DOCUMENT RESUME

ED 373 066 TM 021 786

AUTHOR Roberts, David C.
TITLE Performance-Based Assessment of Software Skills Proficiency: A Demonstration of the "Judd Tests."
PUB DATE May 93
NOTE 7p.; Paper presented at the Annual Meeting of the Research Conference on Human Resource Development (1st, College Station, TX, May 1993).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Competence; Computer Literacy; Computer Oriented Programs; *Computer Software; *Educational Assessment; Evaluation Methods; Human Resources; *Multiple Choice Tests; Occupational Tests; Simulation; *Skills
IDENTIFIERS Concurrent Tests; *Judd Tests; *Performance Based Evaluation

ABSTRACT

The differences between multiple-choice, simulated, and concurrent tests of software-skills proficiency are discussed. For three basic human-resource functions, the advantages of concurrent tests (i.e., those that use the actual application software) include true performance-based assessment, unconstrained response alternatives, and increased job relatedness. The Judd Tests, a recently developed line of software skills proficiency tests that offer assessment in five DOS-platform applications, are described to illustrate the advantages of concurrent testing. These advantages are: (1) concurrent tests are performance based, (2) concurrent tests allow examinees to use all functions normally available in the software application, (3) concurrent tests do not constrain examinees to certain paths of responding, and (4) concurrent tests are defensibly job-related for positions requiring use of computer software. (Author/SLD)

Performance-based Assessment of Software Skills Proficiency:

A Demonstration of the *Judd Tests*

David C. Roberts

The Psychological Corporation

RUNNING HEAD: Performance-based Assessment

Paper presented at the First Annual Research Conference on Human Resource Development.

College Station, Texas. May, 1993

Abstract

The diff rences between multiple-choice, simulated, and concurrent tests of software skills proficiency are discussed. For three basic human resource functions, the advantages of concurrent tests (i.e., those that use the actual application software) include true performance-based assessment, unconstrained response alternatives, and increased job-relatedness. *The Judd Tests*, a line of software skills proficiency tests, are described to illustrate the advantages of concurrent testing.

Performance-based Assessment of Software Skills Proficiency:

A Demonstration of *The Judd Tests*

As a result of increasingly sophisticated technology in the workplace, selection and training of qualified computer users is a growing concern of human resource professionals. Central to this concern is the assessment of computer skills, or those skills necessary to successfully apply the features of hardware or software. Accurate assessment of these skills serves three basic human resource functions:

- *applicant screening* (determination of what skills a prospective employee has acquired)
- *pre-training evaluation* (or diagnosis of skills and prescription of appropriate training)
- *post-training certification* (or assessment of training program effectiveness)

Virtually every job involving the use of a computer necessitates use of software. Software skills proficiency is typically assessed in one of three ways:

- *multiple-choice tests*, which provide on-screen multiple-choice assessment
- *simulations*, which evaluate task performance in a simulation of software
- *concurrent tests*, which evaluate task performance in the actual software application

In methodological terms, multiple-choice tests are knowledge-based assessments, while simulations and concurrent tests are performance-based. That is, multiple-choice tests require examinees to articulate knowledge about performance, while simulations and concurrent tests require examinees to actually perform.

While simulations and concurrent tests are both performance-based, they differ in the degree to which features of the software application are incorporated. Simulations support some software functions, but do not access the actual application; thus, important functions normally available to the user (e.g., on-line help) may not be supported. Concurrent tests, on the other hand, evaluate tasks in the

actual software application, enabling all functions normally available.

Simulations and concurrent tests are further differentiated by the response alternatives available to examinees. In simulations, it is generally the case that two or three "paths" to completion are available and that any deviation from those paths is unacceptable. Because they use the actual software application, concurrent tests do not constrain examinee responses; the examinee may investigate incorrect (and correct) menu choices, make and correct mistakes, and use the on-line help utility of the application without being penalized. The net effect of using a concurrent test is that of having an observer present when the examinee is being tested, then having the observer evaluate the results of examinee responses. In automated testing, that observer is a computer program.

In terms of job-relatedness, concurrent tests are most defensible from a content validity perspective. This is the case because concurrent tests are work samples, in which the major tasks required in a job (e.g., use of a software application) are assembled into a coherent series of test items with standard instructions and scoring procedures. In such tests, the goal is to assess the ability to complete test items, not the level of performance exhibited when completing them. Because they provide a representative sample of behaviors required to successfully use a software application, concurrent tests are virtually identical to observable work behaviors.

To summarize, concurrent software skills tests have a number of advantages over both multiple-choice tests and simulations: 1) concurrent tests are performance-based, asking examinees to perform instead of articulate about performance, 2) concurrent tests allow examinees to use all functions normally available in the software application, 3) concurrent tests do not constrain examinees to certain "paths" of responding, and 4) concurrent tests are defensibly job-related for positions requiring use of computer software.

The *Judd Tests* are a recently developed battery of concurrent software skills tests. *The Judd Test* battery currently offers assessment in five DOS-platform applications (DOS, WordPerfect, Lotus 1-2-3, Paradox, and Microsoft Word), with additional tests (including Windows versions) available in the near future. While each of *The Judd Tests* supports a single software application, all of the tests share a number of features that make them valuable tools for the human resources professional. To illustrate these features, consider *The Judd Test for WordPerfect*:

1. *Test customization:* The Configuration Program allows test administrators to customize the test. The Configuration screen displays each item comprising the test; items contain one or more tasks, each representative of the feature being assessed by the item. The test administrator selects those items which will be administered during the testing session, thus evaluating only those skills which are relevant to the specific job in question. The Configuration Program also allows the test administrator to set-up a number of other test conditions, including time limits, respondent identification, and reporting options.
2. *Complete on-screen administration:* As concurrent assessments, *The Judd Tests* "run in the background" of the actual software application, allowing the examinee access to all functions normally available when using the software. When the test is initiated, the examinee will note that the software application is initiated. All subsequent instructions to the examinee, practice items, and test items appear on-screen. Clear, concise instructions guide the examinee through each task.
3. *Realistic work-related assessment:* The items comprising *The Judd Tests* are tasks that are typically found in the work environment. Throughout the test, the examinee has access to the application's on-line help system, as well as other controls affecting functioning of the test (e.g., Review Directions, Skip Item, etc.).

4. *Performance tracking:* *The Judd Test* monitors all actions/keystrokes used by the examinee during task completion, as well as the time required to complete the task.

5. *Scoring:* Once the examinee completes the test, two types of scores are automatically calculated by *The Judd Test*. *Accuracy* scores represent simply whether or not the examinee completed the task as directed; *Efficiency* scores represent how efficiently the examinee used software features to complete the tasks and are based on number of keystrokes and time. Efficiency score on a task is limited by accuracy score on the task, as efficiency is meaningless unless the task is completed properly.

6. *Multiple reporting levels:* *Judd Test* results can be reported at three levels: *Summary* format details total accuracy and efficiency score for each item, *Condensed* format reports actual and maximum accuracy and efficiency scores for each task, and *Detailed* format provides keystroke-by- keystroke detail for each task.

Given their concurrent nature and scoring method, *The Judd Tests* can be used in all three of the basic human resource functions described previously. Because the tests can be configured for specific jobs (i.e., only those software features used on the job are presented in the test) and thus provide a representative sample of work behaviors, they can be used as defensible, job-related selection tools. *The Judd Tests* can also be used as pre-training assessments to diagnose specific weaknesses in software proficiency; the detailed level of reporting allows focus on the specific areas of software use in which the trainee requires remediation Finally, *The Judd Tests* can serve a post-training evaluation function, allowing both evaluation of training program efficacy and certification of trainee improvement. As concurrent tests like The Judd Tests become more common, it will be interesting to track the criterion-related validity of these products and ascertain whether their performance-based nature and enhanced face validity translate into selection of higher performing employees.