



SkillCheck Professional Test Validation

| | |
|---|----|
| Statement of Validation | 1 |
| Validation Method..... | 1 |
| Further Validation of SkillCheck Tests | 1 |
| Conclusions..... | 2 |
| Validity and the Use of SkillCheck Tests | 3 |
| Considerations in Choosing a Pre-Employment Test | 3 |
| Definitions of Validity..... | 3 |
| Approaches to Demonstrating Validity | 3 |
| Legal Challenges, Adverse Impact, and Validity | 4 |
| Validation of the SkillCheck Product Line | 6 |
| Integrating SkillCheck Products into Current Selection Procedures | 7 |
| Resources for More Information on Test Validation | 10 |

SkillCheck Professional Test Validation

Statement of Validation

Note: This document was prepared by Linkage, Inc., an internationally respected human resource consulting firm and sponsor of the annual Assessment, Measurement, and Evaluation of Human Performance Conference and Exposition.

SkillCheck Standard Tests, the default tests included in *SkillCheck Professional*, *SkillCheck Professional Plus*, and a number of other SkillCheck computer-based testing systems, are valid (i.e., appropriate and meaningful) simulations of the tasks that employees complete while using the software applications. Linkage, Inc. bases this statement of validity on a series of content-oriented validation studies, that used the input of subject matter experts (SMEs). These SMEs have used the various software applications in a variety of clerical jobs as temporary employees.

The SkillCheck Standard Tests are simulations that we found to be direct representations of the tasks involved in the use of the relevant software application (e.g., Use of Word for Windows, Use of Excel, etc.). In addition, we took steps to ensure that the Standard Tests appropriately represent all of the tasks that are encountered in the typical use of the applications. More critical/frequent tasks are represented by more items and less critical/frequent tasks are represented by fewer items. For example, SMEs reported that six types of tasks (customizing, desktop publishing, editing, file management, formatting, and printing) are encountered when using one of the specific word-processing applications. They further reported that printing tasks are more critical and frequent than customizing tasks. Therefore, Standard Tests for word-processing skills contain more items related to printing than to customizing.

Validation Method

In developing SkillCheck Standard Tests, we ensured content-validity in four steps:

1. SMEs identified the major task clusters in the job element associated with use of the application.
2. SMEs rated the criticality and frequency with which each task cluster is used on the job.
3. SMEs indicated the task cluster(s) that each item represents.
4. Using the data generated by the SMEs, we selected a group of items, which is referred to as the Standard Test.

As a result of this four-step process, the Standard Tests appropriately represent the criticality and frequency of the task clusters within the relevant job element. Because the number of test items in each task cluster is approximately proportional to the rated criticality and frequency of each task cluster, we are confident that the items are a representative sample of the tasks in the job element. Additionally, we selected items that the SMEs rated as critical and frequently used.

Further Validation of SkillCheck Tests

Further information on validation and the approach used to develop this statement of validity is provided on page 3. Additional validation of the SkillCheck product line is planned. In the coming months, Linkage, Inc. will conduct a series of studies to demonstrate the predictive strength of the simulations. These studies will investigate the ability of the tests to predict employee performance.

Further Validation of SkillCheck Tests (continued)

One can develop custom tests for a variety of uses with the test makers provided by SkillCheck. The *Professional Plus TestMaker* allows you to create tests by selecting specific items from a database of 100 or more questions for each software product. The SkillCheck *TestMaker* allows you to create automated text-based tests containing multiple-choice and true-false questions. Tests created using these test-making tools are not necessarily valid. (Items in the software question databases have not been individually validated.) Thus, before any custom test that has been created using the test makers can be described as valid, an employer should take steps to substantiate this description. Pages 8–9 of this document include guidelines for validating customized tests.

Conclusions

A limitation to this statement of validity should be made. Our research indicates that the SkillCheck Standard Tests are representative samples of the task clusters in the jobs performed by our sample of SMEs. Common sense suggests that all software end users will perform a mixture of tasks that is similar to the tasks performed by the individuals in our sample. However, differences, sometimes small differences, can usually be found between jobs in the same or different organizations/offices. While the Standard Tests are valid representations of the job elements performed by the SMEs in our sample, we cannot confidently state that they are a valid representation of the tasks in all jobs. Before claiming that the test is a valid representation of a given job, an employer should ensure that the job elements measured by the test are substantially the same as the elements of the job for which an applicant is being selected.

Validity and the Use of SkillCheck Tests

Many SkillCheck customers inquire about how they can integrate the SkillCheck test into their organization's selection process. To answer this question, this section describes the issues that an HR professional would consider in using the test. Specifically, this section outlines:

- The factors that an employment professional would consider before using the test.
- Test validity, the major approaches to validation, and legal challenges to testing.
- The steps that SkillCheck took to ensure the validity of the Standard Tests.
- How the Standard Tests might be integrated into your pre-employment testing procedure.
- Use of custom tests, as developed with the SkillCheck test-making products.
- Resources that you can turn to for further assistance.

Considerations in Choosing a Pre-Employment Test

When choosing an employee-selection testing procedure, an employer should consider three aspects of the test's use: the validity of the testing process, the possibility of legal challenges and adverse impact, and the cost of testing. As will be indicated shortly, if a test is valid, it can withstand legal challenges and it will be cost-effective. Thus, validity is the most critical concern in considering these three aspects of test use. It is important to note that a test itself cannot be considered valid or invalid; rather it is the use of a test in a specific context that should be considered for purposes of validation. SkillCheck has taken steps to demonstrate the validity of its Standard Tests, but ultimately employers must determine whether the validity studies described below are a useful demonstration of validity for their purposes.

Definition of Validity

Validity is the appropriateness, meaningfulness, and usefulness of the inferences made from test scores, and validation is the process of accumulating evidence to support such inferences. Validity in an employee-selection context refers to the ability of a test to select individuals to perform a specific job. Pre-employment selection testing is the primary focus of most test validations. A pre-employment test is not valid in and of itself; rather a testing process is valid in the context of a specific job. Valid testing allows one to select the best employees, which has a direct impact on profitability. Research has consistently shown that the dollar value of an increase in test validity is many times larger than the cost of even the most expensive tests. Thus, any possible increase in cost associated with a more valid test would be small in comparison to the revenue saved or gained by pre-testing employee proficiency.

Approaches to Demonstrating Validity

Many methods, and combination of methods, can be used to demonstrate validity. We have used a combination of methods to validate the SkillCheck test. Below we will discuss the various validation methods in terms of three strategies: content-oriented validation, criterion-oriented validation, and construct-oriented validation.

Approaches to Demonstrating Validity (continued)

Content-Oriented Validation. A test is content valid when it is a representative sample of the content of the job for which applicants are being selected. That is, a test is said to be content valid if it is known to measure the knowledge and skills required for a job or if it is known to simulate the tasks that make up a job. Typically, a content-oriented validation study will begin with a “job analysis” to identify the tasks that make up a job as well as the knowledge/skills required by a job. Then, on the basis of this job analysis, the test developer will design the test so that the test items match the various parts (knowledge/skills or tasks) of the job. Often, critical knowledge/skills and tasks are represented by more items than those that are less critical.

Criterion-oriented Validation. Criterion-oriented validation consists of demonstrating that there is a useful relationship between test scores and one or more measures of job-relevant performance (criterion). That is, a criterion-valid test is able to predict employee performance on some criterion, and the criterion must be relevant to performance on the job. Typical criteria include measures of job performance, training performance, etc.

To validate a test using a criterion-oriented approach, one might first administer the test to a group of job applicants, then randomly select employees from these applicants, and finally (without considering test performance) review the new hire’s performance on the job. The test is valid if the new hire’s job performance could have been predicted by the pre-employment test. The relationship between test scores and performance is usually summarized as a statistical correlation; a significant correlation is a demonstration of validity. Many variants of criterion-oriented approaches to validation exist. For example, if the test is given to current employees and the test’s results are found to be correlated with the employees’ current job performance, the test can be considered valid. This latter approach is referred to as “concurrent validation” because the employees are tested at the same time as their performance is measured.

Construct-oriented Validation. Construct validation is a demonstration that the test measures the human characteristic (or construct) that it is designed to measure. If the test measures a well-defined construct (and only one construct), it is said to be construct valid.

To demonstrate the construct validity of a test, one must first theoretically define the characteristic (construct) that the test measures. This definition should include a list of measures to which the test should relate as well as measures to which it should not relate. Then, one must conduct a series of studies to determine that the test does indeed relate (or does not relate) to the appropriate measures. No single study can demonstrate the construct validity of a testing procedure. Construct validation is not typically used for demonstrating that a test is a meaningful and useful predictor of performance for a specific employment context. Rather, construct validation is typically used to demonstrate that a psychological test measures the construct that it is designed to test. (For example, a test of intelligence must be related to other tests of intelligence, and must not be related to personality.)

Often the differences among the three approaches (content, criterion, and construct) are arbitrary and indistinct, and the three approaches to validation can be combined. None of the three strategies is perfect, and all must be implemented according to professional standards with an eye toward federal and local case law and statutes. To validate the SkillCheck line of products, we relied largely upon content- and criterion-oriented validation strategies. Before describing the process followed to validate the SkillCheck tests, we will briefly review the interplay between test validity and legal challenges as well as adverse impact.

Legal Challenges, Adverse Impact, and Validity

A testing procedure is not likely to be legally challenged unless it is used to make personnel decisions (e.g., select/not select, promote/not promote, retain/lay-off, etc.).¹ If a test is legally challenged, test validity becomes even more critical to the employer. Since most employment tests are used for personnel decisions, it is worth discussing the cause, conditions, and typical process of legal challenges in this description of test validation. However, before describing the role of test validity in legal challenges, we must first discuss the concept of adverse impact.

Nearly all legal challenges to testing revolve around arguments of discrimination against specific minority groups. When a testing procedure screens out a disproportionate number of individuals of a specific race, color, religion, sex, national origin, age, or physical disability, it is said to adversely impact minorities. Adverse impact is a necessary condition for legal challenges of discrimination. In other words, if there is no evidence of adverse impact, state and federal agencies will not pursue claims of discrimination.

Most people agree that an individual should achieve a score on a pre-employment test that represents his or her ability to do the job. An individual should not be screened out of a job based on his or her membership in a particular minority group. Sometimes, however, individuals from a group do, on average, score lower than the average score for another group. In such cases, it can be difficult to know whether there is something wrong with the test or whether individuals in the group have not had the level of experience necessary to perform well on the job. For example, a test may tend to screen out foreign-born job applicants because, as a group, they have less experience or skill in the areas covered by the test. Alternatively, the test might tend to screen out foreign-born applicants because of factors unrelated to job performance (e.g., their inability to read the instructions in written English where the ability to read English is not a job requirement). Regardless of the cause, if testing does screen out minorities at a rate that is meaningfully greater than that of the majority group, the test is said to have adverse impact.

Just as test validity only makes sense in the context of a specific job, so too adverse impact can only be found in the context of a specific job. A test may adversely impact minorities in one situation and not in another situation. One might expect differences in the average skill level of certain groups, but this does not mean that a test of this skill will adversely impact the specific group being tested. For example, male applicants as a group may perform poorly on tests of typing skills. Then again, fewer, but more highly skilled men might apply for a specific typist position. Thus, an employer might have a test that would tend to screen out males from the general population. However, in the specific testing situation described, the employer's test would not adversely impact men. Unfortunately, adverse impact can happen unintentionally and unexpectedly, so an employment professional must continually monitor test use for adverse impact.

Adverse impact is an adequate justification for legally challenging a selection procedure. If it can be demonstrated that a selection process disproportionally screens out minorities, the employer must demonstrate the test's validity for the job. If the test is valid for a job, then there is a business rationale for use of the test and the employer may continue to use the testing procedure despite the adverse impact. In other situations, when an alternative selection device has a demonstrably equal or higher level of validity than the device currently being used, an employer may be legally mandated to use the alternative selection device.

¹ Tests are not always used for personnel decisions. One might use a SkillCheck test to investigate whether a training program caused employees to use a software program more effectively. In such a situation, one might not even record individual scores, because one is really evaluating the training program (not the employees). In this situation, the decisions made based on the test would not affect a specific individual, and thus legal challenges would be extremely unlikely.

Legal Challenges, Adverse Impact, and Validity (continued)

An HR professional may ask, “Do I need to conduct my own validation in case of legal challenges?” While there is no cut-and-dried answer to this question, validation results from a different job or organization may be used if the jobs (i.e., the current job and the job for which the original validation was conducted) are “substantially” the same. The similarity of the jobs can be investigated by conducting a job analysis in one’s organization and comparing the job’s structure to the job for which the original validation was conducted. Obviously, one must eventually make a judgment of whether the jobs are “substantially” the same, and no guidelines have been presented by the courts or by the federal government to assist in this judgment. Such a judgment would need to consider the general milieu of employee/labor/racial relations, as well as likelihood of adverse impact and legal challenges. The extent to which validation studies can be “transported” between jobs is currently being debated by testing professionals and, while there is no clear consensus, it appears that the trend is toward increasing transportability. Although this trend is likely to be reflected in future court cases, at present there are no clear guidelines for identifying whether each organization must conduct its own independent validation study for each specific job.

Validation of the SkillCheck Product Line

The *SkillCheck Professional* and *SkillCheck Professional Plus* products include Standard Tests that have been demonstrated to be valid (i.e., appropriate and meaningful) simulations of the tasks that our sample of clerical employees generally complete while using software applications. This statement of validity is based upon a series of validation studies. The rationale and process underlying this statement of validity is given below.

In validating the SkillCheck line of products, we have demonstrated that the Standard Tests are appropriate and meaningful simulations of the “job element” associated with the use of the software application. Any given job comprises several job elements. For example, a temporary clerical employee hired to do word processing must perform additional job elements, such as following directions (one must produce documents to certain specifications), prioritizing tasks (one must manage to complete several projects or tasks on schedule), and using a micro-computer environment (e.g., Windows 95). Selection procedures are often designed to select individuals to perform all of the duties that comprise a job, and validation is a demonstration that the test or selection system is a meaningful predictor of overall job performance. In the case of SkillCheck Standard Tests, we have validated only the specific job element that each test is designed to measure. For example, SkillCheck’s Word For Windows Standard Test has been demonstrated to be a valid simulation of the “Use of Word for Windows” job element. It is incumbent upon an employer to identify and test the additional elements that comprise a job. Below, we describe the methods used to validate the SkillCheck tests.

Content Validation of the SkillCheck Tests. The SkillCheck products are samples of tasks involved in the use of the various software applications (e.g., Use of Microsoft Word for Windows). Since each question in a SkillCheck Standard Test is a simulation of the actual tasks that comprise the process involved in working with the application, every test item represents a job task. In addition, each Standard Test has been constructed in such a way as to ensure that the mix of test items matches the actual tasks performed on the job (i.e., the Standard Test samples a variety of the most critical and frequent tasks in the job element).

Content validation is really a description of the process used in constructing the SkillCheck Standard Tests. The process that we followed to develop the content-valid simulation is as follows. First, we identified the task clusters as well as the criticality/frequency of the task clusters that are part of the job element. Then, with the assistance of subject matter experts, we identified the task clusters that each test item represents. Finally, we ensured that the mix of items in the Standard Test represents the criticality

and frequency of the tasks to the job. The technical details followed in the process of content validation are available from your SkillCheck representative.

Validation of the SkillCheck Product Line (continued)

Criterion-Oriented Validation. In the next few months, we will validate SkillCheck's tests using a concurrent strategy. We will demonstrate that each test is related to performance in pre-employment training. Specifically, halfway through such training, we will administer the test to trainees. While the test is being administered, the trainer will rate the individual's performance (i.e., job readiness and performance relative to other program participants). Additionally, later in the training program, a skill simulation will be given to program participants. Both sets of scores, the current performance rating and the post-training skill simulation, will be used to assess the relationship between test performance and performance in training.

Construct-Oriented Validation. Since the SkillCheck tests are simulations of job tasks, not measures of a human construct, we have not attempted to demonstrate construct validity. SkillCheck products are life-like simulations and thus represent the same constructs as required for the job. Although each test likely measures several employee characteristics or constructs, we see no utility in identifying and attempting to measure all of the constructs that the simulation represents. For example, one might propose that the SkillCheck tests, among other things, are related to keyboard skills. However, we did not test such a proposal. Regardless of the skills required, if an individual can perform the tasks that are part of the SkillCheck test, they can perform the tasks that make up the job element.

Integrating SkillCheck Products into Current Selection Procedures

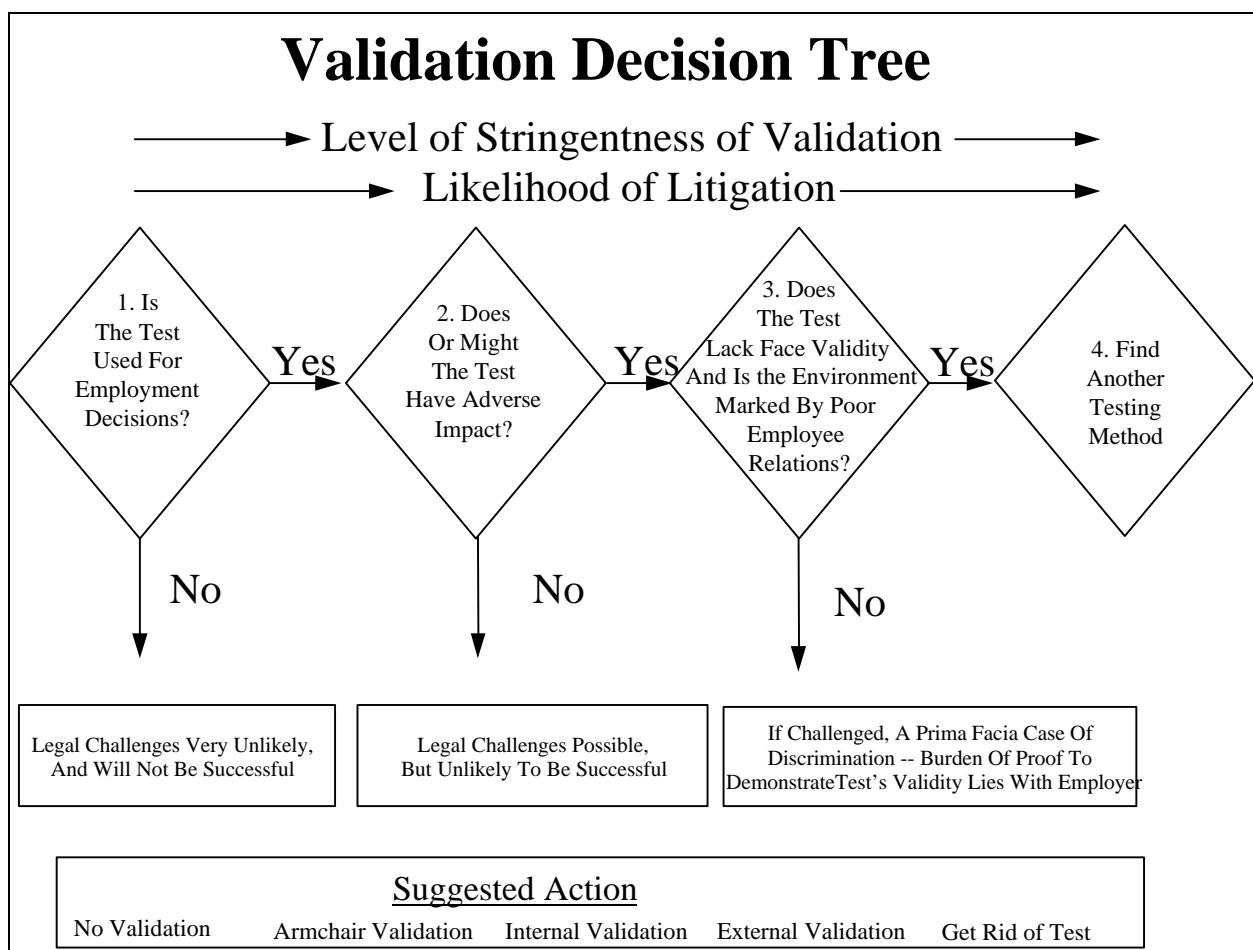
Our research indicates that the SkillCheck Standard Tests are representative samples of the task clusters in the jobs performed by our sample of SMEs. Common sense suggests that the great majority of clerical software end users will perform a mix of tasks that are similar to the tasks performed by the individuals in our sample. Thus, it is likely that the test will be a valid predictor of job performance in many situations. However, without additional information, we cannot state that the Standard Tests will be valid for all specific purposes. The job performed by our subject matter experts (SMEs) may differ from that in another situation. Because of this possibility, an HR professional should determine whether the task clusters represented in the Standard Test are indeed the same as the task clusters in the job element for which he or she is selecting employees. If the job elements are substantially the same, one can use the Standard Test. If the job elements are different, one can develop a custom test using the SkillCheck *Professional Plus TestMaker*. These two options are described in more detail below.

Validation Issues in the Use of the Standard Test. A validation study investigates the usefulness of a test for a specific job. SkillCheck Standard Tests have been demonstrated to be valid because they simulate the job elements that the clerical employees in our sample perform when using various software applications. In general, research suggests that a test that is valid for one job can be valid for other jobs. Thus, the Standard Tests are likely to be a meaningful and useful predictor in your context. However, differences, sometimes small differences, can usually be found between jobs in different organizations/offices. While the Standard Tests are valid representations of the job elements performed by the SMEs in our sample, we cannot confidently state that they are a valid representation of the tasks in all jobs. Before claiming that the test is a valid representation of a given job, an employer should ensure that the job elements that were used to develop the test are substantially the same as the elements of the jobs for which applicants are being selected. That is, if challenged in court, employers must be able to demonstrate that the test is valid in the context of their organization. Because the SkillCheck test is a valid sample of the tasks performed by the SMEs in our focus group, an employer must only demonstrate that the job element that his or her employees perform is substantially the same as the job element performed by our SMEs.

Integrating SkillCheck Products into Current Selection Procedures (continued)

The level of detail and defensibility of a validation should vary according to the audience of the validation report and the context of test use (see Figure 1, the “Validation Decision Tree”). If the test is not being used for personnel decisions, validation is probably not critical. However, if one expects that the test is likely to have adverse impact or if the general organizational context is marked by poor employee relations, you may need a more rigorous validation. If one expects that use of the test will result in adverse impact and legal challenge, it is probably best to conduct your own validation of the test with a qualified consultant. However, if one is confident that the test is a simulation of the tasks required by new employees and one does not expect legal challenges, the validation report, which is available from SkillCheck, should be adequate. Contact your SkillCheck representative for a more complete validation report.

Figure 1: Validation Decision Tree



Validation Issues in the Use of Customized Tests. One of the most useful features of the SkillCheck system is the ability to create one’s own tests (software tests, multiple-choice/true-false tests, typing tests, and data entry tests) to meet one’s own specific needs and the needs of customers. While this capacity allows any number of tests to be developed, a custom test is not necessarily valid. Any test (including custom SkillCheck tests) must go through a process of validation before it can be considered valid.

Integrating SkillCheck Products into Current Selection Procedures (continued)

Of course, there are some situations in which validation is not required (e.g., if the test is not used for personnel decisions). However, in most situations one will want to take some informal steps to ensure validity, and in some situations one may want to conduct a formal and rigorous validation. Such a formal validation process might be conducted by individuals within your organization or through an independent organization (such as SkillCheck, Inc., which offers such services). As with any validation process, the required level of rigor will vary according to the test's use and the context of testing. (Figure 1, "Validation Decision Tree," explains these options.) For example, if the test is likely to have adverse impact, then validation is strongly suggested. A few options for validation of custom tests are described on page 9.

In some situations, there is a low likelihood of litigation and a direct link between test questions and the "job" being measured. In such a situation, one might rely solely upon an "armchair validation." For example, someone may use a custom SkillCheck test to evaluate how well someone has done in a training course for a software product. Using the SkillCheck *Professional Plus TestMaker*, one could select questions that correspond exactly with modules in the course curriculum. In this example, the test is a representation of the training program. Since test questions directly correspond to the training modules, and there is little chance of litigation, there is no need for any separate validation study to validate use of the test in this situation.

In other situations, one would want to conduct an informal validation. An informal validation might be appropriate when there is little likelihood of adverse impact (and thus little likelihood of litigation) but one is making personnel decisions. The goal of such a validation might be to ensure that the test is meaningful in the context of one's organization. In such a situation, one might convene a group of subject matter experts (SMEs) to identify the major tasks of the job (or job element) and then select (or develop) items that represent these critical or frequent tasks. Such steps are not likely to withstand legal challenges, but one would be reasonably sure that the test is measuring important parts of the job.

With regard to creating custom SkillCheck software tests, SkillCheck software test items are simulations of tasks that are part of the job. To demonstrate content validity with a degree of rigor, one must only demonstrate that the contents of these items represent the tasks that are critical to the job or frequently encountered on the job. If one is creating a custom software test for a specific job or a specific client, the key to matching the test with the job is making sure that the test questions accurately reflect the job. SMEs should be included in this process. One might first have SMEs describe the task clusters that make up the job, then have them indicate the criticality and frequency of the tasks, and then finally select items to represent these various aspects of the job.

Validation of Custom Tests Developed with the Multiple Choice/True-False TestMaker. A number of situations arise in which informal validation procedures may be used with the multiple-choice or true-false tests created with the SkillCheck *TestMaker*. An organization may have specific skills it needs to measure (such as company procedures or policies, or use of a specific computer program or piece of machinery). Companies in these situations may have created written tests (such as multiple-choice tests) or other types of procedures (such as verbal or hands-on "question-and-answer" routines that are, in fact, kinds of tests) to screen job applicants. Tools such as the SkillCheck *TestMaker* are available to automate the process of creating simple text-based tests for these custom situations. It is important to reiterate that if these informal testing tools have not gone through a formal validation procedure, the validity of the tests cannot be assured. At the same time, these tests are normally designed by SMEs who both understand the skills they are evaluating and understand the nature of the job in which the skills being tested will be used. In such a situation, if one is not expecting adverse impact or litigation, informal validation procedures may be preferred to expensive and time-consuming professional validation.

Integrating SkillCheck Products into Current Selection Procedures (continued)

Taking the practical steps to complete an informal validation can help prevent one's test from being challenged based on legality or validity. However, it must again be stressed that none of these procedures is a guarantee of test validity. The only way to ensure test validity is to put one's custom test through a validation procedure, such as the content or concurrent validation procedure described elsewhere in this document.

Resources for More Information on Test Validation

Test validation can be quite complex and technical. While the resources listed below can provide some background information on pre-employment testing and validation, there is no substitute for professional training and experience in test validation. If you anticipate any legal difficulties related to the use of your tests, please contact your SkillCheck representative. The resources listed below will provide further information:

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1985). Standards for Educational and Psychological Testing. Washington, DC: American Psychological Association.

Diane Arthur (1994). Workplace Testing: An employers guide to policies and practices. New York: AMACOM.

Robert Gatewood (1990). Human Resource Selection (Second Edition). Chicago: The Dryden Press.

Society for Industrial and Organizational Psychology, Inc. (1987). Principles for the Validation and Use of Personnel Selection Procedures (Third Edition). College Park, MD: author.