of instruction. A good instructional objective might state that upon completion of the training, the trainee must achieve a score of 10 out of 20 on the test or he or she fails the training. Note that for the individual difference and content models, no minimum score is needed to complete the training. The content model merely requires that the training be connected to the KSAs discovered by the job analysis.

Training Goal

Because training is developed to achieve organizational goals, it is a good idea to check whether the training results in the desired outcomes. For example, during a training session, a trainee may have learned how to be more assertive and to show appropriate assertive behaviors, but the trainee might not display these assertive behaviors because work group norms discourage it. Different authors organize training goals differently. Kirkpatrick (1959a, 1959b, 1960a, 1960b) has suggested assessing the following:

- *Reactions:* Trainee reactions to instruction.
- *Learning:* Measures of proficiency taken at the end of training.
- *Behavior:* Measures of performance on the job subsequent to training.
- *Results:* Organizational effectiveness measures that are believed to show the results of increased effectiveness on the job.

If we were training data processing supervisors in skill in dealing with subordinates, for example, we would first ask the trainees at the end of training what they thought of the training. Was it useful? Enjoyable? Second, we would test them on the content of the course to see whether they learned the principles taught in the course (e.g., to balance positive and negative feedback). Third, we would examine performance on the job to determine how well they dealt with their subordinates post training. Fourth, we would look at indices such as data processing service records and complaints about data processing from other departments, which would provide evidence of the functioning of the entire department.

Goldstein and Ford (2002) described four training goals:

1. Training validity.
2. Transfer validity.
3. Intra-organizational validity.
4. Inter-organizational validity.

The first two of these are similar to two of Kirkpatrick's effectiveness categories. Training validity concerns whether anything was learned during training. Training validity is similar to Kirkpatrick's "learning," that is, measures of proficiency taken at the end of training. Training validity can be assessed using the experimental approach or the instructional objectives approach. The most useful information to establish training validity is the quality of trainee performance on tasks that were supposed to be trained. Transfer validity concerns how well the material learned in training applies to the job setting. Transfer validity is analogous to Kirkpatrick's "behavior," or measures of job performance. Here, we are concerned with job performance that should be influenced by the training.

Two of Goldstein and Ford's (2002) goals are different from Kirkpatrick's categories. Intra-organizational validity concerns how well the training evaluation results for one group of trainees applies to another group of trainees in the same organization. When training programs are new or when they are being formally evaluated, there may be pressure on the trainers and trainees to behave in ways that they otherwise would not. When classroom instructors are graded by the test scores that their students receive, for example, the instructors become motivated to "teach to the test." The question asked by intra-organizational validity is whether the results of a training evaluation effort apply to ongoing training programs. In inter-organizational validity, the question is whether training evaluated in one organization is likely to have similar effects when used in another organization. Going back to our data processing supervisor example, suppose we have a new subordinate relations training program and that both the instructors and supervisors know that the program will be evaluated by looking at the supervisors' scores on tests at the end of training and also by looking at supervisors' later job performance. Intra-organizational validity concerns whether the evaluation results for the training will apply when the training no longer involves evaluating trainees' scores at the end of training. Inter-organizational validity would apply if we were borrowing the supervisor training from another company and looking to apply the results of that company's training evaluation study to our company.

We have described training evaluation in terms of a cycle because the results of the evaluation can be incorporated into a new needs assessment in order to improve training. Proceeding through the cycle will result in data that the (a) show that the training is effective, (b) show where the training needs improvement, or (c) show that the training is not the best way to solve the problem.

There is obviously more to the training story than we can get into here. As but one example, our approach to evaluation focuses on final outcomes or comparisons of training and job content. But what if the training decays over time because the trainer begins to leave out big chunks of the program? The website of the Association for Talent Development (www.td.org) is an excellent source.

Key Considerations for Training

Central to training is the content of the work itself, that is, the tasks. KSAs may also be important for training because they support task performance. For example, interviewing patients and measuring their vital signs (a task) is the heart of the medical assistant's job. To carry out the task, the medical assistant has to know how to operate a sphygmomanometer (also known as a blood pressure meter, but the fancy word sounds more impressive). To effectively operate this device, the medical assistant would need to have specific device knowledge, be able to operate it consistently, and so on. The level of detail required for the tasks will depend on the intended purpose. One purpose of task statements is to determine whether the task should be trained, and if so, where. Task statements in such a case can be written fairly broadly, as is done for job description and selection purposes. Such task statements will often be rated by experts, and the ratings will be used to decide, for example, whether the task is to be trained.

When tasks are to be trained (i.e., when job analysis is to provide the basis of the content of training), then the level of task description must be much more detailed. Otherwise, the training must be developed by one or more job experts (which is a good idea anyway, but not always practical). The task list written to assist in training content development for the receiver job is shown in Table 8.5. If you compare the tasks in Table 8.5 to the essential duties and responsibilities in the job description in Table 7.3, you will see that Table 8.5 contains a larger number of tasks and that the tasks tend to be more detailed. Notice that the intent of the job description in Table 7.3 is to communicate the gist of the job. In Table 8.5, however, the job is described in terms of a series of steps as they typically occur on the job. Such a list of tasks could be used to inform training for employees who are new to the receiver job to be sure that the entire job is covered in training.

TABLE 8.5 ■ Example Steps and Detailed Task List for the Receiver Job

Step 1: Unpack the truck.

1. Receive freight bill from trucker.
2. Operate forklift to move pallets from truck to storage.
3. Count pallets and cartons.
4. Sign freight bill if skid count agrees with freight bill.
5. Notify driver if skid count disagrees with freight bill, and he will also recount.
6. Call for exception if driver and receiver agree with new count.
7. Contact supervisor if driver and receiver disagree with the count.
8. Fax notice of overage or underage to shipping company.
9. Call shipping company to notify it of overage or underage.
10. Call shipping company to notify it to search for the missing part if exception is under-received.
11. Search all parts for labels and then call shipping company to ask whether it wants to sign over (receive) the part or send it back if exception is over-received.

Step 2: Unpack the skid.

1. Sign freight bill for piece count.
2. Find shipping company's packing list.
3. Fill out alternate packing list if packing list is lost.
4. Look up PO in computer and print screen to create a packing list if the packing list is lost.
5. Cut shrink wrap with pallet knife.
6. Pull shrink wrap off boxes and discard.
7. Lay out pallets to place unpacked order.
8. Check boxes and products for damage.
9. Put damaged box and product in designated "Bad Box" section to be sent for repacking if only box is damaged.
10. Set product aside for inspection if product is damaged or believed to be damaged.
11. Examine boxes for identification.
12. Move boxes containing the same part numbers to a common location if more than one part is stacked on the pallet.

13. Open boxes with more than one kind of part inside.
14. Multiply and add to determine the total number of each part in the order after parts are separated and collected by number.
15. Underline the vendor part number and circle or write in the quantity received if the packing slip is found.
16. Write the vendor part number and quantity of each part received on the alternate packing slip if the packing slip is not found.
17. Stack boxes so the vendor part number is visible to bar coders.
18. Stack boxes that look alike but have different part numbers in different locations.

Step 3: Receive the parts into the computer system.

1. Enter password to log into system.
2. Punch choice for receiving screen at main menu.
3. Punch PO number to find order of interest at receiving screen.
4. Identify local part numbers from vendor part numbers and other identifying information when order appears.
5. Enter quantities received for each part number received.
6. Compare quantities ordered (Qty Open) to quantity received.
7. Investigate discrepancies between quantity open and quantity received.
8. Fill out alternate packing slip for the discrepancy if parts are found.
9. Compare computer screen to paper to verify that all received entries are correct (verified).
10. Identify (using the computer) current places where product is stored.
11. Troubleshoot any computer error messages (red displays).
12. Enter shipping information into computer (that is, use freight bill screen to enter number of packages, shipper, weight, shipping fee, zip code, and so on).
13. Locate part numbers by using computer searches.
14. Locate part numbers by calling manufacturers.
15. Assign part numbers different POs.
16. Recount items to avoid receiving errors.
17. Troubleshoot parts to be received that do not match part numbers.
18. Call supervisor if troubleshooting is unsuccessful.

(Continued)

TABLE 8.5 ■ (Continued)

Step 4: Do paperwork and labeling.

1. Print barcode labels and put-away document.
2. Reprint mangled bar codes.
3. Color special barcode labels with markers to help bar coders apply labels properly.
4. Collect barcode stickers from printer.
5. Affix order number sticker to packing list.
6. Initial and date sticker.
7. Collect put-away document from printer.
8. Place barcode stickers and associated put-away documents together on the floor in a separate location.
9. Take barcode stickers and put-away documents to parts received on the loading dock (storage area).
10. Replace stickers in barcode printer.
11. Replace paper in put-away document printer.
12. Compare bar code and boxes to verify that barcode stickers are placed on the correct boxes (spot check).

Step 5: Put away parts.

1. Operate forklift to pick up pallets.
2. Operate other machines (for moving pallets) to pick up pallets.
3. Identify place that parts are to be stored from put-away document.
4. Operate forklift to store pallets.
5. Place boxes onto pallet.
6. Remove boxes from pallet and place onto storage rack.
7. Place boxes so that bar code faces pickers.
8. Scan the product number bar code and the storage location bar code to note location.
9. Enter the quantity of the product stored in a particular location in the scan gun.
10. Note new storage locations using the computer.
11. Place scan gun in computer cradle in order to transfer put-away orders from the scan gun to the computer.
12. Print out the transferred put-away orders.
13. Check the original packing slips to the transferred put-away orders for accuracy.
14. Send the put-away orders to corporate in order to validate them.

The training itself typically requires a level of detail that is much greater than is supplied by the task statements. This is because learning task performance requires a linking of specific stimulus conditions to specific desired responses. The specific hazards of the medical assistant job are not apparent from the task statements, but they are readily apparent when you watch someone draw blood or collect various specimens for testing. The number and complexity of computer screens that a receiver utilizes to enter inventory and troubleshoot are not well-represented in the task statements. The detail present in the task statements is not sufficient to enable one to use the screens on the job. In the receiver job, one can learn on the job by watching another person go through the computer screens. One could also learn through a workbook or a computer simulation. If training is done off the job, a great deal of care needs to be taken to ensure that the correct stimulus situations are presented to avoid training the wrong behaviors. We delight in retelling the story of a new cashier in a department store who was having difficulty with a credit card purchase. She told her supervisor that she was doing it exactly as she was taught in training. Her supervisor responded, "Forget everything you learned in training and watch me."

Rating scales. It is often the case that training does not encompass the entire job; rather, only some of the job tasks are trained. Training is expensive, and many tasks can be learned quickly on the job with minimal risk. Rating scales are often used to help decide how to treat different tasks. Examples of commonly used rating scales include the following:

- Importance or criticality.
- Consequence of error.
- Difficulty to learn.
- Frequency of occurrence.
- Location of training.

Tasks are good candidates for training if they are important or critical, occur frequently, are difficult to learn, or have a high consequence of error. Police officers do not often shoot guns at work. However, the task is critical and has serious consequences if there are errors, so shooting must be trained. Firefighters rarely fight fires, but that is the essence of their job. A sales job may involve the use of a computer system to book and track orders that are shipped to customers. If the computer system requires some sophistication to use, it will probably be trained prior to actually doing the job. The location of the training may be asked directly (e.g., "Where should the task be learned?").

Selection Versus Training

Both selection and training tend to focus on tasks that are critical, occur frequently, are difficult to learn or perform, or have a high consequence of error. In short, both are concerned with the job's most important tasks. One major distinction between selection and training has to do with the characteristics of the applicant pool. If most applicants have the knowledge or skill, it makes sense to require the skill upon entry rather than to train for the skill. On the other hand, if most of the applicants do not have the knowledge or skill, or if a company is specifically looking for beginners to learn "the company way," then it makes sense to train for the knowledge or skill. This is the reason job analysts try to establish whether it is practical to expect a skill in the applicant pool and whether new workers need the skill right away. An organization's management and culture will often dictate whether the organization prefers to hire applicants who already possess the needed KSAs or whether they prefer to hire eager but "green" applicants. This of course will affect the focus of the job analysis employed. Remember, purpose is paramount when choosing a job analysis method.

Chapter Summary

Staffing

We described links between job analysis and two of the three main staffing functions, recruitment and selection. Job analysis for recruitment must provide information to communicate the gist of the job and its essential requirements to recruiters and applicants. Job analysis for selection must provide the information needed to infer the knowledge, skills, abilities, and other characteristics (KSAOs) required for successful job performance. Several methods (e.g., C-JAM, PAQ, ARS, job elements, critical incident technique) are useful in providing the worker attributes or KSAOs. Once the KSAOs are determined, they can be used to devise or choose tests or other methods to evaluate the applicants. Validation studies can (should) be conducted when feasible to establish the job relatedness of the chosen tests. Job analysis must provide work-oriented (task) information for this purpose. For content-oriented validation strategies, task information generally provides the strongest basis for inferring validity when it is used to judge the appropriateness or relevance of the test's content. KSAO information, although relevant, requires somewhat more-risky judgments that scores will be valid based on test contents alone. For criterion-oriented validation strategies, tasks provide the basis for establishing criteria used in evaluating the empirical relations between test scores and job performance. For construct-oriented validation strategies, job analysis provides the basis for judgments about relations among measures that are theoretically relevant to the focal test and job performance. Judgments about the quality of tests and assessments, along with the organizational decisions made based on such tests and assessments, are improved when there is more supporting information. This has an ancillary benefit of reducing stress levels in test developers and test takers! This suggests that one should pursue a multipronged strategy for determining test effectiveness and may call for a combination of job analysis methods. Of course, practical concerns will influence the choice of strategy.

Training

Training, like selection, considers jobs to be fixed entities. Both training and selection are most concerned with the central or critical tasks of the job. Unlike selection, in which people are chosen to match the job, training uses instruction to change people to fit the job. The training cycle is a process used to make sure that training is effective in achieving organizational goals. The training cycle tailors the training to fit the organization's objectives. In the needs assessment phase, the requirements of the organization and job are compared to the skills available in the potential trainees. Later phases of the training cycle help evaluate the effectiveness of the training. Job analysis provides input to training by specifying the task content of jobs. It also provides information that is useful in deciding which tasks should and should not be trained and, if the task is to be trained, where the training should occur. In addition to task knowledge, training may also enhance KSAs to support task performance.